

# **Measurement of charged particle distributions in hadronic Z decays and tuning of QCD models**

## **Dissertation**

zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften

an der  
Naturwissenschaftlichen Fakultät  
der  
Leopold-Franzens-Universität Innsbruck

vorgelegt von

**Mag.rer.nat. Andreas Hörtnagl**

Innsbruck, im Juli 1996

# **Measurement of charged particle distributions in hadronic Z decays and tuning of QCD models**

**Doctoral thesis**

submitted at the  
University of Innsbruck, Austria

Faculty of Natural Sciences

by

**Andreas Hörtnagl**

Innsbruck, July 1996

# Preface

The following chapters describe most parts of the work that I carried out at the high energy physics group in Innsbruck. This group is a collaborator of the *ALEPH* experiment at *LEP/CERN*, which was established with the aim of studying the annihilation of electron-positron pairs into photons or *Z*-particles by observing the characteristics of the resulting final states.

We were especially interested in so-called hadronic final states which are built by decays of these intermediate particles into a quark-antiquark pair, plus a succeeding conversion of these colored objects into colorless hadrons.

The analysis is restricted to the contributions of charged particles to hadronic final states. A detailed look is taken on quantities that describe the appearance of the whole set of charged particles in these final states, and to observables that describe some of the kinematical details of their constituents. Charged particles have the advantage that their momenta can be measured to a very high level of accuracy. As a consequence, systematic errors in our data should be less than for example in the data of neutral particles, and our measurements should be able to give very restrictive conditions to the tuning of models as performed in the second part of this work. Before the measurement is done, some detailed introduction into the used methodology is given. The methods, the obtained results and some discussions about the surrounding are given in part I of this thesis, and in appendix A.

In the second part, the *QCD*-related models from chapter 2 (*JETSET 7.4* in the variant of isotropic and for anisotropic decaying gluons, *ARIADNE 4.05* and *HERWIG 5.8*) are tuned to gain optimal agreement with the observed experimental data. For this purpose, an algorithm for tuning models based on the "maximum likelihood principle" and linear parametrizations of the model predictions - LinFit - is used, after it is tested in some detail. The main problem in this tuning of the models is due to the inclusion of systematic errors in the tuning algorithm. The inclusion of systematic deviations between measured data and the predictions of a given model in particular can cause serious problems. As a consequence, the fit is restricted to a so-called "fitable region".

This work is a continuation and an extension of [A1,92]. The main improvements are with respect to the higher statistics of input data (571800 events

taken in the year 1992), the inclusion of other observables and the consideration of more model parameters. The main message of this work is the comparison between the tuned models and the measured data, carried out both in a graphical and a numeric way. These comparisons could give a guide for the regions in which the models "cry out for" improvement.

The work is split in an introduction (consisting of two chapters), two main parts and two appendices. Both main parts start with introductory chapters and conclude with the measurements results. In the appendices some additional information and explanations are given that are not directly related to our analysis, but are worth noting.

The aim of my work at the high energy physics group at the University of Innsbruck was not only to carry out the measurements in the best way possible, but also to go through the final part of preparing myself for becoming a teacher at high school. Because of that I tried to include some examples to illustrate all the topics discussed and to get used to some kind of graphical explanation. It was always very important for me to have some pictorial impressions about things to be understood. They often led the way to ideas that made an explanation and an understanding of difficult topics possible.

It is my opinion that a illustrative example often says "more than thousand words", and therefore it is really important for me to get used to some kind of technique that could open a door to the art of illustrative explanation. I tried to find such a technique by applying detailed examples, and hope, that these attempts will be beneficial for future explanations to future students, even if they somewhat increased the size of this thesis.

While the major contributions to this thesis are derived from studies and theoretical work that I carried out as a doctoral student during the past three years, I did want to accomplish an extra goal with this final report, that seemed especially important to myself. Current regulations do allow Austrian students to publish in English instead of German, however this practice is by no means common at this time, although I personally feel that it is highly desirable and appropriate for scientific work on the Ph.D. level. When I set out to implement this goal I felt some hope that readers who have a more firm command of the English language than I do have at this time, would be willing to apply a bit of extra tolerance in cases where slight grammatical imperfections might have remained.

A.H.

Innsbruck, July 22, 1996

# Danksagung

In erster Linie gilt mein Dank wieder meinem Großvater Ing. Fritz Steiner, ohne dessen vertrauensvolle und großzügige Unterstützung während des Großteils meiner Studienzeit ich wahrscheinlich nie in die Situation gekommen wäre, eine Danksagung für meine Dissertation schreiben zu dürfen.

Für seine immer freundlichen und aufmunternden Unterstützungen, für die Anstellung als Mitarbeiter in seiner Arbeitsgruppe und vor allem für sein großartiges Vorbild als sehr väterlicher Vorgesetzter möchte ich Univ.-Prof. Dr. Dietmar Kuhn, dem Leiter der Hochenergiephysik-Gruppe am Institut für Experimentalphysik der Universität Innsbruck, sehr herzlich danken.

Die stärkste Motivationsquelle in fachlicher Hinsicht, sowohl in guten als auch in weniger guten Zeiten, waren für mich immer wieder die Gespräche mit Univ.-Doz. Dr. Peter Girtler. Seine Fähigkeit Verständnisschwierigkeiten sehr schnell aufzuspüren waren dabei ebenso wesentlich, wie seine Gabe schwierige Sachverhalte geduldig und anschaulich zu vermitteln. Wenn es mir in Zukunft gelingen sollte, meine Schüler annähernd so zu motivieren, wie diese Gespräche mich motiviert haben, dann wird ihnen die Physik sicher Freude bereiten.

Ich möchte mich auch bei meinem Betreuer, Univ.-Doz. Dr. Gerald Rudolph für alle hilfreichen Unterstützungen bedanken.

Es ist mir auch ein Anliegen, mich an dieser Stelle bei all meinen Freunden die mir wissentlich oder auch unwissentlich, vorsätzlich oder auch nur versehentlich immer wieder weitergeholfen haben, ganz herzlich zu bedanken. Vor allem meiner Freundin Maria Garber, meinen Brüdern und meinen Eltern gilt dieser Dank für alle ihren großen und kleinen Hilfen. Dafür, daß sie sich in guten Zeiten mit mir wohlgeföhlt haben, und dafür, daß sie dazu beigetragen haben, die schlechten Zeiten etwas besser zu machen.



# Contents

Preface.....	1
Danksagung.....	3
Contents.....	5
List of Examples.....	10
Chapter 1. An Introduction to the ALEPH experiment.....	11
1.1. Overview.....	11
1.2. The ALEPH subdetectors.....	13
1.2.1. Tracking subdetectors.....	13
1.2.1.1. VDET (minivertex detector).....	13
1.2.1.2. ITC (inner tracking chamber).....	13
1.2.1.3. TPC (time projection chamber).....	14
1.2.2. Calorimeters.....	15
1.2.2.1. ECAL (electromagnetic calorimeter).....	15
1.2.2.2. HCAL (hadron calorimeter) and MUON (muon detector).....	16
1.3. Trigger System.....	16
1.4. Superconducting coil.....	17
1.5. Luminosity monitors.....	17
1.6. Hadronic ALEPH events.....	18
Chapter 2. An introduction to some QCD models.....	21
2.1. Abstract.....	21
2.2. Common ideas.....	21
2.2.1. Structure of an electron-positron annihilation.....	21
2.3. JETSET (version 7.4).....	24
2.3.1. Parton shower.....	24
2.3.2. Fragmentation.....	27
2.3.2.1. Fragmentation into mesons.....	27
2.3.2.2. Fragmentation into baryons.....	29
2.3.2.3. String fragmentation.....	30
2.3.2.4. Remarks.....	32
2.4. ARIADNE (version 4.05).....	33
2.5. HERWIG (version 5.8).....	34
2.5.1. Parton shower.....	34

2.5.2. Fragmentation.....	35
2.5.2.1. The cluster concept .....	35
2.5.2.2. Cluster decay into hadrons.....	36
2.5.2.3. Remarks.....	37
PART I: Measurement of charged particles distributions .....	39
Chapter 3. Distributions and Methods.....	41
3.1. Abstract.....	41
3.2. Measured quantities .....	41
3.2.1. Event properties.....	42
3.2.2. Single particle properties .....	45
3.3. Unfolding procedures .....	47
3.3.1. Treatment of event properties.....	47
3.3.1.1. Introduction and definitions.....	47
3.3.1.2. Behavior at finite statistics .....	51
3.3.1.3. Unfolding of event properties (matrix method).....	59
3.3.2. Treatment of single particle properties .....	62
3.3.2.1. Calculation of the statistical error.....	62
3.3.2.2. Unfolding of single particle properties (factor method) .....	63
3.4. Tests of the simulation.....	65
3.4.1. Accuracy of the full simulation .....	65
3.4.2. A statistical test of unfolded event properties.....	71
3.5. Measurement .....	74
3.5.1. Hadronic events and critical tracks.....	74
3.5.2. ISR and cut corrections.....	76
3.5.2.1. Detector matrix and statistical widths .....	77
3.5.3. Systematic errors.....	79
Chapter 4. Results I: Distributions of charged particles in hadronic ALEPH events .....	81
4.1. Abstract.....	81
4.2. Event properties .....	82
4.2.1. Sphericity.....	82
4.2.2. Aplanarity .....	83
4.2.3. Planarity .....	84
4.2.4. C-Parameter .....	85
4.2.5. 1-Thrust .....	86
4.2.6. Major .....	87
4.2.7. Minor.....	88



4.2.8. Oblateness .....	89
4.2.9. Heavy jet mass normalized to the visible energy.....	90
4.2.10. Light jet mass normalized to the visible energy.....	91
4.2.11. Mass difference normalized to the visible energy .....	92
4.2.12. Total jet broadening .....	93
4.2.13. Wide jet broadening.....	94
4.2.14. Jet resolution parameter .....	95
4.2.15. Negative logarithm of the resolution parameter.....	96
4.3. Single particle properties.....	97
4.3.1. Rapidity with respect to the sphericity-axis .....	97
4.3.2. In-momentum with respect to the sphericity-axis.....	98
4.3.3. Out-momentum with respect to the sphericity-axis .....	99
4.3.4. Negative logarithm of the normalized momentum.....	100
4.3.5. Normalized particle momentum.....	102
4.3.6. Rapidity with respect to the thrust-axis.....	104
4.3.7. In-momentum with respect to the thrust-axis.....	105
4.3.8. Out-momentum with respect to the thrust-axis .....	106
PART II: Model Tuning.....	107
Chapter 5. Statistical methods for model tuning .....	109
5.1. Abstract.....	109
5.2. Maximum likelihood estimation .....	109
5.3. One-dimensional parametrization .....	112
5.3.1. Motivation for parametrizations .....	112
5.3.2. Parametrization of a one-dimensional linear distribution .....	113
5.3.3. Statistical test of the linear parametrization.....	118
5.3.4. Handling of nonlinear dependence (linear range) .....	120
5.3.5. Real parametrisations .....	123
5.3.5.1. Graphical impressions from JETSET.....	123
5.3.5.2. Quantitative examples.....	126
5.4. Multidimensional parameter fitting.....	128
5.4.1. Overview .....	128
5.4.2. Linear parameter fitting.....	128
5.4.2.1. The general way .....	128
5.4.2.2. Fixing of some parameters .....	129
5.4.3. Nonlinear parameter fitting.....	130
5.5. Fitting with fluctuations in the parametrization .....	133
5.5.1. Introductory remark .....	133

5.5.2. One dimensional parametrization by a polynomial.....	134
5.5.3. Including the coefficients widths.....	135
Chapter 6. LinFit - An algorithm for model tuning.....	143
6.1. Overview.....	143
6.2. Introduction of LinFit.....	143
6.3. Fixing of the linear range.....	144
6.3.1. n in the case of a one-parameter fit .....	145
6.3.2. n in the case of many parameters .....	147
6.4. Handling of systematic errors.....	150
6.4.1. Model parameters in the case of imperfect models .....	150
6.4.2. Types of systematic errors.....	150
6.4.3. Treatment of exactly known systematic errors .....	151
6.4.4. Treatment of approximately known systematic errors.....	152
6.4.5. Restriction to a fitable set of bins.....	154
6.5. Additional Tests of LinFit.....	159
6.5.1. Principle.....	159
6.5.2. Test at high simulation statistics .....	159
6.5.2.1. Fit under measurement conditions.....	160
6.5.3. Test at high data statistics .....	163
Chapter 7. Results II: Tuned models and comparison between models and data.	165
7.1. Abstract.....	165
7.2. Model parameters and test quantities.....	166
7.2.1. JETSET 7.4 with anisotropic gluon decay.....	166
7.2.2. JETSET 7.4 with isotropic gluon decay.....	170
7.2.3. ARIADNE 4.05.....	174
7.2.4. HERWIG 5.8.....	178
7.3. Graphical comparison between tuned models and experimental data .....	182
7.3.1. Event properties.....	182
7.3.2. Single particle properties .....	186
7.4. Concluding remarks .....	188
Summary .....	193
Appendices .....	195
Appendix A. Remarks concerning part I.....	197
1. Remarks about the choice of bins .....	197
2. Comparison between event properties for the 1992 and 1993 runs.....	201
3. Remark about normally distributed random variables.....	204

4. Remarks concerning a model independent detector matrix.....	206
5. Values of fitted parameters that were used to calculate the model bias .....	211
Appendix B. Remarks concerning part II.....	213
1. Sources of the distributions used in LinFit.....	213
2. Decoupling of normally distributed random variables.....	215
3. Higher order parametrization in higher dimensions.....	217
4. Parameter values used in example 6.1.....	224
Bibliography.....	225
Lebenslauf.....	228

# List of Examples

Example 3.1: A simple model scenario.....	49
Example 3.2: The model scenario at finite statistics .....	51
Example 3.3: Error propagation in systems of linear equations.....	55
Example 3.4: Matrix correction in the model scenario .....	60
Example 3.5: Correction matrix and correction factors .....	77
Example 5.1: Statistics at the fishpond .....	110
Example 5.2: Generation of a linear distribution .....	113
Example 5.3: Parametrization of a one dimensional linear distribution .....	117
Example 5.4: Linear range .....	120
Example 5.5: Multi parameter fit.....	131
Example 5.6: Estimation with a fluctuating parametrization.....	138
Example 6.1: Estimation with systematic measurement errors .....	153
Example 6.2: Straightforward fits with imperfect models.....	155
Example A1: Distribution of a linear combination of independent and normally distributed random variables.....	204
Example A2: Statistical test in the case of a QCD model-independent detector matrix .....	206
Example A3: Maximum likelihood and minimum curvature.....	207
Example B1: Higher order fit in more than one dimension.....	217

# Chapter 1.

## An Introduction to the ALEPH experiment

### 1.1. Overview

The aim of this introductory chapter is to give a brief description of the *ALEPH* experiment and its supporting infrastructure. The main source for this summary is [AL,94]. More details can be found in [A4,94], [AL,90] and [AL,95].

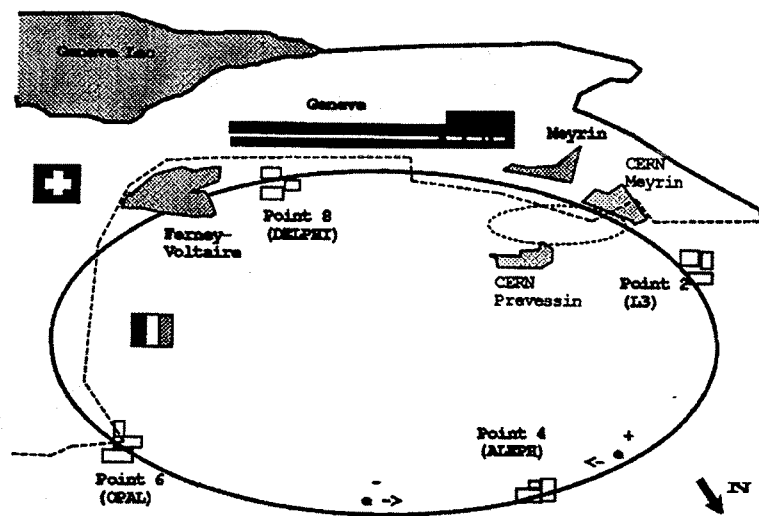


Figure 1.1: The LEP storage ring

*ALEPH* ("a detector for *LEP* physics") is one of the four experiments at the *LEP* storage ring. *LEP* (the "large electron positron storage ring") is situated at *CERN* (the European Laboratory for Particle Physics) in a circular tunnel of  $8.5\text{km}$  diameter, whose depth below ground ranges from  $50$  to  $150\text{m}$  due to the rise and fall of the terrain. A coarse impression is given in the figure above (taken from [AL,94]). The first stage setup of the machine is able to accelerate, store and collide electrons and positrons with a beam energy of up to  $55\text{GeV}$ . At later stages an increase of the c.m. energy up to  $200\text{GeV}$  is planned. The following figure shows a cut-away view of the whole *ALEPH* detector. It is a composition of several independent subdetectors (1 to 8 to be explained later). The detector as a whole is sensitive to almost every known elementary particle.

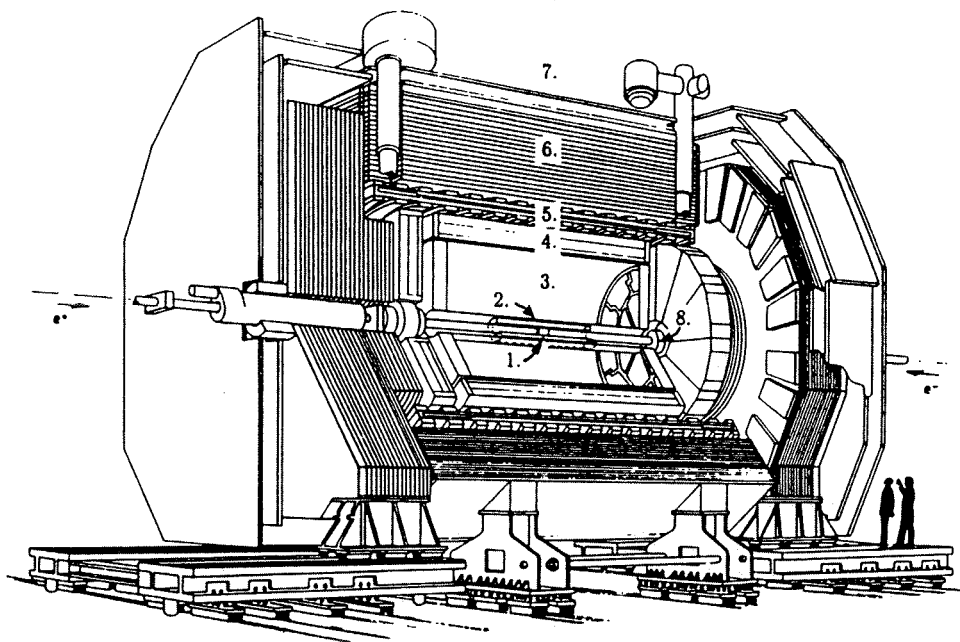


Figure 1.2: The ALEPH detector

- 1 ... Minivertex detector
- 2 ... Inner track chamber
- 3 ... Time projection chamber
- 4 ... Electromagnetic calorimeter
- 5 ... Superconducting magnetic coil
- 6 ... Hadron calorimeter
- 7 ... Muon detection chambers
- 8 ... Luminosity monitors



Figure 1.3: The ALEPH symbol

## 1.2. The ALEPH subdetectors

### 1.2.1. Tracking subdetectors

#### 1.2.1.1. *VDET* (minivertex detector)

This subdetector consists of two layers of silicon wafers that are arranged around the beam pipe in two barrels at radii of  $6.3$  and  $11.0\text{cm}$ . Each wafer has a  $100\mu\text{m}$  strip readout both parallel and perpendicular to the beam direction. The extent along the beam pipe is  $\pm 10\text{cm}$  with respect to the interaction point where electrons and positrons collide. Particles passing through one of the wafers deposit ionization energy, and can therefore be detected by averaging the charge-weighted positions of adjacent strips that have at least three times the mean noise charge. The accuracy of the measured points is about  $12\mu\text{m}$ .

*VDET* hits are used for extrapolating a track that was found by the *ITC* and/or the *TPC* in the direction of the interaction point. By using *VDET* together with the other tracking subdetectors, the spatial coordinates of a given charged particles helix can be measured within about  $30\mu\text{m}$  accuracy.

#### 1.2.1.2. *ITC* (inner tracking chamber)

The *ITC* is a cylindrical multiwire drift chamber, which measures up to eight  $r\phi$  coordinates in the radial region between  $160$  and  $260\text{mm}$ . The drift chamber is filled with gas (a mixture of  $91\%$  Ar and  $9\%$   $\text{CH}_4$ ). A charged particle that is passing through the chamber ionizes this gas along its track and produces pairs of charged particles. The negative electrons are accelerated to the wires. On their way they collide with other molecules of the gas and produce pairs of charged particles again. The resulting avalanche of particles is large enough to be measured as a electric pulse. The time between the beam crossing and rising slope of the pulse in one of the wires is measured. From this quantity the distance between the point where the avalanche started and the wire can be calculated with an accuracy of about  $100\mu\text{m}$ . The coordinate along the beam direction is determined by the difference in arrival time of the signals at each end of the wires. This has a resolution of about  $3\text{cm}$ , and is not used for the standard tracking, but allows the implementation of a three-dimensional first-level track trigger.

### 1.2.1.3. TPC (time projection chamber)

This part of *ALEPH* provides most of the information about charged particle tracks, and is therefore essential for our analysis. It is shown in the following picture, which was taken from [AL,90].

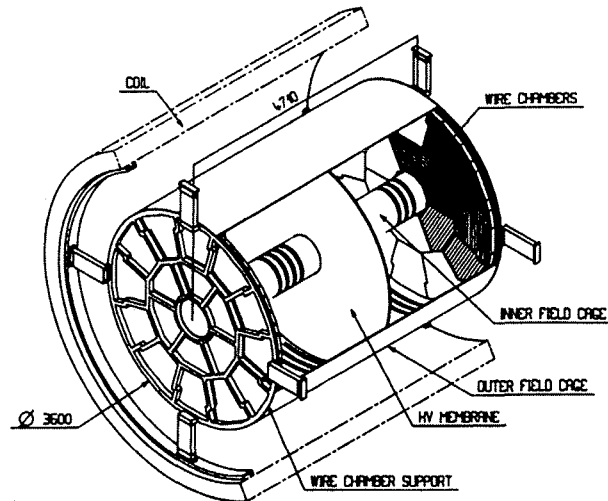


Figure 1.3: TPC overall view

The *TPC* is  $4.7m$  long and extends to a radius of  $1.8m$  measured from the beam pipe. It has magnetic ( $1.5T$ ) and electric ( $115V/cm$ ) fields that are parallel to the beam axis. The electric drift field points from each end plate to the central membrane. As in the *ITC*, a charged particle ionizes the gas (a non-flammable mixture of  $91\%$  Ar and  $9\%$   $CH_4$  at atmospheric pressure) along its track through the chamber. The electrons produced in this ionization drift with a velocity of  $5.2cm/\mu s$  towards one of the end plates, where they induce ionization avalanches in a plane of wire chambers (so-called "sectors"). The ionization density  $dE/dx$  can be measured and used for particle identification. Each end plate consists of 18 sectors.

The *TPC* measures up to 21 three-dimensional points per track. The  $z$ -coordinate is obtained from the drift time and the known drift velocity with about  $800\mu m$  accuracy. The  $r\phi$  coordinate is calculated by interpolating signals induced on cathode pads that are located on the sectors; here the accuracy is about  $180\mu m$ . The  $r$  coordinate is given by the radial position of the pads involved in the measurement. The trajectory of a charged particle inside the *TPC* is a helix, and its projection onto the endplates is an arc of a circle. A measurement of the sagitta



of this arc yields the curvature radius which is proportional to the component of the momentum perpendicular to the magnetic field. Because of that, a measurement of this transverse component of charged particle momenta is possible by using the *TPC*. The measurement can be improved by using information from the *ITC* and the *VDET* to obtain a total accuracy of  $\sigma(1/p_{\perp}) = 0.6 \cdot 10^{-3} \text{ GeV}^{-1}$ .

A problem is the possible presence of positive ions in the drift volume. They are produced near the sense wires of the sectors; if they reach the drift region, they can alter the drift field and cause track distortions. To overcome related problems, the *ALEPH TPC* has a "gating wire grid" situated between the cathode grid at the end plates and the drift region. Normally the gate is closed, that means, the potentials  $V_G \pm \Delta V_G$  are placed on alternate wires of the grid. In the *ALEPH TPC* the values  $V_G \cong -67\text{V}$  and  $\Delta V_G \cong 40\text{V}$  are used, and the resulting dipole fields are sufficient to block the passage of positive ions. The "open" state of the gate is reached by applying the same potential  $V_G$  to each of the wires in the grid. This produces only a parallel addition to the drift field. About  $3\mu\text{s}$  before every bunch crossing, the gate is opened in order to allow electrons to drift into the amplification region of the sections. Only if the first-level trigger is positive, the gate is held open for the maximum drift time of the electrons, which is  $45\mu\text{s}$ .

A laser system is used to provide information on the distortion of particle tracks and to measure the vector of the drift velocity in the *TPC*. Thirty straight ionization tracks are created in the *TPC* by firing two ultraviolet lasers. The measured curvature of these tracks is used to correct the sagitta of particle tracks. The drift velocity is determined from the polar angles and the measured drift times.

## 1.2.2. Calorimeters

### 1.2.2.1. *ECAL* (electromagnetic calorimeter)

The *ECAL* consists of a barrel region which is located inside the magnetic coil surrounding the *TPC*, and two end-caps. Each of these building blocks is divided into 12 modules, each of them is covering an angle of  $30^\circ$ . The modules are built as sandwiches of 45 lead/wire chamber layers. The lead sheets cause electrons, positrons and photons to produce identical showers of particles. The total thickness of the modules is 22 radiation lengths. The energy and position of

each shower is read out via small cathode pads, that are arranged in towers pointing to the interaction point. There are 74000 such towers corresponding to a granularity of  $0.9^\circ \times 0.9^\circ$ . Such a fine granularity in solid angle allows for the distinction of narrowly separated showers. In addition, signals are also available from the wire chambers, providing redundancy in the energy measurement and a low-noise trigger.

*ECAL* can be used to identify electrons, positrons and photons. The energy of photons is measured with an accuracy of about  $18\%/\sqrt{E(\text{GeV})}$ . The efficiency of identifying electrons and positrons is close to 100%.

#### 1.2.2.2. *HCAL* (hadron calorimeter) and *MUON* (muon detector)

The main mechanical support of the *ALEPH* detector is a large iron structure, which serves both as the passive part of the *HCAL* and as the return yoke of the magnet. The magnet iron is instrumented with 23 layers of streamer tubes that are separated by 5cm thick iron slabs. For the barrel region, the total thickness of iron, which forces hadrons to build showers, is 1.2m, corresponding to 7.16 interaction lengths for a hadron passing in a direction perpendicular to these slabs. The energy of the neutral hadrons can be measured with an accuracy of about  $84\%/\sqrt{E(\text{GeV})}$ .

Although hadrons also interact with the *ECAL*, it is only in the *HCAL* where they are fully absorbed by producing showers. *HCAL* is also a part of the muon identification system. Two layers of streamer tubes are installed outside the iron, these layers form the muon chambers. They do not contribute to the measurement of hadronic shower energy, but are used as tracking devices. Each layer reads out two orthogonal coordinates.

### 1.3. Trigger System

The *ALEPH* trigger system is designed to reduce the background (e.g. signals due to beam-gas interactions) to a low level, and to accept all genuine  $e^-e^+$  interactions for disk storage processing and later physics analysis. It is based on three levels:

- The Level-1 trigger initiates the event processing if basic conditions such as *ITC-ECAL* coincidence, or *ITC-HCAL* coincidence are fulfilled.
- The Level-2 trigger checks for the presence of charged particle trajectories that originate from the vertex.
- The Level-3 trigger is applied after data is read out from the detector hardware modules. It is performed entirely in software.

## 1.4. Superconducting coil

The *ALEPH* magnet consists of a superconducting magnet that is cooled by liquid Helium, and an iron yoke. A current of  $5000A$  creates a magnetic field of  $1.5T$ . Its orientation is parallel to the beam pipe. The absolute curvatures of the tracks of charged particles in this magnetic field give their momenta, the signs of the curvatures give the signs of the particles charges.

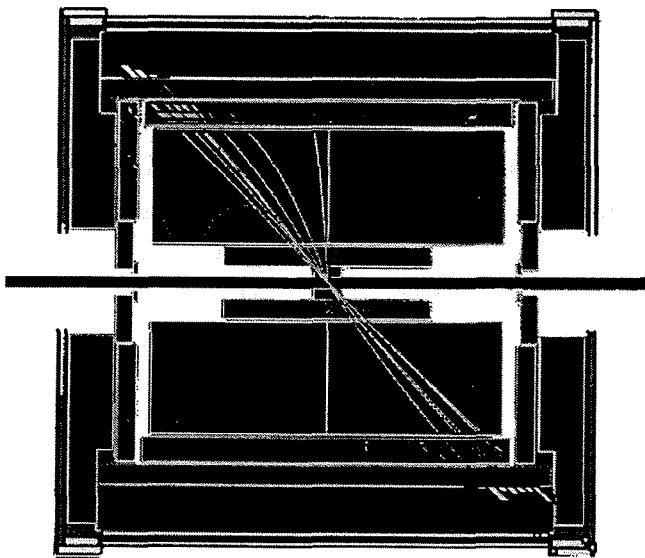
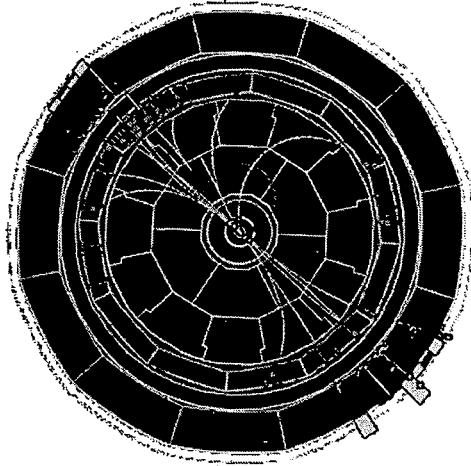
## 1.5. Luminosity monitors

The luminosity  $L := R/\sigma$  has to be known in order to calculate the cross section  $\sigma$  of a given reaction from the measurable rate  $R$  of the same process.  $L$  is a number given by the characteristics of the storage ring. The principle of measuring this quantity is to look at a reaction which is well understood. Usually one looks at Bhabha reactions (elastic scattering of electrons and positrons) which can be measured at small angles with respect to the beam axis.

For example the time-integrated luminosity was about  $22.4pb^{-1}$  in the year 1992. This number corresponds to 687680 collected hadronic events; we used the hadronic cross section at the Z-peak including initial state radiation  $\sigma^{had} = 30.7nb$  to compute this number.

## 1.6. Hadronic ALEPH events

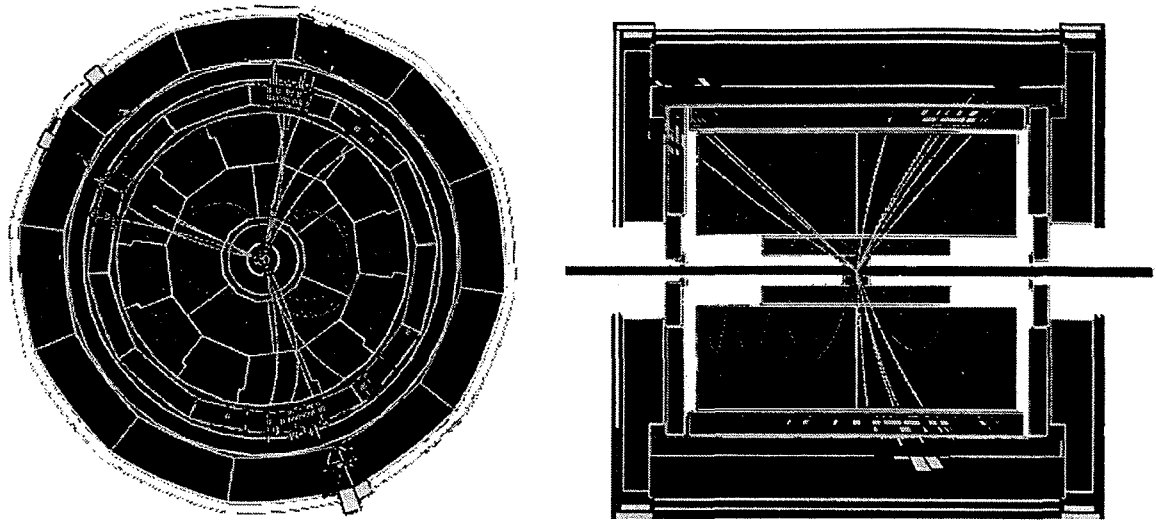
Hadronic events contain on average about 21 charged particles and a similar number of neutrals. They result from the creation of a  $q\bar{q}$  pair by the incoming electron and positron via an intermediate photon or  $Z^0$ . This conversion of a high energetic pair of quarks to a set of hadrons is a very complicated process which is (hopefully) described by QCD. Hopefully means, that we are at the moment not able to calculate details of the reaction  $e^-e^+ \rightarrow \text{hadrons}$  from first principles, and the question if the standard model describes this process in every detail has not yet an answer. Up to now, "only" models that deal with parts of the full solution of this problem, and try to fill the remaining gaps by phenomenological ansatzes are available.



The pictures on this page show two views of the *ALEPH* detector with the tracks of the charged particles in a "two jet event". The straight lines correspond to

charged particles with large momenta, while the strongly curved tracks correspond to particles with very low momentum. In most of the hadronic events the original quark and antiquark hadronize in two sets of particles which are called jets. Jet examples can be seen very clearly in the event illustrated on this page.

Sometimes a very high energetic gluon is radiated by the quark or the antiquark. Because this is also a colored object, it has to hadronize. The result can be a third jet, if the energy of the gluon is big enough, and if it is radiated at wide angle with respect to the momentum of the radiating particle. A very beautiful example of such a three jet event is given in the next figures.



In some cases even more than three jets may be observed. The figures in this section are taken from [AL,95].



## Chapter 2.

# An introduction to some *QCD* models

### 2.1. Abstract

The *QCD* (Quantum Chromo Dynamics) models that are referred to in this chapter lie at the heart of the further analysis, either as a tool for correcting the data or as the object of main interest in the model tuning section. For that reason an introductory chapter seems to be necessary. The main source of the following summary is [L1,89]. A section on common ideas and principles is followed by some sections where each of the three models of interest (*JETSET 7.4*, *ARIADNE 4.05* and *HERWIG 5.8*) is introduced. In these sections all the parameters used for the tuning of the models are discussed in some detail.

### 2.2. Common ideas

#### 2.2.1. Structure of an electron-positron annihilation

The mainly produced particle at the electron positron annihilation at *LEP* is the *Z*-particle. 70% of these particles decay into hadronic final states ([PP, 94]). Today the description of this complicated reaction  $e^-e^+ \rightarrow Z \rightarrow \text{hadrons}$  (or more exactly  $e^-e^+ \rightarrow \gamma^*/Z \rightarrow \text{hadrons}$ ) is done by *QFD* (Quantum Flavor Dynamics or Electroweak Theory) together with *QCD* as far as results are calculable. Up to now it is not possible to calculate expressions for every detail of hadronic final states from first principles. Only parts of the reaction are described up to a given order in some expansion parameter, and the remaining gaps are filled by

phenomenological ansatzes. A schematically overview of the whole reaction is given in the following figure (from [L1,89])

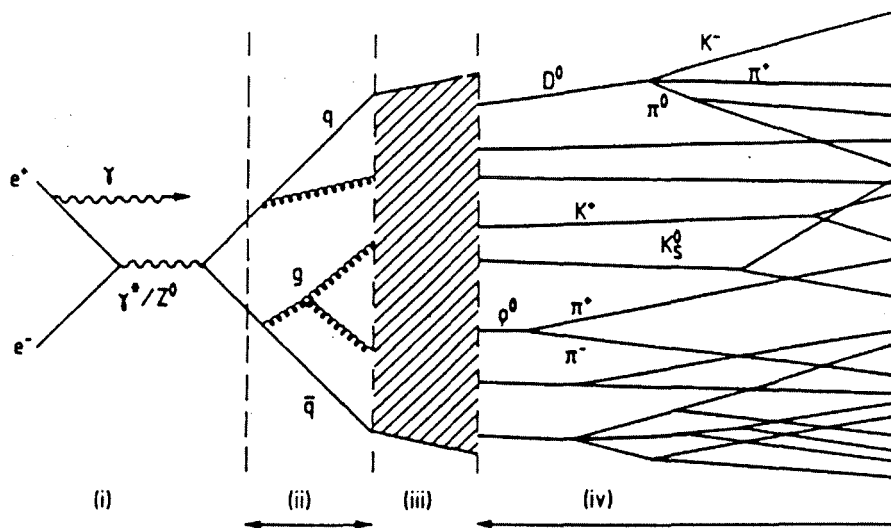


Figure 2.1: Visualization of an electron-positron annihilation event.

The final state emerges after a series of sequential steps as follows

- (i) creation of the primary  $q\bar{q}$  pair by the decay of a  $Z$ -particle or a virtual photon.
- (ii)  $QCD$ -bremsstrahlung and gluon decays
- (iii) fragmentation of quarks and gluons into colorless hadrons
- (iv) decays

In the first step a major contribution to the quantitative predictions comes from  $QFD$ . In addition to the generation of the  $\gamma^*/Z^0$  also a possible initial state bremsstrahlung has to be described. The decay of the  $\gamma^*/Z^0$  into a  $q\bar{q}$ -pair and an eventually appearing radiation of one or more high energetic gluons already belongs to the area that is believed to be described by  $QCD$ , and that is responsible for the observed jet structure of the hadronic final states. Because of the relatively small value of the strong coupling ( $\alpha_s(M_Z) \cong 0.12$ ) due to the very high center of momentum energy of the colliding particles, an expansion of the quantities of interest in a power series seems promising, and can be carried out up



to low orders by the methods of perturbative *QCD* (the expansion parameter is  $\alpha_s(M_Z)/\pi$ ). For the differential cross section full calculations up to  $O(\alpha_s^2)$  are available, for the total cross section the same is true even for  $O(\alpha_s^3)$ .

In the second part, at lower  $Q^2$ , more gluons are created by *QCD*-bremsstrahlung, and some of them decay into  $q\bar{q}$ -pairs. Again it is believed, that a description can in principle be done by the methods of perturbative *QCD*. Because the coupling increases in this part, higher order contributions play a more important role. One possibility that allows for the inclusion of the leading contributions in all orders is the so called "leading logarithm approximation" (*LLA*). Here the leading terms in all orders are taken into account. This kind of approximation makes an iterative ansatz, the *parton shower*, possible, which is used in most of the models for describing this second part. This part ends at the point where the description based upon the *LLA* fails, that means where the contribution of next to leading or higher terms can not be neglected.

In the third part illustrated by figure 2.1, colorless hadrons are built up from the colored partons (quarks and gluons). In principle even this part should be described by *QCD*, but because the mathematical methods fail completely in this highly complicated region, phenomenological *QCD*-inspired ansatzes are used to "jump over the gap". The idea behind these ansatzes are, together with different realizations of the parton shower, the main differences between the different models.

The fourth part describes the decays of unstable particles under the strong, electromagnetic and weak interactions. This is essential for particles with a short lifetime, because their daughters or their daughters' daughters reach the sensitive parts of the detector, and these decay products, not the original particle, are seen in the event. In general an exponential decay law is used together with measured lifetimes and branching ratios to simulate this part of the reaction. The large branching fractions  $\Gamma(Z \rightarrow c\bar{c}, b\bar{b})/\Gamma_{had}$  of 17% and 22% respectively, makes a reliable simulation of charm and bottom hadrons very important.

## 2.3. JETSET (version 7.4)

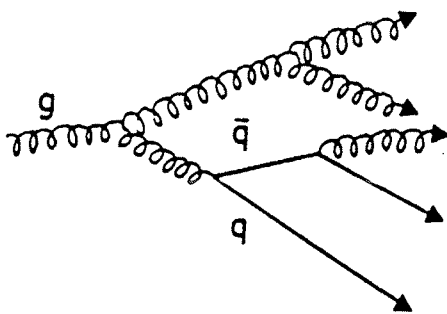
(JETSET: The Lund Monte Carlo for Jet Fragmentation and  $e^+e^-$  Physics)

Characteristics: *JETSET* is a model based upon a coherent parton shower and the concept of string fragmentation. It turns out, that this model is able to fit most parts of the data really well (c.f. chapter 7). This bright advantage is darkened by the fact that the fragmentation part is a function of lots of model parameters, and so this part of the model is a kind of *QCD*-motivated parametrization rather than a model based on calculations from first principles.

The most important of the model parameters are described below.

### 2.3.1. Parton shower

*LLA* plus some simplifications in kinematical variables serve as starting points. These simplifications lead to a limited predictive power for wide-angle parton emission. Consequently a parton shower is believed to give a good description of the substructure of jets, while it is too inaccurate to determine the



global event topology. The figure left (taken from [L1,89]) illustrates how a parton shower is built up from a sequence of branchings. It is possible due to the usage of the *LLA* approximation, to formulate this sequence in an iterative way. Possible types of branchings are  $q \rightarrow qg$ ,  $g \rightarrow gg$  and  $g \rightarrow q\bar{q}$ .

Every branching is characterized by the quantities  $Q$  and  $z$  and an additional azimuth angle that gives the orientation of the plane in which the momenta of the daughters (the products of the branching) lie. This angle is assumed as equally distributed in the *JETSET*-version of "isotropic decaying gluons". In the default option, of "anisotropic decaying gluons", some effects of gluon polarization plus interference with neighboring gluons are included. This second version correlates the production and the decay

plane of gluons. While the branching  $g \rightarrow gg$  tends to take place in the production plane of the gluon, a decay out of the plane is favored for  $g \rightarrow q\bar{q}$ .

$z$  describes the sharing of the four momentum between the daughters  $b$  and  $c$  in the reaction  $a \rightarrow bc$ . In all models  $Q$  has the dimension of a mass (in natural units) and is closely related to the virtual mass of the decaying particle  $a$  or to the transverse momentum of the daughters. The exact definition of this quantity marks one of the differences between the parton showers as used in different models. *JETSET* uses  $Q^2 := m_a^2$ . The quantity  $Q$  represents the evolution parameter of the Parton Shower  $t := \ln(Q^2/\Lambda^2)$ . (The "characteristic scale  $\Lambda$  of QCD" is described later.) In *JETSET* knowledge about the evolution parameter at the decay of a parton corresponds to a knowledge about the mass of the decaying particle.

Even if the evolution parameter  $t$  is an analogy to time, it is really not the same. An ordering in  $t$  (or  $Q$ ) does, in *JETSET*, mean an ordering with respect to the particle masses, and the same ordering does in general not hold in time<sup>1</sup>.  $t_{\max}$  is the starting value of the evolution variable, and corresponds to the  $Q$  value of the very first parton in the shower. The decay probability is set to zero as soon as some lower limit  $t_0$  is reached. Because of the relations of the  $Q$ -value to the mass of the decaying particle, only particles with a mass above this lower limit are able to decay. The  $Q$ -value belonging to the minimal evolution parameter  $t_0$  is one of the free model parameters<sup>2</sup>.

The probability that a branching  $a \rightarrow bc$  happens at a given value  $t$  of the evolution parameter in a small interval  $dt$  around  $t$ , and with a value of the four-momentum sharing in a small interval  $dz$  around  $z$  is given by the *Altarelli Parisi equations*

$$dP_{a \rightarrow bc} = \frac{\alpha_s(\mu^2)}{2\pi} P_{a \rightarrow bc}(z) dt dz$$

The *running coupling* is in first order given by:

$$\alpha_s(\mu^2) = \frac{12\pi}{(33 - 2n_f) \ln\left(\frac{\mu^2}{\Lambda^2}\right)}$$

<sup>1</sup> A similar statement is also true for the other models.

<sup>2</sup>  $Q_0$  is the parameter *PARJ(82)*. Default=1.0GeV "cut-off for the parton shower".

Here  $\mu$  denotes a typical energy scale, which is taken as an approximation of the transverse momentum squared of the branching ( $\mu^2 := p_{\perp}^2 \cong Q^2 z(1-z)$ ). The condition  $Q_0 > 2\Lambda$  avoids the divergence of  $\alpha_s$ . (In JETSET  $Q_0 > 2.2\Lambda$  is used.)

In the expression above  $n_f = 5$  is the number of quark flavors contributing at LEP energies, and  $\Lambda$  is a model parameter that has to be fixed by tuning the model to the data<sup>3</sup>. Roughly speaking,  $\Lambda$  controls the rate of gluon emission. The Altarelli-Parisi splitting kernels are

$$p_{q \rightarrow qg}(z) = C_F \frac{1+z^2}{1-z}, \quad p_{g \rightarrow gg}(z) = N_F \frac{(1-z(1-z))^2}{z(1-z)},$$

$$p_{g \rightarrow q\bar{q}}(z) = T_R (z^2 + (1-z)^2)$$

with  $C_F = 4/3$ ,  $N_C = 3$  and  $T_R = n_f/2$ .

With this input, the probability for a parton branching in an interval  $dt$  around  $t_{decay}$  "after" a value  $t$  of the evolution parameter where it is known to be undecayed can be computed as follows:

$$dP_a = \exp \left[ - \int_{t_{decay}}^t dt' \int dz \sum_{b,c} \frac{\alpha_s(\mu^2)}{2\pi} p_{a \rightarrow bc}(z) \right] \left( \int dz \sum_{b,c} \frac{\alpha_s(\mu_{decay}^2)}{2\pi} p_{a \rightarrow bc}(z) \right) dt$$

The first factor describes the probability that the branching did not happen between  $t$  and  $t_{decay}$ , while the second factor represents the probability that the decay happens in this interval  $dt$ . The values of the evolution parameter are generated according to this distribution starting at the maximum value  $t_{max}$ . After this step the masses of the decaying particles are known. The values of  $z$  are generated by the distributions given by the Altarelli-Parisi splitting kernels. Following the generation of an azimuth angle according to the used option of isotropic or non-isotropic gluon decay, each of the branchings is fixed. The flavor of the quarks is restricted by the phase space.

At the branching of the two initial partons an algorithm is used to match on to the first order three-jet matrix elements.

---

<sup>3</sup>  $\Lambda$  is the parameter PARJ(81). Default=0.29GeV "characteristic scale of QCD"

So far the parton shower is termed as "*conventional shower*". In this state it is still possible that "late" branchings result in partons with large emission angles. Studies beyond the leading-log level show that this is not correct. It turns out that destructive interference effects are large in the region of non-ordered emission angles. As an approximation, these so-called "*coherence effects*" can be taken into account in parton shower programs by requiring a strict ordering in terms of decreasing emission angles. Models that provide such an ordering are called "*coherent models*".

## 2.3.2. Fragmentation

### 2.3.2.1. Fragmentation into mesons

The key to every fragmentation algorithm is again an ansatz that can be used in an iterative way. In addition to an initial quark  $q_1$ , a mechanism that allows for the production of new quark-antiquark pairs  $q_2\bar{q}_2$  is used. In the case of *JETSET* this mechanism is called *string fragmentation mechanism* (it will be discussed later). The new antiquark  $\bar{q}_2$  and the old quark  $q_1$  build up a meson  $q_1\bar{q}_2$ , and the new quark  $q_2$  is ready to take the position of the initial quark  $q_1$ . The proportion of the three quark flavors used in the fragmentation is given by  $u\bar{u}:d\bar{d}:s\bar{s} = 1:1:\gamma_s$ , the production of flavors with higher mass is restricted to the parton shower process<sup>4</sup>. The program accounts for the multiplets given in the following table.  $L$  is the orbital angular momentum,  $S$  the spin of the quark-antiquark system and  $J$  the resulting angular momentum of the meson.

$S$	$L$	$J$	<i>name</i>
0	0	0	<i>pseudo scalar meson multiplet</i>
0	1	1	<i>first axial vector meson multiplet</i>
1	0	1	<i>vector meson multiplet</i>
1	1	0	<i>scalar meson multiplet</i>
1	1	1	<i>second axial vector meson multiplet</i>
1	1	2	<i>tensor meson multiplet</i>

Table 1.1: Meson multiplets in *JETSET*

<sup>4</sup>  $\gamma_s$  is the parameter *PARJ(2)*. Default=0.3 "*s quark suppression*"

First of all a decision has to be made on whether the spin of the  $q\bar{q}$ -system is 0 or 1. There are three parameters that account for different possibilities with  $S=1$ : systems built up from  $u$ - and  $d$ -quarks<sup>5</sup>, systems which contain an  $s$ -quark or are built up from  $s$ -quarks only<sup>6</sup>, and systems which contain  $c$ - and  $b$ -quarks<sup>7</sup>.

The default values for these parameters are motivated by  $q\bar{q}$ -systems with  $L=0$ . Their proportions are estimated by looking at the spin part only. The number of states for  $S=1$  is three times the number of states for  $S=0$ , and therefore a system with  $S=1$  should appear with a probability of  $3/(1+3)=0.75$ . In contrast to this "*spin counting argument*", mesons with  $S=1$  are heavier and should therefore be suppressed by phase space. This difference in mass due to a spin excitation should appear mostly at mesons which are built up from very light quarks. The heavier the constituents, the smaller the mass contribution of the spin excitation. This explains the decreasing values of the default values for decreasing quark masses.

After the value of  $S$  is fixed,  $L$  will be assigned. If  $S=0$  two multiplets are possible; the multiplet of pseudo scalar mesons, and the first multiplet of axial vector mesons. There is one model parameter to regulate the probability for ending up with a meson belonging to the latter<sup>8</sup>. For  $S=1$ , four multiplets are possible. There are three model parameters which determine the probabilities for obtaining with mesons belonging to one of these four multiplets<sup>(9,10,11)</sup>.

The probabilities for producing mesons that belong to one of the multiplets included in *JETSET* are therefore:

$$P_{L=0 \wedge J=0}^{S=0} = \{1 - PARJ(i)\} \{1 - PARJ(14)\}$$

$$P_{L=1 \wedge J=1}^{S=0} = \{1 - PARJ(i)\} PARJ(14)$$

$$P_{L=0 \wedge J=1}^{S=1} = PARJ(i) \{1 - PARJ(15) - PARJ(16) - PARJ(17)\}$$

---

<sup>5</sup>  $P_{u,d}^{S=1}$  is the parameter  $PARJ(11)$ . Default=0.5 "*S=1 probability for a light meson*"

<sup>6</sup>  $P_s^{S=1}$  is the parameter  $PARJ(12)$ . Default=0.6 "*S=1 probability for strange mesons*"

<sup>7</sup>  $P_{c,b}^{S=1}$  is the parameter  $PARJ(13)$ . Default=0.75 "*S=1 probability for higher mesons*"

<sup>8</sup>  $P_{L=1}^{S=0}$  is the parameter  $PARJ(14)$ . Default=0.0

<sup>9</sup>  $P_{L=1 \wedge J=0}^{S=1}$  is the parameter  $PARJ(15)$ . Default=0.0

<sup>10</sup>  $P_{L=1 \wedge J=1}^{S=1}$  is the parameter  $PARJ(16)$ . Default=0.0

<sup>11</sup>  $P_{L=1 \wedge J=2}^{S=1}$  is the parameter  $PARJ(17)$ . Default=0.0

$$P_{L=1 \wedge J=0}^{S=1} = PARJ(i)PARJ(15)$$

$$P_{L=1 \wedge J=1}^{S=1} = PARJ(i)PARJ(16)$$

$$P_{L=1 \wedge J=2}^{S=1} = PARJ(i)PARJ(17)$$

The index  $i$  is used as an abbreviation to stand for  $11$  if the meson is a light one,  $12$  if it is an  $s$ -meson and  $13$  if it is a higher meson. It should be noted that the probabilities for the value of the angular momentum  $J$  (described by  $PARJ(14)$  to  $PARJ(17)$ ) are assumed to be independent of flavor in *JETSET*.

For the flavor-diagonal meson states  $u\bar{u}$ ,  $d\bar{d}$  and  $s\bar{s}$  mixing into physical mesons is included. The  $\eta'$  particles can be suppressed by the factor  $\bar{\eta}'$ <sup>12</sup>.

### 2.3.2.2. Fragmentation into baryons

The parameter  $MSTJ(12)$  allows for a choice of the mechanism for the production of baryons. We use  $MSTJ(12)=3$  (the diquark mechanism), with the possibility of the diquark to be split according to the popcorn scheme. Additionally the production of first rank baryons is suppressed by a given factor<sup>13</sup>. We leave the parameters concerning the popcorn mechanism on their default values, because the data we used for tuning *JETSET* turned out to be insensitive to this parameter.

Baryon production can be achieved by assuming that any flavor represents either a quark or an anti-diquark in a color triplet state. Three parameters are included to cover this phenomenon. The first one is the probability that a  $\bar{q}q$  is produced instead of a  $\bar{q}q$ <sup>14</sup>. The second one is to suppress the production of  $\bar{q}q$  containing  $s$ -(anti)quarks<sup>15</sup>. Only the ground state baryons with  $L=0$  are taken into account. The third parameter decides whether the baryon belongs to the  $J=1/2$  octet or the  $J=3/2$  decuplet<sup>16</sup>.

---

<sup>12</sup>  $\bar{\eta}'$  is the parameter  $PARJ(26)$ . Default=0.4 "Extra  $\eta'$  suppression"

<sup>13</sup>  $\bar{b}_1$  is the parameter  $PARJ(19)$ . Default=1.0 "leading baryon suppression"

<sup>14</sup>  $q\bar{q}/q$  is the parameter  $PARJ(1)$ . Default=0.1 "diquark-antidiquark suppression"

<sup>15</sup>  $(su)/(du)$  is the parameter  $PARJ(3)$ . Default=0.4 "extra strange diquark suppression"

<sup>16</sup>  $(s1)/(s0)$  is the parameter  $PARJ(4)$ . Default=0.05 "spin 1 diquark suppression"

### 2.3.2.3. String fragmentation

Until now we have only mentioned that some mechanism for creating  $q\bar{q}$ -pairs is required in the fragmentation procedure, but the idea behind this mechanism has not been explained. In *JETSET* the so-called *string fragmentation mechanism* is used. This algorithm is very complicated if gluons are radiated from the initial  $q\bar{q}$ -system. A restriction to the case of a pure  $q\bar{q}$ -system makes the discussion rather easy, and because even this simplest case shows the main part of the principle behind string fragmentation, it will be used as a "playground" for the following explanations.

The quark and the antiquark interact via a color field, which can be represented by a 1-dimensional string between them. Consequently the effect of the corresponding string forces can be described by a potential which increases linearly with increasing separation of the quark and the antiquark. This increase of the potential with an increasing distance is the reason for the *confinement* of quarks into hadrons.

If the quark and the antiquark separate, and the distance between them grows, the energy in the color field between them increases too. If this energy exceeds some threshold the possibility for creating a new quark-antiquark pair  $q'\bar{q}'$  is given. After a production of this type we can observe two color neutral systems:  $q\bar{q}'$  and  $q'\bar{q}$  (a remaining string, and a meson) in spite of the initial one  $q\bar{q}$  which no longer interact. If the mass of the remaining string exceeds some threshold<sup>17</sup>, a further production of  $q\bar{q}$ -pairs is possible. If not, the fragmentation is stopped.

The assumption is made, that the probability for creating the pair  $q'\bar{q}'$  is proportional to

$$\exp\left[-\frac{m_{\perp,q}^2}{\sigma^2}\right] = \exp\left[-\frac{m_q^2}{\sigma^2}\right] \exp\left[-\frac{p_{\perp,q}^2}{\sigma^2}\right]$$

Here  $m$  is the rest mass of the produced quark, and  $p_{\perp,q}$  is the transverse momentum of the new quark and antiquark. From this follows, that heavy quarks are suppressed, and the densities for  $m_q$  and  $p_{\perp,q}$  factorize. There is a restriction of heavy quark production to the parton shower and these heavy quarks are actually

---

<sup>17</sup> *PARJ(33)* (Default=0.8GeV) and quark masses are used to define the remaining energy below which the fragmentation of a jet system is stopped and two final hadrons are formed.



not produced by the string fragmentation mechanism. The width  $\sigma$  of this distribution is again a parameter that can be used for tuning *JETSET* to the data<sup>18</sup>.

The longitudinal momentum of the hadron has to be calculated next. In *JETSET* the distribution of the variable  $z$  is given.  $z(E + p_L)$  is the fraction of  $E + p_L$  of the quark (or the antiquark) that goes to the new hadron. In *JETSET* this distribution is given by the "*Lund symmetric fragmentation function*":

$$f(z) \propto \frac{(1-z)^a}{z} \exp\left[-\frac{bm_{\perp}^2}{z}\right], \quad m_{\perp}^2 = m^2 + p_{\perp}^2$$

The parameters  $a$  and  $b$  (they are assumed as flavor independent) determine the shape of the fragmentation function<sup>19</sup>. Here an increase in  $b$  leads to a distribution that favors higher values of  $z$  and therefore leads to a harder hadron spectrum, where  $m$  is the mass of the new hadron.

Since the Lund fragmentation function produces a spectrum for B mesons somewhat harder than observed in real data, the *Peterson et al. fragmentation function* (or *SLAC-formula*) is used in the case of  $c$ - and  $b$ -quarks<sup>20</sup>.

$$f(z) \propto \frac{1}{z \left(1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z}\right)^2}$$

Here  $\epsilon_Q$ ,  $Q \in \{c, b\}$  are two parameters that allow for the tuning of *JETSET* to  $c$ - and  $b$ -data<sup>21</sup>.

---

<sup>18</sup>  $\sigma$  is the parameter *PARJ(21)*. Default=0.35GeV "*transverse momentum width in fragmentation*"

<sup>19</sup>  $a$  is the parameter *PARJ(41)*. Default=0.5 "*First parameter in the Lund string fragmentation function*",  $b$  is the parameter *PARJ(42)*. Default=0.9 GeV<sup>-2</sup> "*Second parameter in the Lund string fragmentation function*"

<sup>20</sup> Different methods exist for changing the fragmentation function. The parameter *MSTJ(11)* is used to switch between this possibilities. We are using *MSTJ(11)*=3, that is the "*hybrid scheme*", where light flavors are treated with the Lund fragmentation function where heavier are described by the Peterson et.al. formula.

<sup>21</sup>  $\epsilon_c$  is the minus the parameter *PARJ(54)*; default of *PARJ(54)*=-0.05;  $\epsilon_b$  is minus the parameter *PARJ(55)*; default of *PARJ(55)*=-0.005.

### 2.3.2.4. Remarks

Not all of the parameters available in *JETSET* have been used to tune the model. The parameters of the heavy flavor (c,b) sector have been fixed before the actual fitting using recent data, since they are almost independent of other parameters. The following assumptions were used:

- (i)  $PARJ(17):PARJ(16):PARJ(15) = 5:3:1$  ("Spin counting")
- (ii)  $PARJ(14)=PARJ(16)$

This assumption leaves only one free parameter, which we choose as  $PARJ(17)$ . We fix it to the value of  $0.2$  to allow for a description of the ratio  $\Gamma(b \rightarrow B^{**})/\Gamma(b \rightarrow B)$  as taken from [A4,95]. The parameter  $a$  is fixed at the value  $0.4$  because it was observed that this restriction of the model parameter does not decrease the quality of the model. In addition  $P_{c,b}^{S=1} := 0.65$  is used as a compromise in the description of  $D^*/D$  (taken from [A1,93]) and  $B^*/B$  production rates (taken from [A4,95]).  $\epsilon_c := 0.04$  is used to get an appropriate description of the  $x_E := 2E/E_{e^+e^-}$  distribution of the  $D^*$  in the region  $x_E > 0.5$ , and  $\epsilon_b := 0.004$  reproduces the mean value  $\langle x_E \rangle$  of  $B$ -mesons.  $\bar{\eta}' := 0.275$  is used to describe the rate of  $\eta'$ -particles. (All quoted values were taken from [Ru,95])

## 2.4. ARIADNE (version 4.05)

(*ARIADNE*: A Monte Carlo for QCD Cascades in the Color Dipole Formulation)

**Characteristics:** *ARIADNE* is a coherent QCD model, which uses the concept of radiating color dipoles. The fragmentation part is the same as in *JETSET*, i.e. the color string fragmentation. Again we have lots of parameters, and again we can observe a good agreement between model predictions and the data.

The QCD cascade is realized by color dipoles (for example the primary  $q\bar{q}$  pair which was created in the  $e^-e^+$  annihilation), these are able to do the following reactions:

$$q\bar{q} \rightarrow q\bar{q}g, \quad qg \rightarrow qgg, \quad gg \rightarrow ggg$$

If a gluon is emitted, the system decays into two independent parts, which are assumed as independent. For example the reaction  $q\bar{q} \rightarrow q\bar{q}g$  ends with the two dipoles  $qg$  and  $\bar{q}g$ . These new dipoles can also radiate gluons and therefore the way to an iterative sequence of branchings is open. Expressions are known that are analogous to the Altarelli-Parisi splitting kernels for the reactions mentioned above. On that basis a parton shower analogous to *JETSET*'s can be generated. The expressions for the individual emissions are based on first order QCD.

In contrast to *JETSET*,  $Q^2 := p_{\perp}^2$  is used in *ARIADNE* as ordering variable, and the two parameters that fix the behavior of the parton shower are "the characteristic scale of QCD"  $\Lambda$  and the transversal momentum  $p_{\perp}^{\min}$  at which the cascade stops. A decrease in the evolution variable does therefore mean a decrease in transverse momenta, and angular ordering is automatically included in the description.

## 2.5. HERWIG (version 5.8)

(HERWIG: Hadron Emission Reaction With Interfering Gluons)

Characteristics: *HERWIG* is a model based on a coherent parton shower that is analogous to the one used in *JETSET*, but that uses a different evolution variable. The fragmentation part of this model is based on the principle of "cluster decays". *HERWIG* does not reproduce the data as well as either *JETSET* or *ARIADNE*. In this work we are using *HERWIG* version 5.8. A really nasty feature of this model is the incomplete and partly incorrect table of particle decay modes. Better agreement with published values are expected for version 5.9, which is unfortunately not available while writing this thesis.

### 2.5.1. Parton shower

The parton shower is done in the same way as discussed in the *JETSET* section, but here the evolution variable

$$Q^2 = E_a^2 \xi_{bc}, \quad \xi_{bc} := \frac{p_b p_c}{E_b E_c}$$

is used, where  $p_b$  and  $p_c$  are the four momenta of the daughters in the reaction  $a \rightarrow bc$ . For the case  $E_{b,c}^2 \gg p_{b,c}^2 = m_{b,c}^2$ , that means for the case where the parton masses can be neglected, the  $\xi_{bc}$  in this evolution variable can be approximated well by  $\xi_{bc} \cong 1 - \cos \vartheta_{bc}$ , and a cascade with decreasing values of the evolution variable corresponds to a series of decreasing angles between the momenta of the daughters. Therefore, angular ordering is included in the parton shower, and *HERWIG* is a coherent QCD model.

Again the shower is stopped as soon as the value of the evolution parameter drops beneath a given limit <sup>22</sup>.

---

<sup>22</sup>  $M_g$  is the parameter *RMASS(13)*. Default=0.75GeV "Gluon effective mass". The condition  $M_g > 4\Lambda - 0.1$  must be fulfilled to avoid divergences in the parton shower.

$\Lambda$  is the parameter *QCDLAM*. Default=0.18GeV

## 2.5.2. Fragmentation

### 2.5.2.1. The cluster concept

In a first step all the double-colored gluons are split up into single-colored quarks and antiquarks.

The "effective gluon mass" is constrained to a value that is bigger than twice the mass of the lightest quark ( $2 \times 0.32 \text{ GeV}$ ), in order to allow for the splitting procedure mentioned above. After splitting, every parton is a single-colored object. Now a definition of a distance between two particles is made, and every quark is combined with its closest antiquark with the corresponding anti-color. Together they form a colorless "cluster". A condition according to the mass of this cluster is introduced as:

$$M_0^n \leq M_{\max}^n + (m_{q_1} + m_{q_2})^n$$

(Here  $M_0$  is the mass of the cluster,  $m_{q_1}$  and  $m_{q_2}$  are the rest masses of the constituents, and  $M_{\max}$  as well as  $n$  are model parameters that allow for a tuning of *HERWIG* to the data<sup>(23, 24)</sup>.)

$M_{\max}$  defines an upper limit for the mass of the clusters. If this condition is not fulfilled, the cluster is split up into two clusters with masses less than  $M_0$ . A new quark-antiquark pair  $q_3 \bar{q}_3$  is generated in order to realize this decay of the cluster  $(q_1, \bar{q}_2)$ , and the clusters  $(q_1, \bar{q}_3)$  and  $(q_3, \bar{q}_2)$  are formed. The flavor of the new quark  $q_3$  is equally distributed over  $u$ ,  $d$  and  $s$ . The masses of the new clusters are computed as  $M_{1,2} := M_0 R_{1,2}^\beta$ . Here  $R_{1,2}$  are two values of a random variable which is equally distributed between 0 and 1, and  $\beta$  again is one of the *HERWIG*-parameters<sup>25</sup>. The directions of the momenta of the decay products are set to be parallel to the direction of the momenta of the primary components  $q_1$  and  $\bar{q}_2$ , and their absolute values are computed from energy-momentum conservation.

---

<sup>23</sup>  $M_{\max}$  is the parameter *CLMAX*. Default=3.35 GeV "maximum cluster mass parameter"

<sup>24</sup>  $n$  is the parameter *CLPOW*. Default=2 "power in Maximum cluster mass parameter"

<sup>25</sup>  $\beta$  is the parameter *PSPLT*. Default=1 "split cluster spectrum parameter"

### 2.5.2.2. Cluster decay into hadrons

HERWIG includes two possibilities that allow for a transition from a cluster to hadrons. They are called one-body mechanism and the two-body mechanism respectively. The decision between them is made by looking at the quark content and the mass of the decaying cluster. The following conditions are used:

$$M_c > m_l(q_1, \bar{u}) + m_l(u, \bar{q}_2)$$

$$M_c > m_l(q_1, \bar{d}) + m_l(d, \bar{q}_2)$$

Here  $M_c$  is the mass of the cluster  $(q_1, \bar{q}_2)$ , and  $m_l(a, \bar{b})$  is the mass of the lightest hadron with quark content  $a\bar{b}$ . Let us for example consider a  $c\bar{s}$  cluster. In this case the conditions above correspond to:

$$M_c > m_l(c, \bar{u}) + m_l(u, \bar{s}) = m_{D^0} + m_{K^+} = 1.86\text{GeV} + 0.50\text{GeV} = 2.36\text{GeV}$$

$$M_c > m_l(c, \bar{d}) + m_l(d, \bar{s}) = m_{D^+} + m_{K^0} = 1.87\text{GeV} + 0.50\text{GeV} = 2.37\text{GeV}$$

If both of them are fulfilled, the cluster decays according to the two-body mechanism. For that reason one quark or diquark flavor  $q_3$  is taken by chance from the set  $\{u, d, s, c, t, b, \bar{u}\bar{u}, \bar{u}\bar{d}, \bar{u}\bar{s}, \bar{d}\bar{s}, \bar{d}\bar{d}, \bar{s}\bar{s}\}$ . By default, all the elements of this set have the same probability, but it is possible to change these settings by changing the values of the parameters  $PWT(1) - PWT(7)$  <sup>(26, 27)</sup>. Now a decay of the cluster into two hadrons

$$C(q_1, \bar{q}_2) \rightarrow h_1(q_1, \bar{q}_3) + h_1(q_3, \bar{q}_2)$$

is forced if the mass condition

$$M_c > m_l(q_1, \bar{q}_3) + m_l(q_3, \bar{q}_2)$$

is fulfilled. This condition suppresses the contribution of quarks with big mass. In the default version the momenta of the decaying particle are generated isotropic

---

<sup>26</sup>  $PWT(1)$ - $PWT(6)$  measures the probability, that the flavor of the new quark  $q_3$  is u, ..., t.

<sup>27</sup>  $PWT(7)$  is the diquark-probability.

in the center of momentum system of the decaying cluster. If this condition is not fulfilled, the cluster "decays" via the one-particle mechanism, that means it is replaced by the lightest hadron with the same quark content. The three momentum is by definition the same as the three momentum of the decaying cluster. A possible overflow in energy is distributed over the surrounding clusters by chance.

### 2.5.2.3. Remarks

There exists a possibility of influencing the behavior of hadrons containing so-called "*perturbative quarks*". These are quarks generated in the primary reaction or during the parton shower, but not in the decay of gluons. The idea behind this is, that these quarks should "remember" their original direction even if they are confined in a cluster. In the default version, the cluster decays isotropic in its center of momentum system ( $CLDIR=0$ ), while the option  $CLDIR=1$  allows for a consideration of perturbative quarks directions. Every decay product (hadron) which contains a perturbative quark inherits the direction of its momentum from this perturbative constituent. After this step, a smearing of the direction of momenta can be carried out according to an exponential distribution in  $\{1 - \cos\vartheta\}$ . Here  $\vartheta$  is the angle between the direction of the perturbative constituent and the direction of the final hadron. The mean value  $\mu$  of this distribution is a parameter of *HERWIG*<sup>28</sup>. This modification leads to an improvement in the description of the distribution of charged particle momenta.

---

<sup>28</sup>  $\mu$  is the parameter *CLSMR*. Default=0.0 "*width of Gaussian angle smearing*"





## PART I:

### Measurement of charged particles distributions



## Chapter 3.

# Distributions and Methods

### 3.1. Abstract

Based on *ALEPH* data recorded in the year 1992 distributions of *event properties* and *single particle properties* of charged particles are measured. The main tool of the measuring procedure is the simulation model *HVFL03*, which consists of the *QCD* generator *JETSET* on one hand and the detector simulation program *GALEPH* on the other. The distributions introduced here provide a major part of the input information for the parameter tuning in part II of this thesis.

Initially all the event and single particle properties will be explained. An unfolding procedure will be discussed in some detail. This unfolding procedure is used to correct the observed data from detector influences. The last part of this chapter describes the restriction to hadronic events and to reliable tracks of charged particles, and therefore opens the way to the measurement.

### 3.2. Measured quantities

Each event can be described by the set of four momenta  $p_i := (E_i, \vec{p}_i)$  of its final state particles. Because we are only interested in properties of charged particles, the index  $i$  numbers the set of charged particles in a given event. The following definitions are taken from [JS,93], [A1.92].

### 3.2.1. Event properties

#### 3.2.1.1. Sphericity

Starting from the three real eigenvalues  $\lambda_1, \lambda_2, \lambda_3$  ( $\lambda_1 \geq \lambda_2 \geq \lambda_3$ ,  $\lambda_1 + \lambda_2 + \lambda_3 := 1$ ) of the "sphericity-tensor"

$$S^{\alpha\beta} := \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}, \quad \alpha, \beta \in \{1, 2, 3\}$$

the quantities *sphericity*  $S := \frac{3}{2}(\lambda_2 + \lambda_3)$ , *aplanarity*  $A := \frac{3}{2}\lambda_3$ , and *planarity*  $P := \lambda_2 - \lambda_3 = \frac{2}{3}(S - 2A)$  are defined.

The value of  $S$  lies between 0 and 1 and is a measure of the shape of the event. Small values correspond to events in which the directions of the particle momenta are roughly parallel to a given axis (defined by the eigenvector corresponding to the biggest of the eigenvalues, also called the "sphericity axis"), while events which large values of  $S$  are isotropic. The former type of events is called "2-jet event".

The value of  $A$  lies between 0 and 0.5 and is a measure of the momentum flow out of the so-called "event plane" as defined by the eigenvectors corresponding to the largest and the middle eigenvalue.

In an analogous way the "generalized sphericity tensor" is defined:

$$S^{(r)\alpha\beta} := \frac{\sum_i |\vec{p}_i|^{r-2} p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}, \quad r \in R$$

With the definition  $r := 1$  the *C-parameter*  $C := 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)$  can be calculated using the Eigenvalues of this generalized tensor.

#### 3.2.1.2. Thrust

The quantity measured is  $1 - T$ , with

$$T := \max_{|\vec{n}|=1} \frac{\sum_i |\vec{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$$

$T$  is called "*thrust*". The direction which corresponds to this maximum value is called "*thrust-axis*", and we use the symbol  $\vec{t}$ . The value of  $1-T$  lies between 0 and 0.5, where the 2-jet events correspond to values close to 0, while isotropic events correspond to values around 0.5. If the maximum  $T$  is calculated from the set of momentum vectors perpendicular to the thrust-axis it is called "*major*"-value  $M$ .

$$M := \max_{|\vec{n}|=1, \vec{n} \cdot \vec{t}=0} \frac{\sum_i |\vec{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$$

The corresponding direction defines the "*major-axis*". A third axis, the minor axis, is defined perpendicular to the thrust and major ones, and a "*minor*"-value is calculated just as thrust and major. The value of the "*oblateness*" is defined as  $O := M - m$ . The closer  $O$  is to zero, the more symmetric is the corresponding event with respect to the thrust-axis. Large values of  $O$  correspond to planar events.

### 3.2.1.3. Jet masses

Events are split into two hemispheres by any separation criterion to define the jet masses. The four momenta in these two hemispheres define the following quantities, called the masses of the hemispheres:

$$m_{\pm}^2 := \left( \sum_{i \in S_{\pm}} E_i \right)^2 - \left( \sum_{i \in S_{\pm}} \vec{p}_i \right)^2$$

Here  $S_{\pm}$  is the symbol for the two hemispheres. If the plane which separates the two hemispheres is constructed in a way that lead to a maximum of  $m_{+}^2 + m_{-}^2$ , the bigger mass is called "*heavy jet mass*"  $M_h^2$  while the other one is the "*light jet mass*"  $M_l^2$ . Sometimes the "*mass difference*"  $M_d^2 := M_h^2 - M_l^2$  is also of interest.

Since the calculation of these jet masses is very time consuming, one has to restrict the analysis to some approximations or these complicated definitions have to be replaced by simpler ones. In our analysis we choose the latter and define "*simplified jet masses*" by dividing the full space into two hemispheres with respect to the thrust-axes. We define the "*simplified heavy jet mass*"  $\overline{M}_h^2$  as the maximum

of the corresponding masses, and the "simplified light jet mass"  $\overline{M}_l^2$  as the minimum. In the remaining chapters we will always leave out the word simplified to be as short as possible, but use the bars in the symbols. In fact not the simplified jet masses, but these masses divided by the "visible energy", that is the sum of the energies of all charged particles in a hadronic final state is measured.

### 3.2.1.4. Jet broadenings <sup>1</sup>

Again a separation of the full space into two hemispheres is made by a plane that is perpendicular to the thrust axis. In both hemispheres  $S_{\pm}$  the following quantity is calculated:

$$B_{\pm} := \frac{\sum_{i \in S_{\pm}} |\vec{p}_i \times \vec{n}_T|}{2 \sum_i |\vec{p}_i|}$$

The sum in the numerator includes the momenta in the hemisphere signed by the + or the - symbol, while in the denominator the sum is done over all momenta of the event. These quantities define the "total jet broadening"  $B_t := B_+ + B_-$  and the "wide jet broadening"  $B_w := \max(B_+, B_-)$ . Both of them are close to zero for 2-jet events.

### 3.2.1.5. Jet resolution parameter $y_3$

As mentioned before, the particles in the final state of an  $e^-e^+$ -annihilation event appear in jets. This property is quantified by so-called cluster algorithms. In this work the "Durham cluster algorithm" is used. For each pair of four momenta  $(p_k, p_l)$  a value of the "resolution variable"

$$y_{kl} := \frac{2 \min(E_k^2, E_l^2) (1 - \cos \theta_{kl})}{E_{vis}^2}$$

can be computed. In this formula we used

$$\cos \theta_{kl} = \frac{\vec{p}_k \cdot \vec{p}_l}{|\vec{p}_k| |\vec{p}_l|}$$

---

<sup>1</sup> Source: [O1,93]

The pair with the smallest value of this variable is replaced by a new pseudo particle. The four momentum of this pseudo particle is the sum of the four momenta of its constituents.  $y_3$  is the smallest value of  $y_{ii}$  when only three pseudo particles are left.

## 3.2.2. Single particle properties

### 3.2.2.1. The normalized particle momentum

We normalize the particle momentum by the momentum of the initial electron:

$$x_p := \frac{|\bar{p}|}{|\bar{p}_{e^-}|}$$

### 3.2.2.2. In-momenta

The "*event plane*" is defined in two ways. In the first way it is based on the eigenvectors corresponding to the first (that is the biggest) and the second eigenvalue of the sphericity tensor. In the second way it is defined by the thrust axis and the analogous axis defined by the major value. The "*in-momentum*" is then the component of the momentum that corresponds to this plane and that is perpendicular to the sphericity axis or the thrust axis respectively.

That means, that the in-momentum is the component of the momentum in the direction corresponding to the second eigenvalue or in the direction corresponding to the major respectively.

$$p_{\perp}^{\text{in}} := |\bar{n}_2 \cdot \bar{p}|$$

### 3.2.2.3. Out-momenta

As in the case of in-momenta, the event plane is defined with the sphericity or the thrust axis. The "*out-momenta*" are the components of the particle momenta that are perpendicular to this event plane. That means that the out-momentum is the component of the momentum in the direction corresponding to the smallest eigenvalue of the sphericity tensor or in the direction that is defined by the minor respectively.

$$p_{\perp}^{out} := |\vec{n}_3 \cdot \vec{p}|$$

#### 3.2.2.4. Rapidities

After the definition of an "*event axis*" by the thrust or the sphericity axis, the rapidity of a particle corresponding to this axis is defined as:

$$y := \frac{1}{2} \ln \left( \frac{E + p_l}{E - p_l} \right)$$

Here  $p_l$  is the longitudinal component of the momentum, that is the component in the direction of the event axis;  $E$  is the energy of the particle. The computation of this quantity requires knowledge of the mass of the particle under observation. Because this is not known in general, normally the  $\pi$  mass is assumed.



### 3.3. Unfolding procedures

#### 3.3.1. Treatment of event properties

##### 3.3.1.1. Introduction and definitions

The true distribution  $t(x)$  of a given quantity  $x \in [x_u, x_o]$  is in general not the same as the distribution  $o(y)$  of the same quantity observed by the detector. Even if the quantity is the same, we use the symbol  $y \in [y_u, y_o]$  for the measured value possibly biased by the detector to avoid confusion in the following discussions. The restricted resolution of the detector as well as secondary reactions and decays are reasons for this bias. Both distributions are connected by the following formula:

$$o(y) = \int_{x_u}^{x_o} D(y, x) t(x) dx \quad (3.1)$$

All detector influences are absorbed in the "*detector function*"  $D: [x_u, x_o] \times [y_u, y_o] \rightarrow \mathbf{R}^+$ . The probability that a value of the quantity of interest which lies in a "very small"<sup>2</sup> interval of width  $\Delta x$  around  $x$  is measured in a "very small" interval of width  $\Delta y$  around  $y$  is given by  $D(y, x) \Delta y$ .

The calculation of this detector function (in fact we calculate a discrete version, not the full detector function) is normally done by using a simulation which consists of a *QCD* model and a detector simulation. The former fills the phase space, while the latter deforms this filling. If one of these constituents of the full simulation is wrong, systematic deviations from the exact detector function can not be excluded. Because of this, we have to introduce the assumption that the simulation that was used for the calculation of the detector function provides a good description of the data.

The measurement is carried out on the basis of the following discretization:

$$x_l \equiv x_{lower} =: x_0 < x_1 < \dots < x_m =: x_{upper} \equiv x_u, \quad y_u =: y_0 < y_1 < \dots < y_n =: y_o, \quad m, n \in \mathbf{N}$$

---

<sup>2</sup> "very small" means that the function  $D$  can be approximated by a constant in this interval.

and on the basis of the discretized versions of the distributions and the detector function which become column vectors  $\bar{t}$  and  $\bar{o}$  and the detector matrix  $D$  (the response matrix) respectively:

$$\Delta x_i := x_i - x_{i-1}, i = 1, \dots, m; \quad \Delta y_j := y_j - y_{j-1}, j = 1, \dots, n$$

$$t_i \Delta x_i := \int_{x_{i-1}}^{x_i} t(x) dx, \quad o_j \Delta y_j := \int_{y_{j-1}}^{y_j} o(y) dy$$

$$D_{ji} \Delta y_j := \frac{1}{t_i \Delta x_i} \int_{y_{j-1}}^{y_j} \int_{x_{i-1}}^{x_i} D(y, x) t(x) dx dy \quad (3.2)$$

It follows from the relation (3.2) that the elements of the detector matrix will in general depend on the true distribution  $t(x)$ , if the width of the bins is that large that neither  $t(x)$  nor  $D(y, x)$  can be approximated by a constant within the bins  $[x_{i-1}, x_i]$ . For that reason the bin widths used in our analysis are small in regions where the slope of  $t(x)$  has large absolute values, while larger bins are allowed in regions where  $t(x)$  has a more constant behavior.

Now the discretization of (3.1) can be done:

$$o_j \Delta y_j = \int_{y_{j-1}}^{y_j} \int_{x_{i-1}}^{x_i} D(y, x) t(x) dx dy = \sum_{i=1}^m \int_{y_{j-1}}^{y_j} \int_{x_{i-1}}^{x_i} D(y, x) t(x) dx dy = \sum_{i=1}^m D_{ji} \Delta y_j t_i \Delta x_i \quad (3.3)$$

Because it is highly improbable that a complete event is lost, the probability of detecting an event property is one<sup>3</sup>. That means:

$$\int_{y_u}^{y_v} D(y, x) dy = 1 \quad \text{or} \quad \sum_{j=1}^m D_{ji} \Delta y_j = 1$$

**Conclusion:** *The discretized versions of the true and the observed distributions are connected via a system of linear equations. The matrix appearing in this system of linear equations describes the influences of the detector. If this matrix and the observed distributions are known with a very high accuracy,*

---

<sup>3</sup> In the case of single particle properties the possibility of annihilation and creation mechanisms has to be taken into account. Because of that new particles will be added, or some of the particles that penetrate the detector will not leave it, and the detection probability for single particles will in general fall below one.

the measurement corresponds to the solution of this system of linear equations<sup>4</sup>.

Before the algorithm for measuring event properties is applied to *ALEPH* data, it will be tested in a very simple model surrounding. This simple model is introduced and illustrated in the following example.

*Example 3.1: A simple model scenario*

*Idea: Before the algorithms discussed below are applied to data, they should be explained and illustrated in detail. For that reason, a detector function together with two "true distributions" and the corresponding "observed distributions" biased by the "detector" are given. The first of them play the role of the data (dat), while the second provide an illustration of the simulation (sim). A dependence of the detector matrix from the model as it would eventually appear is not included in this simple model.*

-----

Numerical values:  $x_u = y_u = 0$ ,  $x_o = y_o = 8$ ,  $m = 12$ ,  $n = 13$ . All bins have the same width..

"data distribution" ( $t_{dat}$ ):

$$t_{dat}(x) := N_d^{-1} f(x), \quad N_d := \int_{x_u}^{x_o} f(x') dx', \quad f(x) := \sum_{i=1}^3 a_i \exp \left[ -\frac{1}{2} \left( \frac{x - \mu_i}{\sigma_i} \right)^2 \right]$$

$$\mu_1 = 1.5, \sigma_1 = 1.5, a_1 = 3.5, \quad \mu_2 = 4.5, \sigma_2 = 0.35, a_2 = 0.8,$$

$$\mu_3 = 5.5, \sigma_3 = 1.5, a_3 = 1.0$$

"simulated distribution" ( $t_{sim}$ ):

$$t_{sim}(x) := N_m^{-1} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu_m}{\sigma_m} \right)^2 \right], \quad N_m := \int_{x_u}^{x_o} \exp \left[ -\frac{1}{2} \left( \frac{x' - \mu_m}{\sigma_m} \right)^2 \right] dx', \quad \mu_m = 2.0, \sigma_m = 2.0$$

---

<sup>4</sup> In fact the accuracy will in general not be sufficiently large. The problems that arise from this point are discussed in the next section.

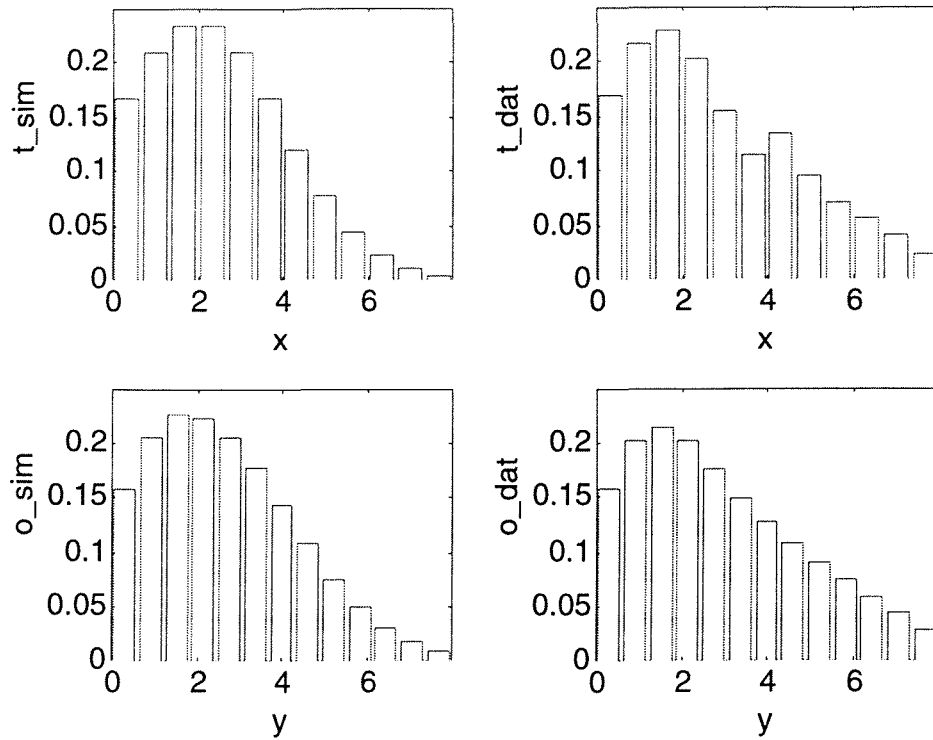


Figure 3.1: Observed and true distributions for "data" and "simulation"

*detector function:*

$$D(y,x) := N^{-1} \exp\left[-\frac{1}{2} \left(\frac{y-x}{\sigma_d}\right)^2\right], \quad N := \int_{y_w}^{y_u} \exp\left[-\frac{1}{2} \left(\frac{y'-x}{\sigma_d}\right)^2\right] dy', \quad \sigma_d := 1$$

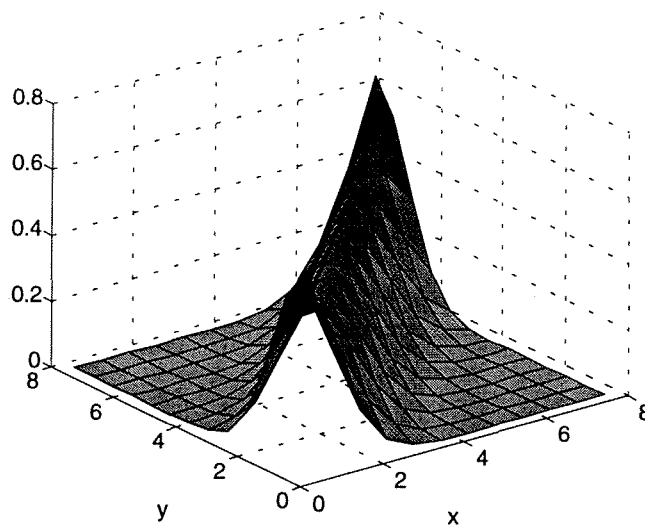


Figure 3.2: Detector function

### 3.3.1.2. Behavior at finite statistics

Naive thinking might suggest, that since the observed distribution and the detector function are known, the true distribution can be calculated by inverting the detector matrix and multiplying this inverse with the vector that represents the observed distribution. Serious problems arise because of the finite number of events we are able to use in the analysis (called the "*finite statistics*"). To demonstrate these problems, the model surrounding is provided with finite statistics in the next example. This example also demonstrates the way to calculate estimates for bin contents in the case of finite statistics and the widths of these quantities.

#### *Example 3.2: The model scenario at finite statistics*

*Idea: Since we have to use finite statistics in the measurement we do only have estimates for the bin contents, and we are able to calculate estimates for their widths. The estimation procedure will be introduced here.*

*These statistical inaccuracies lead to serious difficulties. To get some deeper understanding of this topic, the model scenario should be generated with finite statistics.*

---

#### *(i) Generation of a discrete distribution*

Starting point is the discrete probability distribution  $(p_1, \dots, p_n)$  of a given random variable  $x$  with

$$\sum_{k=1}^n p_k = 1, \quad p_k := \rho_k \Delta x_k$$

Here  $p_k$  is the probability for the case that a value of the random variable  $x$  lies in the  $k$ -th bin. On this basis a separation of the interval  $]0,1]$  into subintervals is made:

$$I_1 := ]0, p_1], \quad I_k := \left] \sum_{l=1}^{k-1} p_l, \sum_{l=1}^k p_l \right], \quad k = 2, \dots, n$$

The width of the  $k$ -th interval is given by the probability  $p_k$ . Normally every programming language used in natural sciences provides a random number generator which produces random numbers equally distributed in the interval  $]0,1]$ . If a value of this equally distributed variable lies in the interval  $I_k$ , a counter that registers the entries in the  $k$ -th bin is increased by unity. The values of these counters are then distributed according to the given discrete probability.

(ii) *Estimation of the bin contents and their widths*

The bin contents  $\bar{B}_k$  of the discrete distribution generated in section (i) are distributed by a *polynomial distribution*<sup>5</sup>. (The bar marks that this is the bin content of a given sample, and it is in general different in another sample.) If samples are generated with  $s$  independent throws ( $s$  marks the used "statistics"), the mean values and the covariance matrix can be calculated (c.f. [Ea,71]). For example we have

$$E(B_k) = sp_k \text{ (mean value), } V(B_k) = \sigma^2(B_k) = sp_k(1 - p_k) \text{ (variance)}$$

Because of this expression for the mean value, the normalized bin contents

$$\bar{\rho}_k := \frac{\bar{B}_k}{s\Delta x_k}$$

are unbiased estimates for the values of the discrete probability distributions  $\rho_k$  (that means  $E(\bar{\rho}_k) = \rho_k$ ), and the standard deviation can be taken as a measure of the distance from  $\bar{\rho}_k$  to  $\rho_k$ <sup>6</sup>. Because of  $V(\alpha X) = \alpha^2 V(X)$ ,  $\alpha \in \mathbf{R}$ , and because  $X$  is any random variable we have

$$\sigma_k = \frac{\sqrt{sp_k(1 - p_k)}}{s\Delta x_k} \equiv \frac{1}{s\Delta x_k} \sqrt{\frac{\bar{B}_k(s - \bar{B}_k)}{s}}$$

<sup>5</sup> The polynomial distribution is also called "multinomial distribution" (c.f. [Br,87]).

<sup>6</sup> The standard deviation marks the mean of the difference of a given value of a random variable from its mean value. Only if this mean value is identical with the quantity of interest (that means if the estimator is unbiased), it is also a measure of the deviation of any output of the random variable from this quantity of interest.

and we end up with

$$\sigma_k \equiv \frac{1}{s\Delta x_k} \sqrt{\frac{\bar{B}_k(s - \bar{B}_k)}{s}} \quad (3.4)$$

We neglect the width of this width due to the statistical fluctuation of the  $\bar{B}_k$  and take the formula above as an expression for the standard deviation.

*(iii) Simulation of detector influences*

Every content of the  $i$ -th bin of the true distribution is "scattered" into the  $k$ -th bin of the observed distribution with a probability  $D_{ki}\Delta y_k$ . To generate the observed distribution, the algorithm introduced in *(ii)* is repeated with

$$s := B_i, \quad p_k = D_{ki}\Delta y_k$$

*(iv) Graphical illustration*

The "simulated distribution" is generated with  $s_{sim} := 4000$ , and the "data distribution" is generated with  $s_{dat} := 1000$ . The following figure illustrates the results. The fact that the normalized bin contents lie inside the error bars (as defined by the standard deviation) with a probability smaller than one can be seen from these graphs in particular.

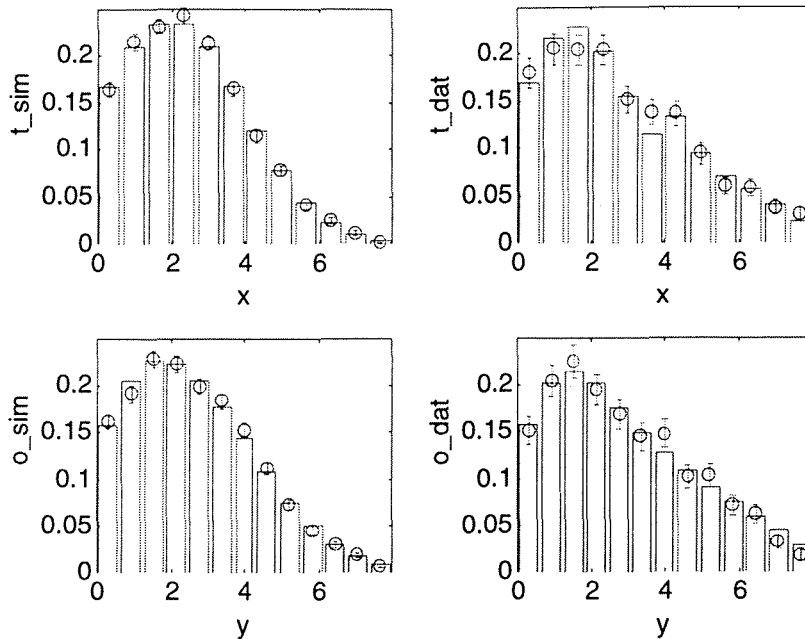


Figure 3.3: Distributions at finite statistics

In fact the estimates  $\bar{p}_k$  for high values of  $s$  are in good approximation distributed by a normal distribution (c.f. *generalized theorem of Moivre and Laplace* [Fi,62]). The true value therefore lies in the region marked by the error bars with a 68% probability. From the expression for the standard deviation (3.4) it is clear that one has to provide a factor 4 in  $s$  to end up with 50% of the error bars. This fact can also be seen in the figure above.

o

Equation (3.3) is a system of linear equations, and it seems to be clear that a solution can be found using standard methods. This is true in principle, but since we only have estimates for the coefficients of this system of linear equations, we will also get estimates for the solutions. The crucial point will be the size of the widths of the solutions. It turns out, that in general the errors of the solutions will be that large, that the result is quite meaningless<sup>7</sup>. The following example illustrates this fact.

---

<sup>7</sup> An example of a meaningless result is the estimate of a probability with an error bigger than one.



*Example 3.3: Error propagation in systems of linear equations*

Idea: In some steps of the solution of a system of linear equations some critical mathematical operations, such as the subtraction of nearly equal numbers can appear. These critical operations can lead to big relative errors in the results. To characterize this critical behavior, normally a number that is called condition is used. This condition is introduced in the following example, and a characterization of the systems of linear equations appearing in our analysis is carried out<sup>8</sup>.

---

Starting point is the following system of linear equations

$$\mathbf{A}\bar{x} = \bar{b}$$

In the case of the measurements of event properties, the matrix  $\mathbf{A}$  will be the detector matrix. In the cases that are interesting for this work, the matrix  $\mathbf{A}$  is quadratic and invertible. Because of this the following discussions are also restricted to this special type of matrices.

In a first step the errors of the matrix coefficients should be zero, only the inhomogenous part  $\bar{b}$  should be biased. This is the same case that can be found in the example "Unfolding of a distribution of a discrete variable" in [Bl,84]. To begin let us define the length of a vector and the generalized "length" of a matrix:

$$\|\bar{x}\| := \sqrt{\sum_{i=1}^n x_i^2} \quad \text{and} \quad \|\mathbf{A}\| := \sqrt{\sum_{i=1}^n \sum_{j=1}^n A_{ij}^2}$$

The definition of the condition of the matrix  $\mathbf{A}$  comes next:

$$\kappa(\mathbf{A}) := \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \geq 1$$

The inequality (c.f. [Sc,93])

$$\frac{\|\Delta\bar{x}\|}{\|\bar{x}\|} \leq \kappa(\mathbf{A}) \frac{\|\Delta\bar{b}\|}{\|\bar{b}\|}$$

---

<sup>8</sup> For details c.f. [Sc,93] or any textbook of numerical mathematics.

holds. This formula makes only predictions about lengths of vectors, but it should be enough to give a first impression about the error propagation behavior. The condition is in this first case of an unbiased matrix  $\mathbf{A}$  an amplification factor for the relative errors. In the example found in [Bl,84], the condition is  $\kappa(\mathbf{A}) \cong 162$ , thus the relative errors of the solutions can be a hundred times larger than the relative errors of  $\bar{b}$ .

If not only  $\bar{b}$  is biased, but also the coefficients of the matrix  $\mathbf{A}$ , the following generalization of the expression above can be used (c.f. [Sc,93]):

$$\frac{\|\Delta\bar{x}\|}{\|\bar{x}\|} \leq \frac{\kappa(\mathbf{A})}{1 - \kappa(\mathbf{A}) \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}} \left\{ \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\Delta\bar{b}\|}{\|\bar{b}\|} \right\}$$

Note that this expression makes only sense if  $1 > \kappa(\mathbf{A}) \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}$ .

If this is not the case, the result of the inversion process should not be trusted. If this condition is fulfilled, the expression above leads to an upper limit that is in general pessimistic. The next figure illustrates this behavior in a rather drastic way.

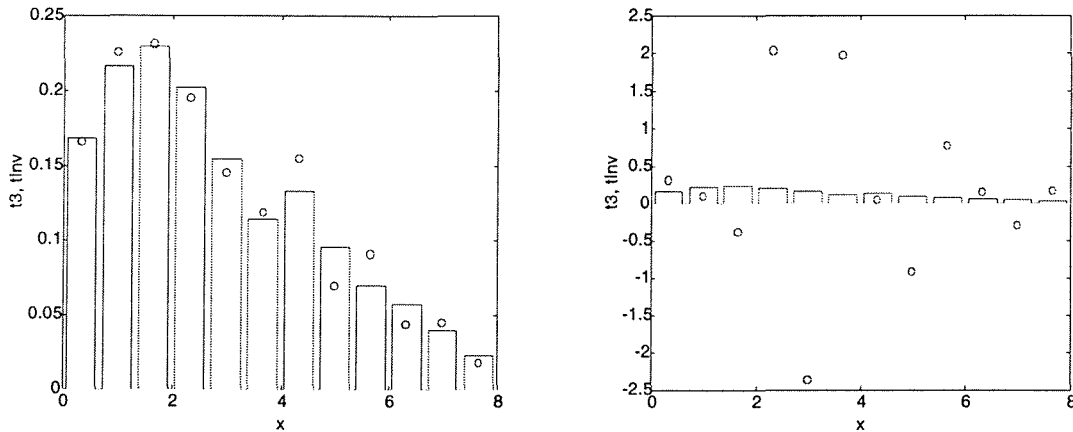


Figure 3.4: Solution of a well conditioned and of a badly conditioned system of linear equations

Again the simple model set up with  $s_{dat} := 5000$  and  $s_{sim} := 50000$  is used to obtain this illustration. The change in the condition is produced by a change of the width  $\sigma_d$  in the detector function. A small width corresponds to a

rather diagonal detector matrix, and therefore to an unproblematic system of equations, while a larger width leads to the difficulties mentioned above. The next table shows more of the details. The following abbreviations are used:

$$\|\Delta\mathbf{A}\|_{rel} := \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|}, \quad \|\Delta\bar{b}\|_{rel} := \frac{\|\Delta\bar{b}\|}{\|\bar{b}\|}, \quad \|\Delta\bar{x}\|_{rel} := \frac{\kappa(\mathbf{A})}{1 - \kappa(\mathbf{A})} \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \left\{ \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\Delta\bar{b}\|}{\|\bar{b}\|} \right\}$$

$\sigma_d$	$\kappa(\mathbf{A})$	$\ \Delta\mathbf{A}\ _{rel}$	$\kappa(\mathbf{A})\ \Delta\mathbf{A}\ _{rel}$	$\ \Delta\bar{b}\ _{rel}$	$\ \Delta\bar{x}\ _{rel}$
0.35	3.34	0.0338	0.1129	0.0427	0.29
1	343.23	0.0721	39.160	0.0432	-

Table 3.1: Matrix conditions in the simple model

Now the interesting question is about the behavior of the detector matrices used in our analysis. To calculate these conditions, we used the bins given in chapter 4. The statistics were  $s_{dat} := 571825$  and  $s_{sim} := 1186173$ . The results are shown in the following table.

quantity	$\kappa(\mathbf{A})$	$\ \Delta\mathbf{A}\ _{rel}$	$\kappa(\mathbf{A})\ \Delta\mathbf{A}\ _{rel}$	$\ \Delta\bar{b}\ _{rel}$	$\kappa(\mathbf{A})\ \Delta\bar{b}\ _{rel}$	$\ \Delta\bar{x}\ _{rel}$
$S$	77.40	0.0073	0.56	0.0045	0.35	2.11
$A$	53.68	0.0071	0.38	0.0029	0.15	0.87
$P$	26.08	0.0091	0.24	0.0027	0.07	0.40
$C$	35.71	0.1012	3.61	0.0043	0.15	-
$T$	90.55	0.0161	1.46	0.0048	0.43	-
$M$	19.27	0.0140	0.27	0.0049	0.09	0.50
$m$	23.88	0.0343	0.82	0.0038	0.09	5.00
$O$	17.44	0.0129	0.23	0.0042	0.07	0.38
$\bar{M}_h^2$	191.26	0.0165	3.15	0.0043	0.82	-
$\bar{M}_l^2$	2281.2	0.0372	849.38	0.0026	5.93	-
$\bar{M}_d^2$	190.90	0.0164	3.13	0.0036	0.69	-
$B_l$	5.42	0.1185	0.64	0.0036	0.02	1.86
$B_w$	6.96	0.0460	0.32	0.0030	0.02	0.50
$y_3$	150.74	0.0186	2.80	0.0019	0.29	-

Table 3.2: Conditions of relevant detector matrices

Conclusion: *Even if the upper limit as calculated by the condition of the detector matrix is pessimistic, the given results clearly indicate that results calculated by inversion of the detector matrix should be mistrusted. If satisfactory results are required, we have to use other methods. If we want to end up with smaller errors we have to provide additional information to the measurement.*

This additional information will be achieved by using of a simulation model. The better the model used for the measurement is, the better the results will be. The discussion of the inclusion for a given model into the analysis is described in the next section.

### 3.3.1.3. Unfolding of event properties (matrix method)

The detector matrix  $D$  that was introduced in the last section is calculated by a Monte Carlo simulation. Finite statistics will in general lead to estimates for the normalized bin contents with errors that are too big for an accurate inversion of the detector matrix. This section introduces a method that allows for an unfolding without any inversion. This big advantage is reached by adding information in form of a simulation model that represents the current "state of the art". If this model describes nature exactly, the analysis is exact too. Because this is not quite the case it will be necessary to introduce an estimation of the error due to the imperfections of this model.

Let  $o_j \Delta y_j =: P(\Delta y_j)$  be the probability for the case that a given event property is measured in an interval  $\Delta y_j$ , and  $t_i \Delta x_i =: P(\Delta x_i)$  the probability for the case that the true value of this quantity lies in the interval  $\Delta x_i$ .  $D_{ji} \Delta y_j =: P(\Delta y_j / \Delta x_i)$  is the probability, that this quantity of interest is measured in the interval  $\Delta y_j$  under the condition that its' true value lies in  $\Delta x_i$ . We can now rewrite formula (3.3) and identify the "*Satz ueber die vollständige Wahrscheinlichkeit*" (c.f. [Bo,91]):

$$P(\Delta y_j) = \sum_{i=1}^n P(\Delta y_j / \Delta x_i) P(\Delta x_i)$$

Because this is true, the expression

$$P(\Delta x_i) = \sum_{j=1}^n P(\Delta x_i / \Delta y_j) P(\Delta y_j) \Leftrightarrow t_i \Delta x_i = \sum_{j=1}^n C_{ij} \Delta x_i o_j \Delta y_j$$

also holds, and the correction can be done by:

$$t_i^{corr} = \sum_{j=1}^n C_{ij} o_j \Delta y_j$$

Here  $C_{ij} \Delta x_i =: P(\Delta x_i / \Delta y_j)$  is the probability that the true value lies in the interval  $\Delta x_i$  under the condition that it is measured in  $\Delta y_j$ . Thus we know the correction matrix  $C$  if we are able to compute the conditioned probabilities  $P(\Delta x_i / \Delta y_j)$ . The key to this calculation will be the *Bayes theorem* (c.f. [Bo,91], [Br,87]):

$$P(\Delta x_i / \Delta y_j) P(\Delta y_j) = P(\Delta y_j / \Delta x_i) P(\Delta x_i) = P(\Delta y_j \wedge \Delta x_i) \Leftrightarrow C_{ij} = \frac{D_{ji}}{o_j} t_i$$

Even in the case of a detector matrix that is independent from the model predictions, this equation shows a linear dependence of the correction matrix from the model distribution. The key for calculating the elements of the correction matrix is given by:

$$C_{ij} \Delta x_i = P(\Delta x_i / \Delta y_j) = \frac{P(\Delta y_j \wedge \Delta x_i)}{P(\Delta y_j)} \quad (3.5)$$

The sometimes very drastic influences of the model predictions on the unfolding procedure should be illustrated (in a pessimistic but nevertheless interesting way) by the following example.

*Example 3.4: Matrix correction in the model scenario*

Idea: In the model scenario we provided a "model" that is somewhat different from the "data". Now we play the following "game". What happens if we use this false model to "correct" the observed distribution with the method introduced above?

---

In the figure right, the result of this "game" is illustrated. The bars represent the given true distribution (dat). A correction matrix that was calculated with this true distribution reproduces the distribution as it should be. The crosses show the different (and false) "model". The circles illustrate the result of the unfolding procedure using this false model. A large bias is the consequence, and indeed the result is not very satisfactory.

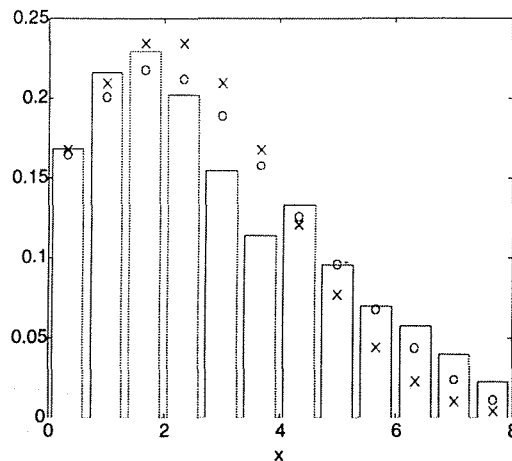


Figure 3.5: - ... t\_dat, o ... t\_corr, x ... t\_sim

We explained how to calculate the elements of the correction matrix. Not only estimates for this elements have to be calculated, but also their widths, and the widths of the corrected distribution. This topic will be dealt with in section 3.5.2.1 where the calculation of errors is introduced.

Conclusion: *It is possible to reproduce the true distribution, if a good simulation is available. The crucial assumption is therefore, that the model reproduces the data "very well". Additional tests should be done to check the quality of the models. Because even the best of the available models does not reproduce data in every detail, an estimation of systematic errors due to this imperfections of the simulation models has to be calculated.*

Before some tests of the model accuracy are done and a method for the estimation of systematic errors is introduced, the discussion about the unfolding procedures should be finished by a section that deals with the single particle distributions.

### 3.3.2. Treatment of single particle properties

There are two important differences between the unfolding of event properties and the unfolding of single particle properties. The first difference is due to the calculation of the statistical error. In the case of event properties we used the fact that each event is generated under the same conditions and that all events are independent. That means: different entries are independent in the case of event properties. On that basis we applied the rules of statistics. In the case of single particle properties this basic assumption does no longer hold. Entries are correlated in general, and we have to use different methods.

The second and even more important difference is that a detector matrix is not available. This is due to the facts that single particles can be annihilated or they can be produced in the detector, and to some technical details in *ALPHA*<sup>9</sup>.

#### 3.3.2.1. Calculation of the statistical error

One possibility is to forget about the correlation mentioned above. In this case one would expect a statistical error that is much smaller than the error that is calculated for event properties, because on average each event contributes roughly 21 charged particles, and so the statistics used for single particle distributions should be 21 times higher than the statistics used for event properties. This calculation can be used for a first approximation.

If one wants to calculate more accurate values for the widths, he has to take the correlations into account. These correlations are due to the fact that a hadronic event is not isotropic, but is arranged in jets. If one knows for example the direction of movement of a given particle in a given jet, he also knows that all the directions of movements of the other constituents of this jet are very similar. The direction of one particle restricts the direction of many others. They are no longer independent and thus correlated. These correlations are especially important for the true distributions. Bins of observed distributions are less correlated because the smearing of the detector decreases the effect.

---

<sup>9</sup> It turned out to be impossible to relate true properties of single particles to observed properties in a unique way. For details see [Ap,95]



For example the rapidity of a particle is essentially a measure of the angle between the direction of movement of this particle and the thrust axis. If the directions of movement are highly correlated, one could argue, that the rapidities are highly correlated too. Indeed discrepancies between the widths calculated with and without including correlations can be observed here.

If one wants to include correlations in the calculation of the widths, one can again use the fact that events are independent, even if single particle contributions are correlated. For each event the full histogram of the single particle distribution is calculated, and on that basis the width in each of the bins can be calculated by the following formula.

$$\sigma_{b,v}^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (t_{b,v}^i - \bar{t}_{b,v})^2 = \frac{n}{n-1} \left\{ \left( \frac{1}{n} \sum_{i=1}^n t_{b,v}^i \right)^2 - \left( \frac{1}{n} \sum_{i=1}^n t_{b,v}^i \right)^2 \right\}$$

Here  $n$  is the number of events,  $t_{b,v}^i$  is the bin content of the  $b$ -th bin of the  $v$ -th distribution in the  $i$ -th event. We are interested in the mean value

$$\bar{t}_{b,v} := \frac{1}{n} \sum_{i=1}^n t_{b,v}^i$$

of a given bin content. The width of this mean value is given by

$$\sigma_{i_{b,v}} := \frac{\sigma_{b,v}}{\sqrt{n}}$$

This can be seen by error propagation. The off-diagonal elements of the covariance matrix are again neglected, even if they are supposed to be bigger than in distributions of event properties.

### 3.3.2.2. Unfolding of single particle properties (factor method)

Since the full detector matrix is not known in the case of single particles, the unfolding is done by the simplest of all unfolding procedures, that is the

### *Unfolding procedures*

application of correction factors. Here the  $j$ -th bin content of the corrected distribution is given by

$$t_j^{corr} = C_j o_j$$

The correction factors are again calculated by using a simulation model. Their estimates are given by

$$C_i := \frac{t_i^{sim}}{o_i^{sim}} \quad (3.6)$$

For the calculation of the width of this expression and of the width of the corrected distribution see section 3.5.2.1. The factor method can be applied if off-diagonal elements of a generalized detector matrix are relatively small. To guarantee this crucial requirement, the width of the bins for single particle distributions is set to be bigger than the resolution of the detector.

### 3.4. Tests of the simulation

We mentioned before that a very important ingredient of our analysis is a simulation that reproduces the data "very well". We will now take a critical look on this assumption by using the very sensitive "eye glasses" of high statistics. If it turns out that all tests are fulfilled, our understanding of hadronic events has reached an optimum and the analysis is simply a test of an optimal model. Because we know that our understanding is not perfect, even if it is really not too bad, we expect that some of the tests will fail. That means that the basic assumption of our analysis is violated, and we should expect a bias of the kind of example 3.4. If a bias appears in the measurement, a corresponding systematic error has to be given. This error is called "*model bias*", its estimation is introduced in section 3.5.3.1.

Two kinds of tests are done. The first one is a very sensitive test. Here a very high statistic is used to check the whole simulation, that means the QCD-part as well as the detector simulation. The deviation of the data from the full simulation is expressed in percent of the data and in units of the calculated statistical error. In addition a  $\chi^2$ -value is computed for each of the distributions. The second test is not as sensitive as the first one, but here it is possible to test the consistency of the corrected data. If one of these tests fail, the basic assumption of a very good simulation was not satisfied, and the calculation of a model bias is necessary.

#### 3.4.1. Accuracy of the full simulation

If the QCD-model and the detector simulation would be perfect, the deviations between the full simulation and the data should be purely statistical, and the statistical widths should have the same size as these deviations. Because of that we look at the behavior of the difference

$$\Delta_j := \frac{r_j^{dat} - r_j^{sim}}{\sqrt{(\sigma_j^{dat})^2 + (\sigma_j^{sim})^2}}$$

*Tests of the simulation*

Here  $r_j^{dat}$  is the data value in the  $j$ -th bin of a given distribution, and  $r_j^{sim}$  is the corresponding value of the full simulation. If the deviations are purely statistical, these values hardly exceed 3. If the simulation is perfect, the sum

$$\eta := \sum_{j=1}^n \Delta_j^2$$

is a sum of squares of  $n$  random variables which are  $N(0,1)$  distributed.  $(n-1)$  of them are independent if bin correlations are neglected<sup>10</sup>, and consequently this sum should be  $\chi^2$  distributed with  $(n-1)$  degrees of freedom in very good approximation<sup>11</sup>. The mean value of this distribution is  $(n-1)$ , while the variance deviation is  $2(n-1)$ . With this method the hypothesis of a perfect simulation can be tested. We know the  $\chi^2$  distribution, and we know a region where the values should lie with a given probability, and can therefore exclude this hypothesis if the value of the test quantity  $\eta$  exceeds some upper limit. The exact values of these upper limits can be found in [Br,87]. For a first impression the deviation from the mean value should not exceed more than two standard deviations, because every  $\chi^2$  distribution with  $m$  degrees of freedom can be approximated by a  $N(m,2m)$  distribution. The accuracy of this approximation increases with increasing  $m$ .

The quantities given in the following tables are calculated with *ALEPH 92* data and the simulation *HVFLO3*. Beneath every symbol of the distribution and the number  $n$  of bins the value of the test quantity  $\eta$  can be seen.

$S$	$A$	$P$	$C$	$1-T$	$M$	$m$	$O$	$\overline{M}_h^2$	$\overline{M}_L^2$	$\overline{M}_d^2$
24	17	21	24	22	22	18	20	23	18	22
157.7	558.8	48.9	165.2	118.1	304.0	436.0	311.8	152.0	160.1	192.7

$B_t$	$B_w$	$-\ln(y_\gamma)$	$x_p$	$y_S$	$p_t^{in}(S)$	$p_t^{out}(S)$	$y_T$	$p_t^{in}(T)$	$p_t^{out}(T)$
20	17	14	46	21	25	19	21	25	19
195.4	228.0	265.9	$10.3 \cdot 10^5$	$6.7 \cdot 10^5$	$5.9 \cdot 10^5$	$21.7 \cdot 10^5$	$8.3 \cdot 10^5$	$3.6 \cdot 10^5$	$25.8 \cdot 10^5$

Table 3.3: Test quantities of the full simulation

<sup>10</sup> In every distribution one bin can be expressed as a linear combination of all the others.

<sup>11</sup> For details about  $\chi^2$  distributions c.f. [Ea,71], [Bo,91].

These numbers are really dramatic in the case of single particle distributions. One should remember, that this is mainly because the  $\chi^2$  test with a statistics as high as for the single particle distributions is very sensitive. In fact most parts of the data are described well (that means better than one percent), as can be seen in the following table. It quotes the differences  $\Delta_j$  and the deviations between data and simulation in percent of the data value (symbol:  $\Delta_j^{\%}$ ).

S		A		P		C		I-T		M	
$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$
-10.0	-6.1	-12.9	-4.1	-0.6	-0.1	-7.0	-6.3	-3.8	-5.6	-12.3	-25.2
2.1	0.9	-0.9	-0.3	3.2	1.2	-1.1	-0.4	-6.0	-4.3	-7.4	-9.4
0.0	0.0	0.1	0.0	-0.5	-0.3	0.5	0.2	-2.7	-1.5	-0.4	-0.4
2.3	1.3	0.7	0.3	2.0	1.3	3.8	1.6	0.5	0.2	0.8	0.6
0.1	0.0	0.1	0.0	-1.4	-1.1	1.4	0.7	-0.3	-0.2	2.7	1.9
0.6	0.4	0.1	0.1	2.8	2.4	1.2	0.7	1.8	1.0	4.1	1.7
-0.2	-0.2	2.4	1.4	-0.3	-0.3	1.6	1.0	2.4	1.5	3.3	1.5
-0.4	-0.3	5.2	4.5	-0.6	-0.7	0.3	0.2	0.9	0.6	2.0	1.0
2.2	1.5	6.8	6.3	0.3	0.3	-0.4	-0.4	1.6	0.8	3.1	1.7
-0.3	-0.2	7.7	10.8	0.1	0.1	-0.4	-0.3	2.7	1.6	0.7	0.3
-2.2	-1.4	8.3	16.1	-1.1	-0.9	0.8	0.8	0.9	0.4	-1.1	-0.6
1.8	1.4	5.5	14.7	-0.2	-0.2	0.7	0.7	0.7	0.4	-1.5	-0.9
0.9	0.8	6.2	21.2	-1.4	-1.6	-0.7	-0.7	2.5	1.9	-1.2	-0.9
-0.3	-0.2	5.0	21.3	-0.3	-0.3	-2.4	-2.5	0.7	0.6	-2.7	-2.3
-0.4	-0.3	7.0	28.5	-2.9	-3.5	-0.5	-0.6	-0.3	-0.3	-1.1	-1.0
0.3	0.3	5.8	33.6	-1.4	-1.8	-1.7	-2.2	-1.9	-2.2	-1.6	-1.7
0.6	0.7	2.5	26.1	-2.2	-3.6	-3.7	-4.9	-0.8	-1.0	-4.4	-5.7
-0.4	-0.6			-1.9	-3.9	-2.8	-3.9	-3.5	-3.4	-2.1	-3.1
1.6	2.4			0.2	0.6	-1.5	-2.2	-0.7	-0.8	-2.2	-4.0
0.0	-0.1			0.6	2.5	-0.3	-0.5	0.6	1.2	0.2	0.3
1.7	2.7			-1.0	-11.8	3.4	8.1	3.8	20.2	-0.8	-2.6
3.3	7.3					4.6	16.5	0.3	8.7	0.5	3.3
3.3	13.9					4.8	26.5				
2.4	39.5					2.1	24.2				

Table 3.4: Comparison between simulation and data

Tests of the simulation

Because the single particle distributions with respect to  $S$  and  $T$  are very similar, only the values for  $S$  are given. One can also see from the tables, that there are regions where the accuracy of the simulation is not so good. The calculation of a "model bias" is therefore essential.

$m$		$O$		$\bar{M}_h^2$		$\bar{M}_l^2$		$\bar{M}_d^2$		$B_i$	
$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$
-6.3	-13.2	3.1	1.2	-4.2	-4.4	6.6	1.6	-8.9	-3.0	-3.0	-9.9
-12.4	-6.9	6.5	2.0	-6.1	-3.1	-4.3	-1.2	-1.4	-0.5	-10.06	-6.5
-1.9	-1.0	5.5	2.9	-3.9	-1.7	-3.9	-1.5	1.7	0.8	3.7	1.3
0.3	0.2	3.2	1.9	-0.1	0.0	-2.8	-1.4	3.6	1.9	4.1	1.4
-0.1	0.0	3.6	2.3	3.9	2.0	-2.5	-1.6	4.0	2.4	3.2	1.3
0.9	0.4	-0.4	-0.3	1.2	0.7	-1.2	-1.0	2.2	1.5	1.2	0.6
0.5	0.7	0.5	0.3	1.1	0.7	0.2	0.2	2.5	1.9	-1.0	-0.6
0.5	0.2	-1.9	-1.3	3.2	2.3	1.3	1.5	5.1	4.3	-1.5	-1.0
0.0	0.0	-2.0	-1.5	4.0	2.2	2.9	4.0	1.4	1.0	-1.2	-0.8
0.8	0.6	-2.4	-2.0	1.1	0.7	3.6	5.9	1.4	1.1	-1.6	-1.3
5.8	4.1	-5.5	-3.8	3.0	1.7	4.3	5.9	1.8	1.2	-1.6	-1.5
6.7	7.5	-4.7	-3.9	0.5	0.4	1.8	3.4	0.7	0.6	-1.0	-1.1
7.1	11.8	-5.4	-5.4	0.9	0.7	2.4	5.7	-2.8	-2.9	-3.9	-4.7
8.0	19.6	-4.9	-5.9	-1.2	-1.2	0.7	2.3	-0.7	-0.9	-1.7	-2.3
4.8	18.4	-5.3	-7.7	0.0	0.0	2.6	10.6	-0.8	-1.2	0.9	1.4
4.1	24.3	-5.6	-9.9	-1.2	-1.7	0.9	4.1	-2.3	-3.8	0.7	1.3
1.8	18.4	-4.2	-9.3	-1.3	-2.1	2.7	24.1	-2.5	-4.9	0.9	2.3
0.9	20.9	-2.4	-6.7	-3.5	-4.7	1.2	37.0	-2.3	-3.6	1.6	6.5
		-1.6	-6.4	-2.2	-4.3			-2.8	-6.4	3.2	28.1
		-1.0	-6.1	-1.5	-4.6			-0.4	-1.3	0.7	23.7
				0.8	3.7			0.5	2.4		
				0.5	3.7			0.9	5.3		
				1.0	11.7						

Table 3.5: Comparison between simulation and data

$B_w$		$-\ln(y_3)$		$x_p$		$y_s$		$p_i^{in}(S)$		$p_i^{out}(S)$	
$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$	$\Delta_i$	$\Delta_i^{\%}$
-10.5	-10.2	-2.2	-7.2	229.0	2.9	32.9	0.3	195.7	0.9	-286.8	-1.0
2.4	0.7	-6.6	-6.2	482.1	4.4	25.3	0.2	291.5	1.3	57.9	0.2
6.2	1.8	-6.1	-4.0	155.7	1.3	-20.4	-0.2	209.0	1.0	-32.6	-0.1
3.0	1.2	-2.4	-1.3	7.1	1.1	-87.2	-0.7	-44.5	-0.2	-141.7	-0.8
0.3	0.2	-2.1	-0.9	-52.4	-0.4	-128.8	-1.0	-167.6	-1.1	-183.0	-1.3
-2.7	-1.6	2.8	1.1	-81.3	-0.7	-171.9	-1.2	-255.0	-1.9	-66.1	-0.6
-0.8	-0.6	2.1	0.8	-204.0	-1.9	-182.1	-1.2	-274.8	-2.3	107.5	1.3
-2.5	-2.0	4.8	1.7	-341.7	-3.3	-193.3	-1.3	-240.0	-2.3	212.7	3.3
-3.6	-3.6	4.8	1.7	-321.9	-2.2	-63.7	-0.4	-224.2	-2.5	326.6	6.4
-2.0	-2.3	0.0	0.0	-211.0	-1.5	62.2	0.4	-162.0	-2.0	435.8	10.5
-3.6	-5.3	-6.8	-8.5	-210.2	-1.7	180.4	1.3	-159.2	-1.7	631.3	14.9
-2.9	-5.2	-7.8	-22.5	-151.7	-1.3	265.4	2.0	-65.5	-0.8	600.3	20.1
-2.4	-5.7	-1.5	-8.8	-127.9	-1.2	310.1	2.6	-70.4	-1.1	519.9	23.8
-2.1	-7.1	0.8	8.8	-60.3	-0.6	213.9	2.1	-6.2	-0.1	442.0	27.1
-0.1	-0.6			-4.8	0.0	234.8	2.6	18.5	0.4	379.2	30.7
0.7	5.3			17.0	0.1	179.4	2.4	56.4	0.1	416.3	33.6
1.1	10.9			116.1	1.0	33.2	0.5	37.1	0.9	272.9	38.1
				104.1	1.0	-53.7	-1.0	75.4	2.2	196.3	46.7
				174.2	1.9	-185.4	-3.5	94.2	3.5	134.4	46.0
				168.7	2.0	-308.5	-9.4	132.1	5.0		
				147.8	1.9	-243.1	-13.2	109.5	6.3		
				145.2	2.0	-109.9	-13.4	122.6	10.6		
				173.0	2.5			110.1	13.9		
				154.4	1.8			79.9	13.0		
				166.5	2.2			60.9	19.9		
				162.8	2.4						
				153.5	2.3						
				75.0	1.3						
				73.2	1.4						
				52.7	1.2						
				4.3	0.1						

Table 3.6: Comparison between simulation and data

For event properties the absolute deviations between data and simulation are smaller, and the statistical errors are bigger than for single particle properties. This leads to a small contribution of systematic uncertainties to the deviation between data and simulation in the first case, while in the latter case the main part of this deviation comes from model imperfections. If one only looks at the  $\Delta_j$ , he can easily get the impression that the simulation is inappropriate. But the quality of a model can be better judged if one looks at the percental deviations  $\Delta_j^{\%}$ . Here a good agreement in the percental range is observed. Indeed the mean value of the absolute values of all  $\Delta_j^{\%}$  given in the three tables is 4.6, and some of the distributions such as  $S$  or  $x_p$  are described much better. In some other distributions a lack of understanding of hadronic events seems to be obvious. Examples are  $A$ ,  $m$  and  $p_i^{out}$ .

Only the first 31 bins are shown for the momentum distribution  $x_p$ . For the calculation of the distributions the bin divisions that were given in chapter 4 and the cuts that were introduced in section 3.5.1 were used.

***Conclusion:*** *The test of the full simulation shows that the simulation describes the data in most of the interesting parts within less than one percent. The usage of this simulation in the measurement therefore makes sense. Errors in the measurements due to the remaining systematic uncertainties of the model will be estimated by introducing a model bias.*

*We also saw that systematic errors are negligible compared to the statistical errors in most part of the event properties, while they exceed the statistical errors dramatically in the case of single particle distributions.*



### 3.4.2. A statistical test of unfolded event properties

In some works about unfolding methods one of the basic assumptions is that the dependence of the detector matrix  $D$  from the underlying QCD part of the simulation can be neglected. This QCD-model is a major ingredient for the calculation of this matrix. It is worth noting that if this assumption is true, it is possible to construct a very sensitive test for the quality of the unfolded distributions.

*Idea: The measured values in each of the bins and the elements of the detector matrix  $D$  are in very good approximation normally distributed, even if the elements of  $D$  are computed as a quotient of normally distributed random variables. Because of this, the differences*

$$d_j := o_j - \sum_{i=1}^m D_{ji} t_i \Delta x_i$$

*are normally distributed too (c.f. appendix A, section 3). Here  $j$  is the index of one bin of a given distribution, and  $(t_1, \dots, t_m)$  as well as  $(\Delta x_1, \dots, \Delta x_m)$  are  $m$ -tuples of real numbers. The former will be a candidate for the unfolded distribution, while the latter is the set of bin widths used for the distribution under observation.*

*If  $(t_1, \dots, t_m)$  are the values of the true distribution, the mean values of these differences are zero. This hypothesis can be tested by a  $\chi^2$  test.*

We neglect correlations between the data values  $o_j$  and the elements of the detector matrix, even if they can appear, because the same model that was used for the calculation of the detector matrix is used for correcting the data. The correlations between the elements of  $D$  are neglected too. On the basis of these approximations, and with the hypothesis that the mean values of all differences vanish, these differences  $d_j$  are independent and normally distributed random variables with known widths and vanishing mean values. Consequently the following random variables are in good approximation  $\chi^2$  distributed with  $n$  (more exactly  $(n-1)$ ) degrees of freedom. (c.f. [Bo,91]):

$$\chi_{+s}^2 := \sum_{j=1}^n \frac{\left\{ o_j - \sum_{i=1}^m D_{ji} t_i \Delta x_i \right\}^2}{\sigma_j^2 + \sum_{i=1}^m (\sigma_{ji} t_i \Delta x_i)^2}, \quad \chi_{-s}^2 := \sum_{j=1}^n \left\{ \frac{o_j - \sum_{i=1}^m D_{ji} t_i \Delta x_i}{\sigma_j} \right\}^2$$

The second expression is applicable if the widths of the elements of the detector matrix are negligible compared to the  $\sigma_j$ , the widths of the data values  $o_j$ .  $\sigma_{j_i}$  are the widths of the elements of the detector matrix  $D_{j_i}$ . (The calculation of the latter is discussed in section 3.5.2.1). For a model independent detector matrix the sensitivity of this test is demonstrated in appendix A, section 4. The following table shows the values of the first test quantity for the event properties that are measured together with an upper limit  $\chi^2_{\max}$  of a 5% confidence level for the  $\chi^2$  distribution with  $n-1$  degrees of freedom (taken from [Br,87]). If the hypothesis is correct, the test quantity should therefore lie beneath this value with a probability of 95%. Again we used *ALEPH* data from 1992 and the simulation *HVFL03*. The statistics were  $s_{sim} := 1186173$  and  $s_{dat} := 571825$ .

<i>distribution</i>	<i>n</i>	$\chi^2_{\max}$	$\chi^2$
<i>S</i>	24	35.2	23.85
<i>A</i>	17	26.3	22.32
<i>P</i>	21	31.4	14.86
<i>C</i>	24	35.2	38.35
<i>T</i>	22	32.7	23.30
<i>M</i>	22	32.7	39.93
<i>m</i>	18	27.6	36.14
<i>O</i>	20	30.1	12.19
$\overline{M}_h^2$	23	33.9	34.09
$\overline{M}_l^2$	18	27.6	49.49
$\overline{M}_d^2$	22	32.7	80.42
$B_l$	20	30.1	34.21
$B_w$	17	26.3	37.56
$-\ln(y_3)$	14	22.4	59.89

Table 3.7: Test quantities for event properties

Table 3.7 indicate the quantities  $\overline{M}_l^2$ ,  $\overline{M}_d^2$ ,  $B_w$  and  $-\ln(y_3)$  as dangerous, because the value of the test quantity is too high. If we remember the very high sensitivity of this test, these deviations are not dramatic, even if they appear. We should also remember that measurements are given with errors, and the test

quantity can be better for other values inside this error bars. An optimization could for example be done by a maximization of the  $\chi^2$  probability inside this error bars.

If the detector matrix is model dependent, this test does not proof the correctness of the unfolded data, but it is only a sort of check for consistency of the unfolding procedure. In the case of entropy methods, a dependence of  $D$  from the *QCD* part of the simulation leads to a dependence of the unfolded data from this *QCD*-model.

We saw that the elements of the correction matrix are dependent on the simulation, even in the case of a vanishing model dependence of the detector matrix  $D$  (c.f. section 3.3.1.3). This might indicate that indeed the dependence of  $C$  is much stronger than that of  $D$ . In fact both matrices, the detector matrix  $D$  and the correction matrix  $C$  seem to be dependent on the used *QCD*-model. If one wants to use the above method, and one is not only interested in presenting some ideas as we were here, he has to test the model independence of  $D$  in detail. Because we do not use this assumption in any other part of this work<sup>1</sup>, we do not go into more details here.

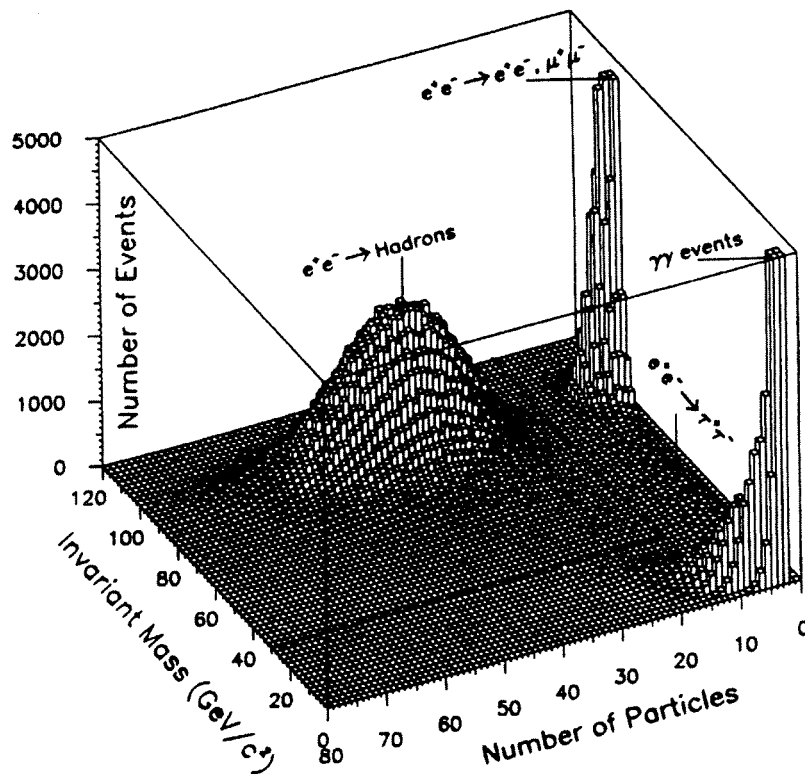
---

<sup>1</sup> Except some discussions in appendix A.

### 3.5. Measurement

#### 3.5.1. Hadronic events and critical tracks

We are only interested in multi-hadronic events, that means, events where the  $Z$  particle decays into a quark-antiquark pair which converts into observable hadrons. Therefore all non-hadronic events must be excluded in a first step. In the figure beneath, a histogram is given which reflects typical properties of hadronic events, and which can therefore be used as a basis for cuts which reject leptonic and  $\gamma\gamma$  events. Note that in this histogram both charged and neutral particles are included, and the that cut-values can therefore not be taken directly. This histogram is taken from [A4,94].



The dashed lines show the preferred region of hadronic events above a number of 10 particles and above an energy of 40 GeV.

A second point that has a negative influence on the measurement is the fact that there exist "blind regions" in every detector. For example it is impossible to detect a particle which propagates along the beam pipe. And it is also impossible to reconstruct the track of a particle with very few hits in the TPC with a

satisfactory accuracy. Tracks that can not be measured, or can only be measured with a bad accuracy are called "*critical tracks*", and are excluded from the analysis.

We are also not interested in particles which do not originate from the main process. These are particles that are generated by the decays of other long-lived particles, and do not directly characterize the hadronic event.

A track of a charged particle is accepted for the further analysis if the following conditions are fulfilled.

$$N_{TPC} \geq 4, \quad d_0 \leq 2cm, \quad z_0 \leq 5cm$$

$$p_t \geq 0.2GeV, \quad 20^\circ \leq \vartheta \leq 160^\circ$$

Here  $N_{TPC}$  is the number of hits in the TPC; this cut makes a good reconstruction of a charged track possible,  $d_0$  is the minimum distance of a track to the beam axis (which is the  $z$ -axis by definition),  $z_0$  is the  $z$  coordinate at the point which corresponds to  $d_0$ . Both cuts eliminate particles which do not come from the interaction point.  $p_t$  is the transverse (that means perpendicular to the  $z$ -axis) component of the particle's momentum and the cut rejects particles which lie in a critical region of the detector, and  $\vartheta$  is the angle between the momentum of the particle and the  $z$ -axis. Only tracks of charged particles are taken into account.

An event is accepted if it passes the following cuts:

$$N_{TRACKS} \geq 5, \quad E_{ch} \geq 15GeV, \quad 35^\circ \leq \vartheta_s \leq 145^\circ$$

Here  $N_{TRACKS}$  is the number of accepted charged tracks, the corresponding cut excludes leptonic events.  $E_{ch}$  is the whole charged energy, the cut rejects  $\gamma\gamma$  events.  $\vartheta_s$  is the angle between the sphericity axis and the beam direction, this cut ensures that the event is well contained in the detector.

Events which are believed to be  $\tau\bar{\tau}$  pairs are excluded in addition. For that reason all the events which contain 5 or 6 charged particles are split into two hemispheres with respect to the thrust axis. In each of the hemispheres the invariant mass is calculated by the resulting four momentum, and the event is rejected from the further analysis if both invariant masses are less than the  $\tau$  mass.

All the event cuts were passed by

571825 hadronic events taken in  
1992

and these were used for the further analysis.

### 3.5.2. ISR and cut corrections

After the correction for detector effects, an additional correction factor analogous to (3.6) was introduced. Two things had to be corrected. We are interested in the reaction  $e^-e^+ \rightarrow Z^0 \rightarrow \text{hadrons}$ , and we should therefore correct for a possible initial state radiation (*ISR*). The second effect is due to the rather arbitrarily chosen event cuts.

The calculation of these correction factors was again done by the simulation model *HVFL03*. "True" distributions  $t_i^{sim}$  were calculated without any cuts, without detector influences, and without initial state bremsstrahlung. Then the corresponding "biased" distributions  $b_j^{sim}$  were calculated by introducing all the cuts behind the detector, and including *ISR*. The correction factors are

$$\bar{C}_j := \frac{t_j^{sim}}{b_j^{sim}}$$

and all unfolded distributions were corrected by

$$t_j = C_j t_j^{corr}$$

### 3.5.2.1. Detector matrix and statistical widths

The principle of the following measurement is demonstrated in the next example using the simple model introduced in example 3.1. The estimation of the statistical widths is given in example 3.2.

#### *Example 3.5: Correction matrix and correction factors*

*Idea: The elements of the correction matrix for event properties, the correction factors for single particle properties and for the correction of ISR as well as cut effects should be calculated. The corresponding widths should be estimated by using linear error propagation*

---

#### *(i) Correction of event properties*

Let us look at one of the distributions of event properties. We use a 2-dimensional histogram where the "true-value" (this is the value without detector influences) is drawn on the x-axis, while the "reconstructed value" (that is the distribution with detector influences) is drawn on the y-axis as starting point. The contents of this histogram are saved in a matrix H. Each event enters once in this histogram, and the following formulas holds:

$$P(\Delta y_j \wedge \Delta x_i) \cong \frac{H_{ji}}{s}, \quad P(\Delta y_j) \cong \frac{1}{s} \sum_{i=1}^m H_{ji}$$

The values of both expressions are normal distributed in good approximation. This is also true for the elements of the correction matrix

$$C_{ij} = \frac{P(\Delta x_i / \Delta y_j)}{\Delta x_i} = \frac{P(\Delta y_j \wedge \Delta x_i)}{P(\Delta y_j) \Delta x_i} \cong \frac{H_{ji}}{\Delta x_i \sum_{k=1}^m H_{jk}}$$

The width of  $P(\Delta y_j \wedge \Delta x_i)$  is given by formula (3.4), the width of  $P(\Delta y_j)$  follows from example 1 in appendix A. The width of the element  $C_{ij}$  is calculated by linear error propagation:

$$x := H_{ji}, \quad y := \sum_{\substack{k=1 \\ k \neq j}}^m H_{jk} \Rightarrow C_{ji} \equiv \frac{x}{\Delta x_i (x+y)} \Rightarrow \sigma_{ji} \equiv \frac{1}{\Delta y_j} \frac{\sqrt{(x\sigma_y)^2 + (y\sigma_x)^2}}{(x+y)^2}$$

The same principle leads to the width of the unfolded distribution:

$$t_i := \sum_{j=1}^n C_{ij} o_j \Delta y_j \Rightarrow \Delta t_i^{stat} \equiv \sqrt{\sum_{j=1}^n \left\{ (\Delta C_{ij} o_j \Delta y_j)^2 + (C_{ij} \Delta o_j \Delta y_j)^2 \right\}}$$

(ii) Correction by correction factors

The correction of single particle distributions, initial state bremsstrahlung and the effect of event cuts is corrected by introducing correction factors. This factor in the  $j$ -th bin is

$$C_j := \frac{n_j}{d_j}$$

where  $n_j$  is used to sign the nominator while  $d_j$  is the denominator of this coefficient. The width of both are calculated in a way that is analogous to formula (3.4), while the width of the factor is calculated by linear error propagation:

$$\Delta C_j = \sqrt{\left( \frac{\Delta n_j}{d_j} \right)^2 + \left( \frac{n_j \Delta d_j}{d_j^2} \right)^2}$$

The correction is done by  $t_j = C_j o_j$ , and the width of the corrected value is again given by linear error propagation:

$$\Delta t_j = \sqrt{(\Delta C_j o_j)^2 + (C_j \Delta o_j)^2}$$

o



### 3.5.3. Systematic errors

Two kinds of systematic errors are calculated. The first one is the "*model bias*", which reflects the fact that the unfolding procedure is slightly model dependent. The second type is the "*cut-systematic*" error. This second type is due to the fact, that the results of the measurements change slightly if different cut values (as introduced in section 3.5.1) are used.

#### 3.5.3.1. Model Bias

In order to estimate the dependence of the corrected distribution on the different *QCD*-parts of the simulation, a simplified variant of the correction procedure is used for different *QCD* models. The deviations in every bin are observed, and the maximum is taken as the "*model bias*".

In this analysis, the models *JETSET 7.4* in the version of anisotropic and in the version of isotropic decaying gluons, *HERWIG 5.8* and *ARIADNE 4.05* with the parameter settings given in appendix A, section 5 were used. In a first step, events were generated by each of the *QCD*-models without initial state radiation. In a next step, the simplified detector simulation, and the same track and event cuts as for real data were applied. This gave true distributions (without *ISR*) and approximations of reconstructed distributions. On that basis, correction matrices, correction coefficients and approximative unfolded distributions were calculated. We used a statistics of  $s_{sim} := 2 \cdot 10^6$  for each of the models.

**Simplified detector simulation:** Because the full detector simulation leads to calculations which are too time consuming, we use a simplified version for the calculation of the model bias. The principle is to provide an algorithm which rejects particles that can hardly be detected by *ALEPH*, and to smear the momenta of the remaining particles according to the resolution of the detector. For this resolution, the value of  $\Delta p \cong 0.0008 p^2$  is used. In a first step all unstable particles are rejected, because they decay before they can be detected. Unstable in this context refers to all particles which decay via the strong or electromagnetic interactions and to weakly decaying charm and bottom hadrons. In a next step, the following cuts are introduced, to give a coarse approximation of the influences of the beam pipe  $p_t < 0.2 GeV$ ,  $\vartheta < 20^\circ$ . In a last step, all particles that were produced by a decay outside the *TPC* are rejected. Finally weak decays of strange

particles are simulated for charged tracks with  $d_0 > 2\text{cm}$  by rejecting them from further analysis. Source of this simplified detector simulation: [GR,95]

### 3.5.3.2. Cut-systematic

The procedure for calculating cut-systematic errors is analogous to the calculation of the model bias. Here not different QCD-models, but different values of the cuts introduced in section 3.5.1 were used ( $N_{TPC} \geq 7$ ,  $d_0 \leq 1\text{cm}$ ,  $p_t \geq 0.3\text{GeV}$ ,  $30^\circ \leq \vartheta \leq 150^\circ$ ,  $N_{TRACKS} \geq 8$ ,  $45^\circ \leq \vartheta_s \leq 135^\circ$ ) one at a time. Again the deviations in every bin were observed, and the maximum values taken as the cut-systematic. This part of the analysis was done by Univ. Doz. Dr. G. Rudolph.

### 3.5.3.3. Remarks about the systematic errors

In most of the bins the systematic error is bigger than the statistical one. The systematic error is typical in the order of a few percent, even if some regions exist, where this error is bigger (for example in the first bin of the normalized particle momentum). For event properties the statistical error is also typical in the size of a few percent, while it is decreased to a few permille in the case of single particle distributions. The dominant part of the systematic error is in most cases the model bias. The main contribution to the model bias comes from the difference between the corrected distributions calculated with the help of *JETSET* and *HERWIG*. Because of all that, it will hardly be possible to reach any further improvement of the measurement by only using higher statistics.

The systematic errors given in the tables of chapter 4 show a fluctuating behavior due to the finite statistics used to calculate them. This is the reason for the peaks appearing in the shape of "(model-data)/error" shown in section 7.3. One possibility is to apply some smoothing criterion to suppress these unwanted fluctuations. Because such a criterion would include arbitrariness into the measurement, and because it turned out that the usage of such a criterion leads to estimations of the model parameters which are in "one sigma agreement" with the results shown in chapter 7, we gave the systematic errors without smoothing.

## Chapter 4.

### Results I: Distributions of charged particles in hadronic *ALEPH* events

#### 4.1. Abstract

In this chapter the results of the first part of this work are given. Here  $g_u$  always indicates the lower limit of a bin, while  $g_o$  is the corresponding upper limit<sup>1</sup>.  $\Delta t$  is the error computed by a quadratic sum of systematic errors and statistical widths.  $\Delta t_{sta}$  is the statistical width,  $\Delta t_m$  the model bias and  $\Delta t_c$  the cut-systematic. The systematic error of the measurement  $\Delta t_{sys}$  is again calculated by a quadratic sum

$$\Delta t_{sys} := \sqrt{\Delta t_m^2 + \Delta t_c^2}$$

In a first part all event properties that were measured in this work are listed. The second part is devoted to single particle distributions.

---

<sup>1</sup> The indices "u" and "o" are abbreviations for the german words "unten" (for the lower border) and "oben" (for the upper border).

## 4.2. Event properties

### 4.2.1. Sphericity

$g_u$	$g_o$	t	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{vis}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	<b>12.362</b>	<b>.410</b>	.082	.402	.396	.068
.0050	.0100	<b>23.328</b>	<b>.248</b>	.109	.223	.213	.066
.0100	.0150	<b>20.227</b>	<b>.158</b>	.100	.122	.114	.043
.0150	.0200	<b>16.691</b>	<b>.125</b>	.091	.085	.075	.041
.0200	.0250	<b>13.410</b>	<b>.103</b>	.081	.063	.058	.024
.0250	.0300	<b>10.788</b>	<b>.097</b>	.072	.065	.052	.039
.0300	.0350	<b>8.870</b>	<b>.135</b>	.066	.118	.118	.010
.0350	.0400	<b>7.408</b>	<b>.081</b>	.060	.054	.053	.012
.0400	.0500	<b>5.922</b>	<b>.115</b>	.038	.109	.101	.041
.0500	.0600	<b>4.508</b>	<b>.053</b>	.033	.041	.040	.010
.0600	.0800	<b>3.258</b>	<b>.023</b>	.020	.012	.004	.011
.0800	.1000	<b>2.317</b>	<b>.023</b>	.017	.016	.013	.009
.1000	.1200	<b>1.742</b>	<b>.040</b>	.015	.037	.036	.006
.1200	.1600	<b>1.211</b>	<b>.014</b>	.009	.011	.010	.003
.1600	.2000	<b>.813</b>	<b>.014</b>	.007	.012	.012	.004
.2000	.2500	<b>.563</b>	<b>.013</b>	.005	.012	.012	.001
.2500	.3000	<b>.397</b>	<b>.010</b>	.004	.009	.009	.003
.3000	.3500	<b>.290</b>	<b>.012</b>	.004	.011	.011	.002
.3500	.4000	<b>.222</b>	<b>.004</b>	.003	.003	.002	.002
.4000	.5000	<b>.148</b>	<b>.005</b>	.002	.004	.004	.001
.5000	.6000	<b>.086</b>	<b>.002</b>	.001	.002	.002	.001
.6000	.7000	<b>.045</b>	<b>.002</b>	.001	.001	.001	.001
.7000	.8000	<b>.012</b>	<b>.001</b>	.001	.001	.000	.000
.8000	.9000	<b>.001</b>	<b>.000</b>	.000	.000	.000	.000

## 4.2.2. Aplanarity

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{mu}$	$\Delta t_{\nu\bar{\nu}}$	$\Delta t_m$	$\Delta t_c$
.0000	.0025	<b>78.493</b>	<b>2.382</b>	.269	2.367	2.352	.271
.0025	.0050	<b>85.976</b>	<b>1.260</b>	.275	1.229	1.220	.151
.0050	.0075	<b>58.231</b>	<b>.497</b>	.231	.440	.433	.081
.0075	.0100	<b>39.489</b>	<b>.362</b>	.192	.307	.302	.052
.0100	.0150	<b>24.016</b>	<b>.222</b>	.106	.196	.191	.042
.0150	.0200	<b>13.464</b>	<b>.179</b>	.080	.160	.155	.040
.0200	.0300	<b>6.912</b>	<b>.098</b>	.041	.089	.087	.018
.0300	.0400	<b>3.285</b>	<b>.044</b>	.028	.034	.019	.028
.0400	.0600	<b>1.438</b>	<b>.024</b>	.013	.020	.019	.007
.0600	.0800	<b>.590</b>	<b>.016</b>	.009	.013	.012	.005
.0800	.1000	<b>.291</b>	<b>.011</b>	.006	.009	.002	.009
.1000	.1200	<b>.156</b>	<b>.008</b>	.005	.006	.006	.002
.1200	.1400	<b>.089</b>	<b>.005</b>	.003	.004	.003	.002
.1400	.1600	<b>.054</b>	<b>.004</b>	.003	.003	.003	.002
.1600	.2000	<b>.028</b>	<b>.003</b>	.001	.003	.002	.002
.2000	.2500	<b>.010</b>	<b>.002</b>	.001	.001	.001	.001
.2500	.3000	<b>.003</b>	<b>.001</b>	.000	.001	.001	.000

### 4.2.3. Planarity

$g_u$	$g_e$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	65.366	.618	.163	.596	.585	.117
.0050	.0100	32.635	.168	.124	.113	.064	.093
.0100	.0150	17.965	.107	.094	.052	.044	.027
.0150	.0200	11.990	.099	.077	.061	.047	.039
.0200	.0250	8.655	.076	.065	.039	.036	.014
.0250	.0300	6.899	.088	.058	.065	.054	.038
.0300	.0350	5.522	.080	.052	.061	.056	.024
.0350	.0400	4.591	.061	.047	.040	.016	.036
.0400	.0500	3.647	.049	.030	.038	.033	.019
.0500	.0600	2.754	.045	.026	.038	.036	.012
.0600	.0800	2.003	.021	.016	.014	.012	.006
.0800	.1000	1.369	.037	.013	.035	.035	.003
.1000	.1200	1.005	.017	.011	.013	.013	.002
.1200	.1600	.678	.014	.006	.012	.012	.003
.1600	.2000	.427	.012	.005	.011	.008	.008
.2000	.2500	.276	.006	.004	.005	.003	.003
.2500	.3000	.175	.007	.003	.007	.006	.002
.3000	.3500	.105	.003	.002	.002	.001	.002
.3500	.4000	.062	.002	.002	.002	.001	.001
.4000	.4500	.026	.001	.001	.000	.000	.000
.4500	.5000	.004	.001	.000	.001	.001	.000

## 4.2.4. C-Parameter

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.0400	.401	.058	.005	.058	.058	.005
.0400	.0800	2.490	.049	.011	.048	.047	.008
.0800	.1200	3.701	.091	.014	.090	.089	.012
.1200	.1600	3.323	.064	.014	.063	.061	.014
.1600	.2000	2.613	.034	.012	.032	.031	.008
.2000	.2400	2.065	.022	.011	.019	.018	.006
.2400	.2800	1.666	.019	.010	.017	.016	.005
.2800	.3200	1.368	.013	.009	.009	.008	.003
.3200	.3600	1.142	.010	.008	.006	.003	.005
.3600	.4000	.981	.014	.008	.012	.012	.003
.4000	.4400	.842	.012	.007	.009	.009	.002
.4400	.4800	.732	.015	.007	.014	.014	.001
.4800	.5200	.639	.018	.006	.017	.017	.002
.5200	.5600	.552	.013	.006	.011	.011	.004
.5600	.6000	.490	.010	.005	.009	.007	.005
.6000	.6400	.432	.013	.005	.012	.012	.005
.6400	.6800	.382	.008	.005	.006	.005	.003
.6800	.7200	.350	.010	.004	.009	.008	.004
.7200	.7600	.316	.009	.005	.007	.007	.003
.7600	.8000	.274	.007	.006	.004	.003	.003
.8000	.8400	.169	.009	.006	.006	.004	.005
.8400	.8800	.080	.006	.005	.004	.003	.002
.8800	.9200	.034	.005	.005	.003	.002	.001
.9200	1.0000	.006	.003	.003	.001	.000	.001

### 4.2.5. 1-Thrust

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	<b>1.017</b>	<b>.180</b>	.022	.178	.177	.018
.0050	.0100	<b>6.035</b>	<b>.277</b>	.054	.272	.263	.068
.0100	.0150	<b>12.437</b>	<b>.182</b>	.077	.165	.147	.076
.0150	.0200	<b>16.071</b>	<b>.268</b>	.086	.254	.250	.041
.0200	.0250	<b>16.454</b>	<b>.272</b>	.087	.257	.253	.047
.0250	.0300	<b>15.246</b>	<b>.300</b>	.084	.288	.286	.037
.0300	.0350	<b>13.380</b>	<b>.161</b>	.079	.141	.132	.050
.0350	.0400	<b>11.582</b>	<b>.144</b>	.073	.123	.109	.057
.0400	.0500	<b>9.346</b>	<b>.076</b>	.047	.059	.051	.031
.0500	.0600	<b>7.159</b>	<b>.124</b>	.041	.117	.112	.035
.0600	.0800	<b>5.088</b>	<b>.065</b>	.025	.060	.059	.012
.0800	.1000	<b>3.427</b>	<b>.025</b>	.020	.015	.012	.009
.1000	.1200	<b>2.482</b>	<b>.024</b>	.017	.017	.016	.006
.1200	.1400	<b>1.847</b>	<b>.046</b>	.015	.043	.043	.005
.1400	.1600	<b>1.390</b>	<b>.035</b>	.013	.033	.032	.005
.1600	.1800	<b>1.072</b>	<b>.019</b>	.011	.016	.011	.011
.1800	.2000	<b>.847</b>	<b>.024</b>	.010	.022	.022	.004
.2000	.2500	<b>.566</b>	<b>.015</b>	.005	.014	.013	.006
.2500	.3000	<b>.307</b>	<b>.005</b>	.004	.004	.002	.003
.3000	.3500	<b>.125</b>	<b>.003</b>	.002	.002	.002	.001
.3500	.4000	<b>.018</b>	<b>.002</b>	.001	.002	.002	.000
.4000	.4500	<b>.001</b>	<b>.000</b>	.000	.000	.000	.000



## 4.2.6. Major

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{su}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0200	.0400	.252	.036	.006	.035	.035	.005
.0400	.0500	1.370	.076	.018	.073	.072	.015
.0500	.0600	2.842	.040	.027	.030	.028	.012
.0600	.0700	4.253	.061	.033	.051	.049	.015
.0700	.0800	5.329	.081	.036	.072	.071	.015
.0800	.1000	6.083	.060	.027	.054	.052	.014
.1000	.1200	5.833	.065	.026	.059	.059	.007
.1200	.1400	4.855	.056	.024	.050	.050	.004
.1400	.1600	3.991	.055	.022	.050	.046	.020
.1600	.2000	3.019	.023	.013	.019	.018	.005
.2000	.2400	2.154	.014	.011	.007	.005	.005
.2400	.2800	1.595	.024	.010	.022	.022	.003
.2800	.3200	1.195	.018	.008	.016	.016	.004
.3200	.3600	.902	.016	.007	.014	.013	.005
.3600	.4000	.683	.018	.006	.017	.017	.003
.4000	.4400	.513	.014	.005	.013	.012	.001
.4400	.4800	.376	.008	.004	.006	.003	.005
.4800	.5200	.271	.006	.004	.004	.003	.003
.5200	.5600	.184	.005	.003	.004	.003	.002
.5600	.6000	.106	.003	.002	.003	.002	.002
.6000	.6400	.042	.003	.001	.003	.003	.001
.6400	.7000	.005	.001	.000	.000	.000	.000

Event properties

4.2.7. Minor

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{sta}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.0200	.178	.030	.005	.030	.029	.006
.0200	.0400	3.237	.153	.020	.151	.149	.026
.0400	.0500	8.107	.101	.044	.091	.087	.025
.0500	.0600	10.451	.097	.049	.084	.083	.010
.0600	.0700	11.274	.100	.051	.086	.085	.014
.0700	.0800	10.892	.066	.050	.044	.043	.011
.0800	.1000	9.003	.059	.031	.050	.048	.015
.1000	.1200	6.208	.063	.026	.057	.056	.014
.1200	.1400	4.043	.031	.021	.023	.023	.004
.1400	.1600	2.536	.027	.017	.021	.021	.004
.1600	.2000	1.299	.020	.009	.018	.017	.005
.2000	.2400	.526	.011	.006	.009	.009	.002
.2400	.2800	.221	.006	.004	.005	.005	.002
.2800	.3200	.095	.005	.002	.005	.003	.003
.3200	.3600	.037	.003	.002	.003	.002	.002
.3600	.4000	.014	.002	.001	.002	.001	.001
.4000	.4500	.004	.001	.000	.001	.001	.000
.4500	.5000	.001	.000	.000	.000	.000	.000

## 4.2.8. Oblateness

$g_u$	$g_v$	$t$	$\Delta t$	$\Delta t_{\text{no}}$	$\Delta t_{\text{pr}}$	$\Delta t_m$	$\Delta t_c$
.0000	.0200	<b>7.480</b>	<b>.056</b>	.030	.047	.042	.020
.0200	.0400	<b>10.798</b>	<b>.072</b>	.035	.063	.056	.029
.0400	.0500	<b>8.679</b>	<b>.072</b>	.045	.057	.051	.024
.0500	.0600	<b>7.107</b>	<b>.064</b>	.041	.049	.044	.021
.0600	.0700	<b>5.816</b>	<b>.059</b>	.037	.046	.039	.025
.0700	.0800	<b>4.776</b>	<b>.042</b>	.033	.026	.022	.013
.0800	.1000	<b>3.775</b>	<b>.026</b>	.021	.015	.013	.008
.1000	.1200	<b>2.841</b>	<b>.042</b>	.018	.038	.037	.008
.1200	.1400	<b>2.243</b>	<b>.023</b>	.016	.017	.016	.005
.1400	.1600	<b>1.783</b>	<b>.023</b>	.014	.018	.018	.001
.1600	.2000	<b>1.315</b>	<b>.020</b>	.009	.018	.015	.009
.2000	.2400	<b>.901</b>	<b>.014</b>	.007	.012	.006	.011
.2400	.2800	<b>.621</b>	<b>.016</b>	.006	.015	.012	.010
.2800	.3200	<b>.430</b>	<b>.014</b>	.005	.013	.013	.005
.3200	.3600	<b>.292</b>	<b>.007</b>	.004	.005	.003	.004
.3600	.4000	<b>.189</b>	<b>.005</b>	.003	.004	.001	.003
.4000	.4400	<b>.117</b>	<b>.005</b>	.002	.004	.003	.002
.4400	.4800	<b>.064</b>	<b>.002</b>	.002	.001	.001	.001
.4800	.5200	<b>.027</b>	<b>.002</b>	.001	.002	.001	.001
.5200	.6000	<b>.004</b>	<b>.001</b>	.000	.001	.001	.000

4.2.9. Heavy jet mass normalized to the visible energy

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{vis}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	1.011	.214	.022	.213	.213	.015
.0050	.0100	7.656	.820	.060	.818	.816	.051
.0100	.0150	15.902	.694	.085	.689	.682	.097
.0150	.0200	19.363	.910	.093	.905	.901	.083
.0200	.0250	18.735	.436	.091	.427	.422	.060
.0250	.0300	16.515	.368	.086	.358	.349	.078
.0300	.0350	14.073	.347	.080	.337	.331	.065
.0350	.0400	11.891	.208	.073	.195	.189	.047
.0400	.0500	9.422	.286	.047	.282	.280	.033
.0500	.0600	7.013	.276	.040	.273	.273	.015
.0600	.0800	4.839	.145	.024	.143	.143	.007
.0800	.1000	3.125	.055	.019	.051	.051	.007
.1000	.1200	2.138	.031	.015	.027	.025	.011
.1200	.1400	1.506	.026	.013	.022	.019	.012
.1400	.1600	1.089	.034	.011	.032	.030	.012
.1600	.1800	.784	.016	.009	.014	.007	.012
.1800	.2000	.579	.015	.008	.013	.009	.010
.2000	.2500	.347	.007	.004	.006	.003	.006
.2500	.3000	.148	.007	.002	.007	.006	.003
.3000	.3500	.055	.003	.001	.003	.003	.001
.3500	.4000	.018	.001	.001	.000	.000	.000
.4000	.4500	.006	.001	.000	.001	.001	.000
.4500	.5000	.002	.001	.000	.001	.001	.000

4.2.10. Light jet mass normalized to the visible energy

$g_u$	$g_o$	$t$	$\Delta t$	$\Delta t_{\mu\mu}$	$\Delta t_{\nu\nu}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	30.888	1.787	.113	1.784	1.778	.148
.0050	.0100	47.599	2.739	.135	2.736	2.735	.054
.0100	.0150	38.540	1.220	.124	1.214	1.213	.049
.0150	.0200	25.948	.908	.105	.902	.901	.046
.0200	.0250	16.777	.576	.085	.570	.569	.034
.0250	.0300	11.127	.505	.070	.500	.500	.022
.0300	.0350	7.606	.431	.058	.427	.427	.012
.0350	.0400	5.401	.418	.049	.415	.415	.007
.0400	.0450	3.909	.227	.042	.223	.223	.012
.0450	.0500	2.917	.205	.036	.202	.201	.019
.0500	.0600	1.921	.102	.021	.099	.097	.021
.0600	.0700	1.131	.040	.016	.036	.034	.012
.0700	.0800	.662	.033	.012	.030	.030	.005
.0800	.0900	.394	.014	.009	.010	.010	.002
.0900	.1000	.242	.016	.007	.014	.014	.002
.1000	.1200	.112	.011	.004	.011	.011	.001
.1200	.1400	.028	.004	.002	.004	.004	.001
.1400	.1600	.003	.001	.001	.001	.001	.000

4.2.11. Mass difference normalized to the visible energy

$g_u$	$g_e$	$t$	$\Delta t$	$\Delta t_{sta}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	31.263	.297	.116	.274	.245	.123
.0050	.0100	27.042	.363	.107	.347	.340	.070
.0100	.0150	21.946	.161	.097	.128	.108	.070
.0150	.0200	17.456	.293	.088	.280	.272	.065
.0200	.0250	13.863	.185	.078	.167	.159	.053
.0250	.0300	11.166	.201	.071	.188	.184	.039
.0300	.0350	9.161	.252	.064	.244	.242	.030
.0350	.0400	7.609	.091	.058	.070	.066	.024
.0400	.0500	5.970	.107	.037	.101	.100	.010
.0500	.0600	4.413	.064	.031	.056	.055	.009
.0600	.0800	3.036	.041	.018	.036	.035	.009
.0800	.1000	1.969	.039	.015	.036	.035	.009
.1000	.1200	1.347	.039	.012	.037	.035	.013
.1200	.1400	.965	.036	.010	.034	.031	.015
.1400	.1600	.707	.025	.009	.024	.020	.013
.1600	.1800	.515	.013	.007	.011	.004	.010
.1800	.2000	.389	.012	.006	.010	.008	.007
.2000	.2500	.238	.010	.003	.009	.009	.003
.2500	.3000	.104	.005	.002	.005	.004	.002
.3000	.3500	.041	.002	.001	.002	.002	.000
.3500	.4000	.015	.001	.001	.000	.000	.000
.4000	.5000	.003	.001	.000	.001	.001	.000

## 4.2.12. Total jet broadening

$g_u$	$g_v$	$t$	$\Delta t$	$\Delta t_{\text{sig}}$	$\Delta t_{\text{sys}}$	$\Delta t_{\text{in}}$	$\Delta t_{\text{c}}$
.0000	.0200	.060	.012	.003	.012	.011	.003
.0200	.0400	2.505	.133	.018	.132	.128	.031
.0400	.0600	8.211	.120	.031	.116	.115	.018
.0600	.0800	9.512	.060	.033	.051	.048	.014
.0800	.1000	7.426	.102	.029	.098	.097	.016
.1000	.1200	5.508	.052	.026	.045	.044	.012
.1200	.1400	4.143	.028	.022	.016	.014	.008
.1400	.1600	3.171	.035	.019	.029	.028	.006
.1600	.1800	2.442	.043	.017	.040	.039	.007
.1800	.2000	1.896	.037	.015	.034	.033	.004
.2000	.2200	1.471	.039	.013	.036	.035	.011
.2200	.2400	1.135	.027	.011	.024	.024	.004
.2400	.2600	.872	.017	.010	.014	.009	.010
.2600	.2800	.644	.015	.008	.012	.010	.007
.2800	.3000	.480	.011	.007	.009	.009	.001
.3000	.3200	.303	.011	.006	.010	.010	.002
.3200	.3400	.159	.006	.004	.005	.004	.001
.3400	.3600	.056	.008	.002	.008	.008	.001
.3600	.3800	.009	.003	.001	.003	.003	.000
.3800	.4000	.001	.001	.000	.001	.001	.000

4.2.13. Wide jet broadening

$g_u$	$g_e$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.0200	<b>1.129</b>	<b>.127</b>	<b>.012</b>	<b>.127</b>	<b>.125</b>	<b>.020</b>
.0200	.0400	<b>12.310</b>	<b>.145</b>	<b>.037</b>	<b>.140</b>	<b>.137</b>	<b>.030</b>
.0400	.0600	<b>11.937</b>	<b>.138</b>	<b>.036</b>	<b>.134</b>	<b>.129</b>	<b>.037</b>
.0600	.0800	<b>7.653</b>	<b>.074</b>	<b>.030</b>	<b>.068</b>	<b>.065</b>	<b>.020</b>
.0800	.1000	<b>5.151</b>	<b>.040</b>	<b>.025</b>	<b>.032</b>	<b>.030</b>	<b>.010</b>
.1000	.1200	<b>3.614</b>	<b>.073</b>	<b>.020</b>	<b>.070</b>	<b>.069</b>	<b>.008</b>
.1200	.1400	<b>2.606</b>	<b>.046</b>	<b>.017</b>	<b>.043</b>	<b>.042</b>	<b>.010</b>
.1400	.1600	<b>1.869</b>	<b>.044</b>	<b>.014</b>	<b>.041</b>	<b>.040</b>	<b>.012</b>
.1600	.1800	<b>1.336</b>	<b>.025</b>	<b>.012</b>	<b>.022</b>	<b>.013</b>	<b>.017</b>
.1800	.2000	<b>.957</b>	<b>.027</b>	<b>.010</b>	<b>.025</b>	<b>.021</b>	<b>.013</b>
.2000	.2200	<b>.643</b>	<b>.015</b>	<b>.008</b>	<b>.013</b>	<b>.008</b>	<b>.010</b>
.2200	.2400	<b>.418</b>	<b>.014</b>	<b>.007</b>	<b>.012</b>	<b>.011</b>	<b>.005</b>
.2400	.2600	<b>.233</b>	<b>.015</b>	<b>.005</b>	<b>.014</b>	<b>.014</b>	<b>.003</b>
.2600	.2800	<b>.111</b>	<b>.008</b>	<b>.003</b>	<b>.008</b>	<b>.007</b>	<b>.003</b>
.2800	.3000	<b>.039</b>	<b>.003</b>	<b>.002</b>	<b>.003</b>	<b>.003</b>	<b>.001</b>
.3000	.3200	<b>.012</b>	<b>.002</b>	<b>.001</b>	<b>.002</b>	<b>.002</b>	<b>.000</b>
.3200	.3600	<b>.001</b>	<b>.000</b>	<b>.000</b>	<b>.000</b>	<b>.000</b>	<b>.000</b>



4.2.14. Jet resolution parameter

$g_u$	$g_e$	$t$	$\Delta t$	$\Delta t_{su}$	$\Delta t_{sy}$	$\Delta t_m$	$\Delta t_c$
.0000	.0050	97.999	.747	.172	.727	.634	.356
.0050	.0100	29.354	.660	.113	.651	.623	.186
.0100	.0150	14.963	.135	.083	.107	.087	.064
.0150	.0200	9.567	.165	.066	.152	.150	.024
.0200	.0250	6.961	.125	.056	.112	.105	.040
.0250	.0300	5.362	.132	.049	.122	.118	.032
.0300	.0350	4.320	.065	.044	.048	.037	.030
.0350	.0400	3.533	.097	.039	.089	.083	.032
.0400	.0500	2.746	.081	.025	.077	.072	.029
.0500	.0600	2.063	.069	.021	.066	.060	.026
.0600	.0800	1.425	.056	.012	.054	.049	.023
.0800	.1000	.944	.026	.010	.025	.020	.015
.1000	.1200	.656	.018	.008	.016	.011	.011
.1200	.1400	.474	.015	.007	.014	.010	.010
.1400	.1600	.346	.013	.006	.011	.007	.008
.1600	.1800	.258	.010	.005	.008	.005	.007
.1800	.2000	.188	.010	.004	.009	.009	.004
.2000	.2400	.110	.005	.002	.004	.004	.002
.2400	.2800	.042	.002	.001	.001	.001	.001
.2800	.3200	.006	.001	.001	.001	.001	.000

4.2.15. Negative logarithm of the resolution parameter

$g_u$	$g_e$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
1.1	1.5	.0084	.0004	.0002	.0003	.0002	.0002
1.5	2.2	.0483	.0015	.0004	.0014	.0011	.0010
2.2	2.9	.0921	.0028	.0005	.0028	.0023	.0016
2.9	3.6	.1306	.0030	.0006	.0029	.0027	.0011
3.6	4.3	.1623	.0026	.0007	.0025	.0025	.0005
4.3	5.0	.1969	.0023	.0008	.0021	.0019	.0009
5.0	5.7	.2318	.0065	.0009	.0064	.0063	.0011
5.7	6.4	.2442	.0028	.0009	.0027	.0023	.0012
6.4	7.3	.1729	.0031	.0007	.0030	.0029	.0008
7.3	8.0	.0721	.0006	.0005	.0003	.0002	.0003
8.0	8.7	.0200	.0015	.0003	.0015	.0014	.0002
8.7	9.4	.0034	.0007	.0001	.0007	.0007	.0001
9.4	10.3	.0004	.0001	.0000	.0001	.0001	.0000

### 4.3. Single particle properties

#### 4.3.1. Rapidity with respect to the sphericity-axis

$g_i$	$g_o$	t	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.2500	6.5602	.0596	.0141	.0579	.0529	.0235
.2500	.5000	6.5779	.0875	.0131	.0865	.0808	.0308
.5000	.7500	6.5785	.1134	.0124	.1127	.1058	.0388
.7500	1.0000	6.5728	.1250	.0118	.1244	.1178	.0400
1.0000	1.2500	6.5580	.1193	.0112	.1188	.1131	.0363
1.2500	1.5000	6.5098	.0954	.0107	.0948	.0933	.0170
1.5000	1.7500	6.4225	.0744	.0102	.0737	.0734	.0062
1.7500	2.0000	6.2541	.0676	.0097	.0669	.0665	.0080
2.0000	2.2500	6.0128	.0287	.0094	.0272	.0264	.0064
2.2500	2.5000	5.5739	.0169	.0089	.0143	.0121	.0077
2.5000	2.7500	4.9393	.0319	.0084	.0308	.0285	.0117
2.7500	3.0000	4.1547	.0739	.0076	.0735	.0729	.0090
3.0000	3.2500	3.3258	.0840	.0067	.0837	.0835	.0053
3.2500	3.5000	2.4971	.0628	.0057	.0625	.0623	.0048
3.5000	3.7500	1.7914	.0225	.0048	.0220	.0216	.0037
3.7500	4.0000	1.2165	.0247	.0038	.0244	.0240	.0042
4.0000	4.2500	.7817	.0176	.0030	.0174	.0172	.0024
4.2500	4.5000	.4878	.0138	.0023	.0136	.0135	.0017
4.5000	5.0000	.2325	.0093	.0010	.0093	.0092	.0005
5.0000	5.5000	.0738	.0055	.0005	.0054	.0054	.0003
5.5000	6.0000	.0212	.0022	.0003	.0022	.0022	.0002
6.0000	7.0000	.0017	.0003	.0001	.0003	.0003	.0001

4.3.2. In-momentum with respect to the sphericity-axis

$g_u$	$g_c$	$t$	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.1000	48.2961	.2756	.0454	.2718	.1680	.2137
.1000	.2000	38.1415	.4282	.0378	.4265	.4059	.1308
.2000	.3000	28.2065	.1519	.0312	.1487	.1384	.0544
.3000	.4000	20.4453	.0697	.0259	.0647	.0550	.0340
.4000	.5000	15.0471	.0562	.0220	.0517	.0425	.0296
.5000	.6000	11.2240	.0376	.0189	.0325	.0264	.0190
.6000	.7000	8.5220	.0439	.0165	.0406	.0359	.0190
.7000	.8000	6.6020	.0394	.0146	.0366	.0283	.0233
.8000	.9000	5.1800	.0253	.0130	.0216	.0177	.0124
.9000	1.0000	4.1556	.0248	.0117	.0218	.0169	.0138
1.0000	1.2000	3.0453	.0135	.0075	.0113	.0102	.0048
1.2000	1.4000	2.0950	.0138	.0062	.0124	.0113	.0049
1.4000	1.6000	1.4732	.0120	.0052	.0108	.0067	.0085
1.6000	1.8000	1.0809	.0086	.0044	.0073	.0068	.0026
1.8000	2.0000	.8036	.0098	.0038	.0090	.0088	.0021
2.0000	2.5000	.5104	.0055	.0021	.0051	.0047	.0019
2.5000	3.0000	.2753	.0031	.0015	.0027	.0021	.0017
3.0000	3.5000	.1601	.0030	.0011	.0028	.0023	.0015
3.5000	4.0000	.0970	.0015	.0008	.0012	.0011	.0006
4.0000	5.0000	.0493	.0009	.0004	.0008	.0006	.0006
5.0000	6.0000	.0202	.0005	.0003	.0004	.0004	.0002
6.0000	7.0000	.0088	.0002	.0002	.0002	.0001	.0001
7.0000	8.0000	.0040	.0002	.0001	.0002	.0002	.0001
8.0000	10.0000	.0012	.0001	.0000	.0001	.0001	.0000
10.0000	14.0000	.0001	.0000	.0000	.0000	.0000	.0000

4.3.3. Out-momentum with respect to the sphericity-axis

$g_u$	$g_v$	t	$\Delta t$	$\Delta t_{sig}$	$\Delta t_{sys}$	$\Delta t_m$	$\Delta t_c$
.0000	.1000	67.6599	.2819	.0507	.2772	.1664	.2217
.1000	.2000	51.2991	.5622	.0426	.5605	.5421	.1426
.2000	.3000	34.3537	.1814	.0343	.1781	.1619	.0742
.3000	.4000	21.3968	.0777	.0274	.0727	.0640	.0347
.4000	.5000	12.8622	.0532	.0214	.0487	.0391	.0291
.5000	.6000	7.7217	.0373	.0169	.0333	.0281	.0179
.6000	.7000	4.6900	.0202	.0133	.0153	.0130	.0080
.7000	.8000	2.8810	.0151	.0105	.0108	.0100	.0043
.8000	.9000	1.8303	.0144	.0085	.0116	.0089	.0075
.9000	1.0000	1.1981	.0221	.0070	.0209	.0201	.0059
1.0000	1.2000	.6731	.0070	.0040	.0058	.0048	.0032
1.2000	1.4000	.3265	.0043	.0028	.0033	.0030	.0015
1.4000	1.6000	.1687	.0030	.0020	.0023	.0015	.0017
1.6000	1.8000	.0922	.0035	.0015	.0032	.0025	.0019
1.8000	2.0000	.0525	.0027	.0011	.0025	.0023	.0010
2.0000	2.5000	.0223	.0007	.0005	.0005	.0002	.0005
2.5000	3.0000	.0069	.0004	.0003	.0003	.0003	.0002
3.0000	3.5000	.0026	.0003	.0002	.0002	.0002	.0001
3.5000	5.0000	.0004	.0001	.0000	.0000	.0000	.0000

#### 4.3.4. Negative logarithm of the normalized momentum

$g_n$	$g_o$	$t$	$\Delta t$	$\Delta t_{su}$	$\Delta t_{sv}$	$\Delta t_m$	$\Delta t_c$
0.2	0.3	.0447	.0016	.0011	.0012	.0009	.0007
0.3	0.4	.0915	.0024	.0017	.0018	.0017	.0005
0.4	0.5	.1483	.0026	.0021	.0016	.0013	.0009
0.5	0.6	.2283	.0050	.0026	.0043	.0040	.0017
0.6	0.7	.3320	.0068	.0032	.0060	.0059	.0010
0.7	0.8	.4514	.0102	.0037	.0095	.0095	.0009
0.8	0.9	.5971	.0185	.0043	.0180	.0179	.0020
0.9	1.0	.7699	.0152	.0049	.0144	.0143	.0015
1.0	1.1	.9605	.0111	.0054	.0097	.0095	.0021
1.1	1.2	1.1693	.0100	.0060	.0079	.0078	.0016
1.2	1.3	1.4127	.0163	.0066	.0148	.0145	.0029
1.3	1.4	1.6530	.0169	.0072	.0153	.0150	.0031
1.4	1.5	1.9045	.0236	.0077	.0223	.0219	.0044
1.5	1.6	2.1989	.0137	.0083	.0109	.0105	.0030
1.6	1.7	2.4727	.0255	.0088	.0239	.0228	.0071
1.7	1.8	2.7568	.0306	.0093	.0292	.0275	.0096
1.8	1.9	3.0433	.0261	.0097	.0242	.0241	.0027
1.9	2.0	3.3371	.0201	.0102	.0173	.0160	.0065
2.0	2.1	3.6529	.0357	.0107	.0340	.0333	.0069
2.1	2.2	3.9125	.0256	.0111	.0231	.0221	.0065
2.2	2.3	4.2109	.0524	.0115	.0511	.0509	.0051
2.3	2.4	4.5090	.0280	.0120	.0253	.0243	.0071
2.4	2.5	4.7471	.0364	.0122	.0343	.0335	.0072
2.5	2.6	4.9748	.0334	.0125	.0309	.0299	.0080
2.6	2.7	5.2399	.0328	.0129	.0302	.0285	.0100
2.7	2.8	5.4311	.0311	.0131	.0282	.0256	.0117
2.8	2.9	5.6685	.0365	.0134	.0339	.0299	.0160
2.9	3.0	5.8628	.0301	.0137	.0268	.0207	.0170

*Negative logarithm of the normalized momentum*

3.0	3.1	<b>6.0539</b>	<b>.0293</b>	.0139	.0258	.0188	.0177
3.1	3.2	<b>6.1244</b>	<b>.0329</b>	.0140	.0298	.0184	.0235
3.2	3.3	<b>6.2687</b>	<b>.0360</b>	.0141	.0331	.0264	.0199
3.3	3.4	<b>6.3769</b>	<b>.0315</b>	.0143	.0281	.0229	.0162
3.4	3.5	<b>6.4744</b>	<b>.0352</b>	.0145	.0321	.0301	.0112
3.5	3.6	<b>6.4548</b>	<b>.0278</b>	.0145	.0237	.0224	.0079
3.6	3.7	<b>6.5032</b>	<b>.0224</b>	.0145	.0170	.0163	.0048
3.7	3.8	<b>6.4969</b>	<b>.0411</b>	.0146	.0385	.0373	.0093
3.8	3.9	<b>6.4360</b>	<b>.0309</b>	.0145	.0273	.0213	.0170
3.9	4.0	<b>6.3381</b>	<b>.0260</b>	.0144	.0216	.0205	.0070
4.0	4.1	<b>6.1493</b>	<b>.0373</b>	.0141	.0346	.0313	.0147
4.1	4.2	<b>6.0619</b>	<b>.0432</b>	.0141	.0409	.0403	.0071
4.2	4.3	<b>5.8987</b>	<b>.0356</b>	.0139	.0327	.0322	.0060
4.3	4.4	<b>5.6985</b>	<b>.0354</b>	.0137	.0326	.0289	.0150
4.4	4.5	<b>5.5540</b>	<b>.0294</b>	.0137	.0260	.0203	.0162
4.5	4.6	<b>5.1274</b>	<b>.0510</b>	.0131	.0493	.0486	.0085
4.6	4.7	<b>4.7782</b>	<b>.0366</b>	.0127	.0343	.0288	.0187
4.7	4.8	<b>4.6405</b>	<b>.0562</b>	.0130	.0547	.0397	.0377
4.8	4.9	<b>4.1903</b>	<b>.0441</b>	.0124	.0424	.0408	.0113
4.9	5.0	<b>3.7719</b>	<b>.0619</b>	.0118	.0607	.0377	.0476
5.0	5.1	<b>3.3580</b>	<b>.0631</b>	.0112	.0621	.0497	.0373
5.1	5.2	<b>2.9515</b>	<b>.0534</b>	.0106	.0524	.0337	.0401
5.2	5.3	<b>2.5649</b>	<b>.2276</b>	.0102	.2274	.0381	.2242
5.3	5.4	<b>2.1851</b>	<b>.6901</b>	.0103	.6900	.0370	.6890

## 4.3.5. Normalized particle momentum

$g_n$	$g_o$	$t$	$\Delta t$	$\Delta t_{su}$	$\Delta t_{vs}$	$\Delta t_m$	$\Delta t_c$
.0040	.0060	<b>478.5686</b>	<b>47.9679</b>	1.1014	47.9552	7.1858	47.4138
.0060	.0080	<b>535.7911</b>	<b>6.5903</b>	1.0374	6.5079	5.6662	3.2010
.0080	.0100	<b>513.0197</b>	<b>4.6930</b>	.9719	4.5911	3.3371	3.1531
.0100	.0120	<b>478.5918</b>	<b>3.4482</b>	.9138	3.3247	3.1994	.9042
.0120	.0140	<b>440.2027</b>	<b>2.4166</b>	.8665	2.2557	2.0692	.8980
.0140	.0160	<b>403.1236</b>	<b>3.0990</b>	.8239	2.9875	2.9781	.2365
.0160	.0180	<b>364.4388</b>	<b>1.6687</b>	.7756	1.4774	1.2842	.7305
.0180	.0200	<b>329.5144</b>	<b>1.3099</b>	.7297	1.0879	.9857	.4602
.0200	.0250	<b>287.2253</b>	<b>1.3731</b>	.4518	1.2965	1.2425	.3700
.0250	.0300	<b>237.3210</b>	<b>.8724</b>	.4047	.7729	.7422	.2157
.0300	.0350	<b>198.0686</b>	<b>.9013</b>	.3642	.8244	.7354	.3727
.0350	.0400	<b>168.6517</b>	<b>.8421</b>	.3322	.7738	.5711	.5221
.0400	.0450	<b>145.4144</b>	<b>.8286</b>	.3065	.7697	.5514	.5371
.0450	.0500	<b>127.0139</b>	<b>.6198</b>	.2850	.5504	.4087	.3687
.0500	.0600	<b>105.3603</b>	<b>.5752</b>	.1863	.5441	.4581	.2936
.0600	.0700	<b>83.4544</b>	<b>.3954</b>	.1635	.3601	.2958	.2053
.0700	.0800	<b>68.1214</b>	<b>.4001</b>	.1472	.3721	.3554	.1101
.0800	.0900	<b>56.2782</b>	<b>.4376</b>	.1332	.4168	.4081	.0849
.0900	.1000	<b>47.6084</b>	<b>.2902</b>	.1230	.2628	.2534	.0698
.1000	.1100	<b>40.3082</b>	<b>.5303</b>	.1124	.5183	.5154	.0543
.1100	.1200	<b>34.4620</b>	<b>.2306</b>	.1039	.2059	.1972	.0593
.1200	.1300	<b>29.8440</b>	<b>.2726</b>	.0967	.2549	.2523	.0362
.1300	.1400	<b>26.0677</b>	<b>.2418</b>	.0906	.2242	.2171	.0558
.1400	.1600	<b>21.2600</b>	<b>.1175</b>	.0573	.1025	.0920	.0452
.1600	.1800	<b>16.6196</b>	<b>.1472</b>	.0508	.1382	.1330	.0373
.1800	.2000	<b>13.2121</b>	<b>.1174</b>	.0455	.1082	.1057	.0233
.2000	.2250	<b>10.3401</b>	<b>.0851</b>	.0359	.0772	.0722	.0272
.2250	.2500	<b>7.9204</b>	<b>.1026</b>	.0313	.0977	.0973	.0086
.2500	.2750	<b>6.1973</b>	<b>.0553</b>	.0278	.0478	.0470	.0088



Normalized particle momentum

.2750	.3000	<b>4.8886</b>	<b>.0607</b>	.0248	.0555	.0539	.0132
.3000	.3250	<b>3.8449</b>	<b>.0295</b>	.0219	.0198	.0196	.0027
.3250	.3500	<b>3.0554</b>	<b>.0355</b>	.0194	.0297	.0275	.0112
.3500	.3750	<b>2.4761</b>	<b>.0385</b>	.0176	.0342	.0338	.0053
.3750	.4000	<b>1.9779</b>	<b>.0415</b>	.0157	.0384	.0374	.0087
.4000	.4300	<b>1.5496</b>	<b>.0430</b>	.0126	.0411	.0409	.0039
.4300	.4600	<b>1.2117</b>	<b>.0391</b>	.0112	.0375	.0375	.0020
.4600	.4900	<b>.9375</b>	<b>.0172</b>	.0098	.0141	.0137	.0034
.4900	.5200	<b>.7318</b>	<b>.0213</b>	.0086	.0195	.0194	.0022
.5200	.5500	<b>.5672</b>	<b>.0110</b>	.0077	.0078	.0077	.0015
.5500	.6000	<b>.4022</b>	<b>.0110</b>	.0049	.0099	.0096	.0023
.6000	.6500	<b>.2602</b>	<b>.0046</b>	.0040	.0023	.0020	.0012
.6500	.7000	<b>.1721</b>	<b>.0056</b>	.0033	.0045	.0043	.0014
.7000	.7500	<b>.1064</b>	<b>.0038</b>	.0025	.0028	.0027	.0010
.7500	.8000	<b>.0587</b>	<b>.0036</b>	.0017	.0031	.0029	.0012
.8000	.9000	<b>.0262</b>	<b>.0011</b>	.0008	.0008	.0007	.0004
.9000	1.0000	<b>.0047</b>	<b>.0010</b>	.0003	.0009	.0008	.0004

4.3.6. Rapidity with respect to the thrust-axis

$g_u$	$g_v$	$t$	$\Delta t$	$\Delta t_{su}$	$\Delta t_{sv}$	$\Delta t_m$	$\Delta t_c$
.0000	.2500	5.8321	.0735	.0121	.0725	.0680	.0251
.2500	.5000	6.4074	.0815	.0127	.0805	.0754	.0281
.5000	.7500	6.6408	.1051	.0128	.1043	.0991	.0324
.7500	1.0000	6.7260	.1180	.0124	.1173	.1082	.0455
1.0000	1.2500	6.7448	.1221	.0117	.1215	.1150	.0392
1.2500	1.5000	6.7030	.0916	.0110	.0909	.0883	.0215
1.5000	1.7500	6.5778	.0684	.0104	.0676	.0674	.0055
1.7500	2.0000	6.3893	.0684	.0099	.0677	.0676	.0024
2.0000	2.2500	6.1346	.0280	.0095	.0264	.0255	.0066
2.2500	2.5000	5.7211	.0188	.0091	.0164	.0069	.0149
2.5000	2.7500	5.0905	.0354	.0085	.0343	.0317	.0132
2.7500	3.0000	4.3070	.0869	.0078	.0866	.0862	.0083
3.0000	3.2500	3.4274	.0783	.0068	.0781	.0780	.0035
3.2500	3.5000	2.5489	.0636	.0058	.0633	.0629	.0068
3.5000	3.7500	1.7492	.0300	.0047	.0296	.0294	.0035
3.7500	4.0000	1.1072	.0234	.0036	.0231	.0229	.0032
4.0000	4.2500	.6623	.0183	.0027	.0181	.0178	.0031
4.2500	4.5000	.3662	.0113	.0019	.0111	.0109	.0021
4.5000	5.0000	.1403	.0071	.0008	.0071	.0070	.0010
5.0000	5.5000	.0286	.0019	.0003	.0019	.0018	.0004
5.5000	6.0000	.0039	.0006	.0001	.0006	.0006	.0001
6.0000	7.0000	.0003	.0005	.0000	.0005	.0005	.0000

**4.3.7. *In-momentum with respect to the thrust-axis***

$g_u$	$g_e$	t	$\Delta t$	$\Delta t_{ga}$	$\Delta t_{sv}$	$\Delta t_m$	$\Delta t_c$
.0000	.1000	43.9759	.3026	.0428	.2995	.2197	.2036
.1000	.2000	39.1145	.4776	.0382	.4760	.4580	.1298
.2000	.3000	29.3200	.1844	.0316	.1817	.1718	.0592
.3000	.4000	21.2891	.0836	.0263	.0793	.0723	.0325
.4000	.5000	15.4467	.0444	.0220	.0385	.0268	.0276
.5000	.6000	11.4536	.0514	.0191	.0477	.0353	.0321
.6000	.7000	8.5950	.0415	.0166	.0380	.0296	.0237
.7000	.8000	6.6393	.0272	.0147	.0228	.0206	.0099
.8000	.9000	5.1976	.0251	.0131	.0214	.0072	.0201
.9000	1.0000	4.1253	.0204	.0117	.0168	.0124	.0112
1.0000	1.2000	3.0460	.0149	.0075	.0129	.0123	.0039
1.2000	1.4000	2.0969	.0209	.0062	.0200	.0193	.0051
1.4000	1.6000	1.4911	.0103	.0052	.0089	.0058	.0068
1.6000	1.8000	1.0880	.0089	.0044	.0078	.0068	.0039
1.8000	2.0000	.8185	.0073	.0038	.0063	.0057	.0027
2.0000	2.5000	.5231	.0059	.0020	.0055	.0049	.0026
2.5000	3.0000	.2937	.0033	.0015	.0029	.0029	.0003
3.0000	3.5000	.1730	.0026	.0011	.0024	.0023	.0006
3.5000	4.0000	.1074	.0015	.0009	.0012	.0005	.0011
4.0000	5.0000	.0587	.0008	.0005	.0007	.0006	.0003
5.0000	6.0000	.0265	.0005	.0003	.0004	.0003	.0002
6.0000	7.0000	.0129	.0004	.0002	.0003	.0002	.0002
7.0000	8.0000	.0065	.0003	.0001	.0003	.0001	.0003
8.0000	10.0000	.0026	.0001	.0001	.0001	.0001	.0000
10.0000	14.0000	.0004	.0000	.0000	.0000	.0000	.0000

4.3.8. *Out-momentum with respect to the thrust-axis*

$g_u$	$g_e$	t	$\Delta t$	$\Delta t_{st}$	$\Delta t_{out}$	$\Delta t_m$	$\Delta t_c$
.0000	.1000	66.3733	.3211	.0510	.3170	.2224	.2259
.1000	.2000	50.2130	.5157	.0424	.5139	.5019	.1102
.2000	.3000	33.6286	.1726	.0338	.1692	.1611	.0516
.3000	.4000	21.2367	.0909	.0267	.0868	.0741	.0452
.4000	.5000	13.1352	.0493	.0211	.0445	.0320	.0309
.5000	.6000	8.1300	.0461	.0168	.0429	.0213	.0373
.6000	.7000	5.1234	.0355	.0136	.0328	.0260	.0201
.7000	.8000	3.3109	.0229	.0111	.0201	.0187	.0072
.8000	.9000	2.1748	.0128	.0091	.0091	.0040	.0081
.9000	1.0000	1.4769	.0179	.0076	.0162	.0133	.0093
1.0000	1.2000	.8756	.0069	.0044	.0053	.0037	.0038
1.2000	1.4000	.4529	.0057	.0032	.0047	.0036	.0030
1.4000	1.6000	.2484	.0048	.0024	.0041	.0031	.0027
1.6000	1.8000	.1407	.0049	.0018	.0045	.0040	.0021
1.8000	2.0000	.0849	.0031	.0014	.0027	.0023	.0014
2.0000	2.5000	.0387	.0010	.0006	.0007	.0007	.0002
2.5000	3.0000	.0139	.0011	.0004	.0011	.0010	.0004
3.0000	3.5000	.0049	.0004	.0002	.0003	.0002	.0002
3.5000	5.0000	.0011	.0001	.0001	.0000	.0000	.0000

PART II:  
Model Tuning



## Chapter 5.

# Statistical methods for model tuning

### 5.1. Abstract

The major goal of the second part of this work is the tuning of some QCD-models to experimental data. We use a fitting procedure based on a "maximum likelihood estimation" of the free model parameters. This chapter will introduce the main ideas behind this fitting procedure. Due to limited computational power we always use (linear) parametrizations of the model predictions. If the statistical fluctuations of the coefficients of these parametrizations can not be neglected, they have to be included into the estimation procedure. This is quite a tricky topic, especially if correlations between these coefficients play a role. First ideas in this direction are presented.

This chapter will only deal with statistical errors. This is a good base for discussing statistical methods. Nevertheless systematic uncertainties have big influences on the results. The inclusion of systematic errors in the estimation of parameters is a part of the next chapter.

### 5.2. Maximum likelihood estimation

Estimation in the statistical sense means that one tries to make quantitative statements about the properties of a set of given objects by using informations about the same properties of a subset. A typical example of an often estimated

quantity is the number of votes a given party will get during an election. Here the number of votes this party gets in a subset of all citizens is known and a "projected result" is computed. Another example is the average number of charged particles in an *ALEPH*-event. Here one knows the number of charged particles in a finite number of events. Because we are using only a subset (that means only a part of the whole information) the estimation will be imperfect, and it is quite important to give a proper error. The principle should be illustrated in more detail by the following example<sup>1</sup>:

*Example 5.1: Statistics at the fishpond*

*Problem: The number of fish in a pond ( $N$ ) is to be calculated. The quantity of fish and the size of the pond does not facilitate an ordinary counting.*

*Idea: First let us catch a number of fish ( $M$ ). After a clear (but undangerous) marking they are put back to their natural surrounding. As soon as a perfect mixture of the marked and unmarked fish is reached,  $n$  of them will again be caught<sup>2</sup>. Now we have a subset where we could count the number of marked candidates ( $m$ ).*

*If nearly all of these fish of the second catch are marked, it looks obvious, that most of the fish in the pond are marked, and  $N$  should be close to  $M$ . If only a few of them are marked, then it seems very probable, that  $N$  is very much higher than  $M$ . This example demonstrates how one could derive quantitative results from these qualitative argumentation.*

---

The probability of picking  $m$  marked and  $n-m$  unmarked things out of a set of  $M$  marked and  $N-M$  unmarked candidates at once is given by the *hypergeometrical probability distribution* (c.f. [Br,87]). That means in our case that the probability of catching  $m$  marked fish at the end of the

---

<sup>1</sup> Found in: [Bo,93]

<sup>2</sup> A possible ban of the marked fish due to their irregular appearance and a resulting imperfect mixture of marked and unmarked fish is a first candidate of a systematic error. If the mixing is not perfect, the second set of fish will not be representative for the whole pond. Another example of a systematic uncertainty is whether there are (unknown) outlets to the pond.



discussed procedure (under the condition that there are  $N$  fish in the pond) is:

$$P(m/N) = \frac{\binom{M}{n} \binom{N-M}{n-m}}{\binom{N}{n}}$$

Using Bayes' Formula it is possible to end up with a formula for the probability for  $N$  under the condition that  $m$  marked fish have been found in the second catch:

$$P(N/m) = \frac{P(N)}{P(m)} P(m/N) \quad (5.1)$$

The favoured  $N$ -value is the one (if it is only one) where the maximum of the probability above is reached ("*maximum likelihood*"). This formula is also the key for marking a *confidence region*, that means a region in which  $N$  lies with the probability  $\alpha$  (the *confidence level*).

The absolute probability for ending up with a number  $m$  of marked fish  $P(m)$  is a fixed factor and plays no role in the rest of the analysis. The remaining problem is to fix the probability  $P(N)$ . Due to the lack of information about this probability<sup>3</sup>, normally one  $P(N) = \text{const.}$  uses. If we combine all fixed factors in one constant  $\lambda$  which is given by the constraint

$$\sum_{N=M}^{\infty} P(N/m) = 1$$

equation (5.1) becomes

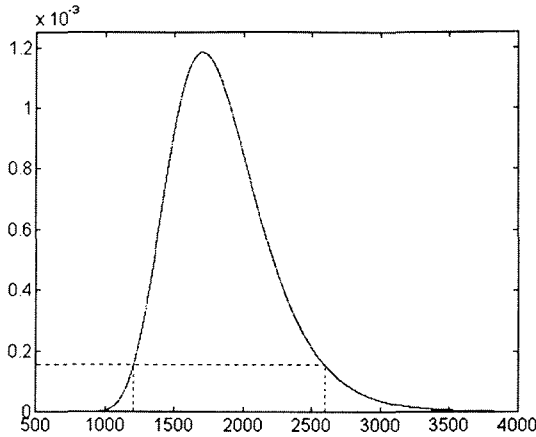
$$P(N/m) = \lambda \binom{N-M}{n-m} \binom{N}{n}^{-1}$$

The values  $M := 250$ ,  $n := 150$ ,  $m := 22$  results in an estimate of  $\bar{N} = 1704$  for the number of fish in the pond. The shape of the probability distribution

---

<sup>3</sup> We only know  $M < N < \infty$  and conclude that the probability outside this region is zero.

and the most obvious example of a  $\alpha := 0.95$  confidence region [1208,2592] is given in the first figure on this next page. We used Stirling's formula (c.f. [Br,87]) for the calculation of the factorial functions appearing in the binomial coefficients.

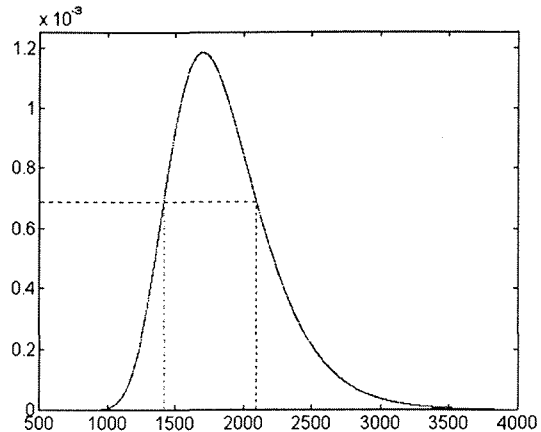


binomial coefficients. After these considerations one could be quite sure, that more than 1200 but less than 2600 fish are in the pond. Even if this is not sufficiently accurate, the principle of a measurement based on statistical considerations is illustrated quite well.

In most cases one uses (motivated by the normal distribution) a 68% confidence level instead of the very strict level of 95% one. That gives the new confidence region [1414,2095], and the final result of the measurement is

$$N = 1704 \begin{matrix} +391 \\ -290 \end{matrix}$$

o



## 5.3. One-dimensional parametrization

### 5.3.1. Motivation for parametrizations

Since QCD-models are implemented as Monte Carlo generators, they provide predictions for measurable quantities only as histograms and not in closed forms. For given parameter values, the bin contents of these histograms are random variables, which are in very good approximation normally distributed with

a known width. For other values of the model parameters, the program has to be executed again, therefore an iterative approach to the maximum of a probability needs a lot of computing time. That is why we are using parametrizations of the full model predictions. The easiest parametrization is the linear one. This linear parametrization is applicable if there is some former knowledge about the values of the parameters of interest. It corresponds to an expansion of the model predictions into a Taylor series and keeping only the terms up to the linear one. The better the former knowledge (that means the expansion point) is, the better this approximation will be. Thanks to previous work we have clear ideas of such expansion points, and a linear approximation thus makes sense.

### 5.3.2. Parametrization of a one-dimensional linear distribution

In the following example, a simple model for illustrating and testing the parametrization of a linear function in one variable will be introduced.

*Example 5.2: Generation of a linear distribution*

*Idea: Normally a computer provides a random number generator which generates equally distributed random numbers  $y$  between 0 and 1. A common method for generating other distributions will now be introduced, and this method will be used for generating the linear distribution*

$$\rho(x) := \frac{a(x-x_0)+b}{\frac{a}{2}(x_0^2-x_u^2)+[b-ax_0](x_0-x_u)}, \quad a, b, x_0 \in \mathbf{R}, x \in [x_u, x_0]$$

*(The denominator is due to the fact that  $\int_{x_u}^{x_0} \rho(x) dx = 1$ .) If the following constraints are fulfilled,  $\rho(x)$  is positive in  $[x_u, x_0]$  (as it should be)*

$$x_u > x_0 - \frac{b}{a} \text{ if } a > 0 \quad \text{and} \quad x_0 < x_0 - \frac{b}{a} \text{ if } a < 0 \quad (5.2)$$


---

The first possibility is to introduce a discretization (a division in several bins) and use the method introduced in example 3.2 to handle the discrete

One-dimensional parametrization

problem. Thanks to the very simple form of the distribution of interest, another more elegant way is possible and, will be used. Let us call the equally distributed variable  $y$  and try to find a function  $f(y)$  such that these function values are linearly distributed. By making a substitution

$$1 = \int_0^1 dy = \int_{x=f(y)}^{f(1)} \frac{dy}{dx} dx = \int_{f(0)}^{f(1)} \frac{df^{-1}(x)}{dx} dx = \int_{x_u}^{x_o} \rho(x) dx$$

we can find the differential equation for the inverse of the unknown function  $f$ :

$$\frac{df^{-1}(x)}{dx} \equiv \frac{dy}{dx} = \rho(x)$$

This equation can be solved by separation of the variables under the conditions  $f^{-1}(x_o) = y_o = 1$  and  $f^{-1}(x_u) = y_u = 0$ . That will lead to:

$$y = \frac{\frac{a}{2}(x^2 - x_u^2) + [b - ax_o](x - x_u)}{\frac{a}{2}(x_o^2 - x_u^2) + [b - ax_o](x_o - x_u)} = f^{-1}(x)$$

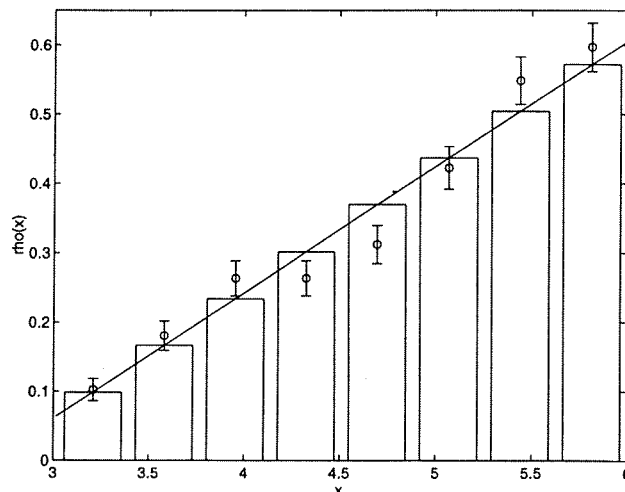
and after inverting

$$x = \left(x_o - \frac{b}{a}\right) + \text{sign}(a) \sqrt{\left(x_o - \frac{b}{a}\right)^2 + \left[\frac{y(x_o^2 - x_u^2) + x_u^2}{x_o - x_u} - 2\left(x_o - \frac{b}{a}\right)\right] \left\{y(x_o - x_u) + x_u\right\}}$$

The sign in front of the square root is due to the condition

$$\left. \frac{dx}{dy} \right|_{y=f^{-1}(x)} = \frac{1}{\rho(x)}$$

in the case of (2). In the diagram to the right, one can see the thrown distribution at a statistic of  $s:=1000$  compared to the given distribution with  $a:=3$ ,  $b:=-2$ ,  $x_o:=2$ ,  $x_u:=3$



and  $x_0 := 6$ . The statistical errors are calculated as discussed in example 3.2. If  $s$  grows, the error bars decrease, and the estimated values for the discrete distribution converge to the unbiased values (as marked by the bars)

o

Armed with this procedure for generating linear distributions it is possible to demonstrate the parametrization in the linear and one-dimensional case. Again the formula of Bayes and the principle of maximum likelihood are the keys to the estimation of a good parametrization.

If we neglect bin correlations,  $n$  of the  $n+1$  bins of a given histogram are independent<sup>4</sup>. We will only look at those  $n$  independent bins for the moment. The linear function  $y(x) := k(x - x_0) + d$ ,  $k, d, x_0 \in \mathbf{R}$ ,  $x \in [x_u, x_o]$  should be fitted to the estimates of these bin contents  $\bar{\rho}_j$ ,  $j = 1, \dots, n$ <sup>5</sup>. Here the formula of Bayes is

$$\begin{aligned} P(k, d / \bar{\rho}_1, \dots, \bar{\rho}_n) &= \frac{P(k, d)}{P(\bar{\rho}_1, \dots, \bar{\rho}_n)} P(\bar{\rho}_1, \dots, \bar{\rho}_n / k, d) = \\ &=: \lambda(\bar{\rho}) P(\bar{\rho} / k, d) \Rightarrow \\ \Rightarrow P(k, d / \bar{\rho}) &= \lambda(\bar{\rho}) \prod_{j=1}^n \left\{ \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left[ -\frac{1}{2} \left( \frac{\bar{\rho}_j - [k(x_j - x_0) + d]}{\sigma_j} \right)^2 \right] \Delta \bar{\rho}_j \right\} \end{aligned}$$

In the second step we use the fact that bin contents are normally distributed in good approximation and that bin correlations can be neglected. Under these conditions,  $P(\bar{\rho} / k, d)$  is the probability of finding bin contents in small regions  $\Delta \bar{\rho}_j$  around  $\bar{\rho}_j$ . The (absolute) maximum of this function is the same as the minimum of the negative "logarithmic likelihood function"

$$-\ln(P(k, d / \bar{\rho})) = \text{const.} + \frac{1}{2} \sum_{j=1}^n \left( \frac{\bar{\rho}_j - [k(x_j - x_0) + d]}{\sigma_j} \right)^2 =: \text{const} + S(k, d)$$

<sup>4</sup> One is given by the condition  $\sum_{i=1}^{n+1} s_i = s$ .

<sup>5</sup> Note that new symbols  $k$  and  $d$  instead of the products of  $a$  and  $b$  with the normalization constant were used.

One-dimensional parametrization

$S(k, d)$  is the basis of the further estimation process. It is called estimation function<sup>6</sup>. The minimum of this function can be found by computing the solution of the following system of linear equations ( $\bar{k}, \bar{d}$ ):

$$\left. \frac{\partial S}{\partial k} \right|_{k=\bar{k}} = 0, \quad \left. \frac{\partial S}{\partial d} \right|_{d=\bar{d}} = 0$$

The result is:

$$\boxed{\bar{k} = \frac{C\xi_1 - B\xi_2}{AC - B^2}, \quad \bar{d} = \frac{A\xi_2 - B\xi_1}{AC - B^2}}$$

with:

$$A := \sum_{j=1}^n \frac{(x_j - x_0)^2}{\sigma_j^2}, \quad B := \sum_{j=1}^n \frac{(x_j - x_0)}{\sigma_j^2}, \quad C := \sum_{j=1}^n \frac{1}{\sigma_j^2}$$

$$\xi_1 := \sum_{j=1}^n \frac{(x_j - x_0)\bar{p}_j}{\sigma_j^2}, \quad \xi_2 := \sum_{j=1}^n \frac{\bar{p}_j}{\sigma_j^2}$$

Because these solutions are linear combinations of independent random variables (bin correlations are neglected) which are normally distributed, they are also normally distributed with the following widths (c.f. appendix A, section 3)

$$\boxed{\sigma_{\bar{k}}^2 = \sum_{j=1}^n \left[ \frac{B - C(x_j - x_0)}{(AC - B^2)\sigma_j} \right]^2, \quad \sigma_{\bar{d}}^2 = \sum_{j=1}^n \left[ \frac{A - B(x_j - x_0)}{(AC - B^2)\sigma_j} \right]^2}$$

Both estimates are unbiased, and the width can therefore be taken as a measure of the distance of the estimates  $\bar{k}$  and  $\bar{d}$  from  $k$  and  $d$ . We will use the following simple example to check the results:

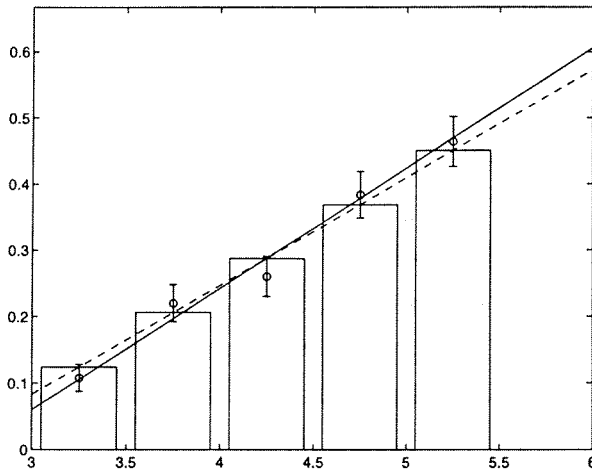
---

<sup>6</sup> The symbol  $S$  comes from the german word "*Schätzfunktion*" for estimation function.

*Example 5.3: Parametrization of a one dimensional linear distribution*

*Idea: The linear distribution of example (3.2) is fixed by  $a:=3$ ,  $b:=-2$  (or  $k \cong 0.1818$ ,  $d \cong -0.1212$ ), then the corresponding histogram with  $(n+1)=6$  using  $s:=500$  bins is generated. On that basis the estimation of  $\bar{k}$  and  $\bar{d}$  is done as described above. By repeating the whole procedure  $s_{rest}:=50000$  times, a distribution of the results is computed which can be compared with the estimated widths.*

The diagram below shows the 5 bins which are used for the estimation of  $k$  and  $d$ . In this example the estimation procedure ends up with  $\bar{k} = 0.1630$  and  $\bar{d} = -0.0792$ . The widths



are  $\sigma_{\bar{k}} = 0.0181$  and  $\sigma_{\bar{d}} = 0.0181$  respectively.

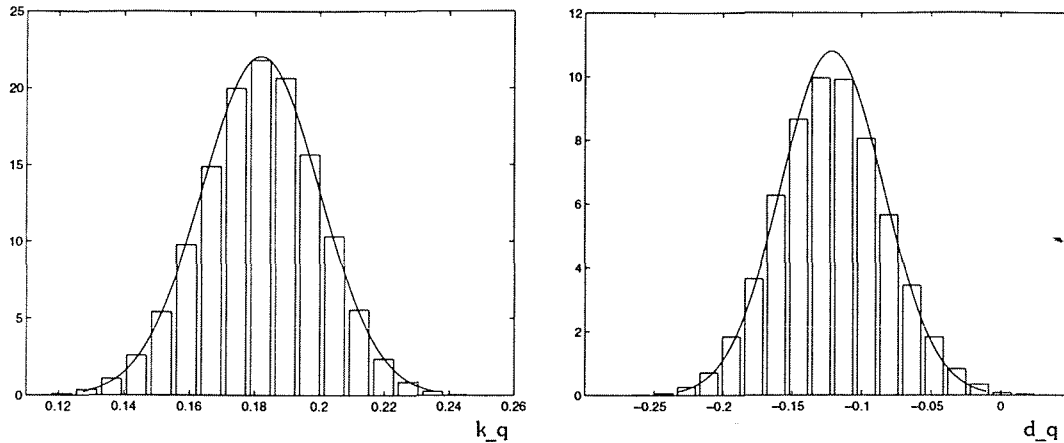
Since we only have estimates  $\bar{\rho}_j$  of the distribution values  $\rho_j$ , it is only possible to calculate estimates for their errors by using formula 3.4. That is the reason why also the errors  $\sigma_{\bar{k}}$  and  $\sigma_{\bar{d}}$  are random variables. The widths of the

quantities in this example are approximately

$$\frac{\Delta\sigma_{\bar{k}}}{\sigma_{\bar{k}}} \cong 0.03, \quad \frac{\Delta\sigma_{\bar{d}}}{\sigma_{\bar{d}}} \cong 0.05$$

If the procedure is repeated 50000 times, the following distributions of the estimated values are seen. The distributed variables are calculated using the estimates for the widths, while the shapes of solid lines are calculated with the exact values ( $k_q$  corresponds to  $\bar{k}$  and  $d_q$  to  $\bar{d}$ ). This is the first reason for the deviations of the histogram from the predicted solid line shape. Other reasons are the small, but existing bin correlations. This deviation is ignored in the parametrization procedure (that means the bin correlations and the random variable behavior of the widths are neglected in

spite of their influence on the result) because they are very small, as expected.



If correlations are known, they could be taken into account by including them in the probability function that is used in Bayes theorem for constructing the estimation function. But this step complicates the parametrization process, and it is not worth doing it.

The second (and even more important) reason for neglecting these bin correlations is that they do not appear in later calculations in this work. Different bins will then mean different points in the space of model parameters, and the model predictions will be calculated independently at each of those points.

o

### 5.3.3. Statistical test of the linear parametrization

Normally we do not know if the linear parametrization is good enough or not. That is why we construct a quantity which is (in good approximation)  $\chi^2$ -distributed with a known degree of freedom if the linear approximation is appropriate<sup>8</sup>. If this quantity lies in the preferred region of the corresponding distribution (for example less than  $2n$ ), we accept the linear parametrization, otherwise we have to use a higher order parametrization. If there are no bin-correlations, the quantity

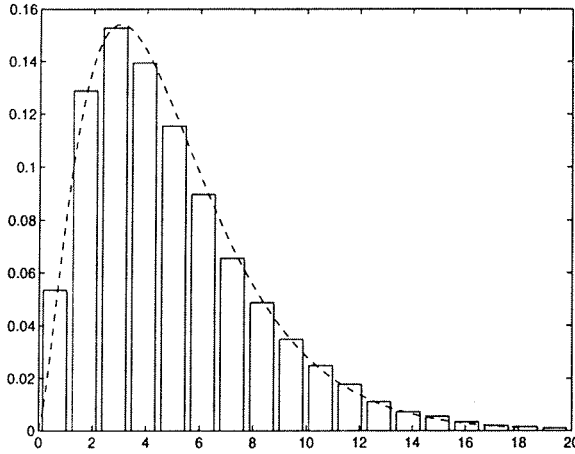
---

<sup>8</sup> For details about  $\chi^2$ -distributions see e.g. [Bo,91], [Ea,71]



$$\eta := \sum_{j=1}^n \left( \frac{\bar{p}_j - [k(x_j - x_0) + d]}{\sigma_j} \right)^2$$

is  $\chi^2$ -distributed with  $n$  degrees of freedom (like every sum of  $n$  squared and independent  $N(0,1)$  random variables).

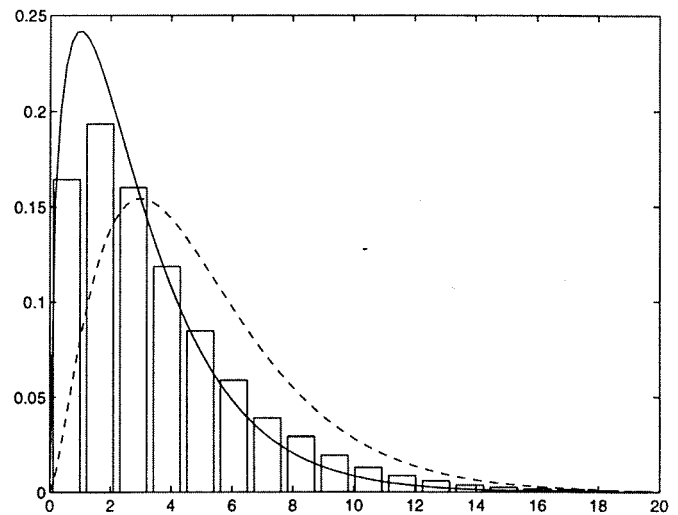


The diagram left shows the distribution of the quantity  $\eta$  for the 5 bins from example 5.3. The deviation from the dashed line (the  $\chi^2$ -distribution with 5 degrees of freedom) is again due to bin-correlations.

By taking bin-correlations into account, and if the exact values of  $k$  and  $d$  are known, it is possible to construct a quantity which is exactly  $\chi^2$ -distributed with 5 degrees of freedom (c.f. appendix B, section 2). In our case the main deviation comes from the fact, that only approximations  $\bar{k}$  and  $\bar{d}$  are known, and so one has to deal with a biased test quantity:

$$\bar{\eta} := \sum_{j=1}^n \left( \frac{\bar{p}_j - [\bar{k}(x_j - x_0) + \bar{d}]}{\sigma_j} \right)^2 \quad (5.3)$$

The distribution of this new quantity is biased to a  $\chi^2$ -distribution with a smaller degree of freedom. This is not surprising, because the estimates  $\bar{k}$  and  $\bar{d}$  correspond to a minimum of exactly this new test quantity. The distribution of the new test quantity is shown in the figure right, it uses the conditions from example 5.3. The dashed line is the  $\chi^2$ -



distribution with  $n=5$ , and the solid line the  $\chi^2$ -distribution with  $n-2=3$  degrees of freedom.

#### 5.3.4. Handling of nonlinear dependence (linear range)

If the test of the linearization fails, one can go to higher order parametrizations or restrict the region in which the linear approximation is trusted. Due to the less demand on computing time we choose the second option and define a *linear range*. For this purpose we compute the *QCD* model prediction at the expansion point with very high statistics. This quantity is used to define the value of the parametrization at the expansion point. Then, model predictions are calculated with smaller statistics at two larger and two smaller values of the parameter under consideration. This is sufficient to compute a first order and a second order parametrization. (That means a linear and a quadratic one.) If the linear parametrization passes the  $\chi^2$ -test, it will be taken, and the linear range is the region in which the parametrization is done. If this test fails, the quadratic part of the second order parametrization is used to calculate an estimate of the error of the linear approximation, and the linear range is then the region in which this quadratic part does not exceed  $n\langle\sigma_s\rangle$ . Here  $\langle\sigma_s\rangle$  is the average statistical error next to the expansion point, and  $n$  can be arbitrarily chosen as e.g.  $n=2$ .

The next example introduces and illustrates the parametrization and the computation of the linear range.

##### *Example 5.4: Linear range*

*Idea: After choosing a nonlinear distribution (an exponential distribution) and generating a discretized version of this distribution, an expansion point is defined, and the first and second order parametrizations are done. Based on those parametrizations, the linear range is calculated and a corresponding graph is plotted.*

---

*(i) Generation of an exponential distribution*

In a way that is analogous to example 5.2, we generate an exponential distributed random variable  $x$  by defining an appropriate function of an equally distributed random variable  $\tilde{y}$ . This is done by solving the differential equation

$$\frac{d\tilde{y}}{dx} = \rho(x)$$

under the conditions  $\tilde{y}(0) = 0$  and  $\tilde{y}(\infty) = 1$ . The solution is

$$\tilde{y}(x) = 1 - \exp\left[-\frac{x}{L}\right]$$

and by the substitution  $y := 1 - \tilde{y}$  we end up with the familiar formula

$$x = -L \ln y$$

As was the case also before,  $n+1$  bins are defined, but only  $n$  are taken into account for the parametrization, in order to have independent model predictions at each point.

*(ii) Parametrization*

The linear parametrization  $l(x) := k(x - x_0) + d$  at fixed  $d := d_0 = m(x_0)$  is found analogous to section 5.3.2. ( $m(x_0)$  is the model prediction at the expansion point.)

$$\bar{k} = \frac{Z}{N}, \quad Z := \sum_{j=1}^n \frac{(x_j - x_0)(\tilde{y}_j - d_0)}{\sigma_j^2}, \quad N := \sum_{j=1}^n \frac{(x_j - x_0)^2}{\sigma_j^2}$$

The estimation function for the quadratic parametrization with a fixed prediction at the expansion point  $q(x) := a(x - x_0)^2 + b(x - x_0) + c$  ( $c := c_0 = m(x_0)$ ) is:

$$S(a,b) := \frac{1}{2} \sum_{j=1}^n \left( \frac{\bar{y}_j - [a(x_j - x_0)^2 + b(x_j - x_0) + c_0]}{\sigma_j} \right)^2$$

The minimum can again be found by solving the system of linear equations

$$\left. \frac{\partial S}{\partial a} \right|_{\substack{a=\bar{a} \\ b=\bar{b}}} = 0, \quad \left. \frac{\partial S}{\partial b} \right|_{\substack{a=\bar{a} \\ b=\bar{b}}} = 0$$

We get

$$\bar{a} = \frac{C\xi_1 - B\xi_2}{AC - B^2} \quad \text{and} \quad \bar{b} = \frac{A\xi_2 - B\xi_1}{AC - B^2}$$

with:

$$A := \sum_{j=1}^n \frac{(x_j - x_0)^4}{\sigma_j^2}, \quad B := \sum_{j=1}^n \frac{(x_j - x_0)^3}{\sigma_j^2}, \quad C := \sum_{j=1}^n \frac{(x_j - x_0)^2}{\sigma_j^2}$$

$$\xi_1 := \sum_{j=1}^n \frac{(x_j - x_0)^2 (\bar{y}_j - c_0)}{\sigma_j^2}, \quad \xi_2 := \sum_{j=1}^n \frac{(x_j - x_0) (\bar{y}_j - c_0)}{\sigma_j^2}$$

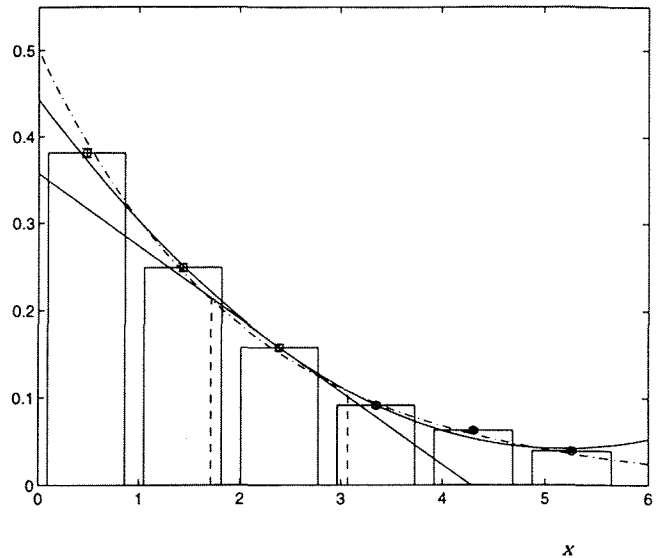
Because the solutions are again linear combinations of normally distributed and independent random variables (in good approximation), the estimation of their widths could be done by using the formula derived in appendix A, section 3 ("linear error propagation"). Because we will not need these widths in the further analysis, we skip this point.

*(i) Computation and illustration of the linear range*

The borders of the linear range can be calculated by the following formula ( $u$  corresponds to the upper, and  $l$  to the lower border):

$$|a(x - x_0)| \leq n \langle \sigma_s \rangle \Rightarrow x_{u,l} = x_0 \pm \sqrt{\frac{n \langle \sigma_s \rangle}{|a|}}$$

For the production of the following picture  $L:=2$ ,  $n:=2$  and  $s:=1 \cdot 10^5$  were chosen. All  $\sigma$  bins that are shown in this diagram are used in the computations. The dashed line marks the resulting linear range. The solid lines represent the linear and quadratic parametrizations, and the dashdotted line show the exponential distribution.



### 5.3.5. Real parametrisations

#### 5.3.5.1. Graphical impressions from *JETSET*

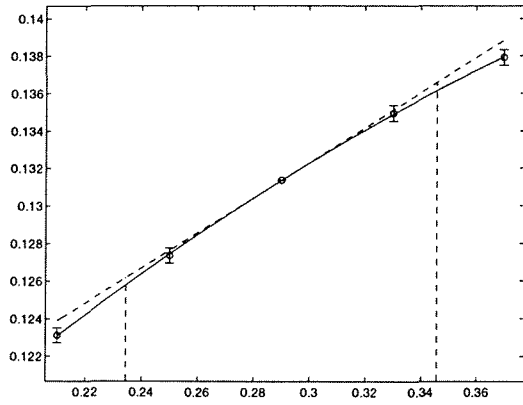
The purpose of this section is to give some graphical impressions of the behavior of *JETSET* in the version of an anisotropic gluon decay. By looking at the *sensitivity*<sup>9</sup>,

$$sens_{b,d,i} := \left| \frac{p_i^0}{x_{b,d}^0} \frac{\partial x_{b,d}}{\partial p_i} \right|$$

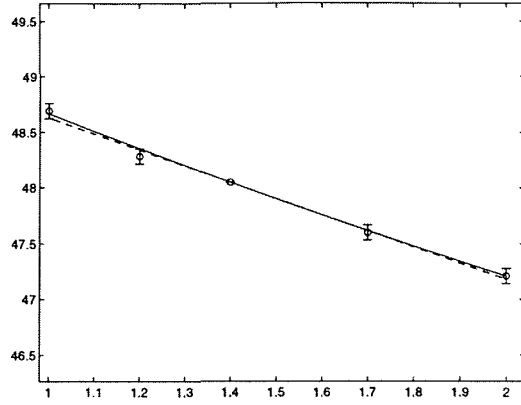
those regions of the observed distributions which give one of the strongest constraints to the *JETSET*-parameters were chosen. The following pictures provide the possibility of getting a feeling about the *JETSET* behavior in the region that will be of interest in the tuning of this model. The dashed lines are the linear approximation and the linear range ( $n=1$ ), while the solid lines are the quadratic approximation.

---

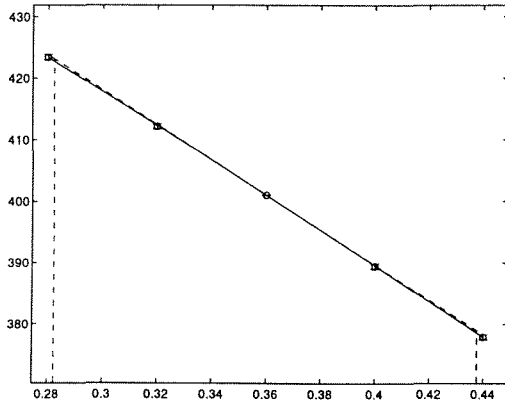
<sup>9</sup> This quantity measures the percental change of the bincontent ( $b$ -th bin,  $d$ -th distribution) from its value at the expansion point due to a change of the  $i$ -th parameter by one percent of its value at the expansion point.



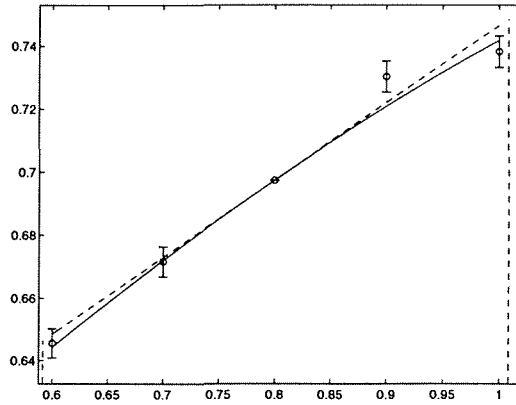
Parameter:  $\Lambda$   
Quantity:  $-\ln(y_3)$ , bin 4: [2.9,3.6]  
 $\chi_{lin}^2 = 9.15, \chi_{qua}^2 = 0.07, sens = 0.20$



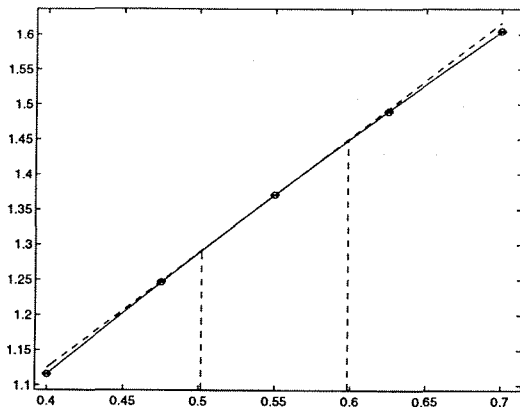
Parameter:  $Q_0$   
Quantity:  $x_p$ , bin 19: [0.09,0.10]  
 $\chi_{lin}^2 = 1.83, \chi_{qua}^2 = 1.29, sens = -0.04$



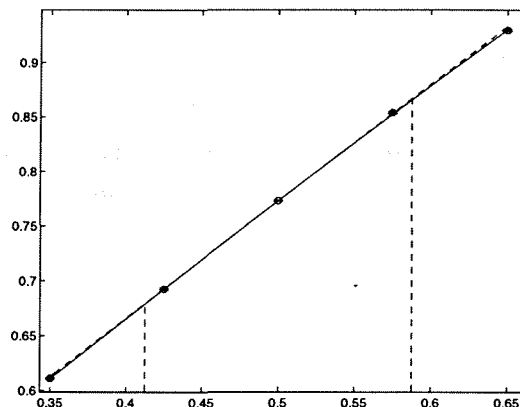
Parameter:  $\sigma$   
Quantity:  $x_p$ , bin 6: [0.014,0.016]  
 $\chi_{lin}^2 = 2.63, \chi_{qua}^2 = 0.23, sens = -0.26$



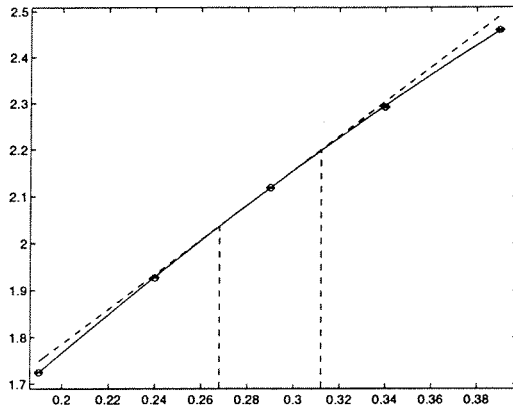
Parameter:  $b$   
Quantity:  $x_p$ , bin 38: [0.49,0.52]  
 $\chi_{lin}^2 = 6.11, \chi_{qua}^2 = 4.30, sens = 0.28$



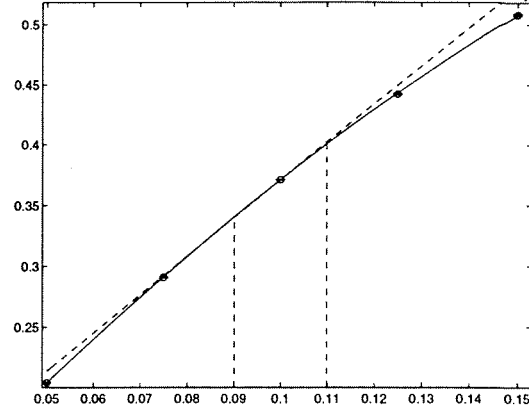
Parameter:  $P_{u,d}^{S=1}$   
Parametrization:  $\langle \rho \rangle$   
 $\chi_{lin}^2 = 198.4, \chi_{qua}^2 = 0.80, sens = 0.65$



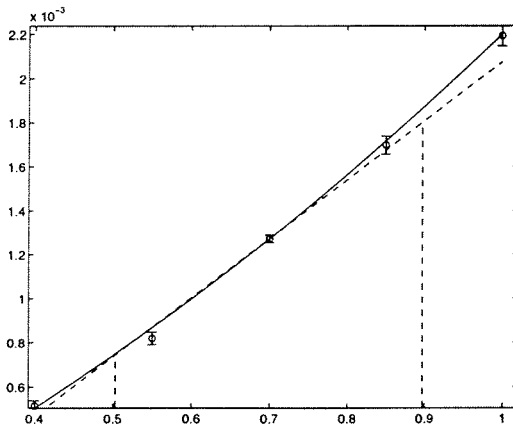
Parameter:  $P_{u,d}^{S=1}$   
Parametrization:  $\langle K^*+ \rangle$   
 $\chi_{lin}^2 = 21.32, \chi_{qua}^2 = 3.04, sens = 0.69$



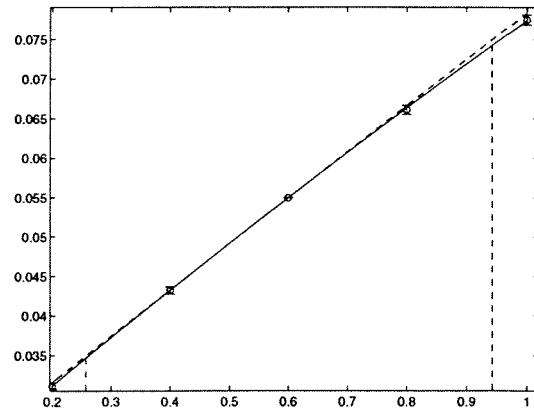
Parameter:  $\gamma_s$   
 Quantity:  $\langle K0 \rangle$   
 $\chi^2_{lin} = 905.2, \chi^2_{qua} = 8.12, sens = 0.50$



Parameter:  $q\bar{q}/q$   
 Quantity:  $\langle \lambda \rangle$   
 $\chi^2_{lin} = 1399, \chi^2_{qua} = 4.01, sens = 0.82$



Parameter:  $su/du$   
 Parametrization:  $\langle \omega \rangle$   
 $\chi^2_{lin} = 14.19, \chi^2_{qua} = 3.44, sens = 1.56$



Parameter:  $b_l$   
 Parametrization:  $-\ln(x_p), \Lambda$   
 $\chi^2_{lin} = 4.16, \chi^2_{qua} = 0.24, sens = 0.63$

**Conclusion:** From the previous pictures we learn that the dependencies are smooth, and that even the linear parametrization gives a good approximation of the model predictions relatively large regions. This statement also holds for the regions that we used for tuning the JETSET variant with isotropic decaying gluons, and for ARIADNE. The regions used for HERWIG fits appear as more non-linear, but since we are only looking at the linear range this is not a big problem. We also see, that the  $\chi^2$ -test provides a very sensitive possibility for testing parametrizations.

### 5.3.5.2. Quantitative examples

The quality of the "low degree approximation" is in this section demonstrated by some numbers. For that purpose we looked at the event properties, and calculated average and maximum test quantities for the linear and the quadratic parametrizations<sup>10</sup>. The values are given in the following table. The JETSET variant with anisotropic decaying gluons was used again.

<i>distribution</i>	$\langle \chi_{lin}^2 \rangle$	$\chi_{lin,max}^2$	$\langle \chi_{qua}^2 \rangle$	$\chi_{qua,max}^2$
<i>S</i>	5.5	189.3	2.4	16.1
<i>A</i>	7.0	161.5	2.3	12.1
<i>P</i>	4.5	148.0	2.2	13.2
<i>C</i>	4.6	110.1	2.1	15.8
<i>1-T</i>	6.4	158.0	2.6	22.7
<i>M</i>	5.6	85.93	2.3	14.4
<i>m</i>	7.8	63.94	2.5	24.5
<i>O</i>	4.2	60.5	2.1	17.7
$M_h^2$	6.3	160.5	2.6	20.3
$B_{tot}$	6.3	218.3	2.4	13.1
$B_w$	6.6	178.2	2.4	21.7
$-\ln(y_3)$	4.2	78.4	2.2	13.8

Table 5.1: Test quantities for the parametrization

The test quantity for the linearization should approximately correspond to 3 because one parameter is fitted to four independent values. For the same reason the test quantity of the quadratic parametrization should be "in the range" of 2.

**Conclusion:** *Even if there are values of the test quantity that clearly exclude a linear and sometimes even the quadratic parametrization, we see from the*

---

<sup>10</sup> Average in this context means an average over all bins of a given distribution and over all parameters. The maximum value is the highest occurring value in all bins of a given distribution and in all parameters.



*table above that on average the quadratic approximation fits the whole range that is used for tuning the parameters. We can also see that the linear approximation does not generally fit the whole region of interest, and that a restriction to a linear range is needed, if one wants to deal with this first degree approximation.*

An analogous statement is also true for the other distributions and multiplicities.

**Remark:** It is also possible to provide parametrizations in higher order (in one dimension or for the case of many dimensions) and to use them in a parameter fit. The crucial point is that correlations between the coefficients of the parametrization normally can not be neglected and has to be introduced into the analysis. To avoid confusion, this topic will be dealt with in appendix *B*, section 3 because in the main part of this work only linear approximations are being used.

## 5.4. Multidimensional parameter fitting

### 5.4.1. Overview

Our first example of a multi parameter fit was the performing of a linear parametrization by estimating the parameters  $k$  and  $d$  (c.f. section 5.3.2). Because the function used was linear in these parameters, we ended up with a system of two linear equations. Here, the same method is used to provide the possibility of performing a fit using such a linear parametrization of the *QCD* model predictions even in more than two dimensions. This method could also be applied for parametrisations in higher order (and for one or more dimensions).

If we are using higher order approximations or if we can not neglect statistical fluctuations of the coefficients in the parametrization, the fitting no longer corresponds to the solving of a system of linear equations. That is why a method for performing a *nonlinear fit* has to be discussed.

### 5.4.2. Linear parameter fitting

#### 5.4.2.1. The general way

If we have a linear model function  $m_b(\bar{p}) := \bar{k}_b \cdot (\bar{p} - \bar{p}^0) + d_b$  with known constants  $k_{b,i}$ , ( $i = 1, \dots, n$ ) and  $d_b$ <sup>10</sup>, that parametrizes the content of a given bin  $b$  as a function of  $n$  model parameters  $\bar{p}$ , and we have also measurements and error bars for the normal distributed bin contents  $\bar{p}_b$ , then we can proceed analogous to section 5.3.2 to construct a probability for the parameter values. We will get:

$$P(\bar{p}|\bar{p}) = N \exp[-S(\bar{p})], \quad S(\bar{p}) := \frac{1}{2} \sum_{b=1}^B \left( \frac{\bar{p}_b - m_b(\bar{p})}{\sigma_b} \right)^2 \quad (5.4)$$

---

<sup>10</sup> In the model tuning, a parametrisation  $m_b(p_i) := k_{b,i} \cdot (p_i - p_i^0) + d_{b,i}$  will be calculated along every parameter axis, and the value  $d_b$  will be calculated as the mean value of the  $d_{b,i}$ :

$$d_b := \sum_{i=1}^n d_{b,i} / n$$

( $N$  is the normalization constant of the probability distribution.) Because of the linear model function, the estimation function is parabolic, and its minimum (the maximum of the probability) is unique. It can be found by solving the following system of linear equations for  $\bar{p}^{\min}$ :

$$\mathbf{M}(\bar{p}^{\min} - \bar{p}^0) = \bar{y}, \quad M_{i,j} = \sum_b \frac{k_{b,i}k_{b,j}}{\sigma_b^2}, \quad y_j = \sum_b \frac{(\bar{p}_b - d_b)k_{b,j}}{\sigma_b^2}$$

Its solution can be found by inverting the matrix  $\mathbf{M}$ :  $\bar{p}^{\min} = \bar{p}^0 + \mathbf{M}^{-1}\bar{y}$ . We obtain the following expression by an expansion of the estimation function at this minimum

$$P(\bar{p}/\bar{\rho}) = \bar{N} \exp\left[-\frac{1}{2}(\bar{p} - \bar{p}^{\min})^T \mathbf{M}(\bar{p} - \bar{p}^{\min})\right],$$

because the first derivative cancels at this minimum, all higher derivatives are zero, and the constant factor  $\exp[-S(\bar{p}^{\min})]$  has been absorbed in the new normalization constant ( $\bar{N} := N \exp[-S(\bar{p}^{\min})]$ ).

That means that the probability of a set of parameters  $\bar{p}$  being the correct one, under the condition that we measured the quantities  $\bar{\rho}$ , correspond to a normal distribution with

$$\sigma_i^2 := C_{ii}, \quad \rho_{ij} = \frac{C_{ij}}{\sigma_i \sigma_j} \text{ (for } i \neq j), \quad \mathbf{C} = \mathbf{M}^{-1}$$

Here  $\sigma_i$  are the widths of the parameter estimates,  $\rho_{ij}$  are the correlation coefficients, and  $\mathbf{C}$  is the covariance matrix.

#### 5.4.2.2. Fixing of some parameters

In the following we sometimes need to fix parameters to values off the expansion point. If the  $m$  ( $m < n$ ) parameters  $(p_{f_1}, \dots, p_{f_m})$  are fixed to the values  $(p_{f_1}^{\text{fix}}, \dots, p_{f_m}^{\text{fix}})$  the fitting of the remaining  $(p_{v_1}, \dots, p_{v_{n-m}})$  corresponds to a changing of the constant  $d_b$  in every bin

$$d_b \rightarrow \bar{d}_b := d_b + \sum_{i=1}^m k_{b,f_i} (p_{f_i}^{fix} - p_{f_i}^0)$$

Using these new constants, the values of the remaining  $n-m$  parameters can be calculated as described above.

### 5.4.3. Nonlinear parameter fitting

For a general estimation function  $S(\bar{p})$ , we could use the formulae derived above only if the higher order contributions to the Taylor expansion around the minimum

$$S(\bar{p}) = S(\bar{p}^{\min}) + \sum_{i=1}^n (p_i - p_i^{\min}) \left. \frac{\partial S}{\partial p_i} \right|_{\bar{p}=\bar{p}^{\min}} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (p_i - p_i^{\min})(p_j - p_j^{\min}) \left. \frac{\partial^2 S}{\partial p_i \partial p_j} \right|_{\bar{p}=\bar{p}^{\min}} + \dots$$

can be neglected, that means, that the estimation function looks like a parabola in the "close surrounding" of its minimum<sup>12</sup>, or equivalently, the model predictions are linear. Because we made sure, that this is true in our case, we can use the formulae derived above, using

$$M_{ij} := \left. \frac{\partial^2 S}{\partial p_i \partial p_j} \right|_{\bar{p}=\bar{p}^{\min}}$$

The following example is a non-QCD illustration and a summary of this rather formal part

---

<sup>12</sup> If the estimation function is not a parabola, more than one minima are possible. In general one has to use additional informations to detect the right one in this case.

**Example 5.5: Multi parameter fit**

Idea: A given model function (the cross-section  $Z \rightarrow \text{hadrons}$  at the  $Z$  resonance) is together with a choice of the free parameters used to generate a histogram. On that basis the method for performing a nonlinear fit that was developed before should be used to estimate these parameters and calculate their errors.

---

**(i) Choice of the model function**

We use the cross-section  $Z \rightarrow \text{hadrons}$  at the  $Z$ -peak as the basis of this example, because it provides a good playground for testing the formulae that we derived above, and illustrates the principle that is used in a very prominent work ([A1,89], [A1,90]). This cross section is near the  $Z$ -resonance dominated by the exchange of a  $Z$ -particle. (The contribution that is due to an exchange of a photon and the interference term can be neglected.) It can be described by the following function:

$$\sigma_h(s) \equiv u \frac{12\pi}{M_Z^2} \frac{s\Gamma_e\Gamma_h}{(s - M_Z^2)^2 + (s\Gamma_Z/M_Z)^2}, \quad s := E_{cm}^2$$

Here  $u := 3.893 \cdot 10^5 \text{ GeV}^2 \text{ nb}$  is the factor for going from natural units ( $\hbar = c = 1$ ) to used ones. The fermionic widths are described by  $s_w^2 := \sin^2 \vartheta_w$  according to:

$$\Gamma_f = \frac{\sqrt{2}}{48\pi} G_F M_Z^3 (v_f^2 + a_f^2) N_c, \quad N_c := \begin{cases} 1, & f \in \{e, \nu\} \\ 3, & f \in \{u, d, c, s, b\} \end{cases}$$

$G_F := 1.1664 \cdot 10^{-5} \text{ GeV}^{-2}$  is the Fermi constant,  $N_c$  is the number of color degrees of freedom and  $v_f$  and  $a_f$  the vector- and axial-vector couplings respectively. They are given by:

$$v_\nu := 1, \quad a_\nu := 1, \quad v_e := -1 + 4s_w^2, \quad a_e := -1$$

$$v_u := 1 - \frac{8}{3}s_w^2, \quad a_u := 1, \quad v_d := -1 + \frac{4}{3}s_w^2, \quad a_d := -1$$

The couplings of the higher generations are defined analogous. Now the contributing widths can be calculated (the top-quark does not contribute at *LEP* energies):

$$\Gamma_h = 2\Gamma_u + 3\Gamma_d, \quad \Gamma_Z = \Gamma_h + 3\Gamma_e + N_\nu\Gamma_\nu$$

We end up with a function consisting of three model parameters:  $p_1 := \sin^2 \vartheta_w$  (the Weinberg angle),  $p_2 := M_Z$  (the mass of the Z-boson) and  $p_3 := N_\nu$  (the number of light neutrino species) for the description of each bin content.

*(ii) Generation of the "experimental" cross-section*

After the choice of a division of the whole energy range into bins, we use the method from example 3.2 to generate the "experimental" distribution. We used  $s := 10^4$  and the following values for the parameters, that are motivated by values obtained from the experiment:

$$p_1 := 0.232, \quad p_2 := 91.182, \quad p_3 := 3$$

*(iii) Fit of the parameters*

The minimum of the estimation function was found by using an iterative approximation algorithm. It yielded:

$$p_1 = 0.259 \pm 0.022, \quad p_2 = 91.200 \pm 0.015, \quad p_3 = 2.907 \pm 0.117$$

which is consistent with our chosen parameter values. The solid line in the following diagram represents our choice while the dashed line is the result that was fitted to the "experimental values".

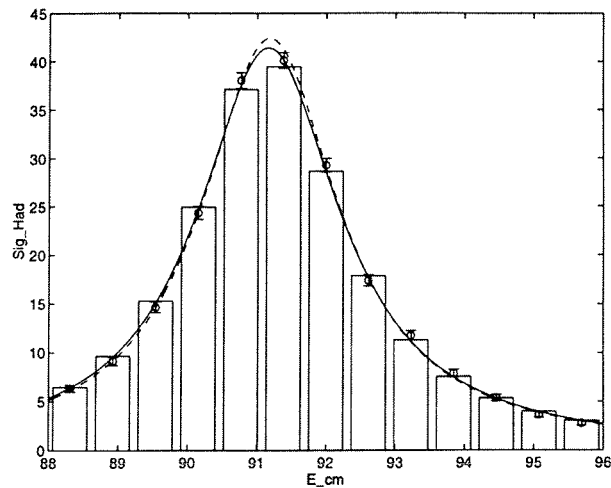


Figure 5.1: Estimation of model parameters

○

## 5.5. Fitting with fluctuations in the parametrization

### 5.5.1. Introductory remark

In the parameter fitting procedure above we have always used (linear) parametrizations of the model predictions without taking their statistical fluctuations into account. This section outlines a procedure which includes such parametrization fluctuations. We do not actually use this method in our parameter tuning procedure, because (as a later test will strongly indicate) the used statistics is high enough to neglect this unwanted complication. In spite of this, this technique could play a role if one wants to go to higher order parametrizations (c.f. appendix *B*, section *3*), because if one deals with more than one parameter, the number of coefficients in the parametrization grows fast with the degree of the parametrization, so that the number of points in the parameter space that is used for the parametrization must be very high, and one could not deal with a very high statistics at each point. As we will see, this method also provides a possibility for taking the correlations between the coefficients of the parametrization into account.

Because this handling of parametrization fluctuations could also play a role in future algorithms for parameter fitting, we describe it in detail, even if it is not a crucial ingredient for our analysis.

We use a polynomial parametrization of degree  $N$  in one dimension to demonstrate the inclusion of statistical fluctuations of the parametrization. This will be introduced in the following section. The one dimensional parametrization will be used to demonstrate the procedure from example 5.3. The case of a polynomial in more than one dimension and a linear parametrization of model predictions is also discussed below.

### 5.5.2. One dimensional parametrization by a polynomial

Every bin content  $\rho_b(\bar{p})$  is parametrized by a polynomial of degree  $N_b$ :

$$m_b(p) \equiv \sum_{n=0}^{N_b} a_{b,n} p^n$$

If we know normally distributed estimates  $\bar{m}_b(p_m)$  for the model prediction at parameter values  $(p_1, \dots, p_M)$ ,  $M > N + 1$ , and the corresponding widths  $\sigma_m^{\text{mod}}$  (the index "mod" is used to avoid confusion with the later used widths for the experimental values  $\bar{p}_j$ ), it is possible to estimate the coefficients  $a_{b,n}$  by using the following estimation function:

$$S_{\text{pol}}(\bar{a}_b) := \frac{1}{2} \sum_{m=1}^M \left\{ \frac{\bar{m}_b(p_m) - \sum_{n=0}^{N_b} a_{b,n} p_m^n}{\sigma_m^{\text{mod}}} \right\}^2$$

We assumed, that the correlations between the estimates  $\bar{m}_b(p_m)$  can be neglected. The minimum  $y$  of the function introduced above corresponds to the solution of the following system of linear equations:

$$\mathbf{M}^{\text{pol}} \bar{a}_b^{\text{min}} = \bar{y}^{\text{pol}}, \quad M_{ij}^{\text{pol}} = \sum_{m=1}^M \frac{p_m^i p_m^j}{\sigma_m^2}, \quad y_i^{\text{pol}} = \sum_{m=1}^M \frac{\bar{m}_b(p_m) p_m^i}{\sigma_m^2},$$

( $i, j = 0, 1, \dots, N$ ) Because the solutions of this system of linear equations are linear combinations of normally distributed and independent random variables, they are as well normally distributed, and similar to section 5.4.2.1 the covariance matrix is given by the inverse of the matrix of second derivatives of the estimation function calculated for  $\bar{a}_b = \bar{a}_b^{\text{min}}$ .



For the case where the widths and correlations of the coefficients  $a_{b,n}$  are negligible, one would perform the fit by using the estimation function

$$S_{par}(p) := \frac{1}{2} \sum_{b=1}^B \left\{ \frac{\bar{p}_b - \sum_{n=0}^{N_b} a_{b,n} p^n}{\sigma_b} \right\}^2$$

### 5.5.3. Including the coefficients widths

We temporarily disregard the correlations between the coefficients. If one wants to include the widths, the most obvious ansatz would probably be

$$S_{par}(p) := \frac{1}{2} \sum_{b=1}^B \frac{(\bar{p}_b - \sum_{n=0}^{N_b} a_{b,n} p^n)^2}{\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p^n)^2} \quad (5.5)$$

Here  $\sigma_{b,n}^a$  is the width of the  $n$ -th coefficient of the parametrization in the  $b$ -th bin, and we used the linear error propagation formula derived in appendix A, section 3 to replace the incomplete widths of the nominators by the correct ones. Nevertheless, this is an ad hoc ansatz, and it must not surprise if it leads into troubles. The first problematic point is that large parameter values correspond to big statistical errors of the nominator (that means a big denominator), which can cause the fit to "run away" to meaningless values (c.f. section 6.5.3). The second point is that we can not hope, that for example the errors calculated by the matrix of second derivatives can be justified by the arguments given above, because we "demolished" the estimation function by an ad hoc ansatz.

A possible way out of this dilemma can be found, by going back to the roots, and by remembering how the estimation function was built in section 5.3.2. For this reason let us write down the probability for the random variable

$$d_b := \bar{p}_b - \sum_{n=0}^{N_b} a_{b,n} p_0^n$$

using fixed values for the parameter  $p := p^0$ . The probability (under the condition  $p := p^0$ ), that this random variable obtains a value belonging to a small interval  $\Delta d_b$  around  $d_b$  is

Fitting with fluctuations in the parametrization

$$P(d_b/p_0) = \frac{1}{\sqrt{2\pi}\sigma_{d_b}} \exp\left[-\frac{1}{2}\left(\frac{d_b}{\sigma_{d_b}}\right)^2\right] \Delta d_b =$$

$$= \frac{1}{\sqrt{2\pi(\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p_0^n)^2)}} \exp\left[-\frac{1}{2} \frac{(\bar{p}_b - \sum_{n=0}^{N_b} a_{b,n} p_0^n)^2}{\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p_0^n)^2}\right] \Delta d_b$$

and according to Bayes theorem we get

$$P(p_0/\bar{d}) \propto \prod_{b=1}^B \frac{1}{\sqrt{\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p_0^n)^2}} \exp\left[-\frac{1}{2} \frac{(\bar{p}_b - \sum_{n=0}^{N_b} a_{b,n} p_0^n)^2}{\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p_0^n)^2}\right]$$

From this probability we conclude that the following expression is a candidate for a estimation function using the widths of the parametrization coefficients

$$S_{par}(p) \propto \frac{1}{2} \left[ \sum_{b=1}^B \log(\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p^n)^2) + \sum_b \frac{(\bar{p}_b - \sum_{n=0}^{N_b} a_{b,n} p^n)^2}{\sigma_b^2 + \sum_{n=0}^{N_b} (\sigma_{b,n}^a p^n)^2} \right] \quad (5.6)$$

Correlations between the coefficients of the parametrization are up to now not included in this function. It can be achieved by a decoupling of random variables as discussed in appendix B and illustrated in example 5.3.

In the tuning of QCD models we have more than one parameters, and we always use a linear parametrization

$$m_b(\bar{p}) \cong \sum_{j=0}^n a_{b,j} p_j + b_b$$

All coefficients  $a_{b,j}$  are independent in our case, and correlations play no role. Because of that, the estimation function (5.6) gets the form

$$S_{par}(p) \propto \frac{1}{2} \left[ \sum_{b=1}^B \log(\sigma_b^2 + \sum_{j=0}^n (\sigma_{b,j} p_j)^2) + \sum_b \frac{(\bar{\rho}_b - (\sum_{j=0}^n a_{b,j} p_j + b_b))^2}{\sigma_b^2 + \sum_{j=0}^n (\sigma_{b,j} p_j)^2} \right] \quad (5.7)$$

Here  $\sigma_b^2 := (\sigma_b^{\text{mod}})^2 + (\sigma_b^0)^2$ ;  $\sigma_b^{\text{mod}}$  is the width of the model prediction in the  $b$ -th bin,  $\sigma_b^0$  is the width of the coefficient  $b_b$ .

#### Observations:

- In the case of vanishing widths of coefficients, the given formula corresponds to the estimation function that was used before up to an irrelevant additive constant.
- The critical behavior for large parameter values is weakened by the logarithmic term.
- New minima can appear due to the rising of the logarithmic term, as is shown in the next example. So one should always check, that he "sits" in the right minimum.

To close and summarize this rather formal chapter, I want to go through the following quite extended example.

*Example 5.6: Estimation with a fluctuating parametrization*

Idea: After the choice of a model distribution, a polynomial parametrization is performed. With this parametrization and by including widths as well as correlations of its coefficients we perform an estimation of the parameter and a calculation of the resulting error.

---

*(i) Choice of the model distribution*

The used model function in one variable  $x$  and with one free parameter  $\alpha$  is:

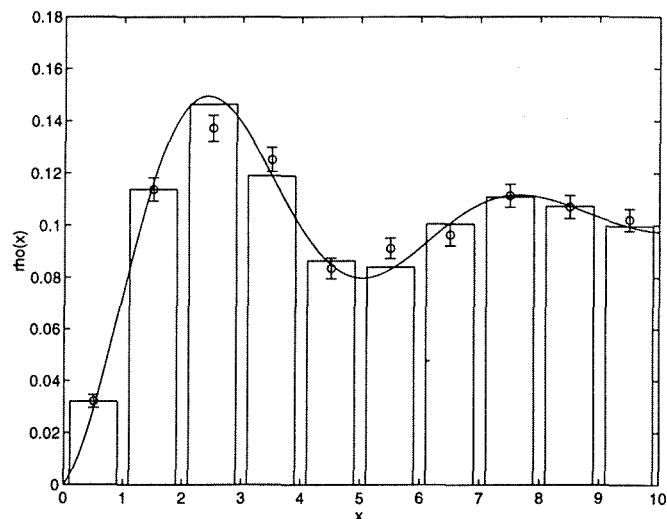
$$\rho(x;(\alpha,\beta)):= Ne^{-\alpha x}[1 - \cos(\beta x)], \quad \beta:= 4\alpha$$

The normalization constant can be evaluated using the condition  $\int_{x_u}^{x_o} \rho(x;(\alpha,\beta))dx = 1$ . The result is

$$N = -\frac{1}{\alpha^2 + \beta^2} \left[ e^{-\alpha x_o} \{ \alpha \cos(\beta x_o) - \beta \sin(\beta x_o) \} - e^{-\alpha x_u} \{ \alpha \cos(\beta x_u) - \beta \sin(\beta x_u) \} \right]$$

The correlations between the coefficients of the parametrizations are very high in this example. That allows us to get a clear impression of the influence of these correlations on the estimation process.

The estimation of the bin contents are calculated as discussed in example 3.2 after the choice of the  $x$ -borders  $x_u := 0, x_o := 10$  and a division of the resulting region into 10 bins.



For  $\alpha := \alpha_0 = 0.3$  and  $s := 5 \cdot 10^3$  the results are shown at the diagram above. Here the solid line is the shape of the model function, while the bars represent the discretized model function.

*(ii) Parametrization*

**Remark:** In the rest of this example, we will use only 6 bins to save computation time.

Every bin content is parametrized as a function of  $\alpha$  in the region  $0.01 \leq \alpha \leq 0.6$  by using the polynomial ansatz that was introduced in section 5.5.2. In a first step 3 different values for  $\alpha$  (the borders of the region of observation and the center of this region) are used to perform a linear parametrization as described in section 5.3.2. The model prediction is estimated at three other points of the chosen parameter region after this step<sup>13</sup>, and together with the previous points, the result has been exposed to the  $\chi^2$  test, using the test quantity (5.3).

If the result of this test is positive (that means  $\chi^2 < \mu + 2.5\sigma \cong n_{DF} + 2.5\sqrt{2n_{DF}}$ , where  $n_{DF}$  is the number of degrees of freedom), the parametrization is accepted. If the test fails, the degree of the parametrization is increased, and the whole procedure repeated.

We discuss two different cases. In the first one, the chosen statistics for the "experimental" distribution is  $s_{dat}^1 := 1 \cdot 10^3$ , and that for generating the model predictions is (at each point)  $s_{mod}^1 := 1 \cdot 10^5$ . In this case the "experimental" errors are much bigger than the widths of the simulation, and their contribution to the estimation of the model parameter can be neglected. Big widths of the simulated bin contents are preferred in the second case, and we used  $s_{dat}^1 := 1 \cdot 10^6$  together with  $s_{mod}^1 := 5 \cdot 10^3$ . The resulting degrees of the parametrizations are shown in the following table.

<i>bin number</i>	1	2	3	4	5	6
<i>high model statistic</i>	2	3	3	4	3	4
<i>low model statistic</i>	2	2	3	2	3	2

<sup>13</sup> The values of  $\alpha$  are place in the middle of the largest free region. If there were equal values, we began from the left.

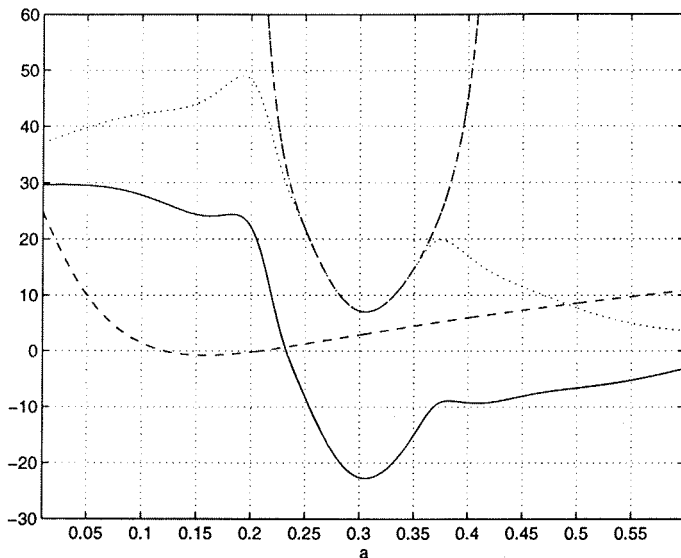
*Fitting with fluctuations in the parametrization*

To get a feeling about the very high correlations appearing in this example, their values for bin 6 are listed below. Similar results appeared in the other bins.

$$(\rho_{ij}) \equiv \begin{pmatrix} 1.0000 & -0.9997 & 0.9989 & -0.9974 & 0.9955 \\ & 1.000 & -0.9997 & 0.9989 & -0.9975 \\ & & 1.000 & -0.9997 & 0.9989 \\ & & & 1.000 & -0.9997 \\ & & & & 1.000 \end{pmatrix}$$

*(iii) Estimation and graphical representation*

The estimation functions of the first case (high model statistics) are drawn in the next diagram. The dashdotted line (coarse parabola) represents the estimation function without consideration of parametrization fluctuations



(similar to (5.3)). The shape near the minimum is the same as for the non-logarithmic part of the new estimation function (including widths as well as correlations of the parametrization similar to (5.4)<sup>14</sup>) shown by the dotted line as expected. The full new estimation function similar to (5.4)

corresponds to the solid line, and the dashed line is the prediction of the new estimation function without consideration of any correlations.

The full new estimation function increases with the parameter values and new local minima appear. The results of the estimation can be seen in the following table:

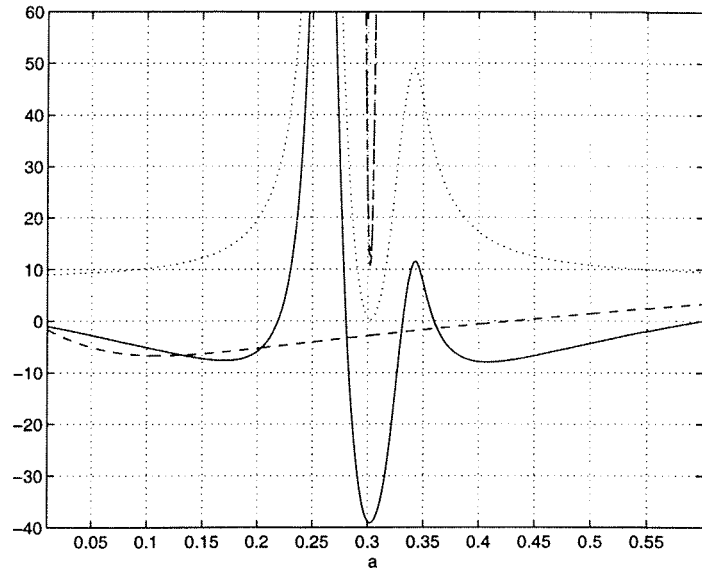
---

<sup>14</sup> For the details of inclusion of the correlations see appendix B. The principle is to write down the correlated random variables as a linear combination of uncorrelated (decoupled) ones, and use derived in appendix A, section 3.

<i>without widths or correlations</i>	$0.3060 \pm 0.0101$
<i>with widths</i>	$0.1580 \pm 0.0339$
<i>with widths and correlations</i>	$0.3060 \pm 0.0101$

The results in the case of neglected widths and correlations are the same as those for the inclusive estimation, as we expected.

The second case (low model statistics) is illustrated by the picture right. The results of the estimation are shown in the next table. It shows significant differences between the results.



<i>without widths or correlations</i>	$0.3020 \pm 0.0005$
<i>with widths</i>	$0.1110 \pm 0.0339$
<i>with widths and correlations</i>	$0.3020 \pm 0.0029$

Even if the estimation without taking the widths or correlations into account are as good as the result of the full estimation, the calculated error is much too small, and the old estimation function should be replaced by the new version.

**Conclusion:** *It is possible to use polynomial parametrizations in fitting procedures even if the correlations and widths of their coefficients can not be ignored. One should however regard, that also correlations could have significant influences on the results of fits.*

o

**Remark:** The correlations in the parametrization were quite strong in the last example, and therefore their influence on the result very prominent. We are only

*Fitting with fluctuations in the parametrization*

dealing with linear parametrisations, and all the coefficients will be calculated independently, and have therefore no correlations. We can therefore forget about contributions from correlations of the parametrization. But if one wants to go to parametrizations with a higher degree, these correlations appear and have to be taken into account.



## Chapter 6.

# LinFit - An algorithm for model tuning

### 6.1. Overview

Starting from the results of the last chapter we introduce and test an algorithm for model tuning based on a linear parametrization of given model predictions. The main difference to the earlier discussions of parameter fitting is the appearance of systematic errors. While it seems possible to include systematic errors of the measurements in the estimation function, the deviations between measured data and model predictions cause a very serious problem. The consequences are systematic differences of the fitted parameters when fitting them to different sets of distributions. These differences can in general not be explained by the errors computed from the matrix of second derivatives. This problem is mainly excluded by restricting the fitting procedure to bins for which the model predictions are in good agreement with the data. The remaining uncertainties are given as systematic errors of the model parameters.

### 6.2. Introduction of LinFit

LinFit consists of the following modules:

- 1. Choice of an expansion point*
- 2. Parametrization of the Monte Carlo predictions and calculation of the linear range*

*3. Restriction to a "fitable region"*

*4. Estimation of the model parameters and calculation of their errors*

*5. Concluding tests*

These modules will now be discussed in more detail. The linear parametrization is done by parametrizing the model predictions by independent linear fits along each of the parameter axes. The results from section 5.3.2 are used for this reason. This simplification only holds if the expansion point is as narrow as possible to the estimated set of parameters. To guarantee this crucial point, we used results from earlier attempts to find the optimal parameter values for setting the expansion point. In addition to the default values of the parameters the main input to this point came from [A1,92] and [Ru,95].

If the expansion point is "too far away" from the fitting result, higher order terms of the parametrization will be necessary. To define the meaning of "too far away" the minimum of the linear ranges of a given parameter among all bins is used. If one of the estimated parameters exceeds a region defined by this *minimal linear range*, the whole set of estimates is rejected, and the fitting procedure is repeated with another expansion point<sup>1</sup>. Instead of attempting to include higher order terms, the linear procedure is iterated, because this is less time consuming. Otherwise, if all the parameters lie in this "*linear region*", the concluding test is passed with a positive result and the estimated parameters are trusted.

## 6.3. Fixing of the linear range

In section 5.3.4 a linear range was introduced, but the factor  $n$  was not fixed. Some graphical impressions were given that led to the obvious choice  $n:=1$ . In this linear region, the deviation of the linear parametrization from the full model prediction is believed to be less than the mean statistical error of the model prediction, and therefore without big influence on the result of the estimation procedure. But these were only graphical impressions, that can or can not be

---

<sup>1</sup> The choice of a new expansion point is a very critical problem because there is no criterion that distinguishes a good expansion point from a bad one. Normally the result of the preceding fit is used as the new expansion point, even if the linear range is exceeded.

verified by using the estimation function<sup>2</sup>. Because of this point of uncertainty some tests are performed in the following in order to calculate an acceptable choice for *n*.

### 6.3.1. *n* in the case of a one-parameter fit

The *JETSET* parameters  $Q_0$  and  $\Lambda$  (one at a time) were put at the ends of different minimal linear ranges (defined by decreasing values of *n*) and model predictions for these parameter settings were calculated<sup>3</sup>. The distribution of  $-\ln(x_p)$  was used to perform fits for  $Q_0$  and the distribution of  $-\ln(y_3)$  was used for fits of  $\Lambda$ . These distributions were also used to define the linear regions. For the "data distribution"  $s_{dat} := 6 \cdot 10^5$  and for the "simulation"  $s_{sim} := 5 \cdot 10^5$  at the points along the parameter axes and  $s_{sim}^0 := 2 \cdot 10^6$  at the expansion point ( $\Lambda = 0.29 GeV$ ,  $Q_0 = 1.4 GeV$ ) was used. These values were chosen in order to have a situation as close as possible to the real parameter tuning scenario. In the fit all parameters shown in table 6.3 (except of  $Q_0$  and  $\Lambda$  respectively) were fixed at the expansion point (which is also shown in table 6.3), and all other parameters were fixed at their default value, or at the value given in appendix B, section 4. The fits were done with and without inclusion of the statistical errors of the coefficients in the parametrization.

If the fit reproduces the given test values of  $Q_0$  and  $\Lambda$ , the corresponding value of *n* is a good choice for the definition of a linear range (for the fit of one parameter).

The following symbols are used:

$\varepsilon_p$  the parameter error calculated by the matrix of second derivatives of the estimation function

$\Delta_p$  the distance between the given value and the result of the fit ( $\Delta_p := p^{fit} - p^{mod}$ ); note, that this distance is negative if the fitted value is too small.

---

<sup>2</sup> Because we know for example that a statistical test of a parametrization is much more sensitive than a graphical one we could argue that the estimation function will be much more sensitive to insufficient parametrizations than we initially expected.

<sup>3</sup> We used the *JETSET* version with isotropic decaying gluons.

*Fixing of the linear range*

$p^{\text{mod}}$  the parameter setting

$n$	$Q_0$			$\Lambda$		
	$p^{\text{mod}}$	$\epsilon_p$	$\Delta_p$	$p^{\text{mod}}$	$\epsilon_p$	$\Delta_p$
1.5	1.7842	0.0090	-0.0074	0.2448	0.0016	-0.0048
1.0	1.7137	0.0090	0.0085	0.2550	0.0016	-0.0028
0.5	1.6218	0.0090	0.0133	0.2682	0.0016	-0.0034
0.1	1.4992	0.0090	0.0010	0.2858	0.0016	0.0020

Table 6.1: Fits without inclusion of fluctuations in the parametrization

$n$	$Q_0$			$\Lambda$		
	$p^{\text{mod}}$	$\epsilon_p$	$\Delta_p$	$p^{\text{mod}}$	$\epsilon_p$	$\Delta_p$
1.5	1.7842	0.0111	-0.0066	0.2448	0.0020	-0.0052
1.0	1.7137	0.0108	0.0091	0.2550	0.0019	-0.0028
0.5	1.6218	0.0105	0.0134	0.2682	0.0018	-0.0034
0.1	1.4992	0.0101	0.0010	0.2858	0.0017	0.0020

Table 6.2: Fits with inclusion of fluctuations in the parametrization

The errors calculated by the estimation function that take fluctuations in the parametrization into account are somewhat larger than the errors calculated without, as was well expected. The method former leads to good agreement between the test values and the results of the fit ( $\Delta_p \leq 2\epsilon_p$  if  $n \leq 1$ ), while the latter seems to produce parameter errors which are somewhat too small.

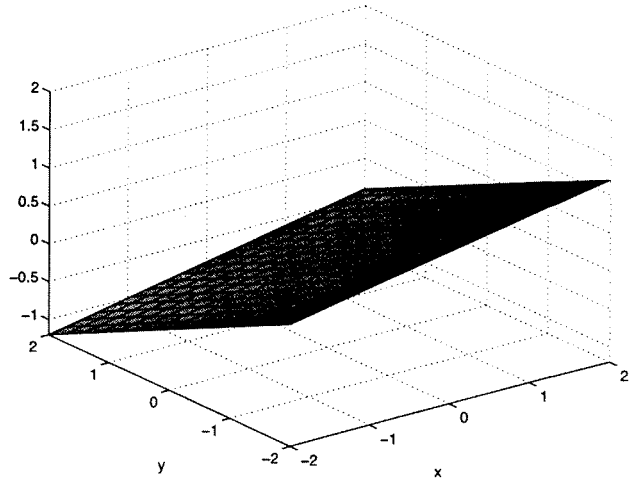
We did only use statistical errors to perform the calculations mentioned above. In the case of real model tuning, the systematic errors of the measurement are added quadratically to the statistical widths (c.f. section 6.4.4), and the contribution of the widths in the parametrization should be even smaller than in the example above. This is a good reason to believe that in our case it is also possible to work with the simpler model function. A time consuming iterative search for a minimum of the estimation function can therefore be replaced by a faster solution of a system of linear equations.

**Conclusion:** *The fits performed in this section seem to confirm the choice of  $n:=1$  for the definition of the linear region in the case of one free parameter.*

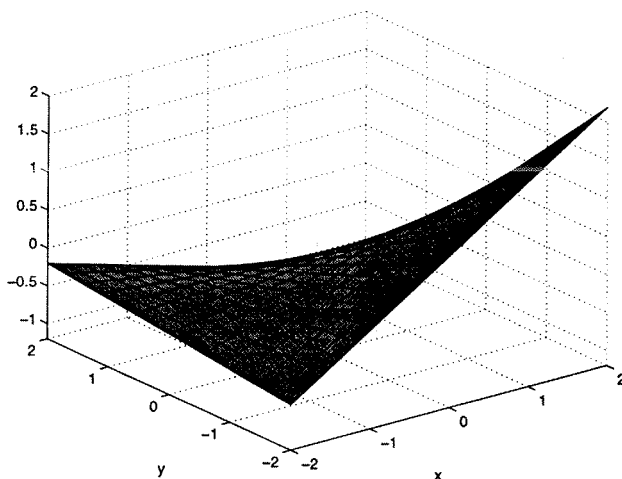
The estimation function (5.6) leads to slightly bigger errors than the estimation function (5.4) without statistical errors of the parametrization.

### 6.3.2. *n* in the case of many parameters

The same procedure as in the last section will be repeated for the case of the 10 parameters used in later *JETSET* fits. Even if the results of the last section allowed us to argue that for a definition of a linear range along the parameter axes a value  $n:=1$  leads to trustable results, the contribution of mixed terms in a Taylor expansion in more than one dimensions could (and will) cause some troubles. The figures on this page should give an impression of this point.



The figure above shows the function  $f(x) := 0.2x - 0.4y$ . This function differs from the second function  $g(x) := 0.2x - 0.4y - 0.25xy$  shown below by the mixed term that has no influence on a parametrization along the parameter axes.



Therefore even if the behavior along the parameter axes is really linear, the model function can be quite nonlinear off these axes. If we want to deal with the linear ansatz in this more-dimensional case, we have to make a tighter restriction than we have done in one dimension. That means we have to use a smaller number for  $n$ .

In our fitting procedure both functions would be treated in the same way (by the parametrization  $f(x)$ ) and so we have to be very careful to get right results. For this reason the same set of distributions was used as in the final model tuning procedure (see chapter 7) together with different values of  $n$  to perform fits to test distributions generated

*Fixing of the linear range*

with *JETSET* (with isotropic gluon decay). The results of the estimation are compared with the parameter values used in the generation of the test distributions. If the difference is less than 2.5 times the calculated error for all of the parameters, the corresponding value of  $n$  is considered as a good candidate for the definition of the linear region in the model tuning procedure.

In the next table the results are summarized. The factors given in this table show the distance between the fitted and the nominal value in units of the error. A factor of 3 means, that this distance is three times the size of the calculated error. The statistical widths of the parametrization are included in the estimation function.

$n$	1.5	1.0	0.5	0.1	0.05	0.01
$\Lambda$	-1.1	0.2	-3.8	4.2	-2.4	2.2
$Q_0$	-3.2	-2.3	-2.9	-1.9	-1.6	-0.4
$\sigma$	-5.9	-4.0	0.1	-3.7	-0.3	-1.7
$B$	3.2	2.4	-0.8	4.1	0.4	1.8
$P_{ud}^{S=1}$	-4.7	3.6	-0.8	0.6	2.0	-0.2
$P_s^{S=1}$	-0.1	-0.2	0.6	-0.4	0.5	0.0
$s/u$	-4.1	-4.7	-2.2	0.0	-0.6	-1.2
$qq/q$	-3.4	-4.0	-1.8	0.0	0.0	0.7
$su/du$	9.1	8.6	4.5	0.4	1.2	1.4
$\bar{b}_1$	4.6	3.6	1.5	1.4	-1.0	-1.9

Table 6.3: Exploring the linear region

Columns with bold faced numbers represent fits where the differences  $\Delta_p$  are less than 2.5 times the error. Note, that the deviations depend on the type of parameter; the worst case occurs for  $su/du$ .

These values for  $n$  initially seem much too small. But we did only use statistical widths for performing this test. We will discuss in section 6.4.4, that an inclusion of systematic errors of the measurement in the estimation function is possible. In this case, the statistical widths of the measurement will be replaced by a quadratic sum of these widths and the corresponding systematic errors. Because of that, systematic deviations in the parametrization are negligible if they are small compared to this quadratic sum, which is bigger than the error used in the previous example. Consequently the  $n$ -value that is used in the real

parametrization is surely allowed to be higher than the value that was calculated in this example. From this point of view even the results for  $n \leq 0.5$  also seem acceptable.

Because of the similarity of the models *JETSET* and *ARIADNE*, this check should also be representative for these models, and  $n:=0.1$  should be a proper value for defining the linear region. *HERWIG* is quite different and uses other parameters. Therefore a similar test series was performed for *HERWIG 5.8*. The results are given in the following table.

<i>n</i>	1.5	1.0	0.5	0.1	0.05	0.01
$\Lambda$	-5.7	-3.3	0.3	-4.0	0.0	-1.0
<i>RMASS(13)</i>	-1.1	-1.1	2.3	-3.1	4.7	-2.7
<i>CLMAX</i>	7.2	3.8	0.6	1.4	-1.5	-2.5
<i>CLSMR</i>	-5.1	-3.2	-2.6	-0.3	-0.7	-1.3
<i>PWT(3)</i>	-1.3	0.2	-1.0	0.4	-1.8	0.7

Table 6.4: Exploring the linear region

This check that was done only with statistical widths indicates that it could be necessary to restrict the linear region to a region which is defined by  $n:=0.01$ . This value is that small, that an iterative search for a proper expansion point will hardly work. The "eye glasses" of the very large statistics that was used can decide between the full model prediction and a linear parametrization up to this very small region. Since we have also systematic errors, this eye glasses are not that sharp, and it makes sense to use  $n:=0.5$ .

**Conclusion:** *The given series shows, that the used Monte Carlo statistics is very appropriate to decide between a correct and an incorrect parametrization. If the full errors would only be of the size of the statistical widths, a usage of a linear parametrization would hardly make sense, because the linear regions that we want to restrict on would be too small. Because we have to include the systematic errors of the measurements in the estimation function, we are able to use a linear approximation, and fix the linear regions for JETSET and ARIADNE with  $n:=0.1$ , and the linear region for HERWIG with  $n:=0.5$ .*

## 6.4. Handling of systematic errors

### 6.4.1. Model parameters in the case of imperfect models

Previously we always dealt with perfect and fitable models. That means it existed one (and only one) set of model parameters which lead to agreement between model predictions and experimental data in the case of vanishing errors. Consequently the final goal of the fitting procedure was to find estimates for these optimal model parameters.

If systematic deviations between model and data appear, no such set of optimal parameters exists. Unfortunately a definition of a reasonable generalization of the optimal model parameters is not unique. The most obvious way to define a preferred choice of parameters is to introduce a distance between model and data in the case of vanishing errors (marked by superscript zeros). This can be done by

$$D(\bar{p}) := \sum_b \{ \rho_b^0 - \bar{m}_b^0(\bar{p}) \}^2$$

The *preferred model parameters* are then defined as the absolute minimum of this function. If the sum is done over different sets of bins, we obtain different *distance functions* and therefore in general different sets of preferred model parameters. A commonly acceptable choice of the set of bins is needed in order to get comparable results. (A commonly accepted definition of a distance measure is also needed.)

### 6.4.2. Types of systematic errors

Two types of systematic errors appear. The first is due to the deviation between the model of interest and the real distributions. The second is because of the systematic errors of the measurement. For the second type of errors estimates of the absolute values are known, while there is no prior knowledge about the model deviations. Thus we try to include all the knowledge about the systematic errors of the measurement in the estimation function, and restrict the tuning to the largest set of bins that can be explained by the model. This set of bins is called the "*fitable region*" (c.f. section 6.4.5).



One needs to be careful with the setting of  $n$  in the definition of the linear region, for otherwise another systematic error due to an imperfect parametrization will appear. Because we are using very small values for  $n$  as motivated by the considerations contained in the last section, these contributions should not be noticeable.

### 6.4.3. Treatment of exactly known systematic errors

In the construction of the estimation function we always used the assumption that the mean value of the quantities  $\bar{d}_b := \bar{\rho}_b - \bar{m}_b(\bar{p}^0)$  is zero<sup>4</sup> (c.f. section 5.3.2). This is only true if there is (at least) one set of parameters  $\bar{p}^0$  that results in mean values for the model predictions  $\bar{m}_b(\bar{p}^0)$  that are equal to the mean values of the data  $\bar{\rho}_b$ . If there are systematic deviations, this basic assumption is not satisfied. Under this circumstances, the mean value of the difference  $\bar{d}_b$  will be different from zero, and if we want to use the same argumentation as in previous sections, the generalized difference

$$\bar{d}_b := \left\{ \bar{\rho}_b - \bar{m}_b(\bar{p}^0) \right\} - \varepsilon_b$$

has to be used. Here  $\varepsilon_b$  is the difference between the mean of the measured data-value and the mean value of the model prediction in bin  $b$  derived at a preferred set of optimal model parameters in the sense of section 6.4.1. Now, the straight forward generalization of the estimation function to the case of systematic errors can be given as

$$S_{\text{sys}}(\bar{p}) := \sum_b \left( \frac{\bar{\rho}_b - \bar{m}_b(\bar{p})}{\sigma_b} - \frac{\varepsilon_b}{\sigma_b} \right)^2$$

From this formula we can see that systematic errors are negligible if the condition  $\sigma_b \gg \varepsilon_b$  holds. In section 6.3.2 we calculated a value  $n$  that was used in the definition of the linear range by taking only the statistical widths into account. Because it seems possible to include the systematic errors of the measurement in the estimation function (c.f. the next section), a comparison to the whole error

---

<sup>4</sup> Apart from the assumption that the data and the model predictions are normally distributed.

should be done, and a somewhat bigger value of  $n$  would also lead to satisfactory results.

Unfortunately usually we cannot know the exact values (with the right signs) of the systematic errors, and so this formula can not be used for parameter estimation.

#### 6.4.4. Treatment of approximately known systematic errors

Because estimations of the systematic errors of the measurements are known, they should be included in the fit according to the general idea that an optimal amount of information should be used. The formula derived above is not usable, because the signs of the errors are not known. Instead of this, a quite obvious method will be introduced and tested which is also used in the literature. The idea is to add the systematic errors to the statistical widths and use this (quadratic) sum:

$$S_{yy}(\bar{p}) := \sum_b \frac{(\bar{p}_b - \bar{m}_b(\bar{p}))^2}{\sigma_b^2 + \varepsilon_b^2} \quad (6.1)$$

That means one wants to "de-weight" the contribution of all the bins that are known to be uncertain. This procedure seems to be the most natural one, and does indeed lead to satisfactory results, as is shown by the following example. Beside this successful test, it is easy to see, that this ansatz follows the right direction. Systematic errors that are not included lead to an average increase of the numerators of the estimation function, while this (quadratic) summation decreases the terms in increasing the denominators. Therefore one could expect that an overall good approximation to the correct estimation function is constructed.

The following example shows, that one could calculate rather senseless results if big systematic errors are neglected, while the results (of this example) are quite satisfactory if one is using the summation ansatz.

*Example 6.1: Estimation with systematic measurement errors*

Idea: We want to demonstrate, that a neglect of systematic errors can cause senseless results while the summation-ansatz leads to satisfactory predictions.

Principle: "Data distributions" are simulated by HERWIG while JETSET provides the "model predictions"<sup>5</sup>. On this basis very good estimates of the systematic deviations are known<sup>6</sup>. Including this systematic errors by the summation ansatz (with quadratic and linear addition), the estimation function is tested by a fit of several JETSET-parameters to "HERWIG-data".

In the following table the results of some fits are shown. The fits were done using the distributions given in table 6.5. The systematic errors are in the range of a few percentage points.

Par.	$p^0$	without $\varepsilon_b^{sys}$		quadratic sum		linear sum	
		$\varepsilon_p$	$\Delta_p$	$\varepsilon_p$	$\Delta_p$	$\varepsilon_p$	$\Delta_p$
$\Lambda$	0.30	<b>0.0009</b>	<b>0.0347</b>	0.0035	-0.0042	0.0044	-0.0026
$Q_0$	1.40	0.0111	0.7664	0.0492	0.0784	0.0642	0.1262
$\sigma$	0.36	0.0005	-0.0324	0.0023	0.0021	0.0027	0.0009
$b$	0.90	<b>0.0040</b>	<b>-0.0576</b>	0.0150	-0.0412	0.0191	-0.0479

Table 6.4: Fitting results

Here  $\varepsilon_p$  is the error that was calculated by the matrix of second derivatives, and  $\Delta_p := p^{fit} - p^0$ . Bold numbers were used if  $|\Delta_p| > 2.5\varepsilon_p$ . Because there exist cases where estimates based on a linear sum of the errors  $\sigma_{tot} = \sigma_b + |\varepsilon_b|$  are worse than estimations based on the quadratic sum  $\sigma_{tot} = \sqrt{\sigma_b^2 + \varepsilon_b^2}$  (even if the calculated errors were large enough) this second

<sup>5</sup> For the values of the model parameter used c.f. appendix B, section 4. JETSET is used in the version of isotropic decaying gluons.

<sup>6</sup> Imperfections due to the limited statistic are small.

one is used. The test quantities for the distributions used in this example are given in the next table.

<i>distribution</i>	<i>bins</i>	<i>without</i> $\epsilon_b^{\text{sys}}$	<i>quadr.</i> <i>sum</i>	<i>sum</i>
<i>S</i>	23	701.7	20.5	13.8
<i>A</i>	16	1245.9	16.1	12.7
<i>1-T</i>	21	1634.8	17.8	14.2
<i>m</i>	18	1524.5	16.2	12.7
$-\ln(y_3)$	14	1480.4	12.0	9.2
$x_p, (x < 0.02)$	8	273.4	6.4	5.3
$x_p, (x > 0.02)$	38	6916.5	42.5	34.8
$p_i^{\text{in}}(S)$	25	2773.7	21.8	16.6
$p_i^{\text{out}}(S)$	19	2251.1	16.8	13.7
$\chi_{\text{ges}}^2/n_{\text{bin}}$	182	18801.9/182	170.2/182	133.1/182

Table 6.5: Test quantities with and without contribution of systematic errors

From the previous table we observe that it seems possible to bring the value of the test quantity into the confidence region, if a (quadratic) sum of statistical and systematic errors is used.

**Conclusion:** *If systematic errors appear, they have to be included in the estimation procedure, if the results need to be trusted. The ansatz of using a quadratic sum of statistical width and systematic error seems to be a good candidate for including these systematic uncertainties.*

o

#### 6.4.5. Restriction to a fitable set of bins

In the last section we arrived at a possible way for taking the systematic errors of measurements into account. The remaining question is how to handle systematic discrepancies between the real distributions and the model under consideration. First of all an example should be introduced that illustrates the problems that appear if one tries to ignore these systematic model inaccuracies.

Even if the systematic errors of the measurements are included in the estimation function, systematic influences from model deviations and imperfect parametrizations remain.

*Example 6.2: Straightforward fits with imperfect models*

*Problem: Big systematic deviations appear in the estimates of the model parameters if the formulas derived in the last chapter are used in the case of systematic deviations between models and data. In other words: Fits to different (sets of) distributions lead to sets of systematically deviating parameters.*

*Principle: The QCD-model JETSET is used because it is well known, that this model describes for example the distribution of  $S$ , while it fails to describe the distributions of  $p_i^{\text{out}}$  and  $p_i^{\text{in}}$  (together with other distributions). Fits of some model parameters are done both in the well described and also in the problematic areas. We use real ALEPH data and include systematic errors of the measurement in the estimation function.*

From earlier fits earlier we know that for example the sphericity  $S$  is a quantity that can be described very well by *JETSET* together with (parts of) other distributions (for values of a global fit c.f. chapter 7). We denote this by saying that the whole  $S$ -distribution is a *fitable range*. A fit of the parameters  $\Lambda$ ,  $Q_0$ ,  $\sigma$  and  $B$  led to the following results. (The other parameters are fixed to the values given in appendix B, section 4)

$p$	$p^0$	$p^{\text{fit}}$	$\varepsilon_p$
$\Lambda$	0.30	0.3058	0.0136
$Q_0$	1.40	1.3118	0.2150
$\sigma$	0.36	0.3958	0.0496
$b$	0.90	1.0188	0.1004

Table 6.6: Fit to  $S$  ( $\eta \cong 6.6$ ,  $n_{\text{bin}} = 23$ )

$p$	$p^0$	$p^{fit}$	$\epsilon_p$
$\Lambda$	0.30	0.3408	0.0057
$Q_0$	1.40	1.2782	0.3631
$\sigma$	0.36	0.2956	0.0070
$b$	0.90	1.2028	0.0613

Table 6.7: Fit to  $p_i^{in}$  ( $\eta \cong 21.1$ ,  $n_{bin} = 25$ )

$p$	$p^0$	$p^{fit}$	$\epsilon_p$
$\Lambda$	0.30	0.5173	0.0086
$Q_0$	1.40	2.0102	0.3178
$\sigma$	0.36	0.2808	0.0044
$b$	0.90	1.6079	0.0699

Table 6.8: Fit to  $p_i^{out}$  ( $\eta \cong 17.1$ ,  $n_{bin} = 19$ )

Even if the test quantities  $\eta$  are in the preferred region, the fits resulted in incompatible values. This discrepancy is dramatic for  $\Lambda$ . The deviation between the  $p_i$ -fits is 12.5 times the sum of their errors, so that these errors are of little usage (and 17.1 times the quadratic sum).

Remark: Here we disregard the linear ranges. Especially the  $p_i^{out}$ -fit is far outside the linear region, so the cause of this big deviation is a composition of imperfect model and the imperfect parametrization of this imperfect model. If one tries to make a fit to all three distributions, the result is:

$p$	$p^0$	$p^{fit}$	$\epsilon^{fit}$
$\Lambda$	0.30	0.3376	0.0032
$Q_0$	1.40	1.9688	0.1572
$\sigma$	0.36	0.3610	0.0022
$b$	0.90	0.9567	0.0296

Table 6.9: Fit to  $S$ , ( $\eta \cong 41.1$ )  $p_i^{in}$  ( $\eta \cong 279.3$ ) and  $p_i^{out}$  ( $\eta \cong 570.8$ )

Now the test quantities indicate well that something is not as it should be, and the obvious assumption is to trust the  $S$ -distribution most.

**Conclusion:** *The parameter errors calculated by the matrix of second derivatives cannot be trusted if systematic deviations between the parametrization of the model and the measured distributions appear. Thus fits to different distributions will in general result in estimations for the parameters which deviate systematically (more than would be expected from the calculated errors).*

o

The result of the last example offers little surprise. If these additional deviations between measurements and model predictions exist, they should be included in the estimation function. If this is not done, it is the same as if they are considered as negligible. In the example above, this is definitely wrong. But if wrong information is included in the estimation process, who would be surprised if the result is wrong, too?

The derived formulae can be used in two possible scenarios. The first is to restrict the fit to regions where these systematic deviations can be neglected. In this case we work on solid ground, because all the assumptions are fulfilled. The second possibility is to include this wrong information, and to calculate additional resulting systematic errors.

In this work an approximation to the first scenario is used. One point is that this scenario is the more honest one, because it points out that the models are able to describe parts of the distributions, while they fail to describe every details of hadronic events. In addition the calculation of the systematic errors from imperfect descriptions cause some rather serious technical problems.

We can for example calculate a systematic error for each of the parameters by making a loop over all possible sets of distributions and by comparing the results. This is not trivial because if we take all possible sets of distributions we should in principle perform a loop over all different combination of bins which require a lot of computing time. Even if we restrict ourselves to all possible combinations of different distributions, their total number is still too large to go through the full loop.

To avoid this problem one could argue, that the main deviation will come from the fits to single distributions. If more than one distribution is used in the fit, the result will always be a compromise between fits to single distributions. Therefore it should be possible to calculate the maximum deviation by looping over all distributions, and by performing a fit to each of them (if this is possible, i.e. if the number of bins is bigger then the number of free parameters) and calculate a systematic error by comparing the result of these fits. Even in this case serious problems remain. The size of the calculated errors will differ if different numbers of parameters are used in the fits. In addition, it will not be possible to perform all these fits by using only one linear parametrization, because the results will in general lie outside the linear range. That means if we try to perform such a calculation of systematic errors of the model parameters, we have the choice to use higher order parametrisations, or to calculate unserious results.

Because of all these difficulties arising in the calculation of systematic parameter errors, we restrict the analysis to a "*fitable region*". A fitable region is a set of bins where all the assumptions introduced while deriving the formulas that were used in the fitting procedure are fulfilled. Especially the deviation between data and model predictions should not be much bigger than the given errors. The remaining problem is to find a good approximation of this fitable region. In this work we use two possibilities to choose a candidate for a the fitable region. Both are based on major deviations from optimal values of test quantities computed in global fits, and both lead to very similar sets of bins. Nevertheless both are approximations, and the need for a systematic error of the fitted parameters due to these approximations is given. Even if they exist, the systematic errors of the parameters due to an imperfect candidate of the fitable region will be much smaller than for the previous discussed case, and therefore a calculation makes sense even if only the linear parametrization is used.

A fit to all bins of a given set of distributions was done to get a first candidate for a fitable region. All bins where the deviation between data and the model prediction exceeds 2.3 times the error of the measurement were rejected<sup>7</sup>. Then, a fit is done using all remaining bins, and the results are taken as the "*best fit parameter values*" ("*s-fit*"). A second candidate is obtained by performing a series of fits. Starting from the same global fit as in the first case, only the bin with the worst (that means biggest) contribution to the test quantity is rejected, and the whole procedure is repeated with all remaining bins. The iteration is stopped as soon as the worst contribution mentioned above is less than  $(2.3)^2$  (or again the deviation between model and data exceeds 2.3 times the expected error). All the bins that pass this series of fits without rejection define the second candidate for an approximately fitable region, and the differences between the parameter values fitted to this second set of bins and the parameters derived from the first set are taken as a systematic errors.

---

<sup>7</sup> The value 2.3 is taken, because it leads to a test quantity of  $\eta/n_{fit} \cong 1$ , if all  $n_{fit}$  bins used in the fit are taken into account.



## 6.5. Additional Tests of LinFit

### 6.5.1. Principle

All the tests in this section are done with the *QCD*-model *JETSET 7.4* in the version of isotropic decaying gluons. Analogous results are expected for the other variant of *JETSET*, and for *ARIADNE*, because of their similarities. The fitting of *HERWIG* was not tested separately. A set of parameter values  $\bar{p}^0$  was chosen to generate test distributions with a statistics  $s_{dat}$ . Because of that, "experimental" distributions were known together with the corresponding values of the parameters of the optimal model. The interesting question dealt with the conditions under which *LinFit* is able to reproduce these data within the calculated errors.

A test of this type can not be a proof of the correctness of the algorithm, but it is an impressive check. Even if it is impossible to make a proof of the correctness using only a few examples, these examples are able to exclude false assumptions, as for example the naive ansatz for including the parametrization width in the estimation function (5.5)

### 6.5.2. Test at high simulation statistics

High simulation statistics means that this statistics is high compared to the data statistics. For example in the case of our measurements the statistics of the simulation was  $s_{sim} := 5 \cdot 10^5$  (and therefore  $s_{sim}^0 := 2 \cdot 10^6$  in the expansion point) while it was  $s_{dat} := 571825$  for the data. The former is available at every point in the parameter space (except the expansion point) that is taken into account, and used to do the parametrization. Because of that one could expect, that the widths of the parametrization can be neglected, and the simple estimation function (5.4) is enough. This expectation is to be tested here.

To avoid complications due to improper parametrisations, the set of parameters  $\bar{p}^0$  that is used to generate the test distributions is used as the expansion point. To get a situation that is similar to the real analysis, the following statistics were used<sup>8</sup>:

---

<sup>8</sup> In fact, the data errors also have a contribution from the correction procedure.

$$s_{dat} := 6 \cdot 10^5, \quad s_{sim} := 5 \cdot 10^5 \quad (\Rightarrow s_{sim}^0 := 2 \cdot 10^6)$$

We especially point out that we only use statistical widths here, and that in the real estimation process also systematic errors appear. The next section illustrates the fact, that a usage of the simplified estimation function (5.4) leads to satisfactory results even in the case of vanishing systematic errors.

### 6.5.2.1. Fit under measurement conditions

In this section we perform a check to see if the simplified estimation function (5.4) can be used to fit the models in the case of our analysis, or if the widths of the parametrization have to be taken into account. For that reason, all parameters that were used in the model tuning of *JETSET* were tuned to the same set of distributions as in the model tuning to real data. In the following tables we compare the results of both estimation procedures.

distribution	$n_{bin}$	$\chi^2_{-sim}$	$\chi^2_{+sim}$
$S$	23	21.0	16.2
$A$	16	21.5	16.6
$1-T$	21	15.5	11.9
$m$	18	15.0	11.6
$-\ln(y_3)$	14	23.1	17.8
$x_p < 0.02$	8	18.5	14.7
$x_p > 0.02$	38	49.7	38.5
$p_t^{out}(S)$	19	16.6	13.5
$p_t^{in}(S)$	25	32.9	26.5
$-\ln(x_p^{K^0})$	28	42.1	32.3
$x_E^\eta$	18	17.1	13.2
$x_E^{\eta'}$	9	11.4	8.8
$x_p^{\rho^0}$	8	12.2	9.4
$\langle K^{*+} \rangle$	1	0.2	0.1

distribution	$n_{bin}$	$\chi^2_{-sim}$	$\chi^2_{+sim}$
$x_p^{K^{*0}}$	8	9.3	7.2
$x_p^{\rho^0}$	8	4.8	3.7
$x_p^{\omega^0}$	6	9.1	7.0
$\langle f_2 \rangle$	1	1.6	1.2
$\langle f_0 \rangle$	1	0.0	0.0
$-\ln(x_p^\Lambda)$	22	25.1	19.3
$\langle \Xi^- \rangle$	1	0.3	0.2
$\langle \Sigma(1385)^+ \rangle$	1	0.1	0.1
$\langle \Omega^- \rangle$	1	0.0	0.0
$\langle \Xi(1530)^0 \rangle$	1	1.1	0.9
$-\ln(x_p^{K^*}) < 0.018$	11	16.4	12.6
$-\ln(x_p^{K^*}) > 0.07$	18	13.4	10.3
$-\ln(x_p^{(\rho,\bar{\rho})}) < 0.018$	6	9.9	7.6
$-\ln(x_p^{(\rho,\bar{\rho})}) > 0.07$	18	6.0	4.6
<b>sum</b>	<b>339</b>	<b>393.9</b>	<b>305.89</b>

Table 6.10: Test quantities for the first fit under measurement conditions

The last table was split up into two parts in order to save space. Even if the test quantities are bigger if parametrization errors are ignored, they fall into a region as expected and can therefore be used to establish the quality of the fit. In the table 6.11 the tuned parameter values are given together with the set of parameters  $\bar{p}^0$ . In this fit, we fixed the following parameters:  $a:=0.4$ ,  $\varepsilon_c := -0.04$ ,  $\varepsilon_b := -0.0035$ ,  $P_{c,d}^{S=1} := 0.65$ ,  $P_{L=1 \wedge J=2}^{S=1} := 0.2$  and  $\bar{\eta} := 0.25$ . These values are used in the generation of test distributions.

Par.	$p^0$	with par		without par	
		$\varepsilon_p$	$\Delta_p$	$\varepsilon_p$	$\Delta_p$
$\Lambda$	0.30	0.0010	-0.0003	0.0008	-0.0004
$Q_0$	1.40	0.0096	-0.0036	0.0086	-0.0038
$\gamma_s$	0.30	0.0005	-0.0007	0.0004	-0.0007
$P_{u,d}^{S=1}$	0.55	0.0010	0.0008	0.0009	0.0008
$P_s^{S=1}$	0.50	0.0009	0.0013	0.0008	0.0013
$qq/q$	0.10	0.0003	-0.0003	0.0003	-0.0003
$(su)/(du)$	0.60	0.0051	0.0000	0.0045	0.0000
$\sigma$	0.36	0.0005	0.0004	0.0005	0.0004
$b$	0.90	0.0039	0.0015	0.0036	0.0014
$\bar{b}_l$	0.6	0.0050	0.0044	0.0044	0.0044

Table 6.11: Estimated parameter values

A first indication that shows that both of them produce good estimates is the fact, that none of the linear ranges was exceeded, even if we used  $n=0.01$ . In table 6.11, "with par" means that the errors of the parametrization are included in the estimation, while they are neglected in the column "without par".  $\varepsilon_p$  is the error that is calculated by the second derivatives of the estimation function, and  $\Delta_p := p^{fit} - p^0$  is (again) the deviation of the fitted value from the nominal value.

In none of the presented estimations, the fitted value differs by more than two sigmas from the nominal value  $p^0$ , even in the case of the simple estimation function. If we include the systematic errors in this function, the error of the parametrization relative to the error of the measurement decreases again, and the accuracy of the approximation by the simple estimation function should increase.

Remark about correlations and choice of distributions: It is clear, that every available piece of information should be used in a fit. Therefore one should expect all the available distributions in the table above. We do not use all of these distributions, because some of them are highly correlated, and we did not include correlations between bins of different distributions in the estimation function.

What will happen if correlated distributions are included? This can be seen most impressively in the case of identical (and therefore maximally correlated) distributions. If we include a distribution twice, it is the same as if we reduce the error used in the estimation function by a factor of  $1/\sqrt{2}$ . In other words: If we use correlated distributions in the fit, we possibly underestimate the parameter errors.

Since the distributions used in this fit led to a good estimation of the nominal parameter values, the correlations for this set of distributions should indeed be negligible. The same can be said about the correlation between bins. This should also be true for single particle distributions because they were also used in the previous fit.

Conclusion: *Both estimation functions produce satisfactory estimates in the case that we are using for the measurement, and we can therefore use the simpler one, especially because an inclusion of the systematic errors of the measurements will make the difference between the predictions even smaller. It seems also to be a good approximation if we neglect bin correlations and correlations between different distributions for the set of distributions used in this section.*

### 6.5.3. Test at high data statistics

This test is done in a way that is similar to the previous section. The following values were used to increase the widths of the parametrization:

$$s_{dat} := 2 \cdot 10^6, \quad s_{sim} := 2.5 \cdot 10^4 \quad (\Rightarrow s_{sim}^0 := 1 \cdot 10^5)$$

The fit results are given in the following tables. The distributions of table 6.13, together with the multiplicities of table 6.10 were used to perform these fits. We also tried a fit by using the naive generalization of the estimation function (5.5), but this estimation failed totally. (For example one value was  $qq/q \cong 56.7$ ).

Par.	$p^0$	without par		with par	
		$\epsilon_p$	$\Delta_p$	$\epsilon_p$	$\Delta_p$
$\Lambda$	0.30	0.0005	-0.0019	0.0023	-0.0021
$Q_0$	1.40	0.0033	-0.0025	0.0185	-0.0023
$\gamma_s$	0.30	0.0002	0.0015	0.0010	0.0016
$P_{u,d}^{S=1}$	0.55	0.0005	0.0029	0.0024	0.0032
$P_s^{S=1}$	0.50	0.0005	0.0018	0.0022	0.0017
$P_{c,d}^{S=1}$	0.65	0.0017	-0.0040	0.0085	-0.0054
$P_{L=1 \wedge J=2}^{S=1}$	0.20	0.0003	-0.0004	0.0013	-0.0004
$qq/q$	0.10	0.0001	-0.0004	0.0006	-0.0004
$(su)/(du)$	0.60	0.0024	-0.0024	0.0118	-0.0030
$\sigma$	0.36	0.0003	0.0006	0.0013	0.0008
$a$	0.40	0.0017	0.0033	0.0083	0.0027
$b$	0.90	0.0020	-0.0090	0.0101	-0.0107
$\epsilon_c$	-0.04	0.0003	0.0006	0.0013	0.0008
$\epsilon_b$	-	0.0000	0.0000	0.0002	0.0000
$\bar{b}_l$	0.6	0.0025	0.0176	0.0123	0.0195
$\bar{\eta}'$	0.25	0.0008	-0.0071	0.0043	-0.0073

Table 6.12: Some fitting results

All deviations bigger than three sigmas are typed in bold letters in the table above. We can see, that the simple estimation is invalid in the case of non-negligible parametrization widths, and that the estimation function (5.6) seems to

be a good candidate for a generalization. For all three estimation procedures, the values of the test quantities are provided in the following table. Only the test quantities for the distributions were listed in order to save space, while the test quantities of the multiplicities were included in the final summation. In the case of neglecting the parametrization errors, the values of the test quantity are much too big, and clearly indicate, that something did not work. It is remarkable that the test quantities for the naive generalization (5.5) are not too bad, even if the estimation failed.

<i>Name</i>	<i>Bins</i>	$\chi^2_{\text{without par}}$	$\chi^2_{\text{naive}}$	$\chi^2_{\text{with par}}$
<i>S</i>	23	452.9	13.6	20.4
<i>A</i>	16	202.9	9.5	9.7
<i>1-T</i>	21	715.2	14.8	33.1
$x_p, (x < 0.02)$	8	103.4	9.1	5.5
$x_p, (x > 0.02)$	38	696.2	30.2	33.9
$p_i^{\text{out}}(S)$	19	304.4	14.5	17.6
$-\ln(x_p^{\pi^\pm}), (x < 0.018)$	16	568.3	13.6	26.6
$-\ln(x_p^{\pi^\pm}), (x > 0.045)$	23	475.1	13.1	21.9
$-\ln(x_p^{K^0})$	28	663.9	13.9	32.0
$-\ln(x_p^{K^\pm}), (x < 0.018)$	11	234.4	9.4	11.4
$-\ln(x_p^{K^\pm}), (x > 0.070)$	18	328.5	11.6	15.4
$-\ln(x_p^{(p,\bar{p})}), (x < 0.018)$	6	49.7	2.4	2.4
$-\ln(x_p^{(p,\bar{p})}), (x > 0.070)$	18	653.9	9.0	30.8
$\chi^2_{\text{ges}}/n_{\text{bin}}$	258	5634.0/258	177.5/258	269.9/258

Table 6.13: Test quantities for all three estimation procedures

**Conclusion:** *If the widths of the parametrization are not negligible, the estimation function (5.6) remains as a candidate for a generalized estimation function. The simple estimation function (5.4) and the obvious generalization (5.5) fail.*

## Chapter 7.

### Results II: Tuned models and comparison between models and data

#### 7.1. Abstract

In section 1.2, all model parameters that were calculated with the fitting algorithm *LinFit* are presented together with the values of the expansion point and the default values. In addition a factor that marks the distance between this expansion point and the result of the fit in units of a linear range defined with  $n:=0.1$  for *JETSET* and *ARIADNE*, and  $n:=0.5$  for *HERWIG* is given.

Some of the model parameters were set to values that differed from their default values. These non-default values are discussed in chapter 2, and they are also listed here to give a better overview. The test quantities for all distributions used in the fit are given both for the "maximum fitable region" defined by one global fit and for the whole distributions. The latter strongly indicate distributions where the models describe the data well, and regions, where the description fails.

Section 1.3 shows some graphical impressions. Here the models are compared with the measurement in a graphical way. In this section details about the quality of the models or about the problems they have in describing the data can be seen.

## 7.2. Model parameters and test quantities

### 7.2.1. JETSET 7.4 with anisotropic gluon decay

Even if the linear range with  $n=0.1$  is exceeded in some cases, the average value of the deviation of the fit result in units of the linear range is  $0.72$  which we take as a sign, that the result can be trusted. The deviation between fitted values and the expansion point in units of the linear range is denoted by

$$f^{lin} := \left| \frac{p^{fit} - p^0}{\Delta_{lin}} \right|$$

where  $\Delta_{lin}$  is the width of the linear range.  $\Delta p_l$  is the distance from the expansion point to the lower (left) end of the parametrized range, while  $\Delta p_r$  is the same for the upper (right) end.  $\Delta p^{sys} := p_i^{fit} - p_s^{fit}$  is the deviation between the estimations found in the fits to the different candidates for a fitable range (s marks the candidate found by a single fit, and i is used for the candidate found by iteration).

<i>name</i>	$p^{default}$	$\Delta p_l$	$p^0$	$\Delta p_r$	$p^{fit}$	$\Delta p^{fit}$	$\Delta p^{sys}$	$f^{lin}$
$\Lambda$	0.29GeV	0.050	0.291	0.050	0.299	0.003	-0.011	1.07
$Q_0$	1.0GeV	0.400	1.520	0.400	1.560	0.050	0.018	0.55
$\sigma$	0.36GeV	0.050	0.370	0.050	0.381	0.002	0.010	1.66
$a$	0.3		0.400		0.400	fixed		
$b$	0.58GeV <sup>2</sup>	0.150	0.805	0.150	0.808	0.014	0.004	0.17
$\epsilon_c$	0.05		0.040		0.040	fixed		
$\epsilon_b$	0.005		0.004		0.004	fixed		
$P_{u,d}^{S=1}$	0.5	0.150	0.558	0.150	0.534	0.019	0.049	1.21
$P_s^{S=1}$	0.6	0.150	0.466	0.150	0.473	0.021	0.086	0.26
$P_{c,b}^{S=1}$	0.75		0.650		0.650	fixed		
$P_{L=1 \wedge J=2}^{S=1}$	0.0		0.200		0.200	fixed		
$\bar{\eta}$	0.4		0.275		0.275	fixed		
$\gamma_s$	0.3	0.050	0.287	0.050	0.292	0.004	0.012	0.74
$qq/q$	0.1	0.030	0.107	0.030	0.106	0.002	0.003	0.28
$(su)/(du)$	0.3	0.300	0.679	0.300	0.659	0.042	-0.063	0.44
$\bar{b}_l$	1.0	0.400	0.564	0.400	0.622	0.029	-0.011	0.84

 Table 7.1: Results from *LinFit*



In addition we used  $P_{L=1 \wedge J=2}^{S=1} : P_{L=1 \wedge J=1}^{S=1} : P_{L=1 \wedge J=0}^{S=1} = 5:3:1$  and  $P_{L=1}^{S=0} = P_{L=1 \wedge J=1}^{S=1}$ . The test quantities corresponding to the parameter values above are listed in the next table.  $\eta_{fit}$  does only include the  $n_{fit}^{bins}$  bins of the fitable region defined by a single global fit, while  $\eta_{all}$  includes all  $n^{bins}$  bins of a given distribution. The former values are given only for the distributions used in the fit, while the latter were calculated for other interesting distributions as well.

<i>distribution</i>	$n^{bins}$	$n_{fit}^{bins}$	$\eta_{all}$	$\eta_{fit}$
<i>S</i>	23	21	23.66	13.1
<i>A</i>	16	9	113.0	29.8
<i>1-T</i>	21	19	26.5	17.9
<i>m</i>	18	14	101.9	24.2
$-\ln(y_3)$	14	13	17.0	9.2
$x_p < 0.02$	8	3	210.6	10.4
$x_p > 0.02$	38	30	124.9	30.3
$p_i^{out}(S)$	19	5	882.2	16.4
$p_i^{in}(S)$	25	12	311.8	16.8
$-\ln(x_p), K^0$	28	28	25.7	25.7
$x_E, \eta$	18	14	62.8	14.9
$x_E, \eta'$	9	9	7.2	7.2
$x_p, \rho^0$	8	7	7.4	4.0
$\langle K^{*+} \rangle$	1	1	0.1	0.1
$x_p, K^{*0}$	8	7	12.1	6.5
$x_p, \varphi^0$	8	3	47.8	3.6
$x_p, \omega^0$	6	4	41.0	2.7
$\langle f_2 \rangle$	1	1	0.2	1.2
$\langle f_0 \rangle$	1	0	17.4	
$-\ln(x_p), \Lambda$	22	20	34.8	12.2
$\langle \Xi^- \rangle$	1	0	5.6	
$\langle \Sigma(1385)^+ \rangle$	1	1	0.1	0.1
$\langle \Omega^- \rangle$	1	1	0.5	0.5
$\langle \Xi(1530)^0 \rangle$	1	1	0.9	0.9
$-\ln(x_p), K^+(x < 0.018)$	11	11	4.0	4.0
$-\ln(x_p), K^+(x > 0.07)$	18	17	26.2	20.4
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	6	6	1.1	1.1
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	18	18	17.1	17.1
<b>sum</b>	<b>349</b>	<b>275</b>		<b>289.1</b>

<i>distribution</i>	$n^{bins}$	$\eta_{all}$
<i>P</i>	21	36.6
<i>C</i>	24	32.2
<i>M</i>	22	39.8
<i>O</i>	20	158.8
$\bar{M}_h^2$	21	15.9
$B_i$	17	12.1
$B_w$	17	18.4
$-\ln(x_p)$	52	556.2
$p_i^{out}(T)$	19	964.8
$p_i^{in}(T)$	25	164.7
$y_T$	21	53.0
$\langle K^0 \rangle$	1	0.5
$p_i(T), K^0$	25	90.5
$x_E, D^{*\pm}$	15	21.5
$\langle \Lambda \rangle$	1	0.1
$p_i(T), \Lambda$	25	60.4
$-\ln(x_p), \pi^+(x < 0.018)$	16	116.5
$-\ln(x_p), \pi^+(x > 0.07)$	23	3.6

$$\Rightarrow \eta_{fit} / n_{fit}^{bins} \cong 1.05$$

 Table 7.2: Test quantities corresponding to best fit parameters found by *LinFit*

Model parameters and test quantities

In the next two tables, bins that were used in both fits are marked by an "×". All the bins removed from the candidate for a "fitable region" as defined by a single global fit are signed by an "S". The bins which are rejected by the iterative sequence of fits are signed by an "I".

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$S$	×	$S$	×	×	×	×	×	×	×	×	$S$	×	×	×
$A$	×	×	$I$	$S,I$	$S,I$	×	×	$I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$I$
$I-T$	×	×	$S,I$	×	×	×	×	×	$S$	×	×	×	$I$	×
$m$	×	×	×	×	$I$	×	×	×	×	$I$	$S,I$	$S,I$	$S,I$	$S,I$
$-\ln(y_3)$	×	×	×	×	×	×	×	×	×	$S$	×	×	×	×
$x_p < 0.02$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$I$	×						
$x_p > 0.02$	×	×	×	×	×	×	×	$S$	×	×	×	×	×	×
	×	×	×	×	×	×	×	$S,I$	$S,I$	×				
$p_i^{out}(S)$	$S,I$	$I$	$S,I$	×	$S,I$	$S,I$	×	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$
$p_i^{in}(S)$	$S,I$	$S,I$	$S,I$	$S,I$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S$
$-\ln(x_p), K^0$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$x_E, \eta$	$S$	×	×	×	$I$	×	×	×	×	×	$I$	×	$S$	×
$x_E, \eta'$	×	×	×	×	×	×	×	×	×					
$x_p, \rho^0$	×	×	×	×	×	×	×	$S$						
$\langle K^{*+} \rangle$	×													
$x_p, K^{*0}$	×	×	×	×	$S$	×	×	$I$						
$x_p, \varphi^0$	×	$S$	$S$	$S$	$S$	×	$I$	$S,I$						
$x_p, \omega^0$	×	×	×	×	$S,I$	$S,I$								
$\langle f_2 \rangle$	×													
$\langle f_0 \rangle$	$S,I$													
$-\ln(x_p), \Lambda$	×	$S,I$	$S,I$	×	×	×	×	×	×	×	×	×	×	×
$\langle \Xi^- \rangle$	$S,I$													
$\langle \Sigma(1385)^{\pm} \rangle$	×													
$\langle \Omega^- \rangle$	×													
$\langle \Xi(1530)^0 \rangle$	×													
$-\ln(x_p), K^{\pm}(x < 0.018)$	×	×	×	×	×	×	×	×	×	×	×			
$-\ln(x_p), K^{\pm}(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	×	×	×	×	×	×								
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Table 7.3: Accepted bins in the "fitable range" (first part)

15	16	17	18	19	20	21	22	23	24	25	26	27	28	
×	×	×	×	×	×	×	×	×						$S$
$I$	×													$A$
×	×	×	×	×	×	×								$I-T$
$I$	×	×	×											$m$
														$-\ln(y_3)$
														$x_p < 0.02$
×	×	×	$S$	$S,I$	×	$S,I$	$S$	$S$	×	×	×	×	×	$x_p > 0.02$
$S,I$	$S,I$	$S,I$	$S,I$	$I$										$p_i^{out}(S)$
×	×	×	×	×	×	×	×	×	×	×				$p_i^{in}(S)$
×	×	×	×	×	×	×	×	×	×	×	×	×	×	$-\ln(x_p), K^0$
×	×	$S,I$	$S,I$											$x_E, \eta$
														$x_E, \eta'$
														$x_p, \rho^0$
														$\langle K^{*+} \rangle$
														$x_p, K^{*0}$
														$x_p, \phi^0$
														$x_p, \omega^0$
														$\langle f_2 \rangle$
														$\langle f_0 \rangle$
×	×	×	×	×	×	×	×							$-\ln(x_p), \Lambda$
														$\langle \Xi^- \rangle$
														$\langle \Sigma(1385)^+ \rangle$
														$\langle \Omega^- \rangle$
														$\langle \Xi(1530)^0 \rangle$
														$-\ln(x_p), K^+(x < 0.018)$
$S,I$	×	×	×											$-\ln(x_p), K^+(x > 0.07)$
														$-\ln(x_p), (p, \bar{p})(x < 0.018)$
×	×	×	$I$											$-\ln(x_p), (p, \bar{p})(x > 0.07)$

Table 7.4: Accepted bins in the "fitable range" (second part)

### 7.2.2. JETSET 7.4 with isotropic gluon decay

The average value of the deviation of the fit result in units of the linear range is 1.17 which we take as a sign that the result can be trusted. Again we calculated

$$f^{lin} := \left| \frac{p^{fit} - p^0}{\Delta_{lin}} \right|$$

where  $\Delta_{lin}$  is the width of the linear range.  $\Delta p_l$  is the distance from the expansion point to the lower (left) end of the parametrized range, while  $\Delta p_r$  is the same for the upper (right) end.  $\Delta p^{sys} := p_i^{fit} - p_s^{fit}$  is the deviation between the estimations found in the fits to the different candidates for a fitable range (s marks the candidate found by a single fit, and i is used for the candidate found by iteration).

<i>name</i>	$p^{default}$	$\Delta p_l$	$p^0$	$\Delta p_r$	$p^{fit}$	$\Delta p^{fit}$	$\Delta p^{sys}$	$f^{lin}$
$\Lambda$	0.29GeV	0.050	0.315	0.050	0.324	0.003	-0.009	1.06
$Q_0$	1.0GeV	0.400	1.482	0.400	1.487	0.040	-0.020	0.06
$\sigma$	0.36GeV	0.040	0.364	0.040	0.373	0.002	0.009	1.36
$a$	0.3		0.400		0.400	fixed		
$b$	0.58GeV <sup>2</sup>	0.100	0.895	0.100	0.931	0.014	-0.007	2.17
$\epsilon_c$	0.05		0.040		0.040	fixed		
$\epsilon_b$	0.005		0.004		0.004	fixed		
$P_{u,d}^{S=1}$	0.5	0.100	0.519	0.100	0.564	0.015	0.006	2.39
$P_s^{S=1}$	0.6	0.100	0.511	0.100	0.474	0.018	0.028	1.88
$P_{c,b}^{S=1}$	0.75		0.650		0.650	fixed		
$P_{L=1 \wedge J=2}^{S=1}$	0.0		0.200		0.200	fixed		
$\bar{\eta}^1$	0.4		0.250		0.250	fixed		
$\gamma_s$	0.3	0.050	0.288	0.050	0.292	0.004	0.005	0.54
$qq/q$	0.1	0.025	0.108	0.025	0.109	0.002	0.001	0.23
$(su)/(du)$	0.3	0.200	0.686	0.200	0.652	0.035	-0.005	0.92
$\bar{b}_l$	1.0	0.300	0.528	0.300	0.592	0.027	0.002	1.15

Table 7.5: Results from *LinFit*

In addition we used  $P_{L=1 \wedge J=2}^{S=1} : P_{L=1 \wedge J=1}^{S=1} : P_{L=1 \wedge J=0}^{S=1} = 5:3:1$  and  $P_{L=1}^{S=0} = P_{L=1 \wedge J=1}^{S=1}$ . The test quantities corresponding to the parameter values above are listed in the next

table.  $\eta_{fit}$  does only include the  $n_{fit}^{bins}$  bins of the fitable region defined by a single global fit, while  $\eta_{all}$  includes all  $n^{bins}$  bins of a given distribution. The former values are given only for the distributions used in the fit, while the latter were calculated for other interesting distributions as well.

<i>distribution</i>	$n^{bins}$	$n_{fit}^{bins}$	$\eta_{all}$	$\eta_{fit}$
<i>S</i>	23	22	12.0	9.0
<i>A</i>	16	14	51.4	36.5
<i>1-T</i>	21	18	61.2	23.9
<i>m</i>	18	16	41.3	23.3
$-\ln(y_3)$	14	14	15.4	15.4
$x_p < 0.02$	8	3	182.5	9.2
$x_p > 0.02$	38	31	86.5	27.2
$p_t^{out}(S)$	19	5	676.9	19.2
$p_t^{in}(S)$	25	13	217.1	11.5
$-\ln(x_p), K^0$	28	27	26.2	21.0
$x_E, \eta$	18	15	50.0	21.5
$x_E, \eta'$	9	9	6.3	6.3
$x_p, \rho^0$	8	8	5.3	5.3
$\langle K^{*+} \rangle$	1	1	0.1	0.1
$x_p, K^{*0}$	8	7	12.3	6.8
$x_p, \varphi^0$	8	3	52.1	3.5
$x_p, \omega^0$	6	4	47.3	3.3
$\langle f_2 \rangle$	1	1	0.1	0.1
$\langle f_0 \rangle$	1	0	16.5	
$-\ln(x_p), \Lambda$	22	20	38.5	11.3
$\langle \Xi^- \rangle$	1	0	5.0	
$\langle \Sigma(1385)^+ \rangle$	1	1	0.1	0.1
$\langle \Omega^- \rangle$	1	1	0.6	0.6
$\langle \Xi(1530)^0 \rangle$	1	1	0.9	0.9
$-\ln(x_p), K^+(x < 0.018)$	11	11	3.7	3.7
$-\ln(x_p), K^+(x > 0.07)$	18	18	24.6	24.6
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	6	6	2.8	2.8
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	18	17	18.1	10.9
<i>sum</i>	349	286		297.9

<i>distribution</i>	$n^{bins}$	$\eta_{all}$
<i>P</i>	21	31.0
<i>C</i>	24	43.1
<i>M</i>	22	41.8
<i>O</i>	20	125.9
$\overline{M}_h^2$	21	30.8
$B_t$	17	17.9
$B_w$	17	36.7
$-\ln(x_p)$	52	485.1
$p_t^{out}(T)$	19	771.5
$p_t^{in}(T)$	25	139.8
$y_T$	21	54.3
$\langle K^0 \rangle$	1	0.5
$p_t(T), K^0$	25	94.8
$x_E, D^{*+}$	15	22.4
$\langle \Lambda \rangle$	1	0.0
$p_t(T), \Lambda$	25	75.3
$-\ln(x_p), \pi^+(x < 0.018)$	16	107.7
$-\ln(x_p), \pi^+(x > 0.07)$	23	3.9

$$\Rightarrow \eta_{fit} / n_{fit}^{bins} \cong 1.04$$

 Table 7.6: Test quantities corresponding to best fit parameters found by *LinFit*

Model parameters and test quantities

In the next two tables, bins that were used in both fits are marked by an "×". All the bins removed from the candidate for a "fitable region" as defined by a single global fit are signed by an "S". The bins which are rejected by the iterative sequence of fits are signed by an "I".

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$S$	×	$S$	×	×	×	×	×	×	×	×	×	×	×	×
$A$	×	×	×	×	×	×	×	×	$I$	$I$	$S,I$	×	$S,I$	$I$
$I-T$	×	$S,I$	$S,I$	×	×	×	×	×	$S,I$	×	×	×	×	×
$m$	×	×	×	×	×	×	×	×	×	×	$I$	$S,I$	$I$	$S,I$
$-\ln(y_3)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$x_p < 0.02$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$I$	×						
$x_p > 0.02$	×	×	×	×	×	×	×	$S$	×	×	×	×	×	×
	×	×	×	×	×	×	×	$S,I$	$S,I$	×				
$p_i^{out}(S)$	$S,I$	$I$	$I$	×	$S,I$	$S,I$	$I$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$
$p_i^{in}(S)$	$S,I$	$S,I$	$S,I$	$S,I$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S$	$S$	×
$-\ln(x_p), K^0$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$x_E, \eta$	$S$	×	×	×	$I$	×	×	×	×	×	×	×	×	×
$x_E, \eta^i$	×	×	×	×	×	×	×	×	×					
$x_p, \rho^0$	×	×	×	×	×	×	×	×						
$\langle K^{**} \rangle$	×													
$x_p, K^{*0}$	×	×	×	×	$S$	×	×	×						
$x_p, \varphi^0$	×	$S$	$S$	$S$	$S$	×	×	$S,I$						
$x_p, \omega^0$	×	×	×	×	$S,I$	$S,I$								
$\langle f_2 \rangle$	×													
$\langle f_0 \rangle$	$S,I$													
$-\ln(x_p), \Lambda$	×	$S,I$	$S,I$	×	×	×	×	×	×	×	×	×	×	×
$\langle \Xi^- \rangle$	$S,I$													
$\langle \Sigma(1385)^+ \rangle$	×													
$\langle \Omega^- \rangle$	×													
$\langle \Xi(1530)^0 \rangle$	×													
$-\ln(x_p), K^+(x < 0.018)$	×	×	×	×	×	×	×	×	×	×	×			
$-\ln(x_p), K^+(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	×	×	×	×	×	×								
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Table 7.7: Accepted bins in the "fitable range" (first part)

15	16	17	18	19	20	21	22	23	24	25	26	27	28	
×	×	×	×	×	×	×	×	×						$S$
×	×													$A$
×	×	×	×	×	×	×								$I-T$
×	×	×	×											$m$
														$-\ln(y_3)$
														$x_p < 0.02$
×	×	×	×	$S$	×	$S$	$S,I$	$S$	×	×	×	×	×	$x_p > 0.02$
$S,I$	$S,I$	$S,I$	$S,I$	$S,I$										$p_i^{out}(S)$
×	×	×	×	×	×	×	×	×	×	×				$p_i^{in}(S)$
×	×	×	×	×	×	×	×	×	$S$	×	×	×	×	$-\ln(x_p), K^0$
×	×	$S,I$	$S,I$											$x_E, \eta$
														$x_E, \eta'$
														$x_p, \rho^0$
														$\langle K^{*+} \rangle$
														$x_p, K^{*0}$
														$x_p, \phi^0$
														$x_p, \omega^0$
														$\langle f_2 \rangle$
														$\langle f_0 \rangle$
×	×	×	×	×	×	×	×							$-\ln(x_p), \Lambda$
														$\langle \Xi^- \rangle$
														$\langle \Sigma(1385)^+ \rangle$
														$\langle \Omega^- \rangle$
														$\langle \Xi(1530)^0 \rangle$
														$-\ln(x_p), K^+(x < 0.018)$
×	×	×	×											$-\ln(x_p), K^+(x > 0.07)$
														$-\ln(x_p), (p, \bar{p})(x < 0.018)$
×	×	×	$S,I$											$-\ln(x_p), (p, \bar{p})(x > 0.07)$

Table 7.8: Accepted bins in the "fitable range" (second part)

### 7.2.3. ARIADNE 4.05

The average value of the deviation of the fit result in units of the linear range is  $0.78$  which we take as a sign that the result can be trusted. Again we calculated

$$f^{lin} := \left| \frac{p^{fit} - p^0}{\Delta_{lin}} \right|$$

where  $\Delta_{lin}$  is the width of the linear range.  $\Delta p_l$  is the distance from the expansion point to the lower (left) end of the parametrized range, while  $\Delta p_r$  is the same for the upper (right) end.  $\Delta p^{sys} := p_i^{fit} - p_s^{fit}$  is the deviation between the estimations found in the fits to the different candidates for a fitable range (s marks the candidate found by a single fit, and i is used for the candidate found by iteration).

<i>name</i>	$p^{default}$	$\Delta p_l$	$p^0$	$\Delta p_r$	$p^{fit}$	$\Delta p^{fit}$	$\Delta p^{sys}$	$f^{lin}$
$\Lambda$	0.22GeV	0.050	0.228	0.050	0.241	0.002	-0.009	1.80
$p_l^{min}$	0.60GeV	0.000	0.749	0.300	0.754	0.024	-0.065	0.19
$\sigma$	0.36GeV	0.060	0.359	0.060	0.362	0.003	0.011	0.51
$a$	0.3		0.400		0.400	fixed		
$b$	0.58GeV <sup>2</sup>	0.150	0.831	0.150	0.861	0.016	0.036	1.73
$\varepsilon_c$	0.05		0.040		0.040	fixed		
$\varepsilon_b$	0.005		0.004		0.004	fixed		
$P_{u,d}^{S=1}$	0.5	0.100	0.566	0.100	0.540	0.018	0.034	1.49
$P_s^{S=1}$	0.6	0.100	0.468	0.100	0.464	0.019	0.060	0.20
$P_{c,b}^{S=1}$	0.75		0.650		0.650	fixed		
$P_{L=1 \wedge J=2}^{S=1}$	0.0		0.200		0.200	fixed		
$\bar{\eta}'$	0.4		0.286		0.286	fixed		
$\gamma_s$	0.3	0.050	0.286	0.050	0.287	0.004	0.012	0.20
$qq/q$	0.1	0.040	0.114	0.040	0.113	0.002	0.003	0.19
$(su)/(du)$	0.3	0.300	0.658	0.300	0.636	0.041	-0.049	0.52
$\bar{b}_l$	1.0	0.400	0.515	0.400	0.581	0.027	-0.014	0.98

Table 7.9: Results from *LinFit*

In addition we used  $P_{L=1 \wedge J=2}^{S=1} : P_{L=1 \wedge J=1}^{S=1} : P_{L=1 \wedge J=0}^{S=1} = 5:3:1$  and  $P_{L=1}^{S=0} = P_{L=1 \wedge J=1}^{S=1}$ . The test quantities corresponding to the parameter values above are listed in the next



table.  $\eta_{fit}$  does only include the  $n_{fit}^{bins}$  bins of the fitable region defined by a single global fit, while  $\eta_{all}$  includes all  $n^{bins}$  bins of a given distribution. The former values are given only for the distributions used in the fit, while the latter were calculated for other interesting distributions as well.

<i>distribution</i>	$n^{bins}$	$n_{fit}^{bins}$	$\eta_{all}$	$\eta_{fit}$
<i>S</i>	23	20	44.1	16.5
<i>A</i>	16	16	23.0	23.0
<i>1-T</i>	21	21	14.8	14.8
<i>m</i>	18	18	30.1	30.1
$-\ln(y_3)$	14	14	9.8	9.8
$x_p < 0.02$	8	2	242.9	9.0
$x_p > 0.02$	38	29	88.1	31.4
$p_i^{out}(S)$	19	5	489.7	12.4
$p_i^{in}(S)$	25	13	292.6	18.2
$-\ln(x_p), K^0$	28	28	20.4	20.4
$x_E, \eta$	18	15	53.5	16.2
$x_E, \eta^i$	9	8	9.0	4.5
$x_p, \rho^0$	8	7	7.7	3.5
$\langle K^{**} \rangle$	1	1	0.0	0.0
$x_p, K^{*0}$	8	7	12.8	6.3
$x_p, \varphi^0$	8	3	48.2	4.0
$x_p, \omega^0$	6	4	38.7	2.3
$\langle f_2 \rangle$	1	1	0.2	0.2
$\langle f_0 \rangle$	1	0	17.3	
$-\ln(x_p), \Lambda$	22	20	42.0	17.6
$\langle \Xi^- \rangle$	1	0	3.9	
$\langle \Sigma(1385)^{\pm} \rangle$	1	1	0.1	0.1
$\langle \Omega^- \rangle$	1	1	0.2	0.2
$\langle \Xi(1530)^0 \rangle$	1	1	1.5	1.5
$-\ln(x_p), K^{\pm}(x < 0.018)$	11	11	8.5	8.5
$-\ln(x_p), K^{\pm}(x > 0.07)$	18	17	23.4	18.3
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	6	6	1.1	1.1
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	18	17	16.9	10.8
<i>sum</i>	349	286		280.1

<i>distribution</i>	$n^{bins}$	$\eta_{all}$
<i>P</i>	21	29.8
<i>C</i>	24	19.9
<i>M</i>	22	48.0
<i>O</i>	20	91.8
$\overline{M}_h^2$	21	22.6
$B_t$	17	18.1
$B_w$	17	6.2
$-\ln(x_p)$	52	605.8
$p_i^{out}(T)$	19	533.9
$p_i^{in}(T)$	25	182.9
$y_T$	21	49.1
$\langle K^0 \rangle$	1	0.5
$p_i(T), K^0$	25	77.3
$x_E, D^{*\pm}$	15	23.4
$\langle \Lambda \rangle$	1	0.1
$p_i(T), \Lambda$	25	64.8
$-\ln(x_p), \pi^{\pm}(x < 0.018)$	16	128.5
$-\ln(x_p), \pi^{\pm}(x > 0.07)$	23	5.8

$\Rightarrow \eta_{fit} / n_{fit}^{bins} \cong 0.98$

Table 7.10: Test quantities corresponding to best fit parameters found by *LinFit*

Model parameters and test quantities

In the next two tables, bins that were used in both fits are marked by an "×". All the bins removed from the candidate for a "fitable region" as defined by a single global fit are signed by an "S". The bins which are rejected by the iterative sequence of fits are signed by an "I".

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$S$	×	$S$	$S$	$S,I$	×	×	×	×	×	×	×	×	×	×
$A$	×	×	×	×	×	×	×	×	$I$	$I$	×	×	×	×
$1-T$	×	×	×	×	×	×	×	×	×	×	×	×	$I$	×
$m$	×	×	×	×	×	×	×	×	×	$I$	$I$	$I$	$I$	×
$-\ln(y_3)$	×	×	×	×	×	×	×	×	×	$I$	×	×	×	×
$x_p < 0.02$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$I$	$S$						
$x_p > 0.02$	×	×	$S$	$S$	×	×	×	$S$	×	×	×	×	×	×
	×	×	×	×	×	×	×	×	$S$	×				
$p_i^{out}(S)$	$S,I$	$I$	$S,I$	×	$S,I$	$S,I$	×	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$
$p_i^{in}(S)$	$S,I$	$S,I$	$S,I$	$S,I$	×	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S,I$	$S$	$S,I$	×
$-\ln(x_p), K^0$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$x_E, \eta$	$S$	×	×	×	×	×	×	×	×	×	×	×	×	×
$x_E, \eta'$	×	×	×	$S$	×	×	×	×	×					
$x_p, \rho^0$	×	×	×	×	×	×	×	$S$						
$\langle K^{*+} \rangle$	×													
$x_p, K^0$	×	×	×	×	$S$	×	×	×						
$x_p, \varphi^0$	×	$S$	$S$	$S$	$S$	×	$I$	$S,I$						
$x_p, \omega^0$	×	×	×	×	$S,I$	$S,I$								
$\langle f_2 \rangle$	×													
$\langle f_0 \rangle$	$S,I$													
$-\ln(x_p), \Lambda$	×	$S,I$	$S,I$	×	×	×	×	×	×	×	×	×	×	×
$\langle \Xi^- \rangle$	$S$													
$\langle \Sigma(1385)^+ \rangle$	×													
$\langle \Omega^- \rangle$	×													
$\langle \Xi(1530)^0 \rangle$	×													
$-\ln(x_p), K^+(x < 0.018)$	×	×	×	×	×	×	×	×	×	×	×			
$-\ln(x_p), K^+(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	×	×	×	×	×	×								
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Table 7.11: Accepted bins in the "fitable range" (first part)

15	16	17	18	19	20	21	22	23	24	25	26	27	28	
×	×	×	×	×	×	×	×	×						<i>S</i>
×	×													<i>A</i>
×	×	×	×	×	×	×								<i>I-T</i>
×	×	×	×											<i>m</i>
														$-\ln(y_3)$
														$x_p < 0.02$
×	×	×	<i>S</i>	<i>S,I</i>	×	<i>S</i>	<i>S</i>	<i>S</i>	×	×	×	×	×	$x_p > 0.02$
<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×										$p_i^{out}(S)$
×	×	×	×	×	×	×	×	×	×	×				$p_i^{in}(S)$
×	×	×	×	×	×	×	×	×	×	×	×	×	×	$-\ln(x_p), K^0$
×	×	<i>S,I</i>	<i>S,I</i>											$x_E, \eta$
														$x_E, \eta'$
														$x_p, \rho^0$
														$\langle K^{**} \rangle$
														$x_p, K^{*0}$
														$x_p, \varphi^0$
														$x_p, \omega^0$
														$\langle f_2 \rangle$
														$\langle f_0 \rangle$
×	×	×	×	×	×	×	×							$-\ln(x_p), \Lambda$
														$\langle \Xi^- \rangle$
														$\langle \Sigma(1385)^+ \rangle$
														$\langle \Omega^- \rangle$
														$\langle \Xi(1530)^0 \rangle$
														$-\ln(x_p), K^+(x < 0.018)$
<i>S</i>	×	×	×											$-\ln(x_p), K^+(x > 0.07)$
														$-\ln(x_p), (p, \bar{p})(x < 0.018)$
×	×	×	<i>S,I</i>											$-\ln(x_p), (p, \bar{p})(x > 0.07)$

Table 7.12: Accepted bins in the "fitable range" (second part)

### 7.2.4. HERWIG 5.8

The average value of the deviation of the fit result in units of the linear range ( $n:=0.5$ ) is  $0.53$  which we take as a sign that the result can be trusted. Again we calculated

$$f^{lin} := \left| \frac{p^{fit} - p^0}{\Delta_{lin}} \right|$$

where  $\Delta_{lin}$  is the width of the linear range.  $\Delta p_l$  is the distance from the expansion point to the lower (left) end of the parametrized range, while  $\Delta p_r$  is the same for the upper (right) end.  $\Delta p^{sys} := p_i^{fit} - p_s^{fit}$  is the deviation between the estimations found in the fits to the different candidates for a fitable range (s marks the candidate found by a single fit, and i is used for the candidate found by iteration).

name	$p^{default}$	$\Delta p_l$	$p^0$	$\Delta p_r$	$p^{fit}$	$\Delta p^{fit}$	$\Delta p^{sys}$	$f^{lin}$
LAMQCD	0.18GeV	0.02	0.15	0.02	0.151	0.001	0.007	0.14
RMASS(13)	0.75GeV	0.00	0.65	0.20	0.684	0.006	-0.023	1.27
CLMAX	3.35GeV	0.30	3.65	0.30	3.703	0.018	-0.066	0.55
CLSMR	0.0	0.20	0.70	0.20	0.704	0.039	-0.126	0.06
PWT(3)	1.0	0.20	0.80	0.20	0.830	0.014	0.069	0.61

Table 7.13: Results from *LinFit*

The test quantities corresponding to the parameter values above are listed in the next table.  $\eta_{fit}$  does only include the  $n_{fit}^{bins}$  bins of the fitable region defined by a single global fit, while  $\eta_{all}$  includes all  $n^{bins}$  bins of a given distribution. The former values are given only for the distributions used in the fit, while the latter were calculated for other interesting distributions as well.

<i>distribution</i>	$n^{bins}$	$n_{fit}^{bins}$	$\eta_{all}$	$\eta_{fit}$	<i>distribution</i>	$n^{bins}$	$\eta_{all}$
$S$	23	13	152.9	27.8	$P$	21	133.5
$A$	16	11	114.1	17.8	$C$	24	67.5
$1-T$	21	10	209.6	18.1	$M$	22	212.9
$m$	18	10	251.0	15.6	$O$	20	83.7
$-\ln(y_3)$	14	13	26.4	21.5	$\overline{M}_h^2$	21	45.4
$x_p < 0.02$	8	5	91.2	15.4	$B_t$	17	157.4
$x_p > 0.02$	38	13	540.1	34.4	$B_w$	17	65.0
$p_i^{out}(S)$	19	10	251.6	15.6	$-\ln(x_p)$	52	710.8
$p_i^{in}(S)$	25	13	169.4	23.9	$p_i^{out}(T)$	19	417.6
$-\ln(x_p), K^0$	28	20	106.4	20.1	$p_i^{in}(T)$	25	148.2
$x_E, \eta$	18	16	41.4	23.4	$y_T$	21	31.1
$x_p, \rho^0$	8	7	8.4	3.1	$\langle K^0 \rangle$	1	0.0
$\langle K^{*+} \rangle$	1	1	0.0	0.0	$x_E, \eta^i$	9	29.8
$x_p, K^{*0}$	8	6	31.3	10.1	$p_i(T), K^0$	25	726.6
$x_p, \varphi^0$	8	4	51.4	7.6	$x_E, D^{*\pm}$	15	41.9
$x_p, \omega^0$	6	4	24.1	9.5	$\langle \Lambda \rangle$	1	62.4
$\langle f_2 \rangle$	1	1	0.0	0.0	$-\ln(x_p), \Lambda$	22	1790.2
$-\ln(x_p), K^\pm(x < 0.018)$	11	6	58.2	16.0	$\langle \Xi^- \rangle$	1	382.4
$-\ln(x_p), K^\pm(x > 0.07)$	18	10	89.0	22.6	$p_i(T), \Lambda$	25	461.5
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	6	6	3.9	3.9	$\langle \Sigma(1385)^+ \rangle$	1	128.3
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	18	13	141.8	3.7	$\langle \Omega^- \rangle$	1	1519.4
					$\langle \Xi(1530)^0 \rangle$	1	1819.4
					$-\ln(x_p), \pi^\pm(x < 0.018)$	16	40.1
					$-\ln(x_p), \pi^\pm(x > 0.07)$	23	18.0
<i>sum</i>	313	192		310.2			

$$\Rightarrow \eta_{fit} / n_{fit}^{bins} \cong 1.62$$

Table 7.14: Test quantities corresponding to best fit parameters found by *LinFit*

Model parameters and test quantities

In the next two tables, bins that were used in both fits are marked by an "×". All the bins removed from the candidate for a "fitable region" as defined by a single global fit are signed by an "S". The bins which are rejected by the iterative sequence of fits are signed by an "I".

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>S</i>	<i>S,I</i>	×	<i>S,I</i>	<i>S,I</i>	×	<i>S</i>	<i>S</i>	<i>S</i>	×	<i>S</i>	<i>S,I</i>	×	×	<i>S,I</i>
<i>A</i>	<i>I</i>	<i>S,I</i>	<i>S,I</i>	<i>I</i>	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×	×	×	×	×	×
<i>1-T</i>	<i>S,I</i>	<i>S,I</i>	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×	×	<i>S</i>	×	×	×	<i>S,I</i>	<i>S</i>
<i>m</i>	<i>S,I</i>	<i>S,I</i>	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×	<i>S</i>	<i>S</i>	×	×	×	×
$-\ln(y_3)$	<i>I</i>	×	×	<i>S</i>	×	×	×	×	×	×	×	×	×	×
$x_p < 0.02$	×	<i>S,I</i>	<i>S,I</i>	×	×	×	<i>I</i>	<i>S,I</i>						
$x_p > 0.02$	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>I</i>	×	×	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>
	<i>S,I</i>	×	<i>S,I</i>	×	×	<i>S</i>	<i>S,I</i>	<i>S</i>	<i>S,I</i>	<i>S,I</i>				
$p_i^{out}(S)$	<i>S,I</i>	×	×	×	×	×	×	×	×	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>
$p_i^{in}(S)$	<i>S,I</i>	×	×	×	<i>S</i>	<i>S</i>	×	×	<i>I</i>	<i>I</i>	<i>S,I</i>	×	×	<i>S</i>
$-\ln(x_p), K^0$	×	×	×	×	×	×	<i>S</i>	<i>S</i>	<i>S,I</i>	<i>S</i>	×	×	×	×
$x_E, \eta$	<i>S,I</i>	×	<i>I</i>	×	×	×	×	×	×	×	<i>S,I</i>	×	×	×
$x_p, \rho^0$	×	×	×	×	×	×	×	<i>S</i>						
$\langle K^{++} \rangle$	×													
$x_p, K^{*0}$	×	×	×	×	<i>S,I</i>	<i>S,I</i>	×	×						
$x_p, \varphi^0$	×	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×	×	<i>S,I</i>						
$x_p, \omega^0$	×	×	<i>S,I</i>	×	×	<i>S,I</i>								
$-\ln(x_p), K^+(x < 0.018)$	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>I</i>	<i>S,I</i>	<i>I</i>	<i>S,I</i>	<i>I</i>	×	×	×			
$-\ln(x_p), K^+(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>
$-\ln(x_p), (p, \bar{p})(x < 0.018)$	×	×	×	×	×	×								
$-\ln(x_p), (p, \bar{p})(x > 0.07)$	×	×	×	×	×	×	×	×	×	×	×	×	×	<i>S</i>

Table 7.15: Accepted bins in the "fitable range" (first part)

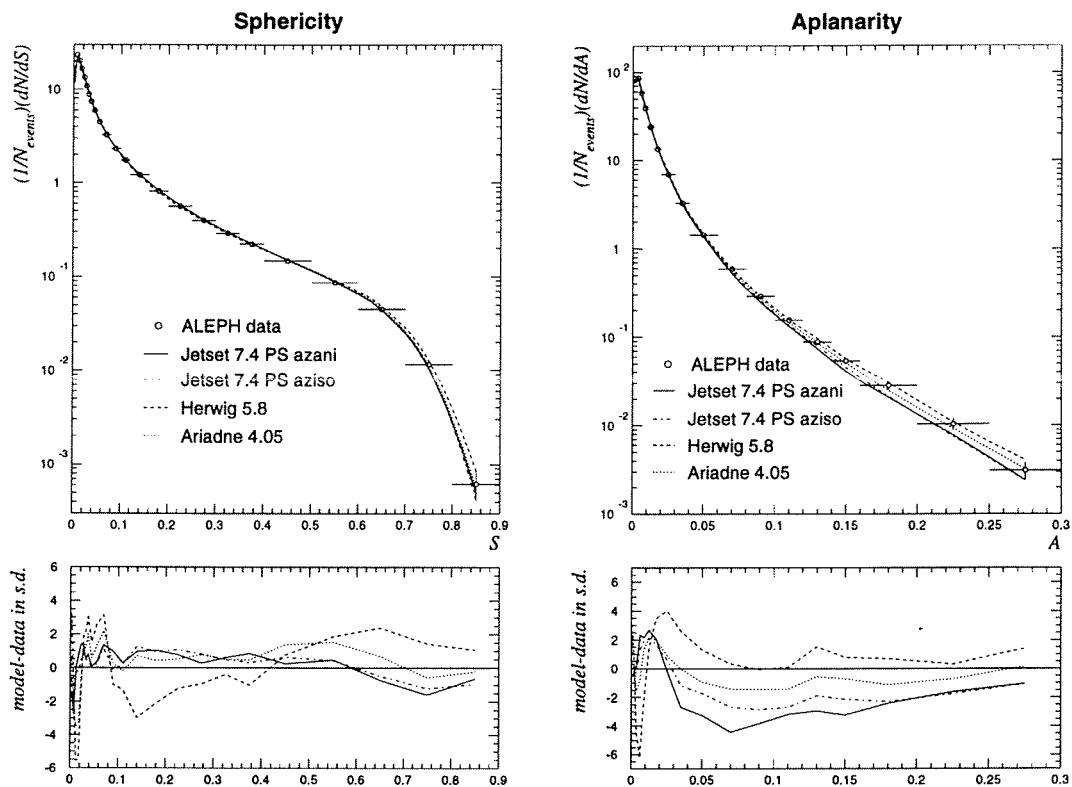
15	16	17	18	19	20	21	22	23	24	25	26	27	28	
×	×	×	×	×	×	<i>I</i>	<i>S,I</i>	×						<i>S</i>
×	×													<i>A</i>
×	×	×	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>								<i>I-T</i>
×	×	×	×											<i>m</i>
														$-\ln(y_3)$
														$x_p < 0.02$
<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×	×	×	×	<i>S,I</i>	×	<i>S,I</i>	<i>S,I</i>	×	$x_p > 0.02$
<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>I</i>										$p_i^{out}(S)$
×	<i>S</i>	×	×	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×				$p_i^{in}(S)$
×	×	×	×	×	×	×	<i>I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>	×	×	$-\ln(x_p), K^0$
×	×	×	×											$x_E, \eta$
														$x_p, \rho^0$
														$\langle K^{**} \rangle$
														$x_p, K^{*0}$
														$x_p, \phi^0$
														$x_p, \omega^0$
														$-\ln(x_p), K^+(x < 0.018)$
<i>S,I</i>	<i>S,I</i>	<i>S</i>	<i>S</i>											$-\ln(x_p), K^+(x > 0.07)$
														$-\ln(x_p), (p, \bar{p})(x < 0.018)$
<i>S</i>	<i>S,I</i>	<i>S,I</i>	<i>S,I</i>											$-\ln(x_p), (p, \bar{p})(x > 0.07)$

Table 7.16: Accepted bins in the "fitable range" (second part)

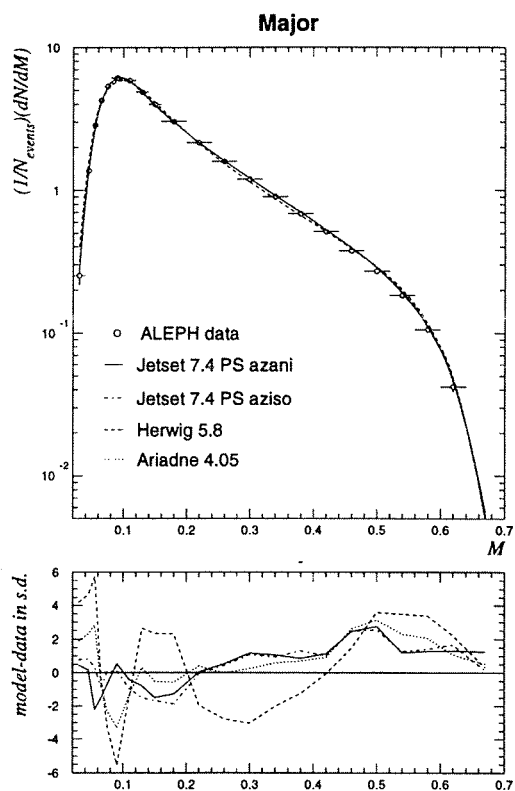
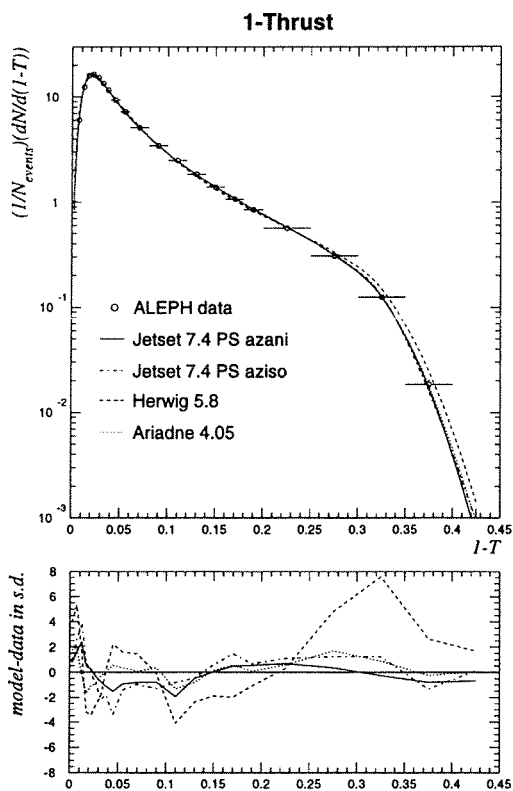
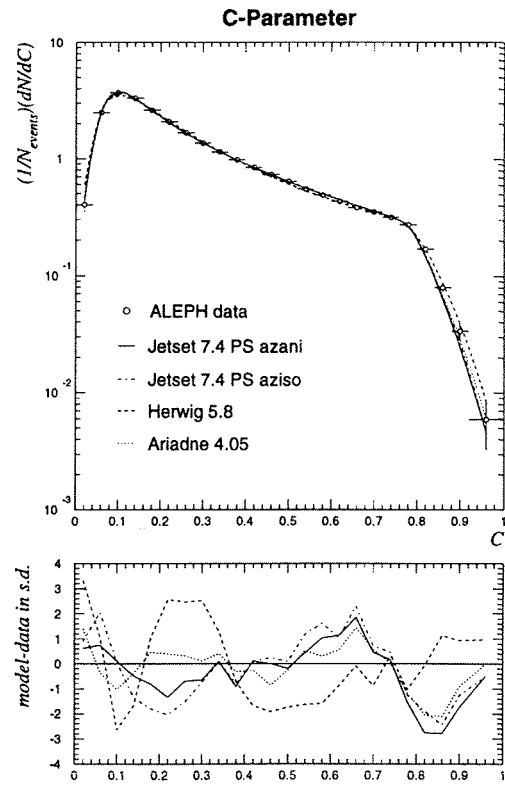
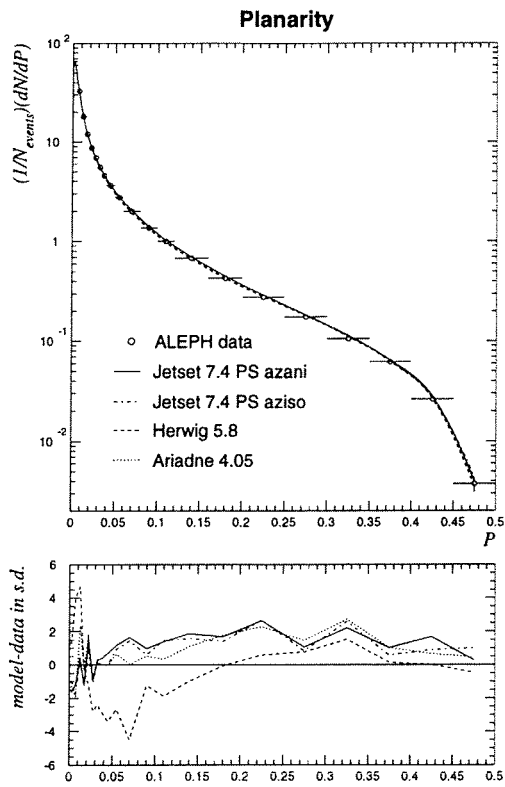
## 7.3. Graphical comparison between tuned models and experimental data

The aim of this section is to compare the predictions of the tuned models with the measured distributions listed in chapter 4. All the figures below are divided into two parts. The first part shows the shape of the measured distribution together with the prediction of the tuned models (calculated with a statistics of  $s_{sim} = 2 \cdot 10^6$ ), while the second part shows the deviation between the models and the data in units of the error of the measurement. The symbol  $E_{vis}$  is used for the "visible energy" in an event, that means the sum of the energies of all charged particles of the final state.

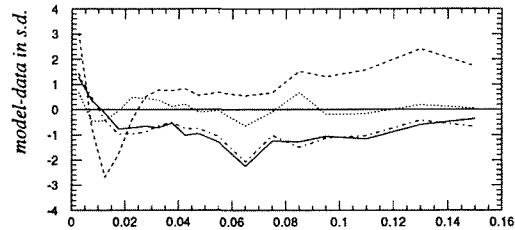
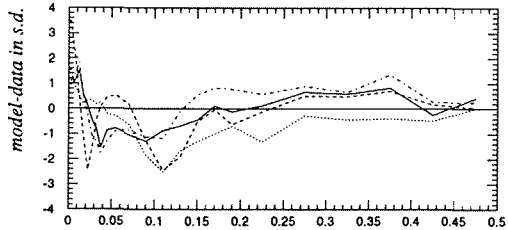
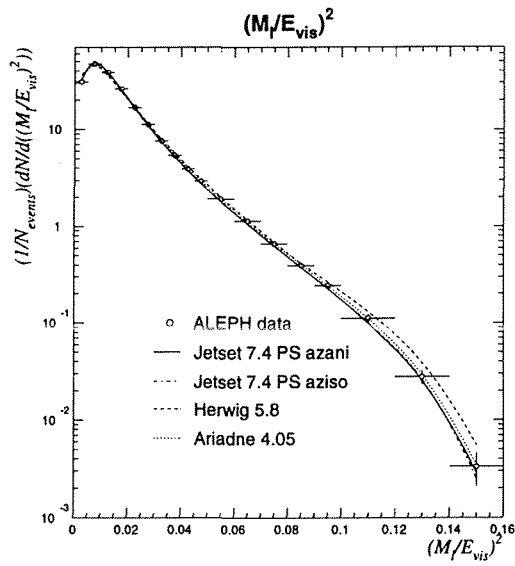
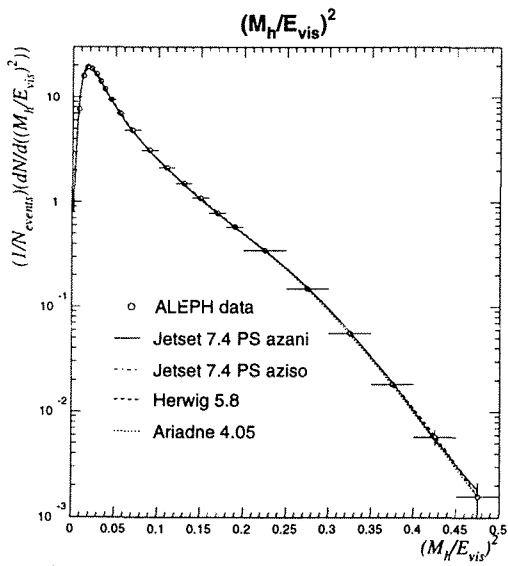
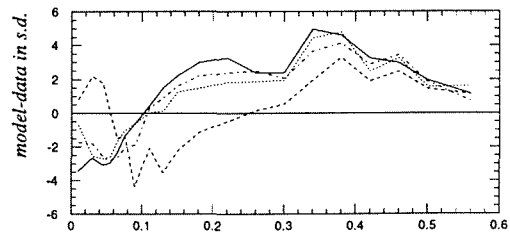
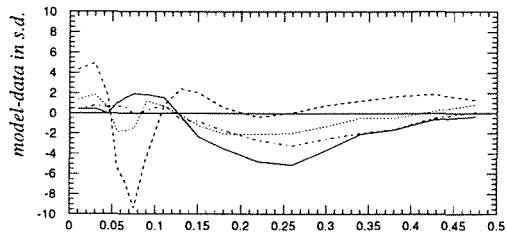
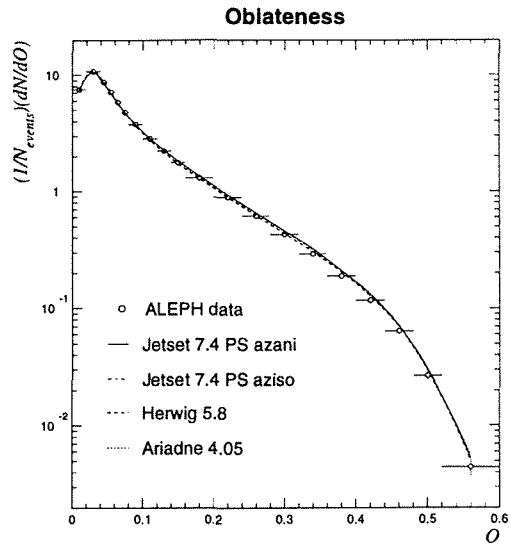
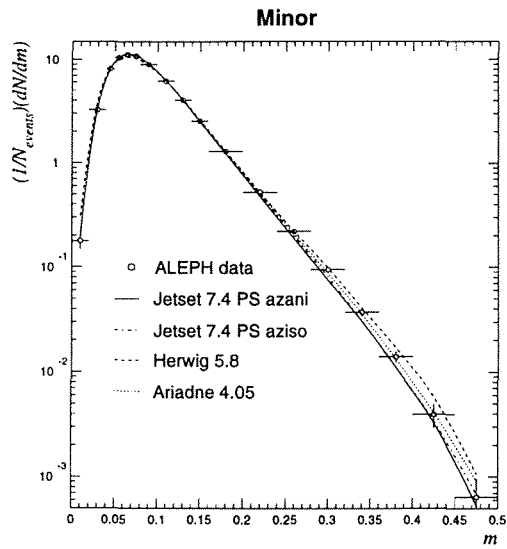
### 7.3.1. Event properties

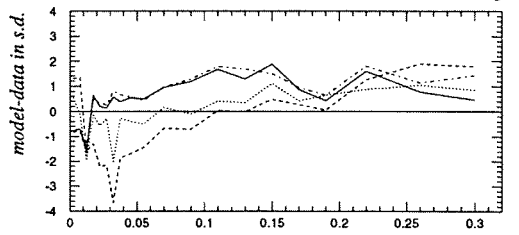
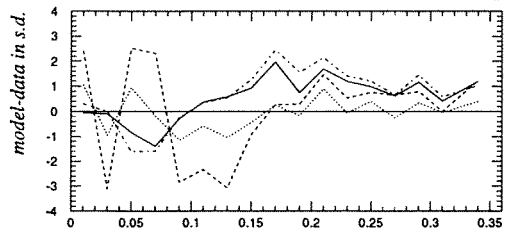
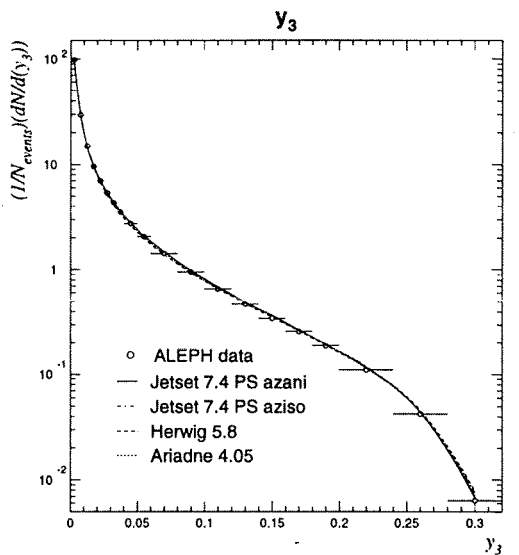
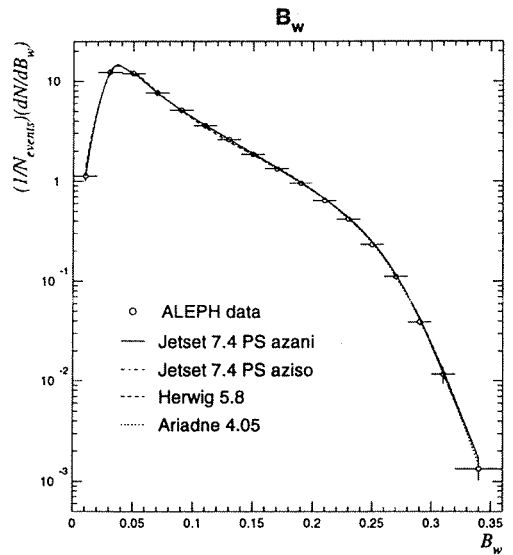
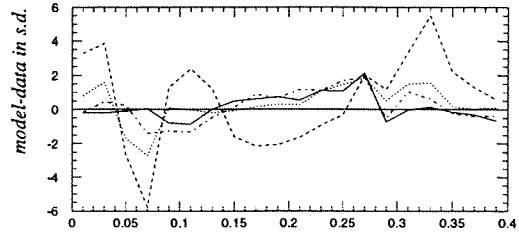
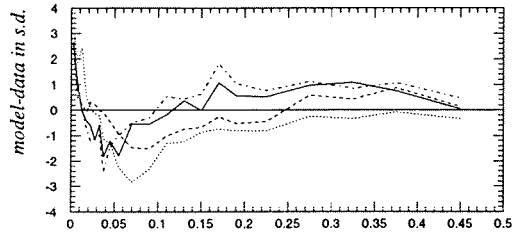
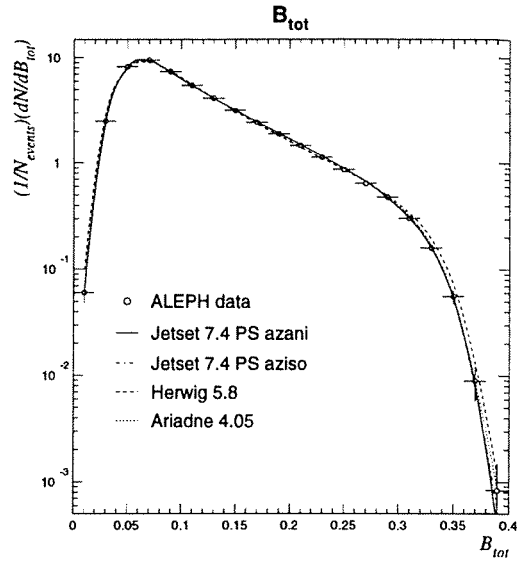
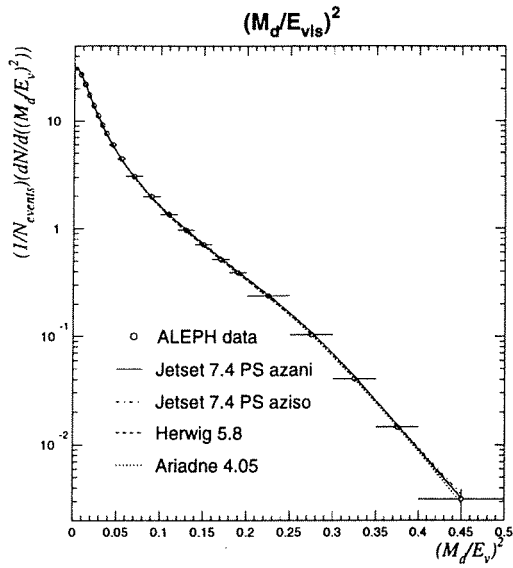




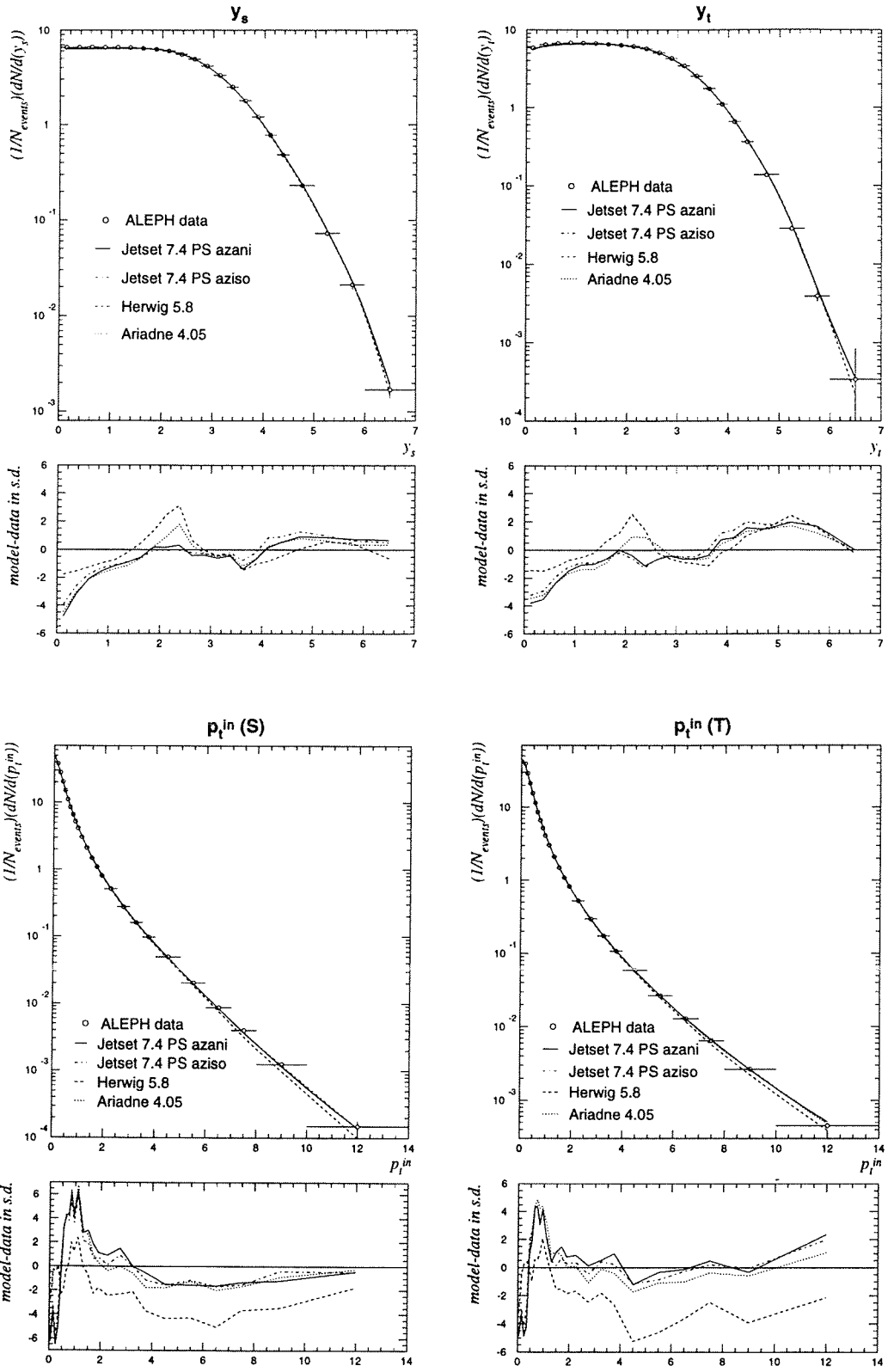


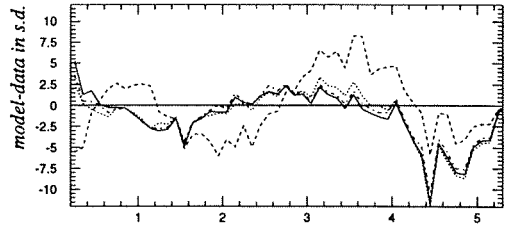
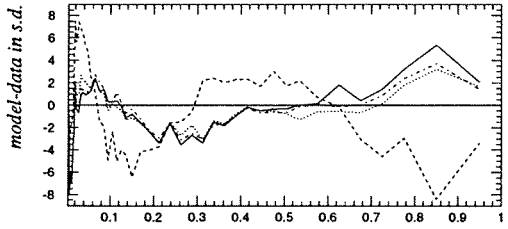
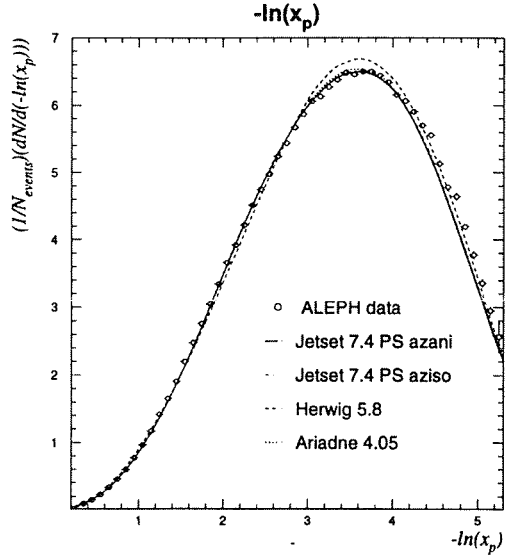
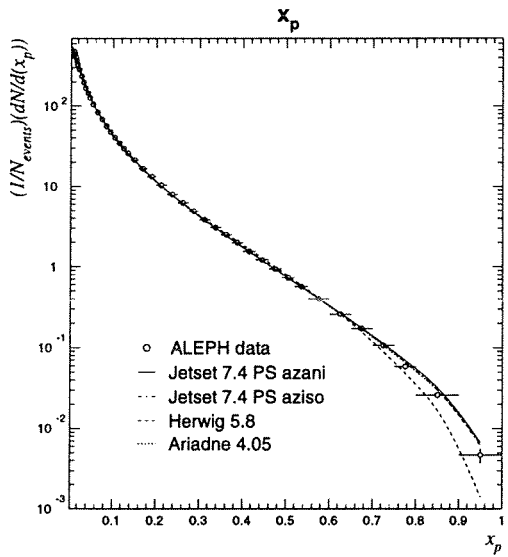
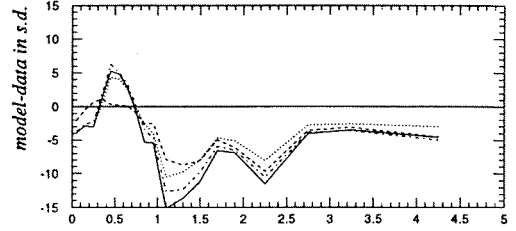
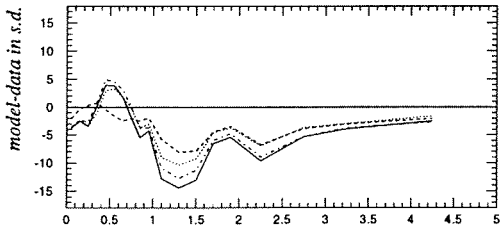
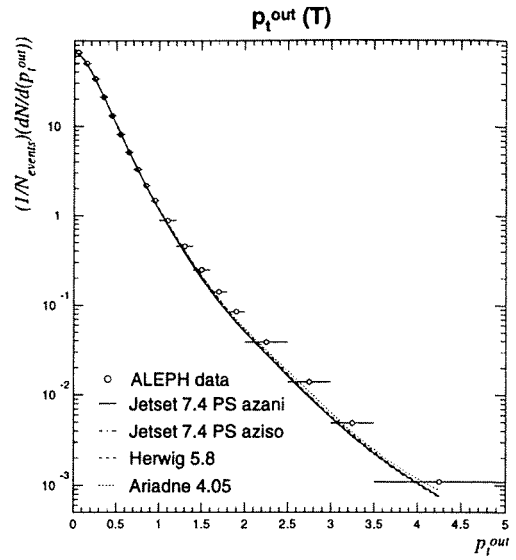
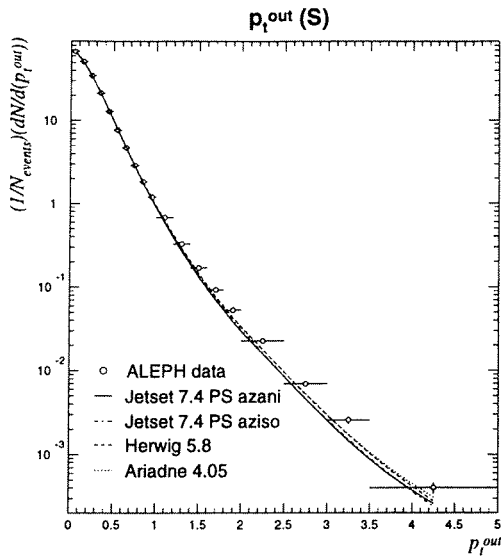
Graphical comparison between tuned models and experimental data





### 7.3.2. Single particle properties





## 7.4. Concluding remarks

A test quantity based on the whole set of distributions listed in section 7.2 leads to the following numbers:

<i>Model</i>	<i>n</i>	$\eta_{all}$	$\eta_{all}/n$
<i>JS 7.4 azani</i>	714	4469.2	6.26
<i>JS 7.4 aziso</i>	714	3737.3	5.23
<i>Ariadne 4.05</i>	714	3449.0	4.83
<i>Herwig 5.8</i>	713	11455.3	16.07

In this sense, *ARIADNE* is the best model, followed by *JETSET* with isotropic gluon decay, *JETSET* with anisotropic gluon decay and *HERWIG*. The large value of the *HERWIG* test quantity is mainly due to the strange baryon sector (c.f. table 7.12). Also for event properties, *HERWIG* has problems to describe the data, while the other models are in better agreement with the data (c.f. table 7.2, 7.6, 7.10 and 7.12).

It is worth noting, that the *JETSET* version with isotropic decaying gluons is in better agreement with the measured event properties, than the version with anisotropic gluon decay (c.f. tables 7.2 and 7.6). All the models have problems with the description of  $p_i^{out}$  and  $p_i^{in}$ . Especially the distributions of  $p_i^{out}$  show dramatic deviations from the data in the region above  $1\text{GeV}$  up to 15 times the error of the measurement. The distribution of the normalized particle momentum is underestimated in the region of very low momenta ( $x_p \leq 0.014$ ). The description of  $\langle f_0 \rangle$  seems to be impossible. Apart from this problem, multiplicities of identified hadrons are in good agreement with the model predictions of *JETSET* and *ARIADNE*.

An interesting quantity, that is not directly used in the fit, is the mean charged multiplicity. This quantity is for example measured in [A2,95] as  $\langle n_{ch} \rangle = 20.91 \pm 0.22$ . By integrating any single particle distribution, the following values are obtained:

<i>Model</i>	$\langle n_{ch} \rangle$	$\eta$
<i>JS 7.4 azani</i>	20.53	2.98
<i>JS 7.4 aziso</i>	20.59	2.12
<i>Herwig 5.8</i>	20.75	0.53
<i>Ariadne 4.05</i>	20.54	2.83

In the table above, also the test quantity

$$\eta := \left( \frac{\langle n_{ch} \rangle - \langle n_{ch} \rangle_{\text{model}}}{\sigma} \right)^2$$

is given. One can see, that all the models describe the charged multiplicity within two sigmas although systematically below the data. The description of *HERWIG* is the best.

The next tables contain the correlation coefficients obtained from the matrix of second derivatives of the estimation functions used for parameter tuning. Because of the similarity of the models *JETSET* and *ARIADNE* only the correlation coefficients for *JETSET* with anisotropic decaying gluons and for *HERWIG* are given.

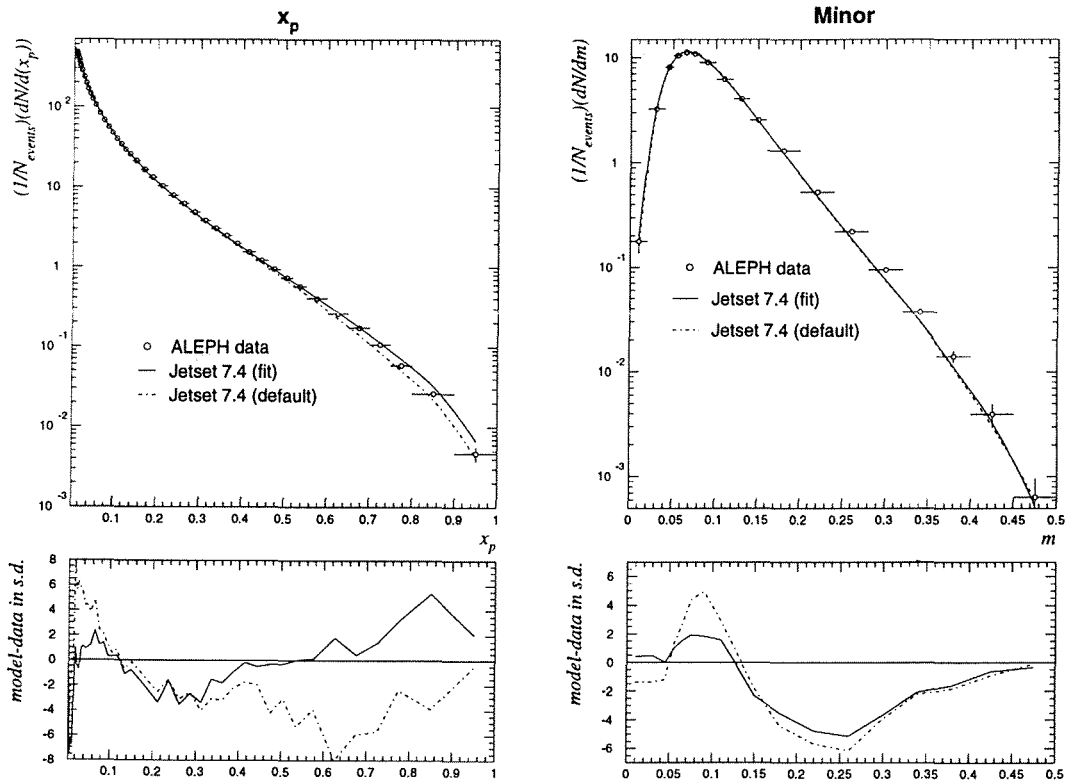
	$\Lambda$	$Q_0$	$\sigma$	$b$	$P_{u,d}^{S=1}$	$P_s^{S=1}$	$\gamma_s$	$qq/q$	$su/du$	$\bar{b}_1$
$\Lambda$		0.20	-0.60	0.48	0.02	-0.01	0.08	0.25	-0.15	-0.15
$Q_0$			0.12	-0.42	0.19	0.36	0.25	0.26	-0.23	-0.13
$\sigma$				-0.30	0.35	0.26	0.19	0.08	-0.03	0.00
$b$					0.32	-0.05	-0.01	-0.10	0.18	0.01
$P_{u,d}^{S=1}$						-0.11	0.28	0.43	-0.22	-0.14
$P_s^{S=1}$							0.22	0.02	-0.05	0.01
$\gamma_s$								0.30	-0.61	-0.03
$qq/q$									-0.60	-0.51
$su/du$										-0.01
$\bar{b}_1$										

Concluding remarks

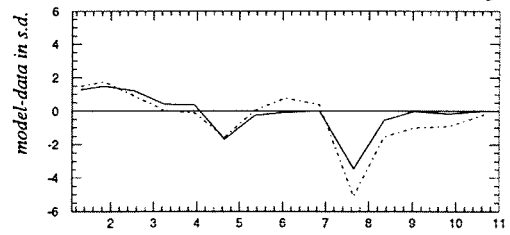
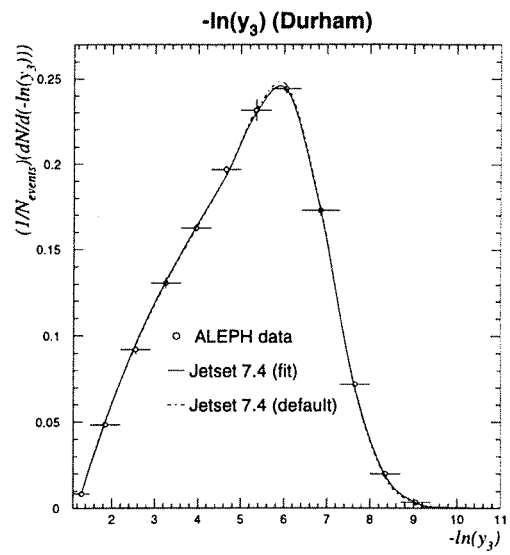
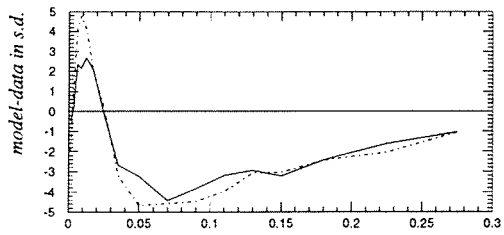
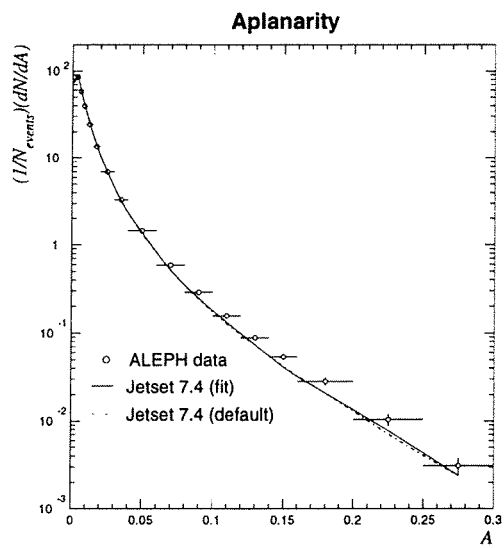
	LAMQCD	RMASS(13)	CLMAX	CLSMR	PWT(3)
LAMQCD		-0.25	-0.48	-0.42	-0.25
RMASS(13)			0.25	-0.06	0.39
CLMAX				0.03	0.17
CLSMR					-0.10
PWT(3)					

These tables show that correlations are in general not negligible. Because of that, one is forced to use multi parameter fitting procedures if estimations of the parameter values are to be calculated. Only if correlations can be neglected, an independent estimation of parameters is possible.

To conclude, a comparison between fitted and default distributions should be given to show the improvement of the description due to the fitting procedure. For that reason the model predictions for four representative distributions are calculated using the estimated set of best-fit parameters and the default values for the model parameters respectively. We used the model *JETSET 7.4* in the version of anisotropic gluon decay.









# Summary

In the first part of this work, we studied some characteristics of charged particles of hadronic final states in *ALEPH* events, while some methods for tuning models with free parameters to experimental data were discussed in detail in the second part. The main ingredient in the first part was a good simulation for the QCD process as well as for the detector effects. After we performed some tests to make sure that the available simulation is indeed good enough, standard methods for unfolding of detector effects and to correct for ISR and cut-influences were used to perform the measurement.

The main results of this measurement are shown in the tables of chapter 4. They lead to the following conclusions:

- In most of the bins the systematic error is bigger than the statistical one.
- The systematic error is typical in the order of a few percent, even if some regions exist, where this error is bigger.
- For event properties the statistical error is also typical in the size of a few percent, while it is decreased to a few permille in the case of single particle distributions.
- The dominant part of the systematic error is in most cases the model bias. The main contribution to the model bias comes from the difference between the corrected distributions calculated with the help of *JETSET* and *HERWIG*.

Detailed discussions of methods for model tuning were presented in the second part of this work. Starting point for all of them was the maximum likelihood principle. These methods together with linear parametrizations of the model predictions were the basis of an algorithm that we used for the estimation of the model parameters. We called this algorithm "LinFit". In order to minimize the contribution of systematic deviations between model predictions and experimental data to results of the fitting procedure, a restriction to a so-called "fitable region" was introduced.

In addition a new method to include fluctuations of parametrization coefficient in the estimation process was developed.

The main results of the parameter tuning are:

- *ARIADNE* is the best model, followed by *JETSET* with isotropic gluon decay, *JETSET* with anisotropic gluon decay and *HERWIG*, because it leads to the best overall description.
- The *JETSET* version with isotropic decaying gluons is in better agreement with the measured event properties, than the version with anisotropic gluon decay.
- All the models have problems with the description of  $p_i^{out}$  and  $p_i^{in}$ . Especially the distributions of  $p_i^{out}$  show dramatic deviations from the data in the region above  $1\text{GeV}$  up to 15 times the error of the measurement.
- The distribution of the normalized particle momentum is underestimated in the region of very low momenta ( $x_p \leq 0.014$ ).
- The description of  $\langle f_0 \rangle$  seems to be impossible using the spin counting simplification introduced in section 2.3.2.4.
- Apart from this  $\langle f_0 \rangle$ -problem, multiplicities of identified hadrons are in good agreement with the model predictions of *JETSET* and *ARIADNE*.
- All the models describe the charged multiplicity within two sigmas although systematically below the data. The description of *HERWIG* is the best.

## Appendices



## Appendix A.

### Remarks concerning part I

#### 1. Remarks about the choice of bins

In the predecessor of this work [A1,92] unfolding was done even for event properties by the factor method explained in 3.3.2.2. This method should be a good approximation if the off-diagonal elements of the detector matrix are small compared to the diagonal elements. In some works (for example [O1,90]), the separation into bins is chosen such that more than 60% of the bin contents remain "in a given bin", while the rest is smeared over the other bins by detector influences to force this special form of the detector matrix, .

Here an analysis is performed, that provides such a bin separation for most of the distributions observed in the main part of this work. For that reason, a starting bin-width is set to be  $\frac{1}{40}$  (which is a arbitrary choice) of the full range of the distribution. If less than 40% of the contents of this bin is smeared into others, the width is kept, and the next bin is calculated in a similar way. If more than 40% are smeared into other bins, the width is increased by  $\frac{1}{40}$  of the full range, and this new width is checked by the 40% criterion.

The results are shown in the following table. The numbers beneath the symbols of the distributions mark the number of resulting bins. If this number is big, one can expect that the detector is able to measure the corresponding quantity with a high accuracy, and a neglect of off-diagonal elements of the detector matrix makes sense. If only a few bins are observed, the smearing of the detector is big, and a matrix unfolding has to be done.

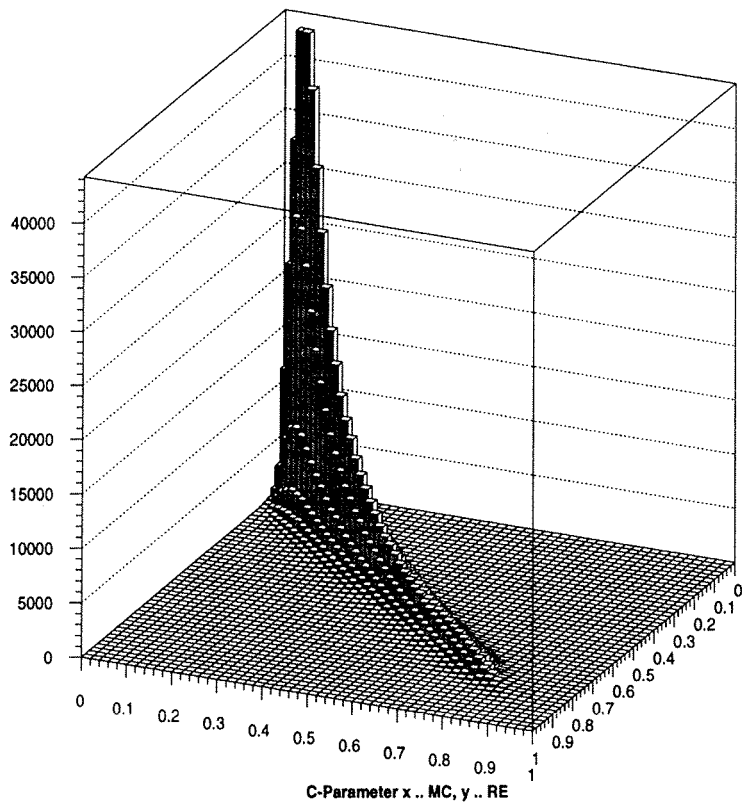
The first and the last number in the columns of bin borders mark the whole observed range. The analysis is based on the simulation *HVFL04*, it uses a statistics of  $s_{sim} := 30000$ .

<i>S</i>	<i>A</i>	<i>P</i>	<i>C</i>	<i>1-T</i>	<i>M</i>	<i>m</i>	<i>O</i>	$\overline{M}_h^2$	$\overline{M}_l^2$	$\overline{M}_d^2$	$y_3$
17	13	15	16	12	19	14	13	7	4	6	6
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
.018	.005	.013	.025	.010	.018	.010	.015	.010	.003	.010	.009
.035	.010	.025	.050	.020	.035	.020	.030	.021	.008	.035	.021
.053	.015	.038	.075	.032	.053	.030	.047	.046	.024	.084	.046
.070	.020	.050	.105	.048	.070	.040	.068	.090	.120	.161	.097
.088	.028	.066	.145	.070	.088	.051	.093	.163		.289	.219
.108	.036	.083	.192	.099	.105	.065	.122	.294		.400	.350
.135	.047	.107	.250	.130	.124	.081	.159	.400			
.166	.063	.135	.320	.172	.175	.102	.211				
.207	.092	.164	.392	.220	.147	.128	.270				
.247	.118	.204	.475	.262	.205	.162	.329				
.289	.168	.250	.562	.325	.242	.206	.408				
.359	.194	.304	.665	.400	.285	.261	.483				
.429	.200	.362	.752		.334	.327	.600				
.507		.427	.822		.392	.400					
.588		.500	.875		.448						
.695			1.00		.511						
.700					.577						
					.635						
					.700						

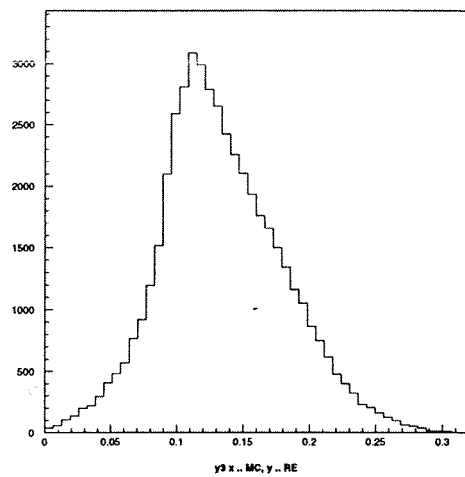
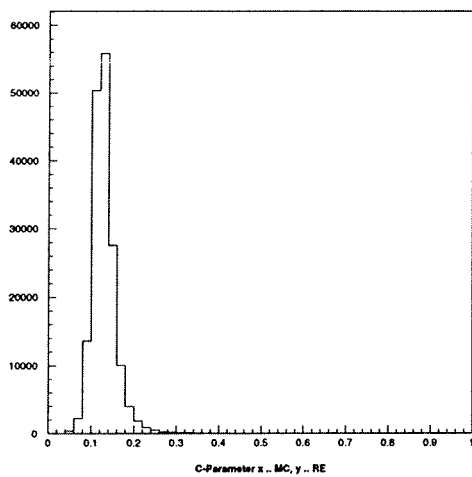
Table A1: Bin borders according to the 40% criterion

The jet masses and the resolution parameter are smeared most by the detector according to this analysis. One should therefore check at least for these distributions if the unfolding by correction factors leads to satisfactory values. We will do this by comparing the factor-unfolded by the matrix-unfolded values. For every distribution the histogram *H* (c.f. example 3.5) forms the basis of this analysis. In the next figure, this histogram is shown for the "C-Parameter".





A slice was cut out of the histogram H for two special cases. These cases that are shown in the next figures illustrate the *40%* smearing quite well.



The figure on the left side shows the smearing of the 5-th C-Parameters bin ([0.105,0.145]), while the picture right shows the smearing in the 5-th  $y_3$  bin ([0.097,0.219]).

In the next table, a comparison between an unfolding with factors and an unfolding with matrices is shown. Since [A1,92] uses correction factors, even if the bin widths are smaller than in the previous table, one could expect errors due to neglecting the off-diagonal elements.

$S$	$A$	$P$	$C$	$1-T$	$M$	$m$	$O$	$\overline{M}_h^2$	$\overline{M}_l^2$	$\overline{M}_d^2$	$B_l$	$B_w$	$y_3$
.19	.09	.08	.06	.01	.24	.08	.27	.01	.04	1.42	.04	.17	.00
.69	.34	.56	.30	.01	.07	.10	.07	.02	.22	.30	.15	.32	.04
.24	.09	.60	.03	.54	1.45	1.18	.73	.13	.23	1.30	.58	.22	.50
.55	.11	.62	.29	.62	.34	.75	.10	.27	.10	.87	.09	.03	.51
.36	.16	1.00	.16	.01	.42	.06	.79	.97	.13	1.08	.03	.18	.56
.11	.21	.96	.07	.35	.11	.41	1.05	.10	.02	.12	.13	.33	.10
.13	.03	.29	.21	.57	.01	.05	.23	.00	.16	.15	.59	.21	.19
.25	.50	.32	.14	.24	.16	.01	.29	.74	.21	2.30	.20	.12	.01
.32	.08	.11	.35	.03	.29	.37	.05	.27	.56	.35	.01	.46	.02
.26	.40	.08	.22	.30	.32	.20	.09	.19	.60	.11	.10	.17	.13
.81	.70	.27	.35	.29	.41	.63	.35	.11	.58	.12	.06	.55	.27
.72	.13	.09	.29	.19	.08	.50	.02	.35	.11	.07	.13	.24	.01
.05	.56	.29	.03	.81	.08	.57	.14	.13	.38	.78	1.01	.16	.36
.20	.11	.22	.48	.10	.32	.90	.02	.70	.53	.05	.06	.28	.70
.12	.38	.35	.45	.15	.17	.07	.18	.12	1.02	.04	.62	.26	.56
.02	.41	.08	.12	.64	.11	.38	.56	.40	.14	.68	.10	.26	.74
.06	.07	.17	.83	.11	.88	.09	.27	.22	1.28	.69	.14	.33	.22
.17		.30	.54	.40	.09	.01	.12	.95	.94	.17	.12		.59
.51		.33	.60	.27	.32		.04	.22		.54	.46		.70
.22		.48	1.71	.23	.61		.01	.26		.18	.08		1.63
.12		.31	.83	1.11	.11			.90		.45			
.56			.52	.24	.25			.16		.17			
.67			.26					.22					
.69			.02										

Table A2: Comparison between factor and matrix method

The bin widths used in this work (and for the calculation of the values in the table above) are very similar to the bin widths used in [A1,92]. The same conditions as in the measurement are used. All the calculations are done with 1992 data. The table above shows the deviations between the matrix-corrected and the factor-corrected values in units of the error of the measurement. We can see from the table above, that the deviations lie mainly below one sigma. This is indeed a sign, that the usage of the simple factor method makes sense. This is especially true because we are using a higher statistics, and therefore the errors are smaller than in [A1,92].

## 2. Comparison between event properties for the 1992 and 1993 runs

Generally simulation models become better every year, and also the reconstruction of particle tracks changes with time. Because of that, a different simulation is available for each year. We saw, that a change in the simulation will in general lead to a change of the measured distributions. In the following tables the differences between the corrected 92- and 93-data are given in units of the statistical error.

$$\Delta_i := \frac{(t_i^{92} - t_i^{93})}{\sqrt{(\Delta t_i^{92})^2 + (\Delta t_i^{93})^2}}$$

The measurement is based on  $s_{dat} := 571825$  and  $s_{sim} := 1186173$  for 1992 as well as  $s_{dat} := 375696$  and  $s_{sim} := 1185262$  for 1993. The table shows some deviations which cannot be explained as purely statistical fluctuation.

In table A4 the same deviation is given in units of the total measurement error, i.e. the quadratic sum of the statistical error, the cut-systematic and the model bias. No further deviations can be observed in units of this error, and the measurements are compatible in that sense<sup>1</sup>.

---

<sup>1</sup> Note, that the model bias is the same for both years, and therefore fully correlated.

<i>S</i>	<i>A</i>	<i>P</i>	<i>C</i>	<i>1-T</i>	<i>M</i>	<i>m</i>	<i>O</i>	$\bar{M}_h^2$	$\bar{M}_l^2$	$\bar{M}_d^2$	<i>B<sub>t</sub></i>	<i>B<sub>w</sub></i>	<i>y<sub>3</sub></i>
-.18	-3.31	-1.59	-3.36	-1.07	-.99	-.97	-1.04	-2.34	-9.91	-2.84	-.41	-1.82	-2.41
-1.80	1.76	.27	-4.33	-3.15	-.56	-2.65	-.31	-7.06	-3.86	-2.14	-.65	-1.60	2.18
-1.02	1.23	.01	-2.15	-2.84	.38	-.81	1.23	-5.76	1.92	.05	-2.06	-.02	1.56
.65	-.13	.30	.33	-.94	-.86	-.50	.33	-2.91	4.06	1.03	-.87	.92	-.67
-1.08	.75	-.94	1.09	-.59	-1.50	-.08	.95	.21	3.01	.20	1.32	.17	.12
-.52	-.12	.52	1.59	-.14	-1.05	-.40	-1.27	1.29	3.09	1.23	.71	.71	.00
.72	.52	-.18	.68	.39	-.02	1.29	-.78	2.01	2.11	.44	-.17	.94	.25
1.21	.17	-.15	.78	.13	-.42	-.08	-1.29	1.89	1.72	.99	.87	.48	.62
.89	-1.03	1.32	-.68	.91	1.59	1.49	.01	1.87	1.49	1.40	.15	-.25	.34
.25	-.31	.27	1.18	1.25	.61	.57	-.16	1.95	2.26	.66	.24	.73	.27
-.93	.71	.48	1.33	.99	-.89	1.06	.27	1.93	2.14	.39	.74	.02	-.49
1.26	.17	.46	1.13	.12	.72	.06	1.01	1.00	1.62	.06	.34	.69	.43
1.83	-.66	.13	.88	.99	-.21	.51	.67	1.48	.84	.29	.70	-.62	-.35
-.08	.15	.52	.22	1.27	1.10	.67	.86	.42	.34	.59	-.18	-.96	.11
-.45	.42	-.46	.72	.70	.31	-.42	.55	.58	1.38	.92	.81	-.54	.22
-.18	.41	.46	-.11	.28	.96	1.14	-.21	.52	.58	-.44	-.63	.34	.15
-.43	-.29	-.05	-.54	-.33	.30	-.12	-.31	.92	-.08	.19	-.41	.17	.81
.37		-.84	1.26	.33	.70	-.38	-.70	.94	.04	.50	1.20		.03
1.26		1.24	.53	.27	-.47		-.45	-.17		-.14	.24		-.31
-.29		1.11	.15	-.07	-.59		-.08	-.37		-.59	.62		-.25
-.65		-.22	1.06	1.25	.36			.07		.06			
1.26			.68	.63	.19			-.05		.37			
1.20			.47					1.23					
.54			-.09										

Table A3: Differences between corrected distributions from 92 and 93 in units of the statistical error

<i>S</i>	<i>A</i>	<i>P</i>	<i>C</i>	<i>1-T</i>	<i>M</i>	<i>m</i>	<i>O</i>	$\bar{M}_h^2$	$\bar{M}_l^2$	$\bar{M}_d^2$	$B_l$	$B_w$	$y_3$
-.01	-.10	-.67	-.35	-.18	-.06	-.05	-.18	-.35	-.92	-.68	-.02	-.10	-.52
-1.19	.37	.05	-1.52	-.76	-.05	-.13	-.04	-.76	-.28	-.47	-.04	-.35	1.04
-.27	.14	.00	-.35	-1.22	.08	-.18	.40	-.84	.25	.02	-.54	.00	1.37
.15	.02	.29	.11	-.27	-.68	-.37	.35	-.33	.67	.44	-.12	.48	-.51
-.22	.07	-1.19	.61	-.15	-.46	-.03	.61	.04	.64	.12	.40	.12	.06
-.12	-.02	.38	1.42	-.06	-.17	-.08	-.62	.39	.65	.92	.35	.29	.00
.15	.09	-.17	.82	.16	.00	.17	-.22	.66	.40	.18	-.15	.49	.11
.38	.07	-.15	.93	.12	-.10	-.01	-.42	.95	.29	.86	.71	.21	.42
.36	-.84	1.17	-.81	.61	.76	.38	.00	.53	.38	.53	.09	-.08	.25
.08	-.18	.22	.91	.99	.53	.30	-.05	.48	.52	.25	.14	.47	.16
-.36	.36	.27	.74	.84	-.62	.69	.07	.36	.48	.10	.37	.01	-.27
.72	.14	.45	.70	.10	.43	.02	.27	.22	.53	.02	.20	.54	.23
.83	-.71	.09	.42	.49	-.11	.20	.18	.44	.26	.11	.30	-.44	-.23
-.07	.12	.39	.14	.59	.43	.33	.54	.12	.18	.24	-.08	-.62	.07
-.33	.23	-.27	.48	.37	.15	-.17	.28	.27	.62	.45	.48	-.45	.15
-.15	.26	.26	-.06	.14	.73	.58	-.10	.21	.21	-.24	-.20	.16	.11
-.26	-.25	-.06	-.23	-.19	.11	-.07	-.20	.53	-.04	.15	-.17	.10	.63
.17		-.82	.78	.18	.30	-.33	-.56	.37	.03	.25	.31		.02
.76		1.23	.42	.10	-.23		-.29	-.11		-.12	.08		-.34
-.28		1.25	.14	-.03	-.29		-.03	-.25		-.50	.47		-.24
-.58		-.14	.65	.54	.16			.06		.04			
.85			.62	.66	.15			-.02		.13			
1.12			.48					.45					
.45			-.11										

Table A4: Differences between corrected distributions from 92 and 93 in units of the whole error

### 3. Remark about normally distributed random variables

In this section we prove a fact that is important to several parts of this work. This also serves as an example that shows the handling of probability distributions.

*Example A1:* Distribution of a linear combination of independent and normally distributed random variables

*Fact:* If  $X_j, j \in \{1, \dots, N\}$  are normally distributed and independent random variables, that means if their probability density is given by

$$\rho(x_1, \dots, x_N) := \prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{1}{2}\left(\frac{x_j - \mu_j}{\sigma_j}\right)^2\right], \quad x_j, \mu_j \in \mathbf{R}, \quad \sigma_j \in \mathbf{R}^+$$

then each linear combination  $Y := \eta + \sum_{j=1}^N \xi_j X_j, \eta, \xi_j \in \mathbf{R}$  is again normally distributed, and the mean value and width are given by:

$$\mu = \eta + \sum_{j=1}^N \xi_j \mu_j, \quad \sigma = \sqrt{\sum_{j=1}^N (\xi_j \sigma_j)^2}$$


---

The proof is done in two steps. In the first step, the calculation of this density is explicitly done for the case  $N=2$ . The second step is a generalization by induction with respect to  $N$ .

*Step 1:*

The probability density of a random variable  $Y_2 := \eta + \xi_1 X_1 + \xi_2 X_2$  can be calculated as:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \rho(x_1) \rho(x_2) \Big|_{y_2 = \eta + \xi_1 x_1 + \xi_2 x_2} = \int_{-\infty}^{\infty} \left\{ \frac{1}{|\xi_2|} \int_{-\infty}^{\infty} dx_1 \rho(x_1) \rho\left(\frac{y_2 - \eta - \xi_1 x_1}{\xi_2}\right) \right\} dy_2 = \\ &= \int_{-\infty}^{\infty} \rho(y_2) dy_2 \Rightarrow \end{aligned}$$

$$\Rightarrow \rho(y_2) = \frac{1}{2\pi\sigma_1\sigma_2} \frac{1}{|\xi_2|} \int_{-\infty}^{\infty} dx_1 \exp\left[-\frac{1}{2}\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2\right] \exp\left[-\frac{1}{2}\left(\frac{y - \eta - \xi_1 x_1 - \xi_2 \mu_2}{\sigma_2 \xi_2}\right)^2\right] =: (*)$$

It follows that:

$$\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{y - \eta - \xi_1 x_1 - \xi_2 \mu_2}{\sigma_2 \xi_2}\right)^2 = \frac{(\xi_1 x_1)^2 + (\xi_2 \mu_2)^2}{(\xi_1 x_1)^2 (\xi_2 \mu_2)^2} (\xi_1 x_1 - \kappa)^2 + \left(\frac{y - [\eta + \xi_1 x_1 + \xi_2 \mu_2]}{\sqrt{(\xi_1 x_1)^2 + (\xi_2 \mu_2)^2}}\right)^2$$

with:

$$\kappa := \frac{[y - \eta - \xi_2 \mu_2](\xi_1 \mu_1)^2 + \xi_1 \mu_1 (\xi_2 \mu_2)^2}{(\xi_1 \mu_1)^2 + (\xi_2 \mu_2)^2}$$

$$\Rightarrow (*) = \frac{1}{2\pi\sqrt{(\xi_1 \mu_1)^2 + (\xi_2 \mu_2)^2}} \exp\left[-\frac{1}{2}\left(\frac{y - [\eta + \xi_1 x_1 + \xi_2 \mu_2]}{\sqrt{(\xi_1 x_1)^2 + (\xi_2 \mu_2)^2}}\right)^2\right]$$

Now the proof for  $N = 2$  is done, and we continue with

**Step 2:**

The starting point for the proof by induction is the assumption, that the proof for the case  $Y_N := \eta + \sum_{j=1}^N \xi_j X_j$  is already done, and concluding from this, we prove that this assumption does also hold for  $Y_{N+1} := \eta + \sum_{j=1}^{N+1} \xi_j X_j$ . This can be seen as follows:

$Y_{N+1} := Y_N + \xi_{N+1} X_{N+1}$  again is a sum of two normally distributed and independent random variables, and we can write down their mean value and width, using the result of step 1:

$$\mu_{Y_{N+1}} = \mu_{Y_N} + \xi_{N+1} \mu_{N+1} = \eta + \sum_{j=1}^{N+1} \xi_j \mu_j$$

$$\sigma_{Y_{N+1}}^2 = \sigma_{Y_N}^2 + (\xi_{N+1} \sigma_{N+1})^2 = \sum_{j=1}^{N+1} (\xi_j \sigma_j)^2$$

o

## 4. Remarks concerning a model independent detector matrix

The  $\chi^2$ -test of section 3.4.2 can be used to check directly the correctness of the unfolded data if the detector matrix D (c.f. section 3.3.1.1) is (in good approximation) independent of the QCD-part of the simulation. The sensitivity of this test will be illustrated in the next example.

*Example A2: Statistical test of unfolded data in the case of a QCD model-independent detector matrix*

*Idea: The increase of the sensitivity of the  $\chi^2$ -test introduced in section 3.4.2 with an increasing statistics will be demonstrated using the matrix unfolding procedure from section 3.3.1.3, .*

-----

The test quantities with and without inclusion of the statistical widths of the parametrization are given in the following table in the context of the simple detector model that was introduced in example 3.1. The tests are applied to the "data" distribution  $t_{dat}$ , to the "incorrect simulation"  $t_{sim}$ , and to the result obtained by an unfolding with this "incorrect simulation"  $t_{mat}$ . (For details see example 3.1.) The second and third of these distributions are clearly excluded by this test, while the correct distribution passes it. The high sensitivity of this test, which is given by the big values of the test quantity, reflects the fact that it is applicable even in cases where the simulation does not fail as dramatic as in the case of this simple model.



$S_{sim}$	$S_{dat}$	$\chi^2_{+s}(t_{mat})$	$\chi^2_{-s}(t_{mat})$	$\chi^2_{+s}(t_{dat})$	$\chi^2_{-s}(t_{dat})$	$\chi^2_{+s}(t_{sim})$	$\chi^2_{-s}(t_{sim})$
10000	1000	16.24	17.25	6.61	7.54	43.59	46.20
50000	5000	69.88	75.99	8.89	9.80	240.18	252.60
100000	10000	127.28	138.19	11.70	12.99	479.50	504.25
500000	50000	502.62	546.14	13.07	15.89	2032.07	2132.80
1000000	100000	1071.91	1164.47	7.29	8.54	4289.13	4503.45
1500000	150000	1519.56	1650.25	21.53	25.53	6067.65	6367.23
2000000	200000	2080.18	2258.93	16.67	20.20	8157.90	8559.11

With  $n=13$  the hypothesis of a vanishing mean value (and thus the hypothesis of a correct simulation) can be excluded with a probability of 95% (99%) if the value of the test quantity exceeds 22.4 (27.7) (c.f. [Br,79]).

o

If an additional information is given, this test can be expanded to a full unfolding procedure. The maximum of the  $\chi^2$  probability is not unique without this information, and therefore we can not apply the methods of parameter estimation introduced in chapter 5. Many so-called "oscillatory solutions" have the same or even a higher probability as the true distribution in this case. One preferred maximum is taken by introducing new information. This point is illustrated by the next example.

*Example A3: Maximum likelihood and minimum curvature*

Idea: The  $\chi^2$ -test that was discussed in the last example will be expanded to a "pseudo unfolding algorithm". For that reason the principle of "minimum curvature" is introduced. On that basis the unfolding of a distribution is equal to the solution of a minimization problem that is restricted by one condition.

-----

$\bar{t}$  signs the unknown true distribution (the unfolded distribution). Analogous to section 3.4.2 the  $\chi^2$ -distributed quantity

$$\chi_{+s}^2(\bar{t}) := \sum_{j=1}^n \frac{\left\{ o_j - \sum_{i=1}^m D_{ji} t_i \Delta x_i \right\}^2}{\sigma_j^2 + \sum_{i=1}^m (\sigma_{ji} t_i \Delta x_i)^2}$$

is constructed. In addition the following measure of the "overall curvature" of  $\bar{t}$

$$k(\bar{t}) := \sum_{i=2}^{m-1} \left[ 2 \frac{\Delta x_{i-1} (t_{i+1} - t_i) - \Delta x_i (t_i - t_{i-1})}{\Delta x_{i-1} \Delta x_i (\Delta x_{i-1} + \Delta x_i)} \right]^2$$

is introduced. It is motivated by the definition of second derivatives. In the case of bins with equally widths this expression becomes simpler in this form

$$k(\bar{t}) := \sum_{i=2}^{m-1} \frac{(t_{i+1} - 2t_i + t_{i-1})^2}{\Delta x^4}$$

The most probable value of  $\chi_{+s}^2$  is  $n-2$ . Consequently every vector  $\bar{t}$  which accounts for the condition  $\chi_{+s}^2(\bar{t}) = n-2$  should be a favorite candidate for the unfolded distribution. The one which minimizes  $k(\bar{t})$  is taken from this set of vectors. This conditioned minimum is found by iteration.

A vector  $\bar{t}_0$  which satisfies the condition  $\chi_{+s}^2(\bar{t}_0) - (n-2) = 0$  is the starting point. To find such a vector, a maximum of the  $\chi^2$ -probability

$$P(\chi_{+s}^2) := \frac{\exp\left[-\frac{\chi_{+s}^2}{2}\right] (\chi_{+s}^2)^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}$$

is calculated using standard minimization methods. The next step is to calculate an normalized and orthogonal basis  $B := \{\bar{b}_1, \dots, \bar{b}_{n-1}\}$  in the vector space that is perpendicular to  $\bar{V}\chi_{+s}^2(\bar{t}_0)$ , using the standard basis of  $\mathbf{R}^n$  and an orthogonalization algorithm ("Schmidt'sches Orthogonalisierungsverfahren" c.f. [Br,87]). Now the direction

$$\bar{v} := \sum_{i=1}^{n-1} \lambda_i \bar{b}_i \quad \text{with} \quad \sum_{i=1}^{n-1} \lambda_i^2 = 1$$

in this space is evaluated, in which the slope of the function  $k$  reaches its minimum. For this reason, the solution of the system of equations

$$\frac{\partial a_v(\bar{t}_0)}{\partial \lambda_j} = \frac{\partial}{\partial \lambda_j} \left\{ \bar{\nabla}k(\bar{t}_0) \cdot \left[ \sum_{i=1}^{n-2} \lambda_i \bar{b}_i + \sqrt{1 - \sum_{i=1}^{n-2} \lambda_i^2} \bar{b}_{n-1} \right] \right\} = 0$$

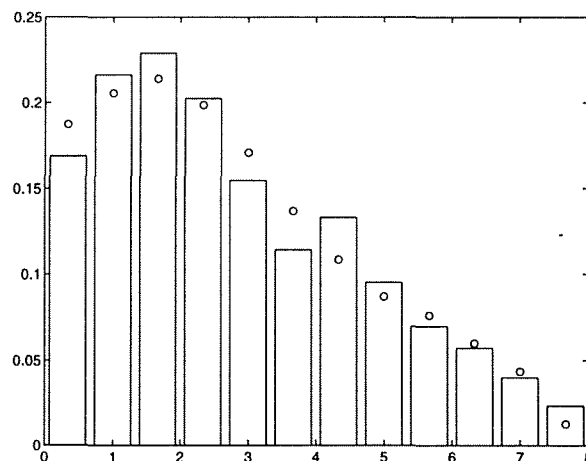
$$\Rightarrow \bar{\nabla}k(\bar{t}_0) \cdot \left[ \bar{b}_j - \frac{\lambda_j \bar{b}_{n-1}}{\sqrt{1 - \sum_{i=1}^{n-2} \lambda_i^2}} \right] = 0$$

has to be found.

With  $|\lambda_{n-1}| = \sqrt{1 - \sum_{i=1}^{n-2} \lambda_i^2}$  we observe that  $\lambda_j = \frac{\bar{\nabla}k(\bar{t}_0) \cdot \bar{b}_j}{\bar{\nabla}k(\bar{t}_0) \cdot \bar{b}_{n-1}} |\lambda_{n-1}| =: \xi_j |\lambda_{n-1}|$  and thus

$$\lambda_{n-1} = \pm \frac{1}{\sqrt{1 - \sum_{i=1}^{n-2} \xi_i^2}}$$

Now it is possible to make a small step  $\varepsilon$  in the direction of the minimum value of the slope of  $k$  without changing the value of  $\chi_{+s}^2$ . The latter has to be checked, because we cannot really make infinitesimal steps. If the  $\chi_{+s}^2$  is in a given tolerance region, this new vector  $\bar{t}_1$  is accepted as a starting point for the next iteration step. If the check fails, the correction of  $\bar{t}_1$  is done parallel to  $\bar{\nabla}\chi_{+s}^2(\bar{t}_0)$ , or if this is not possible, the  $\varepsilon$ -value is decreased. In this example we used  $s_{sim} := 5 \cdot 10^5$  and  $s_{dat} := 5 \cdot 10^4$  and got the following unfolded distribution:



The nominal values of the true distribution is given in form of bars. Even if the result seems to be better than the oscillating "distributions" given by matrix inversion (c.f. example 3.3), a bias through a straight line can be seen.

o

The main problems appear in the calculation of error bars. Even if it seemed possible to get better results by using the ad hoc principle of "minimal curvature", it is not possible to calculate the bias introduced by such an ad hoc ansatz. If a proper estimation of errors is regarded, one has to use well understood algorithms based on first principles. Otherwise one could hardly say more than he believes, that the calculated results are "not too bad". A very interesting idea that starts from first principle is given by "*entropy methods*" (for example see [Sc,94])

## 5. Values of fitted parameters that were used to calculate the model bias

Results from simplified corrections of distributions, using different *QCD* models are compared to calculate an estimate for the model bias. The parameter values which are different from the default values are given in the following table. For a detailed explanation of the parameters see [JS,93], [AR,92] and [HE,92].

<i>JS 7.4 (aniso)</i>		<i>JS 7.4 (iso)</i>		<i>ARIADNE 4.05</i>		<i>HERWIG 5.8</i>	
<i>Par.</i>	$p^0$	<i>Par.</i>	$p^0$	<i>Par.</i>	$p^0$	<i>Par.</i>	$p^0$
<i>PARJ(1)</i>	0.094	<i>PARJ(1)</i>	0.095	<i>PARA(1)</i>	0.218	<i>QCDLAM</i>	0.172
<i>PARJ(2)</i>	0.289	<i>PARJ(2)</i>	0.289	<i>PARA(3)</i>	0.580	<i>RMAS(13)</i>	0.727
<i>PARJ(3)</i>	0.566	<i>PARJ(3)</i>	0.571	<i>PARA(5)</i>	0.580	<i>CLMAX</i>	3.700
<i>PARJ(11)</i>	0.538	<i>PARJ(11)</i>	0.533	<i>MSTA(1)</i>	1	<i>VPCUT</i>	1.000
<i>PARJ(12)</i>	0.499	<i>PARJ(12)</i>	0.543	<i>MSTA(2)</i>	1		
<i>PARJ(13)</i>	0.600	<i>PARJ(13)</i>	0.600	<i>MSTA(3)</i>	0		
<i>PARJ(14)</i>	0.096	<i>PARJ(14)</i>	0.096	<i>MSTA(5)</i>	1		
<i>PARJ(15)</i>	0.032	<i>PARJ(15)</i>	0.032	<i>MSTA(20)</i>	1		
<i>PARJ(16)</i>	0.096	<i>PARJ(16)</i>	0.096	<i>MSTA(30)</i>	0		
<i>PARJ(17)</i>	0.160	<i>PARJ(17)</i>	0.160				
<i>PARJ(21)</i>	0.373	<i>PARJ(21)</i>	0.363				
<i>PARJ(26)</i>	0.400	<i>PARJ(26)</i>	0.400				
<i>PARJ(41)</i>	0.500	<i>PARJ(41)</i>	0.400				
<i>PARJ(42)</i>	1.008	<i>PARJ(42)</i>	1.030				
<i>PARJ(54)</i>	-0.050	<i>PARJ(54)</i>	-0.050				
<i>PARJ(55)</i>	-0.0045	<i>PARJ(55)</i>	-0.0045				
<i>PARJ(81)</i>	0.297	<i>PARJ(81)</i>	0.324				
<i>PARJ(82)</i>	1.330	<i>PARJ(82)</i>	1.300				
<i>MSTJ(11)</i>	3	<i>MSTJ(11)</i>	3				
<i>MSTU(41)</i>	1	<i>MSTU(41)</i>	1				
		<i>MSTJ(46)</i>	0				



## Appendix B.

### Remarks concerning part II

#### 1. Sources of the distributions used in *LinFit*

In addition to the distributions that was presented in preceding parts of this thesis, the following measured distributions and multiplicities are used in the tuning procedure:

##### (i) Baryon data

<i>distribution</i>	<i>bins</i>	<i>source</i>	<i>experiment</i>
$-\ln(x_p^{(p,\bar{p})}), x < 0.018$	6	[A1,94]	ALEPH
$-\ln(x_p^{(p,\bar{p})}), x > 0.070$	18	[A1,94]	ALEPH
$\langle \Lambda \rangle$	1	[A2,94]	ALEPH
$-\ln(x_p^\Lambda)$	22	[A2,94]	ALEPH
$p_i^\Lambda(T)$	30	[A2,94]	ALEPH
$\langle \Xi^- \rangle$	1	[A3,95]	ALEPH
$\langle \Xi(1530)^0 \rangle$	1	[A3,95]	ALEPH
$\langle \Sigma(1385)^\pm \rangle$	1	[A3,95]	ALEPH
$\langle \Omega^- \rangle$	1	[A3,95]	ALEPH

(ii) Meson data

<i>distribution</i>	<i>bins</i>	<i>source</i>	<i>experiment</i>
$-\ln(x_p^{\pi^{\pm}}), x < 0.018$	16	[A1,94]	ALEPH
$-\ln(x_p^{\pi^{\pm}}), x > 0.045$	23	[A1,94]	ALEPH
$\langle K^0 \rangle$	1	[A2,94]	ALEPH
$-\ln(x_p^{K^0})$	28	[A2,94]	ALEPH
$p_i^{K^0}(T)$	30	[A2,94]	ALEPH
$-\ln(x_p^{K^{\pm}}), x < 0.018$	11	[A1,94]	ALEPH
$-\ln(x_p^{K^{\pm}}), x > 0.070$	18	[A1,94]	ALEPH
$x_p^{K^{*0}}$	8	[A1,95]	ALEPH
$\langle K^{*+} \rangle$	1	[D1,94]	DELPHI
$x_E^{\eta}$	18	[A3,94]	ALEPH
$x_E^{\eta'}$	9	[A3,94]	ALEPH
$x_p^{\phi}$	8	[A1,95]	ALEPH
$x_p^{\rho_0}$	8	[A1,95]	ALEPH
$x_p^{\omega}$	6	[A1,95]	ALEPH
$\langle f_0 \rangle, 0.05 \leq x_p \leq 0.6$	1	[D1,94]	DELPHI
$\langle f_2 \rangle, 0.05 \leq x_p \leq 1.0$	1	[D1,94]	DELPHI
$x_E^{D^{*\pm}}$	15	[A1,93]	ALEPH



## 2. Decoupling of normally distributed random variables

*Idea: The following decoupling procedure, allows for a construction of a  $\chi^2$ -distributed random variable, and for the generalization of the linear error propagation from example A1, to the case of correlated random variables.*

---

The distribution of  $n$  normally distributed random variables is given by:

$$\rho(\bar{x}) = \frac{1}{(\sqrt{2\pi})^n \det(\mathbf{C})} \exp\left[-\frac{1}{2}(\bar{x} - \bar{\mu})^T \mathbf{C}^{-1}(\bar{x} - \bar{\mu})\right]$$

Here  $\mathbf{C}$  is the covariance matrix, which is a quadratic and symmetric matrix. Because of these characteristics, even the inverse of  $\mathbf{C}$  is symmetric, and it is possible to find a matrix  $\mathbf{U}$  that diagonalizes this inverse of  $\mathbf{C}$ :  $\mathbf{U}\mathbf{C}^{-1}\mathbf{U}^T = \text{diag}(\lambda_1, \dots, \lambda_n)$ . The substitution  $\bar{x} = \mathbf{U}(\bar{x} - \bar{\mu})$  allows for a calculation of the probability density of this new set of random variables:

$$\rho(\bar{x}) = \frac{1}{(\sqrt{2\pi})^n \det(\mathbf{C})} \exp\left[-\frac{1}{2}\bar{x}^T \bar{\mathbf{C}}^{-1}\bar{x}\right] = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{\bar{x}_i}{\sigma_i}\right)^2\right]$$

With  $\sigma_i := \frac{1}{\sqrt{\lambda_i}}$  and  $\mathbf{U}\mathbf{C}^{-1}\mathbf{U}^T =: \bar{\mathbf{C}}^{-1}$

Therefore the new random variables  $\bar{x}_i := \sum_{j=1}^n U_{ij}x_j$  are normally distributed and independent, and their mean values are  $\bar{\mu}_i := \sum_{j=1}^n U_{ij}\mu_j$ .

This decoupling will now be illustrated in constructing a  $\chi^2$  distributed quantity out of a set of normally distributed but correlated random variables. The covariance matrix of the normally distributed bin contents  $\rho_i$ , which is also the covariance matrix of the differences  $x_i := \bar{\rho}_i - \{k(x_i - x_0) + d\}$  is:

$$C_{ij} = \begin{cases} \frac{sp_i(1-p_i)}{(s\Delta x_i)^2}, & \text{if } i \neq j \\ \frac{-sp_i p_j}{(s\Delta x_i)^2}, & \text{else} \end{cases}$$

These differences were decoupled by the method developed above, and the test quantity was constructed

$$\eta_{uc} := \sum_{i=1}^n \left( \frac{\bar{x}_i}{\sigma_i} \right)^2$$

which is  $\chi^2$  distributed with  $n (= 5)$  degrees of freedom. Similar to the procedure in example 5.3, this distribution is compared to the generated distribution of the test quantity. The agreement with the  $\chi^2$  is indeed better than in this example which ignored correlations.

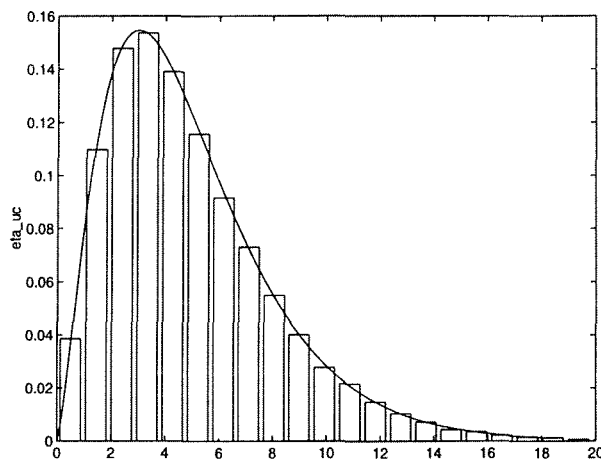


Figure B1: Test quantity for the decoupled random variable

o

### 3. Higher order parametrization in higher dimensions

We always used linear parametrisations for the model predictions in the parameter tuning described in the main part of this thesis. Because of this we always had to be very careful to stay inside a linear region. This was possible in our case, because some former knowledge about the parameter values was available. It also turned out, that some kind of successive approximation to an optimal expansion point of the linear parametrization worked. But what is the right proceeding in the case of absence of such information? Is it possible to use higher order parametrizations, or does the limited computing power that is available today restrict us to linear approximations?

The main problem in using higher order parametrisations is the number of their coefficients. For each of them, at least two points in the parameter space should be calculated to fix this parametrization. And for each of these points, a statistics as high as possible should be used, to avoid contributions to the errors from insufficient parametrization accuracy. All in all the demands of computing resources are generally much too high to perform simultaneous fits for many parameters.

The following example demonstrate a very simple kind of generalization of the parametrization to higher orders. This is primarily thought as a summary of the methods that were derived in chapter 5, and as an illustration of their usage in more complicated cases of parameter estimations. It is definitely not a try to give a generalization usable in any possible case, even if it lead to results much better than the estimations based upon a linear parametrization.

#### *Example B1: Higher order fit in more than one dimension*

*Idea: The model function for the hadronic cross-section of an electron-positron annihilation at the Z-peak given in 5.5 will be used to illustrate a simple generalization of the parametrization to higher orders. The conditions of this example are chosen in a way that leads to non-negligible contributions of the parametrization fluctuations to the estimations function. The estimation is done with the simple estimation function (5.4), and with the generalization (5.6), which is used here in a nontrivial way, because the*

full covariance matrix of the parametrization coefficients has to be taken into account. The results of both estimation processes will be compared.

---

*(i) Preparation of the surrounding*

We use the hadronic cross section in the range of  $E_{cm} \in [88, 96] GeV$ , and divide this region into 13 bins with equally widths. After the generation of a discrete test distribution in a way that is analogous to example 3.2, and by using  $s := 5 \cdot 10^5$  and the parameter values  $\sin^2 \vartheta_w =: p_1 = 0.232$ ,  $N_v =: p_2 = 3$  and  $M_z =: p_3 = 91.182$ , a linear parametrization that is similar to the one used in LinFit is calculated. The region taken into account was  $p_1 = 0.2 \pm 0.15$ ,  $p_2 = 3 \pm 2$  and  $p_3 = 91 \pm 1.5$ . Here the center values correspond to the expansion point, and we used four points on every parameter axis off the center point to do the calculation. These points are located at the ends of the parameter intervals, and on the half way between these ends and the center. We used  $s := 10^4$  in every point.

The expansion point is slightly apart from the chosen set of parameters, because we want to demonstrate, that even the simple choice of a higher order parametrization that we use in this example leads to better agreement between estimations and the nominal parameter values, than the linearized version. The results of the estimation based on a linear parametrization are:

$$p_1 := 0.3240 \pm 0.0092, \quad p_2 := 1.8880 \pm 0.1560, \quad p_3 := 91.2729 \pm 0.0247$$

These estimates are clearly not compatible with the nominal values, and we conclude, that the usage of the expansion point given above leads to systematic errors in the results of the fits.

*(ii) Parametrization by a higher order polynomial*

The model prediction  $m_b(\bar{p})$  of a given bin content  $\rho_b$  is parametrized by the following polynomial in  $n$  dimensions and with a degree of  $d$

$$m_b(\bar{p}) := \sum_{\substack{i_1, \dots, i_n=1 \\ i_1 + \dots + i_n = d}}^d c_{i_1, \dots, i_n}^b p_1^{i_1} \cdots p_n^{i_n} =: \sum_{n=1}^N a_n X_n(\bar{p})$$

The second way of writing this polynomial is a sum over all  $n$ -tuples  $(i_1, \dots, i_n)$  of powers which fulfills the condition  $i_1 + \dots + i_n = d$ . The bin index is left out in this case to avoid too dense notations. We will use this second expression in the realization of the parametrization. If estimates for the model predictions  $\bar{m}(\bar{p}_j)$  and the corresponding widths  $\sigma_j$  are known in  $M > N$  points of the region in the parameter space that one wants to parametrize, the coefficients of this parametrization can be estimated in a way analogous to section 5.4.2.1 in every bin of the distribution of interest. The minimum of the estimation function

$$S_{pol}(\bar{a}) := \frac{1}{2} \sum_{j=1}^M \left\{ \frac{\bar{m}_j - m(\bar{p}_j)}{\sigma_j} \right\}^2$$

is found by solving the following system of linear equations:

$$\mathbf{M}^{pol} \bar{\mathbf{a}}^{\min} = \bar{\mathbf{y}}^{pol},$$

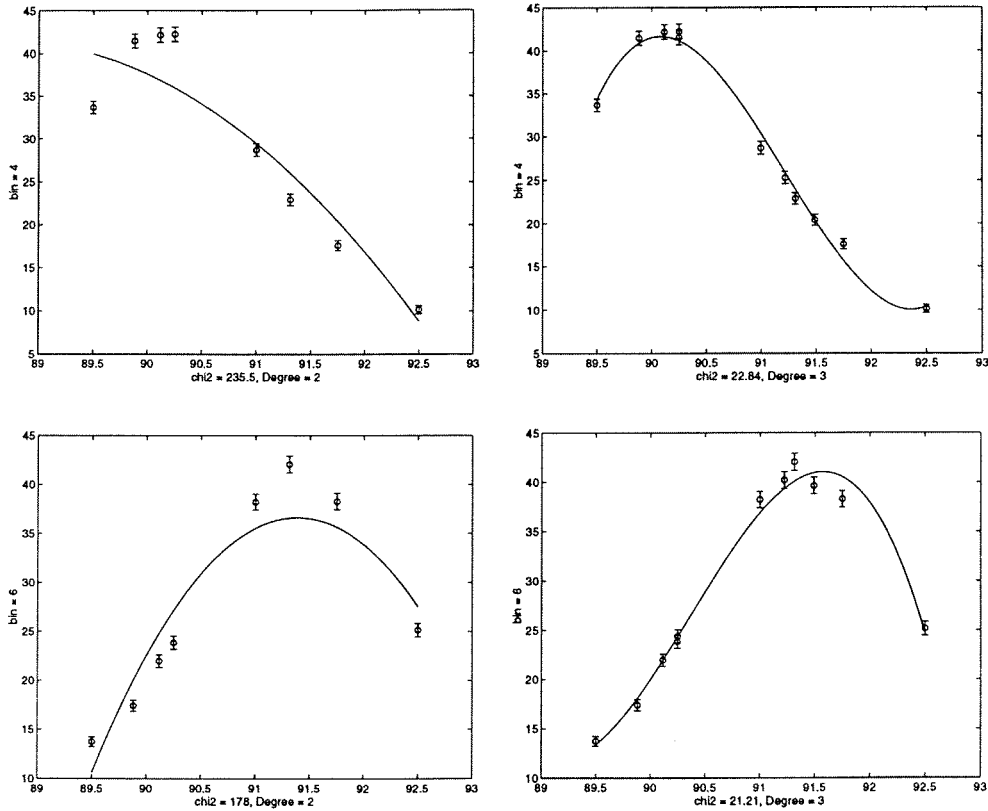
$$M_{kl}^{pol} = \sum_{j=1}^M \frac{X_k(\bar{p}_j) X_l(\bar{p}_j)}{\sigma_j^2}, \quad y_k^{pol} = \sum_{j=1}^M \frac{\bar{m}(\bar{p}_j) X_k(\bar{p}_j)}{\sigma_j^2}$$

( $k, l = 0, 1, \dots, N$ ), and the covariance matrix of the coefficients is again given by the inverse of  $\mathbf{M}$ .

### (i) A simple higher order parametrization

It is definitely not the best way to simply increase the degree of the polynomial if a test of the linear parametrization fails, because in general not all the new terms will contribute to a better parametrization. To get a better understanding for a minimal parametrization, we introduce fits along the parameter axes again, and increase the degree of the one-dimensional polynomial used to parametrize the behavior of the model predictions along these parameter axes, until a  $\chi^2$  test quantity lies in a preferred region. In the next pictures the degrees of these one dimensional parametrizations define an upper limit for the powers that are used in the parametrization of the

model function. The approximation of the parametrization along the  $p_3$ -axis are given for the 4-th and for the 6-th bin.

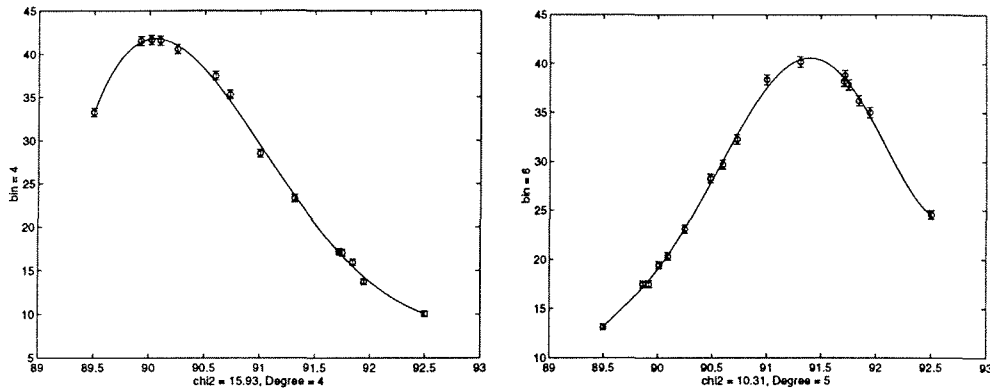


The parametrization along the different axes were accepted if the distance between the test quantity  $\eta$  and their mean value was less than three times the standard deviation of  $\eta$ . This condition led to the following degrees of the polynomial parametrizations in the 13 bins:

<i>bin</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
$d_{p_1}$	2	2	2	2	2	2	2	2	2	2	2	1	2
$d_{p_2}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$d_{p_3}$	3	2	4	3	4	3	3	2	2	2	2	2	2

Here  $d_{p_i}$  signs the degree of the polynomial that passed the  $\eta$ -condition in the  $i$ -th bin. A main problem can be seen in the parametrization of bin 6. Even if the  $\eta$ -condition is fulfilled, the parametrizations of this bin obviously requires a higher degree. If one uses a higher statistics in every point of the parameter space (for example  $s := 2.5 \cdot 10^5$  was used to produce

the following pictures), this can be seen as well. Because of the higher sensitivity of the statistical test in this latter case, the following degree of the parametrization in these two bins are needed to pass the test.



That means, that an incorrect parametrization could be used even if the test quantity lies in the preferred region, because the statistics used in every point is not high enough.

After these one dimensional parametrizations were done, the number of coefficients  $N$  was calculated, and the model predictions for  $3N$  additional points were generated. The points were taken at random out of the observed region in the parameter space. After this step, the parametrization was done as discussed in (i).

#### (iv) Estimation of the parameters

Up to now we calculated a parametrization of the given model predictions, which we do not expect to be perfect, but is still better than the linear parametrization. The estimation of the parameters was now done with the simple estimation function (5.4) and with the new one (5.6). Here, the calculation of the width of a linear combination of normally distributed random variables will be discussed in the case of non-negligible correlations. We use this to calculate the width of

$$m_b(\bar{p}) = \sum_{n=1}^N a_n X_n(\bar{p}) =: \sum_{n=1}^N Y_n \lambda_n$$

Here the symbols  $X_n(\bar{p}) := \lambda_n$  for the real coefficients and  $a_n := Y_n$  for the normally distributed random variables are used. We calculate the matrix

U that diagonalizes the inverse of the given covariance matrix, as discussed in section 2 of this appendix, and construct the uncorrelated random variables

$$Z_i := \sum_{j=1}^n U_{ij} Y_j$$

By take the inverse of the equation above, every linear combination of correlated random variables can be expressed as a linear combination of uncorrelated variables by substitution, and consequently the linear error propagation of example A1 can be applied.

The estimations led to the following results:

$$p_1 := 0.2381 \pm 0.0016 (3.7), \quad p_2 := 3.0995 \pm 0.0147 (6.8), \quad p_3 := 91.1829 \pm 0.0022 (0.4),$$

$$p_1 := 0.2407 \pm 0.0038 (2.3), \quad p_2 := 3.0882 \pm 0.0345 (2.6), \quad p_3 := 91.1829 \pm 0.0051 (0.2),$$

The results in the first line are computed with the simple estimation function (5.4), while the results in the second line are based on the generalization (5.6) The values in brackets are the absolute values of distances between these estimations and the nominal values. Even if the parametrization was not optimal, it led to satisfactory results, and not only to results that are better than the estimates based on a linear parametrization (there the distances were (10.0), (7.1), (3.7)). The estimation with the simple estimation function gave good estimations, but the errors calculated from it are too small, as should be expected.

o

**Remark:** We must observe the starting point, that is used in the iterative approximation to the minimum of the estimation function as a point where complications might arise. Especially the generalization is highly non-linear, and has therefore in general more than one local minimum. A first test to see if the right minimum is reached can again be done by a  $\chi^2$  test.

If a proper degree of the polynomial is known, the generalized estimation function can be used to realize this parametrization even if the statistics in every point of the parameter space is not very high. For example it should be possible to



use a slightly generalized algorithm to realize a higher order parametrization. The fit along the axes, as discussed in the previous example should be done with high statistics at every point, in order to avoid false parametrizations as far as possible. But the  $3N$  additional points in parameter space can also be calculated by using a lower statistics, because their widths should be included in the estimation function. Therefore the last example and this remark can be used as a starting point for the generation of higher order parametrizations in future parameter fitting procedures.

#### 4. Parameter values used in example 6.1

To demonstrate the inclusion of systematic errors in the estimation function, the following parameter settings were used. Values that are not given in this table, and that are not set to certain values in chapter 2, are fixed at their default values.

<i>JS 7.4 (iso)</i>		<i>HERWIG 5.8</i>	
<i>Par.</i>	$p^0$	<i>Par.</i>	$p^0$
$\Lambda$	0.30	<i>QCDLAM</i>	0.15
$Q_0$	1.40	<i>RMASS(13)</i>	0.70
$\sigma$	0.36	<i>CLMAX</i>	3.50
$P_{u,d}^{S=1}$	0.55	<i>CLSMR</i>	0.50
$P_s^{S=1}$	0.50	<i>PWT(3)</i>	1.00
$qq/q$	0.10	<i>PWT(7)</i>	1.00
$(su)/(du)$	0.60		
$\gamma_s$	0.30		
$b$	0.90		
$\bar{b}_l$	0.60		

# Bibliography

- [A1,89] *ALEPH* Collaboration: Determination of the number of light Neutrino species. *Phys.Lett. B 231 (1989) 519-529*.
- [A1,92] *ALEPH* Collaboration: Properties of hadronic Z decays and test of QCD generators. *Z.Phys.C - Particles and fields 55, 209-234 (1992)*.
- [A1,90] *ALEPH* Collaboration: A precise determination of the number of families with light Neutrino and of the Z Boson partial width. *Phys.Lett. B 235 (1990) 399-411*.
- [A1,93] *ALEPH* Collaboration: Production of Charmed Mesons in Z Decays. *CERN-PPE/93-208*.
- [A1,94] *ALEPH* Collaboration: Inclusive  $\pi^\pm$ ,  $K^\pm$  and  $(p, \bar{p})$  Differential Cross-section at the Z Resonance. *CERN-PPE/94-201*.
- [A2,94] *ALEPH* Collaboration: Production of  $K^0$  and  $\Lambda$  in hadronic Z decays. *CERN-PPE/94-074*.
- [A3,94] *ALEPH* Collaboration: Production Rates of  $\eta$  and  $\eta'$  in Hadronic Z Decays. Contribution to *ICHEP94* Glasgow, Ref. 0535.
- [A4,94] *ALEPH* Collaboration: Performance of the *ALEPH* detector at *LEP*. *CERN-PPE/94-170*.
- [A1,95] *ALEPH* Collaboration: Inclusive Production of Neutral Vector Mesons in Hadronic Z Decays. *CERN-PPE/95-100*.
- [A2,95] *ALEPH* Collaboration: Measurements of the Charged Particle Multiplicity Distribution in Restricted Rapidity Intervals. *CERN-PPE/95-82*.
- [A3,95] *ALEPH* Collaboration: Hyperon Production in Z Decays. Contribution to *HEP 95*, Brussels: Ref. *EPS0419*.
- [A4,95] *ALEPH* Collaboration: Production of Excited Beauty States in Z Decays. *CERN-PPE/95-108*.
- [AL,90] *ALEPH* Collaboration: *ALEPH*: A Detector for Electron-Positron Annihilations at *LEP*. *A294 (1990) 121-178*.

- [AL,94] *ALEPH* Collaboration: *ALEPH 101*. An Introduction to the *ALEPH* Offline System. *ALEPH 94-150*, *SOFTWR 94-012*.
- [AL,95] *ALEPH* Collaboration: The *ALEPH* Handbook 1995, Volume 1. *CERN 1995*.
- [Ap,95] H.Albrecht, E.Blucher, J.Boucrot: ALPHA User's Guide. Version 120/121. *ALEPH 95-112*. *SOFTWR 95-006*.
- [AR,92] L.Lönnblad: *ARIADNE* - A Program for Simulation of *QCD*-Cascades Implementing the Colout Dipole Model. *Comput. Phys. Commun.* 71 (1992) 15.
- [Bl,84] V.Blobel: Unfolding methods in high-energy physics experiments. Lectures given at the 1984 *CERN* school of computing. Aigublava, Spain 1994.
- [Bo,91] K.Bosch: Elementare Einführung in die Wahrscheinlichkeitsrechnung. 5. Auflage, Vieweg 1991.
- [Bo,93] K.Bosch: Aufgaben und Lösungen zur angewandten Statistik. 2. Auflage, Vieweg 1993.
- [Br,87] I.N.Bronstein, K.A.Semendjajew: Taschenbuch der Mathematik. 23. Auflage, Deutsch 1987.
- [Br,92] S.Brandt: Datenanalyse (mit statistischen Methoden und Computerprogrammen). 3. Auflage, BI Wissenschaftsverlag 1992.
- [D1,94] *DELPHI* Collaboration: Production Characteristics of  $K^0$  and Light Meson Resonances in Hadronic Decays of the  $Z^0$ . *CERN-PPE/94-130*.
- [D1,95] *DELPHI* Collaboration: Tuning and Test of Fragmentation Models Based on Identified Particles and Precision Event Shape Data. *CERN-PPE/95-80*.
- [Di,93] G.Dissertori: Eine Untersuchung von Parton Schauer Modellen der perturbativen Quantenchromodynamik (Diplomarbeit). Innsbruck 1993.
- [Ea,71] W.T.Eadie, D.Drijard, F.E.James, M.Roos, B.Sadoulet: Statistical Methods in Experimental Physics. North-Holland 1971.
- [Fi,62] M.Fisz: Wahrscheinlichkeitsrechnung und mathematische Statistik. Deutscher Verlag der Wissenschaft 1962.
- [Fi,89] R.D.Field: Applications of Perturbative *QCD*. (Frontiers in Physics) Addison Wesley 1989.
- [Ha,84] F.Halzen, A.D.Martin: Quarks and Leptons (An Introductory Course in Modern Particle Physics). Wiley & Sons 1984.

- [HE,92] G.Marchesini, B.R.Webber, G.Abbiendi, I.G.Knowles, M.H.Seymour and L.Stanco: *HERWIG* - a Monte Carlo event generator for simulating Hadron Emissions With Interfering Gluons. Computer Physics Communication *67 (1992) 465*.
- [JS,93] T. Sjöstrand: *PYTHIA 5.7* and *JETSET 7.4* (Physics and Manual). *CERN-TH.7112/93*.
- [Kn,95] E.Kneringer: String-effekt in 3-Jet Ereignissen am Z-Pol (Dissertation), Innsbruck 1995
- [L1,89] R.Kleiss: Z Physics at *LEP 1*, Volume 3: Event Generators and Software, *CERN 89-08*
- [Lo,90] E.Lohrmann: Einführung in die Elementarteilchenphysik. Teubner 1990.
- [O1,90] *OPAL* Collaboration: A measurement of global event shape distributions in the hadronic decays of the  $Z^0$ .
- [O1,93] *OPAL* Collaboration: A Determination of  $\alpha_s(M_{Z^0})$  at *LEP* using Resummed *QCD* Calculatons. *CERN-PPE/93-38*
- [PA,95] *ALEPH* Collaboration: Performance of the *ALEPH* detector at *LEP*. Nuclear Instruments & Methods in Physics Research *A 360 (1995) 481-506*.
- [PP,94] Particle Data Group: Review of Particle Properties, Physical Review D *50/3, 1994*
- [Ru,95] Univ.Do. Dr. Gerald Rudolph: private communications
- [Sc,93] H.R.Schwarz: Numerische Mathematik. Teubner 1993.
- [Sc,94] M.Schmelling: The method of Reduced Cross-Entropy. A General Approach to Unfold Probability distributions. Nucl. Instrument. and Meth. Phys. Res. A *340 (1994)*.
- [Vo,95] R.Vogl: Hadronization and Azimuthal Angle Shift in 3-Jet Events at  $Z^0$ , An Alternative Approach to the String Effect, Dissertation, Innsbruck 1995

# Lebenslauf

- 20.2.1996 Geboren in Matri a./Br. als Sohn des Polizeibeamten Andreas Hörtnagl und dessen Frau Ingrid Hörtnagl geb. Steiner (Beruf: Hausfrau)
- 1972-1976 Besuch der Volksschule in Matri am Brenner
- 1976-1980 Besuch der Hauptschule in Matri am Brenner
- 1980-1985 Besuch der Höheren Technischen Bundes Lehr- und Versuchsanstalt in Innsbruck (Höhere Lehranstalt für elektrische Nachrichtentechnik und Elektronik)
- 20.6.1985 Reifeprüfung
- 1.10.1985-
- 30.9.1986 Präsenzdienst als Einjährig-Freiwilliger
- 1.10.1986 Beginn des Physik-Studiums mit Erweiterungsstudium Physik-Mathematik-Lehramt an der Leopold-Franzens-Universität Innsbruck;  
ab Herbst 1991 Diplomarbeit am Institut für Theoretische Physik
- 4.3.1993 Diplomprüfung
- 20.3.1993 Sponson zum Magister der Naturwissenschaften
- 1.4.1993-
- 31.8.1993 Akademikertraining am Institut für Experimentalphysik.  
Vorbereitende Arbeiten zu einer Dissertation
- 1.9.1993 Beginn des Unterrichtspraktikums (Physik: HTL II, höhere Abteilung für Tiefbau, Mathematik: HTL I, höhere Abteilung für Nachrichtentechnik und Elektronik)
- 1.7.1994 Anstellung als wissenschaftlicher Mitarbeiter am Institut für Experimentalphysik.
- 4.8.1995 Anstellung als Vertragsassistent am Institut für Experimentalphysik