ETH Institute for
Particle Physics

ETHZ-IPP Internal Report 2008-14
August 2008

Study of the $b \to \mu D^0 X$ Production Cross-Section at the CMS Experiment

DIPLOMA THESIS

presented by
Benjamin Stieger

under the supervision of
Prof. Dr. Urs Langenegger

ETH Zürich, Switzerland

Abstract

An analysis for the measurement of the differential cross-section of $b\bar{b}$-production at the CMS experiment is presented. Signal events are extracted by fitting a D^0 mass-peak from reconstruction of the decay channel $B \to \mu D^0 X$ with $D^0 \to K\pi$. Covering about 0.47 % of all B^+ and B^0 decays the analysis was developed on a data sample corresponding to only about $0.5\,\mathrm{pb}^{-1}$ and is thus feasible with startup data of the LHC collider. The B-production cross-section is measured in bins of $p_T(\mu D^0)$ and subsequently unfolded to bins of $p_T(B)$. The results are compared to the generated cross-sections.

Contents

1 Introduction

1.1 Motivation

The production of heavy flavor hadrons, that is hadrons containing a charm or bottom quark, will provide a substantial background to almost any measurement performed at a high energy hadron collider. With the startup of the LHC we will advance to a new level of available energy for the production of known and – hopefully – yet unknown particles. To quantify background spectra, measurements of their production cross-sections are needed. As current simulations are based on results from previous experiments at lower energies, a measurement of the B-meson production cross-section at 14 TeV will provide more accurate input parameters and help reduce systematic uncertainties.

This thesis presents a technique for gaining that knowledge in the startup phase of the LHC machine, such that it could be utilized in forthcoming analyses.

1.2 Heavy Flavor Production

Unlike the situation at an e^+e^--collider, there is a great variety of initial states involved in the particle production at a hadron machine. At the energies available to the LHC the constituents of the protons form a veritable soup of quarks and gluons, far beyond a uud valence-quark combination. The various channels for the production of heavy quarks can be classified in three general categories.

Gluon-gluon fusion / quark-antiquark fusion These are the leading order processes, where two constituents (either two gluons or a quark-antiquark pair) of the initial protons directly produce a heavy quark-antiquark pair. The corresponding Feynman diagrams are depicted in fig. 1 and 2. Characteristically the two final states involving the heavy quarks are observed 'back-to-back', i.e. with little to no combined transverse momentum.

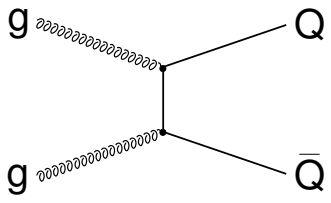

Figure 1: Gluon-gluon fusion

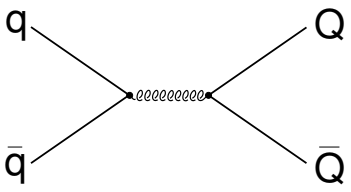

Figure 2: $q\overline{q}$ fusion

Flavor excitation In these 'next to leading order' processes, an off-shell heavy flavor quark constituent of the proton, necessarily originating from a $g \to Q\overline{Q}$ process within the initial hadron, is put on-shell by scattering with a gluon or quark of the opposing proton, see fig. 3. In this case the final states need not be back-to-back, as a third final state (from the scattered object) can carry away some transverse momentum.

Gluon splitting After scattering off a hard gluon or quark, a final state gluon can split up into a $Q\overline{Q}$ pair, see. fig 4. The resulting heavy flavored final states can carry a large combined transverse momentum and thus be concentrated within a comparatively small cone of angular separation.

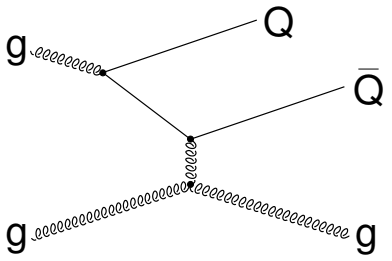

Figure 3: Flavor excitation

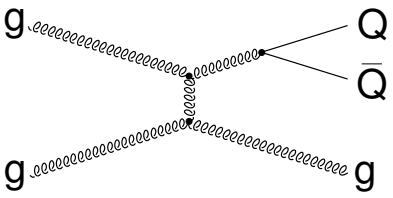

Figure 4: Gluon-splitting

There are of course many variations of these processes, e.g. involving gluon emission in initial or final states. The contributions of these three classifications to the total cross-section depend on the center-of-mass energy. At 14 TeV the prevalent mechanism for heavy-flavor production should be flavor-excitation, followed by pair creation and gluon splitting.[1]

In the analysis presented here we will focus on the production of strange-less b-mesons, i.e. combinations of a heavy b quark and either $\overline{u}$ or $\overline{d}$ partners. A produced $\overline{b}$-quark will fragment into a B_u (B^+) or a B_d (B^0) with a probability of roughly 40 % each, the remaining 20 % being shared by the production of B_s and b-baryons.[2]

1.3 Signal Channels

In order to reduce the amount of data necessary for a measurement of the B-production cross-section, several different B decay modes should be included in the analysis. Performing only a partial event reconstruction further extends the number of processes covered.

The signal channels chosen in this analysis are the semileptonic decays of (strange-less) B-mesons involving a D^0 charmed meson and a muon. Exhibiting a rather distinct signal in the detector, the muon is preferred over its other leptonic counterparts. The D^0 can be produced either directly in the decay of a charged B^- or indirectly over D^{*0} and D^{*-} intermittent states, resulting in turn from B^- and $\bar{B}^0$ decays. The combined branching fractions of charged and neutral B mesons to μD^0 make up roughly 12 % of all $B^\pm$ and B^0 decays. The different contributions are shown in table 1, the corresponding Feynman diagrams are displayed in figures 5 to 8.

Table 1: The three main contributions to a μD^0 signal.

	Decay	$\mathcal{B}$	Decay	$\mathcal{B}$
Direct	$B^- \to D^0 \mu^- \bar{\nu}_\mu$	2.15 %		
Indirect	$B^- \to D^*(2007)^0 \mu^- \bar{\nu}_\mu$	6.5 %	$D^*(2007)^0 \to D^0 \pi^0/\gamma$	100 %
Neutral	$B^0 \to D^*(2010)^- \mu^+ \nu_\mu$	5.35 %	$D^*(2010)^- \to D^0 \pi^-$	67.7 %
Total	$B^\pm/B^0 \to \mu D^0 X$	12.27 %		

As indicated, in order to maximize the effective branching fraction, no distinction is made between the different final states, and only the muon and D^0 decay products are reconstructed in this analysis. To be able to identify these events a suitable decay channel of the D^0 has to be chosen. Involving only two relatively stable particles and featuring a fairly large branching fraction of 3.80 %, the process $D^0 \to K^- \pi^+$ presents a prime candidate for the D^0 reconstruction. The

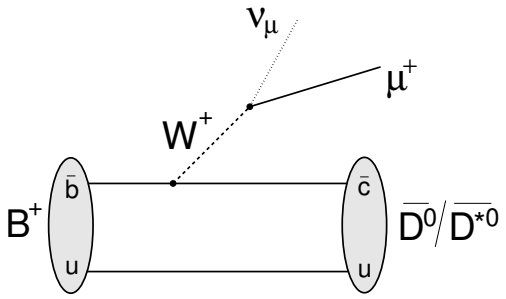

Figure 5: Feynman diagram for the processes $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$ and $B^+ \to \bar{D}^{*0} \mu^+ \nu_\mu$

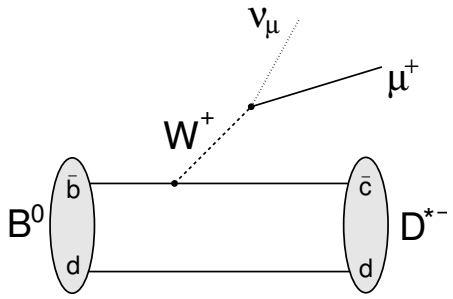

Figure 6: Feynman diagram for the process $B^0 \to D^{*-} \mu^+ \nu_\mu$

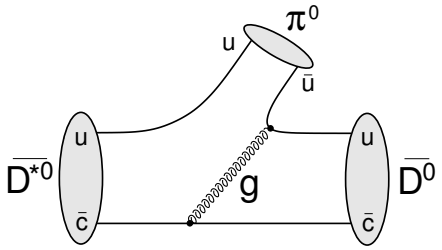

Figure 7: Feynman diagram for the processes $\bar{D}^{*0} \to \bar{D}^0 \pi^0$

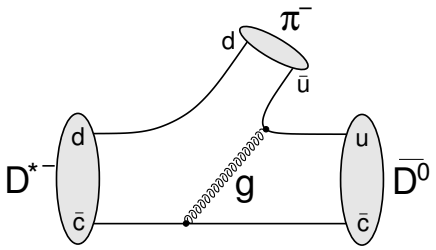

Figure 8: Feynman diagram for the process $D^{*-} \to \bar{D}^0 \pi^-$

corresponding Feynman diagram is depicted in figure 9. Including the branching fraction of the selected D^0 decay, the overall fraction of $B^\pm/B^0$ decays covered in this analysis is 0.47 %.

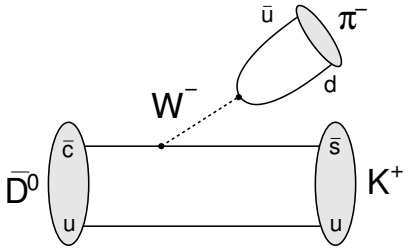

Figure 9: Feynman diagram for the process $\bar{D}^0 \to K^+ \pi^-$

Following the quark decay chain of $b \to c \to s$ and the intermediate emission of two oppositely charged W bosons, one can deduce that the sign of the charge of the muon from the B-decay and of the K from the D^0 decay have to coincide. The combination of oppositely charged leptons and kaons involves the decay $c \to d$ and the creation of a $\bar{s}u$ quark pair and is therefore doubly Cabibbo suppressed.

Possible backgrounds to these signals include the charm production of a D^0 with the semileptonic decay of a $\bar{D}^0$ from the associated $\bar{c}$, and the mixing and subsequent decay of a B^0 both providing a correct μD^0 combination.[1]

[1]See sec. 4.4 for a more detailed treatment of background effects.

1.4 The CMS Experiment

The Large Hadron Collider (LHC) at CERN is destined to become the new frontier of particle physics and cosmology when it starts operation later this year. The proton-proton accelerator, constructed in the tunnel of the old LEP machine, has a design center-of-mass energy of 14 TeV, about a factor 7 more than the currently most powerful machine. The beam is expected to reach an instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}s^{-1}$ or $10\,\mathrm{nb}^{-1}$ per second after proper calibration. At the beginning of data taking however the total rate of recorded data will more likely be in the range of $5\,\mathrm{pb}^{-1}$ per month.[2]

One of the two general purpose detectors built for the LHC is the Compact Muon Solenoid, or CMS, which is currently in the final stages of installation at LHC point 5 near Cessy, France. Its main purpose will be to search for and investigate yet unknown physical phenomena like the Higgs field and its relation to electroweak symmetry breaking or extensions of the standard model such as supersymmetry. Discovering the unknown is a challenge that requires versatility – consequently CMS was designed to be able to handle a large variety of different physical situations, without sacrificing precision regarding the key points of interest.

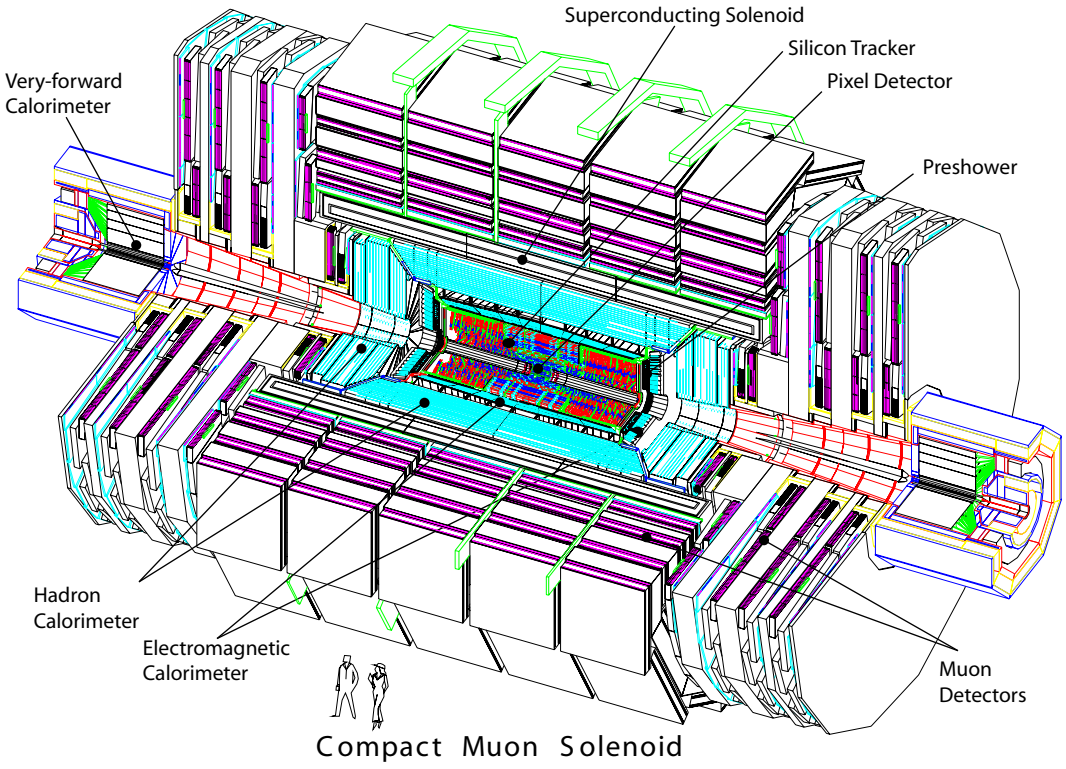

Figure 10: The CMS detector [3]

The detector is constructed around the largest superconducting magnet ever built (see fig. 10). This solenoid, 12.5 m long and roughly 6 m in diameter, produces a magnetic field of about 4 T

[2] At $2 \times 10^{30}\,\mathrm{cm}^{-2}s^{-1}$. As an illustration: this analysis was developed on a set of data that corresponds to less than $0.5\,\mathrm{pb}^{-1}$.

and gives CMS its name. The magnetic flux is returned by a massive iron yoke which defines the dimensions of the whole detector, standing at 21.5 m length and 15 m diameter while contributing the largest part of the overall mass of 12′500 t. The magnetic field outside the solenoid still measures about 2 T. Interleaving the iron return yoke are the muon chambers, consisting of drift tubes in the barrel region, cathode strip chambers in the endcaps, and fast resistive plate chambers for triggering, see figure 11. Other than the muon system, every component of the detector is

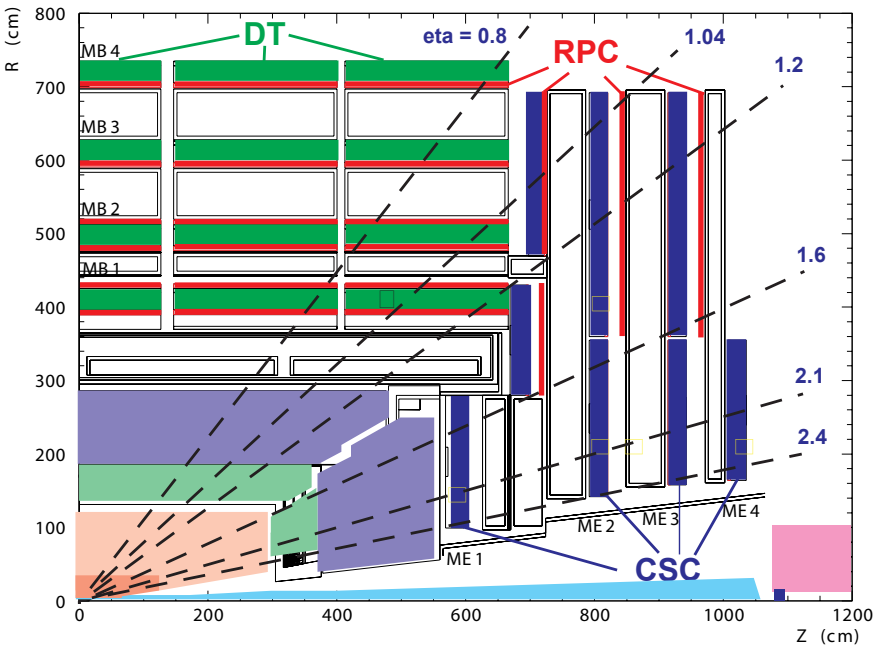

Figure 11: A longitudinal cut showing one quarter of the CMS detector, the different muon chambers are tagged: DT are drift tubes, CSC are cathode strip chambers, RPC are resistive plate chambers. Some η values are indicated, the beamline is along the bottom axis.

constructed inside the solenoid. Closest to the interaction point is the silicon pixel detector, which concentrates almost 50 million data taking elements in a volume of less than 0.2 m^3. Together with the surrounding silicon microstrip detector, and the muon chambers outside the magnet, this forms a powerful tracking system, capable of reconstructing the momenta of charged particles with a precision of below ten percent up to high transverse momenta.[3] Enclosing the tracker is an electromagnetic calorimeter, consisting of 80′000 oblong lead-tungstate ($PbWO_4$) crystals, facing the interaction point. The density and radiation length of these crystals are such that more than 99 % of the energy of electromagnetic showers is absorbed and measured on a length of only about 25 cm. The outermost and largest component situated inside the solenoid is the hadronic calorimeter. Consisting of layers of brass interleaved with plastic scintillators, it is a conventional sampling calorimeter with a depth of about 5 nuclear interaction lengths. Including the very forward calorimeter it covers the entire spatial angle down to about $|\eta| = 5$.[4]

[3]The resolution at low transverse momenta of about 10 GeV is below 1 %. At 100 GeV the resolution is still better than 10 %, depending on the η.

[4]The pseudo-rapidity η is defined as $-\ln\tan\theta/2$ and is therefore a measure for the angle θ between the track and the beamline. $\eta = 0$ corresponds to a 90° angle, $|\eta| \to \infty$ for $\theta \to 0°$. The CMS muon system covers an $|\eta|$-range of 0 to 2.4 or $\theta = 90°$ to $\sim 10.37°$, see fig. 11

Even at initial luminosity the LHC beam will produce upwards of 10^4 pp collisions per second; at design luminosity this number is pushed up to about 10^9 collisions per second, with more than 20 pp-interactions per bunch-crossing, corresponding to a data rate on the order of 10^3 Tbytes/s.[5] This outrageous amount of data has to be reduced to the maximum data acceptance rate of the CMS data acquisition system of about 100 Gbytes/s, and then passed on for further processing and selection. To achieve this rejection power while at the same time preserving the most interesting events, CMS uses a sophisticated system of trigger mechanisms. The first stage of selection (the L1 or level 1 trigger) utilizes the fastest components of the detector to quickly decide which events are to be passed on to the higher level triggers. There are various possible selection criteria applied to decide whether an event is considered worthy of recording, e.g. a high transverse momentum muon or missing energy in the calorimetry above a certain threshold value. Once accepted for processing by the L1 trigger, all the information from slower detector components is gathered and passed on to more complex selection algorithms in the high level trigger (HLT). After all selections CMS will produce about 10 Tbytes per day for offline processing. As might be expected the exact triggering thresholds will be adjusted to the instantaneous luminosity, such that the maximum amount of data is recorded at all times.

2 Monte Carlo Samples

2.1 Production

All Monte Carlo samples used in this study were produced using the CMS software (CMSSW), version 1.6.11. Unbiased pp collisions were generated by PYTHIA[4] (v.6.409), using the MSEL = 1 setting, the decays being governed by EvtGen[5] (v8.16). The detector reaction to these physical processes was reproduced using the fast parameterized simulation, FAMOS, avoiding a time-consuming full-detector simulation in GEANT.

2.2 'Data' Sample

A large number of unbiased events was produced to simulate real physics data. For that purpose the only preselection of events consisted of a generator-level filter requiring the existence of muons fulfilling certain transverse momenta criteria. An event was accepted if either one muon with $p_T > 4.5\,\text{GeV}$ or two muons each with $p_T > 2.5\,\text{GeV}$ were found among the generated particles. These thresholds were chosen to be slightly below the lowest L1 trigger cuts for single muon and dimuon respectively. Events that fulfilled these constraints were then passed on to the reconstruction modules, described in sec. 3.2.

About 2.46×10^{10} generated pp events were processed by this filter resulting in 9.04×10^6 accepted events, corresponding to a selection efficiency of about 3.68×10^{-4}. At a total pp-cross-section of 54.71 mb as assumed by PYTHIA this amount of data corresponds to an integrated luminosity of $0.4489\,\text{pb}^{-1}$. Recall that this corresponds to less than 6 days of data taking at an initial instantaneous luminosity of $2 \times 10^{30}\,\text{cm}^{-2}s^{-1}$. At design luminosity this amount of data will be gathered in less than one minute.[6]

2.3 Signal Sample

For determination of the reconstruction efficiency and extraction of the generated B-production cross-section an extensive sample containing only signal events was produced as well. Apart from

[5] Assuming about 1Mbyte of data per interaction.
[6] $10^{34}\,\text{cm}^{-2}s^{-1} = 10\,\text{nb}^{-1}$ per second

using a different filter the setup was identical to the one used for the 'data'-sample. The signal filter required a muon emanating directly from a B-meson (either B^0 or $B^{\pm}$), and a D^0 coming directly or indirectly from the same B-meson. The sample therefore only contains events in which a process $B \to \mu D^0 X$ had been generated. However the D^0 decay was not enforced, to be able to extract a generator-level cross-section in bins of $p_T(\mu D^0)$. Another advantage of this circumstance is that the signal sample still contains 'background' to our reconstructed channels, thereby providing a realistic environment for the fitting algorithms used to determine the overall efficiency. Such as it is, the exact same fitting procedures as in the 'data' sample can be performed on the signal sample, which would not be possible had it only contained $B \to \mu K \pi X$ events. No cut whatsoever on kinematical properties was applied.

This filter was applied to about 3.04×10^{10} generated pp events and produced 3.81×10^7 signal events, corresponding to an efficiency of 1.25×10^{-3}. The efficiency reflects the probability of producing a B-meson (roughly 1 %) with subsequent decay into μD^0 (the above mentioned 12.27 %).

3 Event Reconstruction

3.1 Signal Signature

To find a signal event in a set of data, one has to study the properties of its decay. We are looking for a muon from the B-decay with two additional charged tracks from a D^0-decay (see fig. 12). The two hadronic tracks should have roughly the same direction as the muon and their invariant mass should be compatible with the rest-mass of a D^0-meson. Furthermore the track with the same charge as the muon can be assumed to be the K (see sec. 1.3).

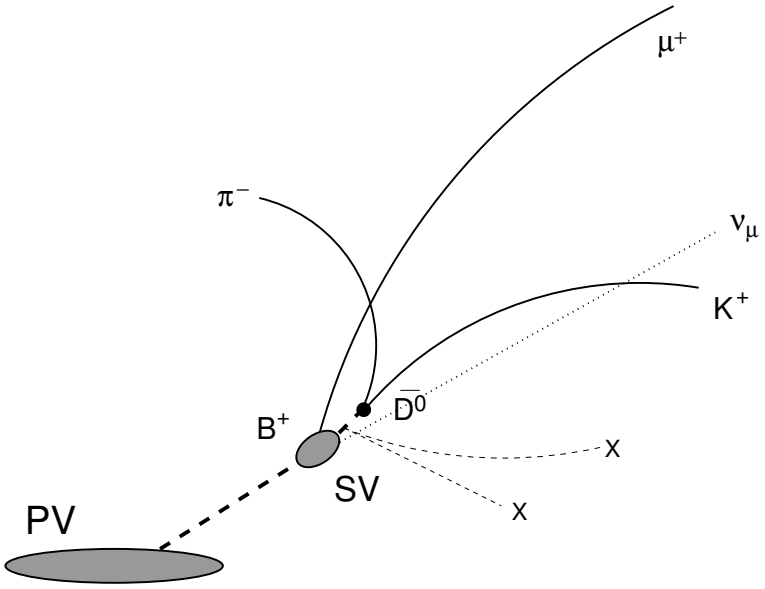

Figure 12: Decay topology of $B^+ \to \mu^+ \bar{D}^0 X$ with $\bar{D}^0 \to K^+ \pi^-$

3.2 D^0 Reconstruction

To reconstruct the D^0 from the semileptonic decay of a B-meson, we start with the muon. For each muon track (as identified by the muon system) in a given event we first look for a nearby track having the same charge. Assuming this is the K we search for another nearby track of the opposite charge and assume it to be the π. Given that the alleged K and π tracks fulfill some loose requirements on transverse momenta and angular separation we build a vertex from their tracks. Again we apply a loose cut on the quality of this vertex fit, and store all relevant information in a D^0-candidate data-structure. Note that there can be several such candidates per event.

The exact requirements for building a D^0-candidate were the following[7]:

- Muon track:
  - $p_T(\mu) > 3.0\,\mathrm{GeV}$
- Kaon track:
  - $p_T(K) > 1.0\,\mathrm{GeV}$
  - $Q(K) = Q(\mu)$
  - $\Delta R(\mu, K) < 1.5$
- Pion track:
  - $p_T(\pi) > 1.0\,\mathrm{GeV}$
  - $Q(\pi) \neq Q(K)$
  - $\Delta R(\mu, \pi) < 1.5$
- K/π system:
  - $m_{inv}(K, \pi)$ between 1.6 and 2.1 GeV
  - χ^2/N_{dof} of vertex fit < 10

The information stored in the D^0-candidate structure were the charges, three-momenta and indices of each of the μ, K and π tracks and the mass, three-momentum and vertex-properties of the D^0. Before storing the tracks they were refitted using the D^0-vertex as an additional constraint. Additionally for each D^0-candidate a μD^0-candidate was built holding the mass, three-momentum and vertex-properties of the μD^0-system. For building the vertices a simple Kalman vertex-fitter algorithm was utilized.

3.3 Candidate Selection

As previously mentioned the D^0-reconstruction allowed for multiple candidates per event. Since the probability for the occurrence of two simultaneous signal events is of order $(0.47\,\%)^2 \simeq 0.002\,\%$, these multiple signal events were discarded and only one candidate per event was processed in the analysis. This makes the selection of a valid candidate for each event an important issue to reduce background. To investigate this task and find a suitable algorithm for selection, all the candidates were examined on a generator level and tagged as 'signal' or 'background', 'signal' corresponding to the case where the μ, K and π tracks came from a $B \to \mu D^0 X$ event. The comparison of some physical properties for good and bad candidates are displayed in figures 13 - 26.

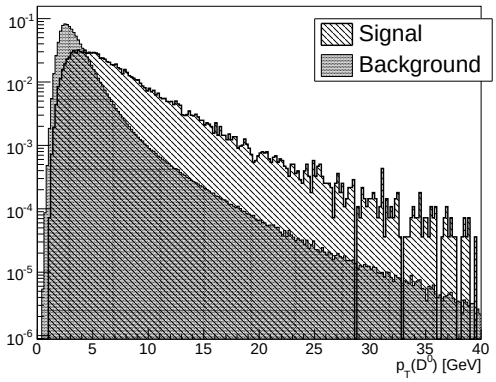

Figure 13: $p_T(D^0)$

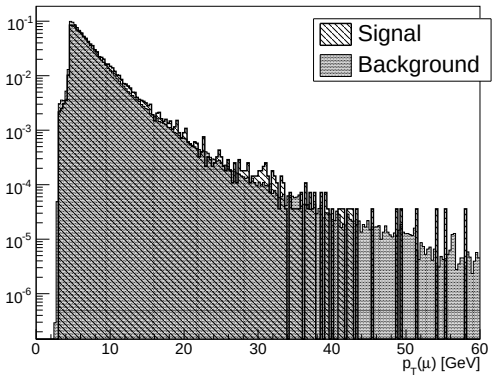

Figure 14: $p_T(\mu)$

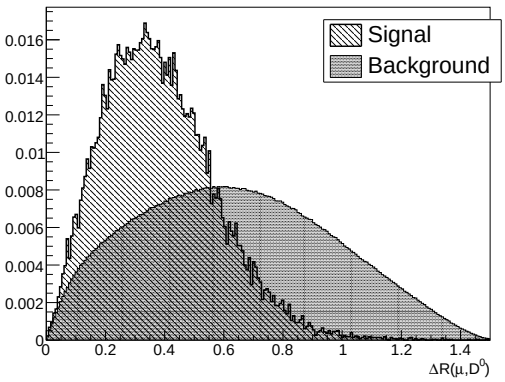

Figure 15: $\Delta R(\mu, D^0)$

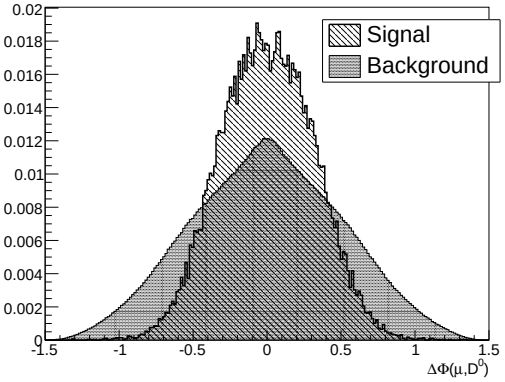

Figure 16: $\Delta\Phi(\mu, D^0)$

The final selection algorithm was based on a cut on the angular separation between the μ and K tracks and subsequent maximization of the transverse momentum of the D^0. Additionally candidates in which the μD^0-system had an invariant mass larger than the B-meson rest mass were discarded. To summarize, a candidate was preferred over another given that:

- the angular separation of its μ and K tracks was smaller than 1.0,
- the invariant mass of its μD^0 system was smaller than 5.0 GeV,
- the combined transverse momentum of its K and π tracks was greater than that of the other candidates.

Using this recipe, a candidate selection purity of 88.66 % was achieved, meaning out of 100 events in which a good candidate was present, 89 times the right one was selected.

[7] where Q is the sign of the charge of a track and ΔR is defined as $\sqrt{\Delta\eta^2 + \Delta\phi^2}$

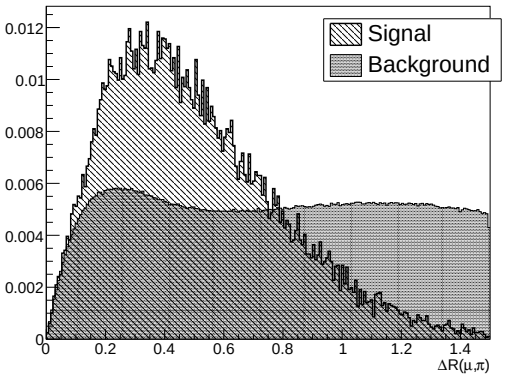

Figure 17: $\Delta R(\mu, \pi)$

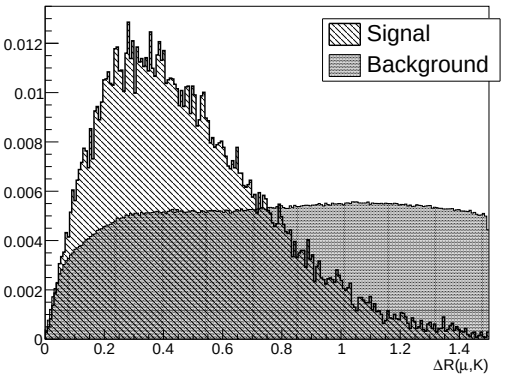

Figure 18: $\Delta R(\mu, K)$

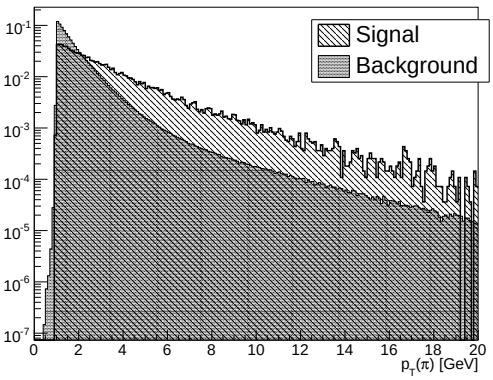

Figure 19: $p_T(\pi)$

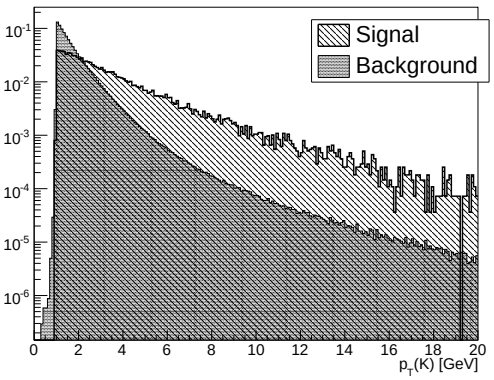

Figure 20: $p_T(K)$

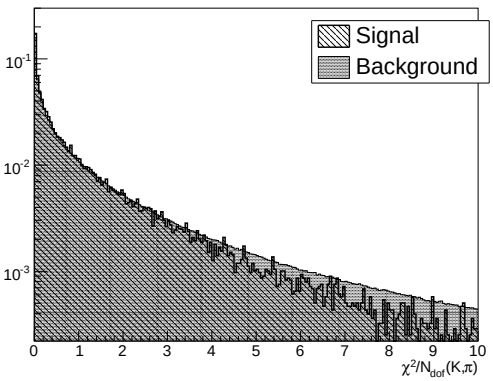

Figure 21: χ^2/N_{dof} for D^0 vertex

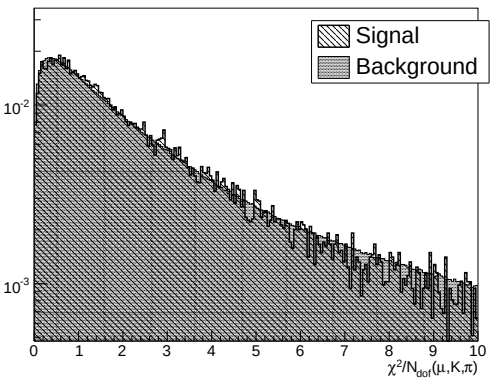

Figure 22: χ^2/N_{dof} for μD^0 vertex

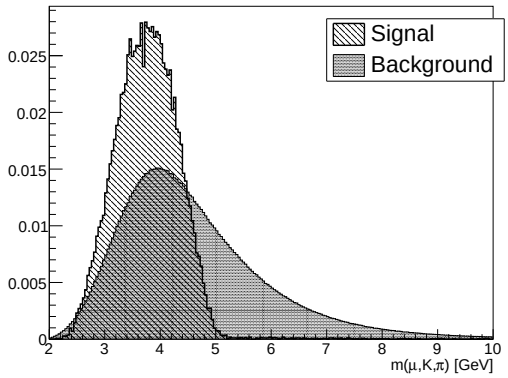

Figure 23: $m(\mu, D^0)$

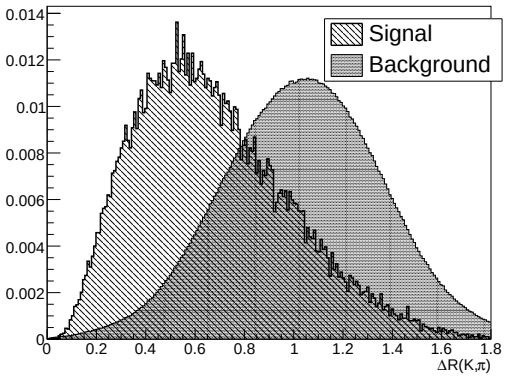

Figure 24: $\Delta R(K, \pi)$

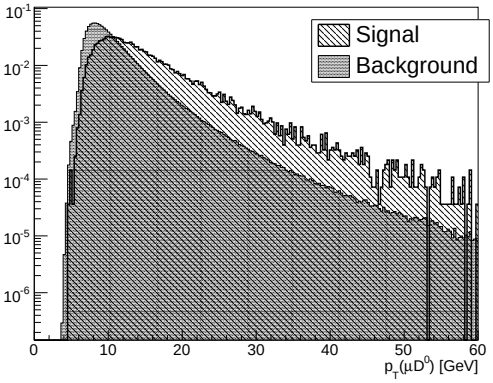

Figure 25: $p_T(\mu D^0)$

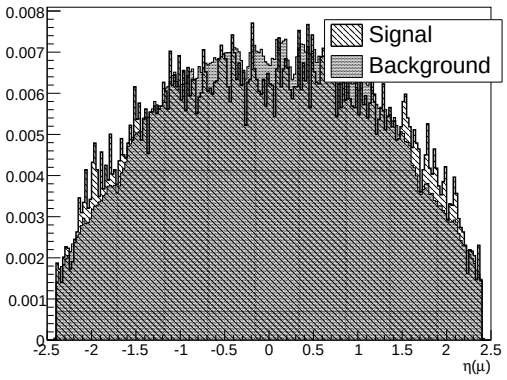

Figure 26: $\eta(\mu)$

3.4 Fitting Signal Peaks

After selecting a candidate for each event, and dumping its relevant data to a ROOT tree, one is faced with the task of deciding which of the candidates were actual signal candidates. The method applied in this analysis is to cast the masses of these alleged D^0 candidates into a histogram, and to determine the number of events in the peak surrounding the D^0 rest mass. Applying only the single muon trigger at 5 GeV, as used for the entire analysis, this histogram (produced on the 'data' sample) is depicted in figure 27. The histogram was chosen to include 60 bins between 1.7 and 2.0 GeV, implying a bin-width of 5 MeV. The vast background below the signal peak stems from events in which there was no signal to begin with.[8]

In the following, three different approaches are presented on how to determine the number of events in the peak, all including some form of fitting. What all methods have in common is the assumption of an exponential form for the background function. The choice of an exponential form, rather than polynomial, was made on account of a slight curvature in the background shape that appears when widening the mass window to $1.6 - 2.1$ GeV. However in the narrower mass

[8]Recall that if there had been a signal candidate it would have been selected in 89 % of cases.

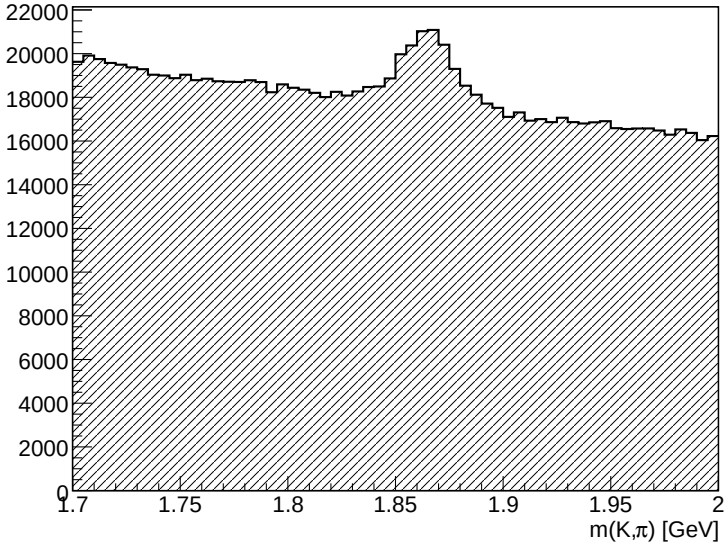

Figure 27: Mass of all D^0 candidates, peaking at the D^0 rest-mass of about 1.865 GeV.

window of $1.7-2.0$ GeV that was finally used, the difference resulting from the two different shapes of background is insignificant. The simplest way – and first approach – is to quantify the signal component as a single gaussian peak. However, the mass resolution of the detector and therefore also the width of this gaussian depend on the η of the tracks, resulting in a somewhat poor fit quality. The two subsequent methods presented take account of this fact by utilizing a superposition of several gaussian peaks and an additional binning in $\eta(\mu D^0)$ respectively. In all cases the number of signal events is extracted by integrating the signal component of the fit-function in the range 1.84 - 1.89 GeV.

3.4.1 Gaussian Fitting

The most straightforward attempt to quantify the number of signal events is made using the following function for a fit:

$$f_{fit}(x) = f_{backgr.}(x) + f_{signal}(x) = p_0 \exp(p_1 x) + p_2 \exp\left(-\frac{(x-p_3)^2}{2\,p_4^2}\right),$$

where $p_0, \ldots p_4$ are the fit parameters. It consists of an exponential part describing the background and a gaussian signal peak. An example fit using this function is shown in figure 28. To ensure the stability of the fitting procedure, i.e. that the fit always produces meaningful results, the parameters were confined to suitable ranges and initialized at estimated starting values:

- p_0 and p_1 (Scale and decay constant of exponential background)
  - The scale and decay constant can be estimated from the first and last five bins of the histogram, where the contribution from a signal component is to be neglected. These estimations were used as starting values for the fit algorithm.
  - The scale was additionally restricted to be positive.

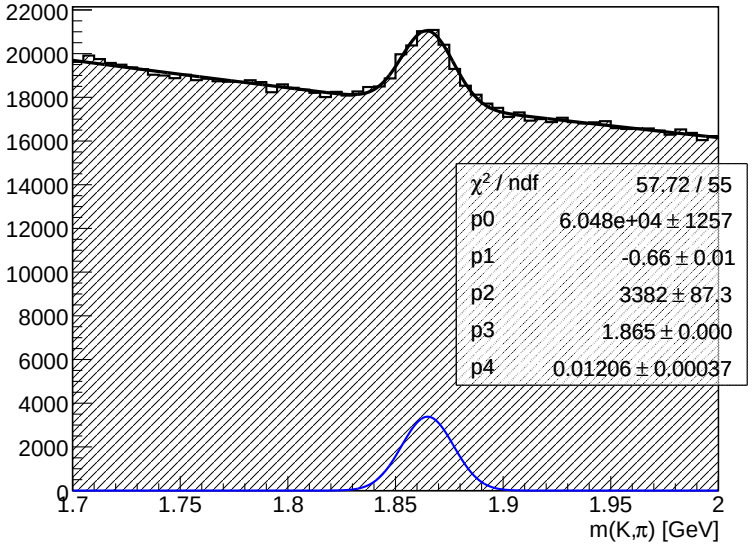

Figure 28: Example fit using a single gaussian peak over exponential background

- p_2 (Peak height of gaussian)
  - The starting value of the signal height was set to be the maximum entry of the histogram minus the value of the background shape (calculated with the starting values from above) at $x = 1.87\,\text{GeV}$.
  - The gaussian height was also constrained to positive values.
- p_3 (Mean of gaussian)
  - Naturally the initial value of the signal mean was chosen to be roughly the D^0 rest-mass of $1.87\,\text{GeV}$.[9]
  - The range of the mean was restricted to be between $1.84\,\text{GeV}$ and $1.89\,\text{GeV}$
- p_4 (Width of gaussian)
  - The signal width starting value was chosen to be $30\,\text{MeV}$.
  - To reject fitting of statistically insignificant fluctuations, the width was constrained to values greater than $5\,\text{MeV}$.

The problem with this rather naive approach shows when comparing the resulting number of events from integrating the gaussian part of the fitting function, with a reference value obtained by counting the signal events on a generator level in the same mass range. The fitting values are systematically too low by about 10 %.[10] Plotting only the candidates tagged as signal, i.e. contributing to the peak (fig. 29 upper left) and fitting them shows that a single gaussian peak is an inadequate approximation to the peak, see figure 29 upper right. The underlying reason appears

[9] Recall the precise value of the D^0 rest-mass of $1.8645\,\text{GeV}$

[10] When integrating over the entire mass range of $1.7 - 2.0\,\text{GeV}$. This number decreases to about $2 - 3\,\%$ when narrowing the signal mass window to the above mentioned $1.84 - 1.89\,\text{GeV}$.

to be the fact that the mass resolution of the detector and thereby the width of the gaussian signal depends on the pseudo-rapidity η of the tracks used. The problem at hand is therefore that the signal peak does not consist of a single gaussian, but in fact various gaussians with various widths. Two different techniques taking this circumstance into account are presented in the next two sections.

3.4.2 Triple Gaussian Fitting

One possibility to account for multiple widths is to use a superposition of different gaussians as a fit-function. We begin by adding one additional gaussian peak to the previous attempt, so the new (signal) fitting function is:

$$f_{signal}(x) = p_2 \exp\left(-\frac{(x-p_3)^2}{2\,p_4^2}\right) + p_5 \exp\left(-\frac{(x-p_6)^2}{2\,p_7^2}\right),$$

where we now have 6 fitting parameters (p_0 and p_1 are reserved for the exponential background). Using this function we again perform our example fit and compare the result to the previous version, see figure 29 lower left. As can be seen from the figure, the value resulting from a double gaussian fit, as well as the χ^2/N_{dof} of the fit, are, even though improved compared to the single gaussian, still unsatisfactory. The next step is to add another gaussian function to the fit. The performance of this triple gaussian fit on the same example histogram is shown in figure 29, bottom right panel. The deviation from the generator level counting (top left panel), is now on the order of 1 %. On a purely superficial comparison the fit is clearly better than the one using the double gaussian, which is also reflected in a χ^2/N_{dof} value below one. One could now proceed to add a fourth gaussian function, however in the interest of fit stability, and seeing how more than 11 parameters are on the verge of over-parameterizing the fit, we draw the line at 3 gaussians. The fit function used for the analysis then takes on the following form:

$$f_{fit}(x) = p_0 \exp(p_1 x) + p_2 \exp\left(-\frac{(x-p_3)^2}{2\,p_4^2}\right) + p_5 \exp\left(-\frac{(x-p_6)^2}{2\,p_7^2}\right) + p_8 \exp\left(-\frac{(x-p_9)^2}{2\,p_{10}^2}\right),$$

with 11 parameters. The parameters were again constrained to meaningful values:

- p_0 and p_1 (Scale and decay constant of exponential background)
  - For the background parameters, the same estimation as mentioned in the previous section were used, see sec. 3.4.1
- p_2, p_5, p_8 (Peak heights of gaussians)
  - The starting values of the signal heights was set to be one third of the peak value of the histogram minus the value of the background shape at $x = 1.87\,\text{GeV}$.
  - The heights were restricted to positive values.
- p_3, p_6, p_9 (Means of gaussians)
  - The initial values of all three signal means were chosen to be roughly the D^0 rest-mass of 1.87 GeV
  - The ranges of the means were closely constrained to be between 1.86 GeV and 1.87 GeV, as more lenient restrictions invariably lead to the occurrence of grossly misshapen fit results.

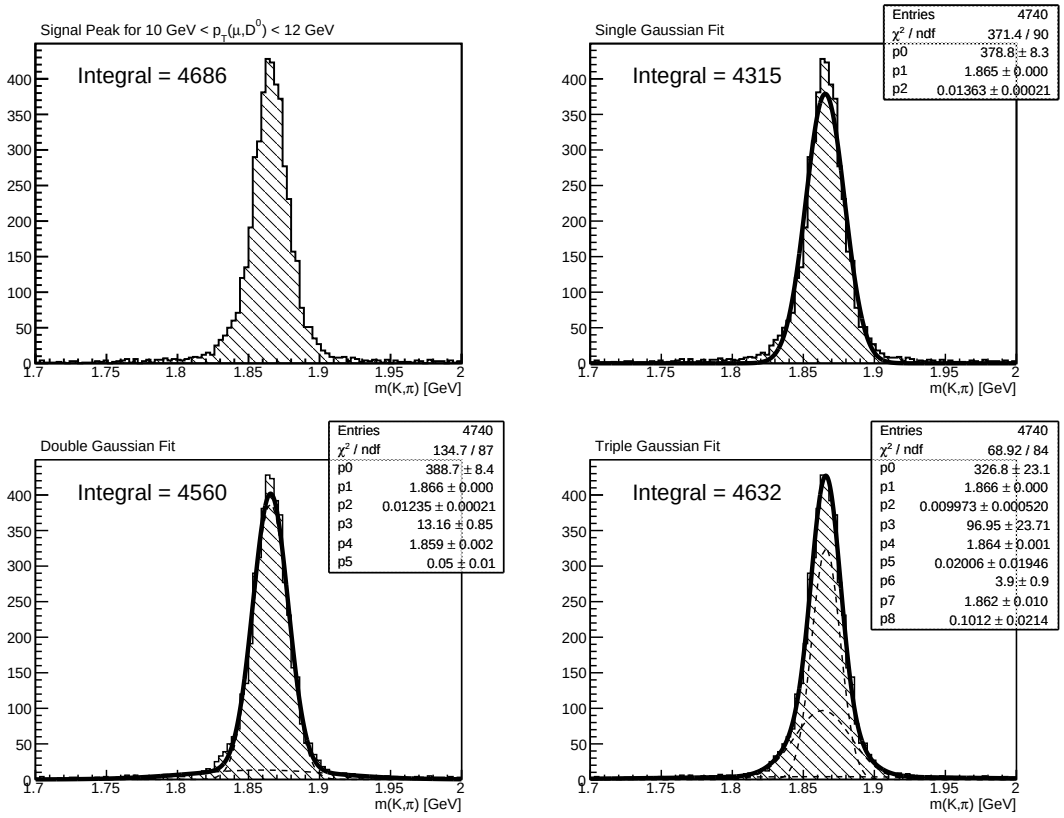

Figure 29: Comparison of different fitting functions for an example bin. Top left is the initial histogram, top right is fitting with a single gaussian, bottom left is a double gaussian fit, bottom right is a triple gaussian fit. The corresponding integral results are given for each panel.

- p_4, p_7, p_{10} (Widths of gaussians)
  - To force the fitting algorithms into using three distinct gaussian peaks, the allowed values for the widths were chosen to be disjoint.
  - The three initial values used were 7 MeV, 20 MeV and 50 MeV
  - The first width was constrained to the range of $5 - 10$ MeV, the second to $10 - 30$ MeV and the third to $30 - 60$ MeV.

An example fit using this function with the indicated constraints on the parameters is shown in figure 30. As can be seen from the resulting parameters, only two of the gaussian signals are actually put to use, the third being suppressed to a vanishing scale (p_8). This is apparently a result of the interference of the background on the fitting procedure. For this reason the improvement of the resulting number of signal events as observed for the signal-only fits, is substantially lessened, though still present when including the background.

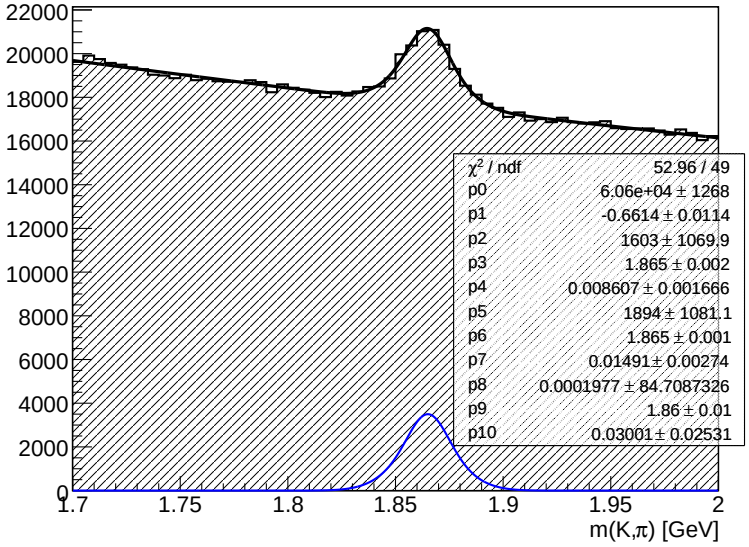

Figure 30: Example fit using a triple gaussian signal peak over an exponential background.

3.4.3 Eta Binning and Single Gaussian Fitting

The natural way of handling a dependence on η is to include an additional[11] binning in $\eta(\mu D^0)$, the downside of course being a reduction of available data for each bin. The histograms so produced are then fitted using a single gaussian with exponential background, as described above. Three bins of $\eta(\mu D^0)$ were chosen between $\eta = 0$ and 2.5 (compare figure 11).

Bin	1	2	3
η-Range	0.0-0.5	0.5-1.0	1.0-2.5
Bin width	0.5	0.5	1.5

To be able to compare the η-binned results to the previous attempts, the bins are summed over after the fitting procedure. As the measurement of an η-binned production cross-section is in itself an interesting result, they are also quoted in the next section. This additional binning now allows for the observation of the above mentioned effect of the η-dependence of the mass-resolution of the detector. We plot the average fit results[12] for the width of the signal gaussian (p_4) for each of the three $\eta(\mu D^0)$ bins, see figure 31.

3.4.4 Consistency Check and Error Estimation

To investigate whether the fitting algorithms and the shape of the implemented fitting functions introduced a bias, and to estimate the uncertainty of the integral results, a toy simulation was performed. Using the fitting result to define a probability density function for signal and background components, a new histogram was filled. The number of simulated signal and background

[11] See sec. 4.1 for a description of the chosen p_T binning.

[12] Averaged over the $p_T(\mu D^0)$ binning, described in 4.1. The value from the last bin was omitted, as it was dominated by the uncertainty, owing to poor statistics. The average runs therefore over $p_T(\mu D^0)$ between 5 and 45 GeV.

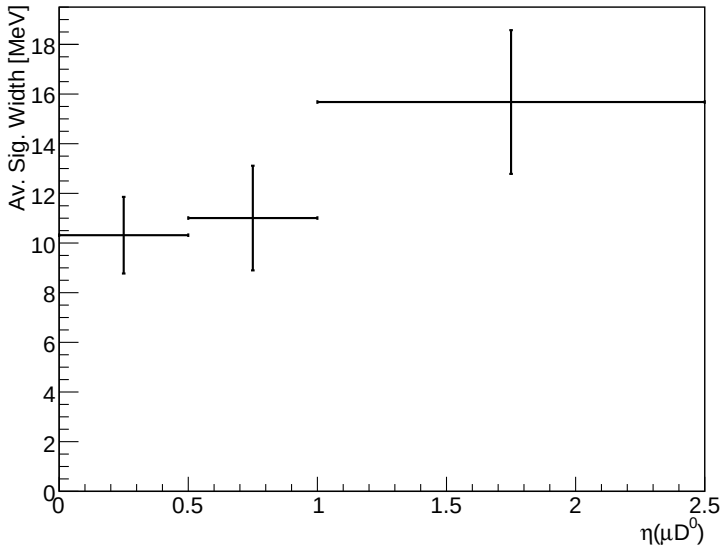

Figure 31: Fit result for p_4 (width of signal gaussian), averaged over all but last $p_T(\mu D^0)$ bin, i.e. $p_T(\mu D^0)$ values from 5-45 GeV. Note the increasing signal width with increasing $\eta(\mu D^0)$, caused by a deterioration of the mass-resolution of the detector towards high $|\eta|$-values. See appendix for all the corresponding fit plots.

events was taken to be the integral (over the entire mass window) of the signal and background parts of the original fit result, thereby producing a simulation of a new measurement. This simulated histogram was then fitted using the identical fitting algorithm as used for the original result, resulting in a new value for the number of signal events,[13] to be compared with the initial result. This procedure was repeated a number of times, the various values being saved in a histogram. The standard deviation of these values was then used as the statistical uncertainty of the fitting procedure.

If the fitting shape introduced a systematical bias, e.g. reproducing too few events, the results of the fitting on these toy simulations should display a systematical shift towards lower values. For both the single gaussian and triple gaussian signal shapes, no such shift could be detected in the simulation results. The observed tendency to produce lower values from fitting than from truth counting, can therefore either be attributed to tailing signal peaks, caused by detector effects and not picked up by the fitting procedure, or simply to statistical fluctuations.

Entering into both the measurement of the number of signal events and the efficiency determination as described further on, a systematical bias in the fitting procedure will however not show in the cross-section results, as a deviation in the number of signal events would cancel with an equal deviation in the efficiency.[14]

See figure 32 for the output of such a toy simulation, performed using the single gaussian fitting procedure on the 'data' sample. For each $p_T(\mu D^0)$ bin (see sec. 4.1), the mean values of the fits and their RMS are printed.

[13] Recall that the number of signal events is extracted by integrating in the range of $1.84 - 1.89$ GeV, rather than the entire mass window.

[14] Assuming the bias is identical for data and Monte Carlo.

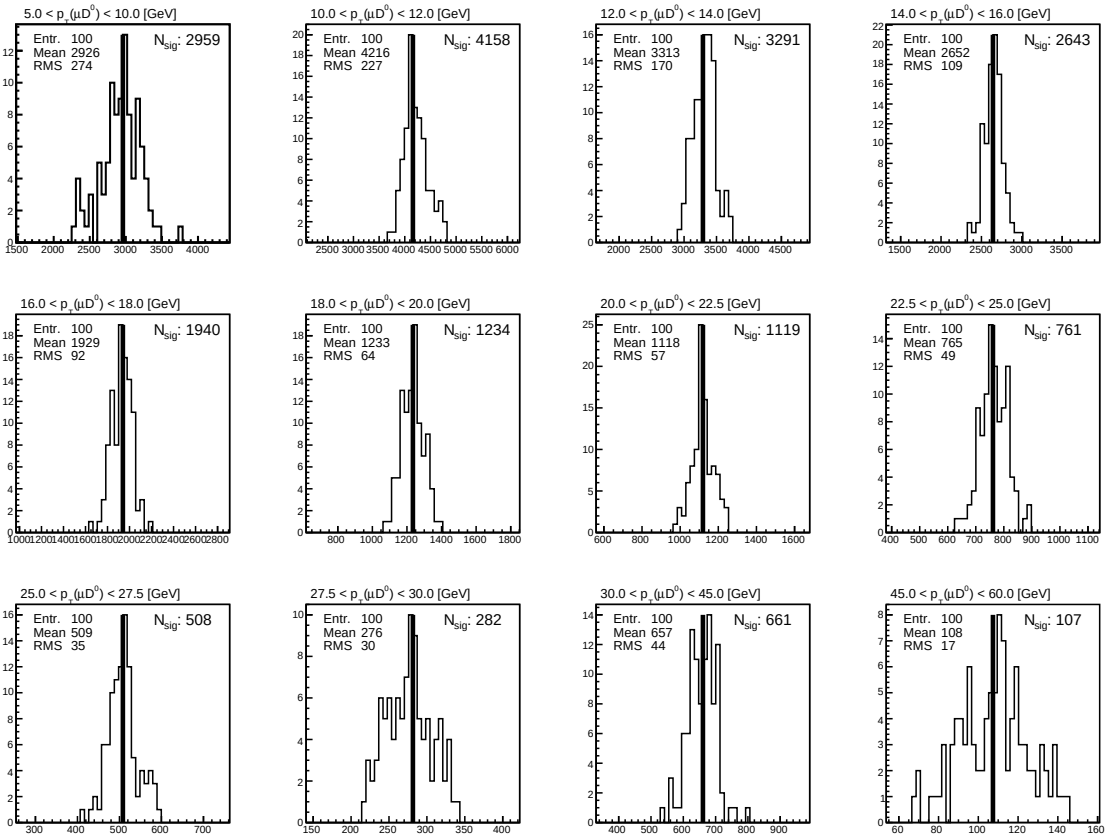

Figure 32: Results of toy fittings for each of the $p_T(\mu D^0)$ bins, for single gaussian fitting on the 'data' sample. 100 simulations are performed for each bin. The solid vertical line depicts the number of signal events from the original 'data'-fit, its value is printed in the top right corner of each pad.

4 Measurement of Differential Cross-Section

4.1 Transverse Momentum Binning

To arrive at a differential cross section one has to cast the D^0 candidates in some form of binning, and determine the number of signal events per bin. Since we reconstruct the μD^0 system in this analysis, a binning in $p_T(\mu D^0)$ was chosen.[15] Because of the overall shape of the $p_T(B)$ spectrum and the reconstruction performance of the detector for small p_T, the number of D^0 candidates per bin decreases towards higher and lower transverse momenta, respectively. To account for this changing amount of available statistics per bin a variable bin width was chosen. The following table displays the range and bin width for each bin:

[15] See section 5 for a means of estimating the cross-section in bins of $p_T(B)$.

Bin	1	2	3	4	5	6
Range / GeV	5-10	10-12	12-14	14-16	16-18	18-20
Bin width / GeV	5	2	2	2	2	2
Bin	7	8	9	10	11	12
Range / GeV	20-22.5	22.5-25	25-27.5	27.5-30	30-45	45-60
Bin width / GeV	2.5	2.5	2.5	2.5	15	15

4.2 Fitting Results

Using the three fitting methods described in sec. 3.4, we extract the number of signal events in each $p_T(\mu D^0)$ bin for both the 'data' and signal samples. The resulting values for N_{sig} from each different fitting algorithm and the reference value from truth counting (also integrated in the signal mass-window of 1.84 - 1.89 GeV) are printed in tables 2 and 3, the fits themselves are displayed in the appendix, page 42ff. The quoted errors are the statistical uncertainties, determined using the toy simulations described in sec. 3.4.4.

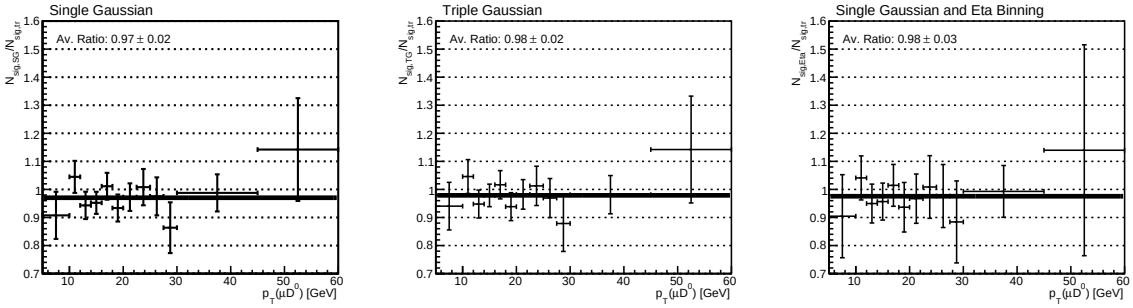

Figure 33: Ratio of reconstructed signal events to generated signal events (from truth counting). Results for the 'data' sample.

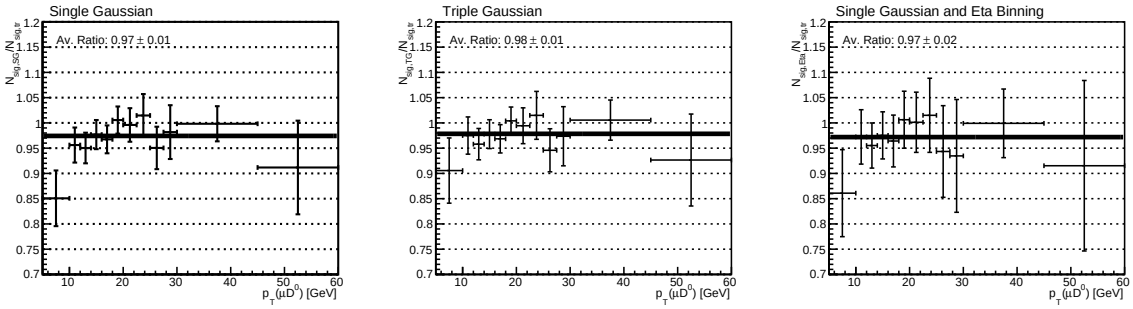

Figure 34: Ratio of reconstructed signal events to generated signal events (from truth counting). Results for the signal sample.

As already mentioned the ratio of reconstructed events to generated events from truth counting

Table 2: Number of signal events for each $p_T(\mu D^0)$ and $\eta(\mu D^0)$ bin, reconstructed using the different fitting methods as described in sec 3.4, performed on the 'data' sample. The errors are estimated by performing 100 toy simulations, as described in the text.

Bin	1	2	3	4
	5-10 GeV	10-12 GeV	12-14 GeV	14-16 GeV
Truth Counting	3260	3980	3489	2775
Single Gaussian	2959 ± 274	4158 ± 227	3291 ± 170	2643 ± 109
Triple Gaussian	3066 ± 276	4163 ± 240	3307 ± 174	2716 ± 111
$\eta(\mu D^0)$-summed	2949 ± 482	4143 ± 312	3313 ± 241	2655 ± 183
η bin 1	1013 ± 198	989 ± 171	990 ± 123	761 ± 100
η bin 2	813 ± 284	1164 ± 149	857 ± 128	645 ± 91
η bin 3	1123 ± 336	1990 ± 215	1467 ± 164	1248 ± 124
Bin	5	6	7	8
	16-18 GeV	18-20 GeV	20-22.5 GeV	22.5-25 GeV
Truth Counting	1918	1322	1151	755
Single Gaussian	1940 ± 92	1234 ± 64	1119 ± 57	761.2 ± 48.8
Triple Gaussian	1950 ± 97	1241 ± 66	1131 ± 61	764.5 ± 52.9
$\eta(\mu D^0)$-summed	1946 ± 143	1238 ± 117	1113 ± 101	762 ± 84
η bin 1	539 ± 74	383 ± 59	254 ± 54	196 ± 41
η bin 2	478 ± 71	345 ± 63	300 ± 54	191 ± 51
η bin 3	929 ± 99	510 ± 79	560 ± 66	375 ± 53
Bin	9	10	11	12
	25-27.5 GeV	27.5-30 GeV	30-45 GeV	45-60 GeV
Truth Counting	521	326	669	94.0
Single Gaussian	508.1 ± 35.3	282 ± 30	660.7 ± 44.3	107 ± 17
Triple Gaussian	505.0 ± 36.2	286 ± 32	656.5 ± 45.6	107 ± 18
$\eta(\mu D^0)$-summed	509 ± 58	288 ± 48	664.3 ± 61.8	107 ± 35
η bin 1	181 ± 32	88.1 ± 25.0	202 ± 33	34.0 ± 12.8
η bin 2	142 ± 31	96.2 ± 25.2	170 ± 32	32.6 ± 10.0
η bin 3	186 ± 38	104 ± 32	292 ± 41	40.6 ± 31.4

shows a slight tendency towards underestimating the number of events in the peaks, see figures 33 and 34.

4.3 Efficiency Determination

To find the number of produced B-events[16] from the number of reconstructed signal events, one needs the efficiency of the reconstruction mechanism. This includes a wide range of effects such as the physical probability for a B-meson to decay into our signal channel (the branching fraction) or the kinematic possibility for the event to be recorded in the detector (the acceptance). Ideally one would like to determine as much as possible from the data itself, such as to reduce systematic uncertainties from hypotheses put into Monte Carlo studies. However, since at this point no data is available yet, one can only try to find ways to measure detector efficiencies using other simulations.

[16]By 'B-event' we refer to an event containing a $B^{\pm}$ or B^0; i.e. the probability for a b-quark to fragment into one of these is already included.

Table 3: Number of signal events for each $p_T(\mu D^0)$ and $\eta(\mu D^0)$ bin, reconstructed using the different fitting methods as described in sec 3.4, performed on the signal sample. The errors are estimated by performing 200 toy simulations, as described in the text.

Bin	1	2	3	4
	5-10 GeV	10-12 GeV	12-14 GeV	14-16 GeV
Truth Counting	3568	4455	4053	2985
Single Gaussian	3035 ± 197	4260 ± 155	3852 ± 123	2916 ± 86
Triple Gaussian	3231 ± 230	4344 ± 164	3883 ± 126	2919 ± 85
$\eta(\mu D^0)$-summed	3072 ± 307	4331 ± 240	3872 ± 181	2911 ± 139
η bin 1	882 ± 135	1222 ± 111	1056 ± 93	822.9 ± 73.0
η bin 2	818 ± 132	1041 ± 118	1009 ± 92	727 ± 74
η bin 3	1372 ± 242	2068 ± 177	1806 ± 125	1361 ± 93
Bin	5	6	7	8
	16-18 GeV	18-20 GeV	20-22.5 GeV	22.5-25 GeV
Truth Counting	2164	1509	1223	777
Single Gaussian	2094 ± 60	1518 ± 41	1218 ± 41	788.4 ± 33.0
Triple Gaussian	2096 ± 61	1515 ± 41	1216 ± 43	788.6 ± 36.9
$\eta(\mu D^0)$-summed	2086 ± 111	1519 ± 85	1225 ± 73	788.7 ± 57.0
η bin 1	628.6 ± 62.6	419 ± 45	373 ± 40	223 ± 31
η bin 2	537 ± 56	419 ± 47	338 ± 38	210 ± 33
η bin 3	920.7 ± 72.5	680.5 ± 55.6	514.0 ± 47.6	355.4 ± 35.0
Bin	9	10	11	12
	25-27.5 GeV	27.5-30 GeV	30-45 GeV	45-60 GeV
Truth Counting	510	324	725	114
Single Gaussian	484.7 ± 21.4	318.1 ± 17.3	723.8 ± 25.2	104 ± 11
Triple Gaussian	482.4 ± 21.7	315.4 ± 19.0	729.0 ± 28.8	105.6 ± 10.4
$\eta(\mu D^0)$-summed	481.1 ± 46.3	303 ± 36	724.4 ± 49.2	104 ± 19
η bin 1	138 ± 27	67.5 ± 22.3	221 ± 29	40.6 ± 10.1
η bin 2	120 ± 21	98.1 ± 18.7	171 ± 25	36.8 ± 10.8
η bin 3	223 ± 31	137 ± 22	332.7 ± 31.5	27.0 ± 12.3

4.3.1 Overview

We attempt to give a complete list of effects that act on the number of produced B-events and thereby determine the number of reconstructed signal events in our analysis method:

- Out of a certain number of pp events produced by PYTHIA we select those that contain either a muon with transverse momentum greater than 4.5 GeV or two muons with transverse momenta each greater than 2.5 GeV. This was accomplished using a filter on generator level, i.e. no reconstruction effects affect the number of accepted events. The filter was included to reduce the amount of data written to disk, just as a trigger reduces the amount of data recorded by the data acquisition system. The values were chosen such that a trigger can be simulated in a later stage, i.e. events that would pass the lowest single muon trigger at 5 GeV or the dimuon trigger at 3 GeV[17] should be included in the written data. Since the trigger

[17] see sec. 4.3.3.

acts on less precise data slightly lower values for the thresholds were chosen. The events satisfying the filter are then passed onto the reconstruction modules.

- In our reconstruction module we require certain conditions for the building of a D^0-candidate, as described in sec. 3.2. These are now applied on a reco level, i.e. we use tracks and vertices from the reconstruction procedures of CMSSW rather than generator information from PYTHIA. This implies that we have successfully reconstructed tracks and e.g. correctly identified muons in the muon system, i.e. it includes effects such as the muon ID efficiency (see sec. 4.3.4) or the tracking efficiency. It also includes the kinematic acceptance of the detector, i.e. the spatial angle covered by the various detector subsystems. The thresholds at this stage of the analysis data flow were chosen as low as possible without building too many candidates per event. Ideally every signal decay that could potentially be recorded by the detector would pass these thresholds, and they only serve the purpose of suppressing false candidates. All candidates built by the reconstruction module are then written into a ROOT TTree that is processed in the secondary analysis stage.

- The next step is to reduce the number of candidates per event to one, i.e. select the 'best' one. The criteria for the selection of a candidate were described in detail in sec. 3.3. The data used for the selection is the data written to the TTree by the reconstruction modules, i.e. it includes the reconstruction effects mentioned under the previous point. The candidate selection being far from perfect with a purity of about 88.66 %, some good candidates are lost in the process which leads to fewer events in the signal peak.

- Before the candidate is written, we also perform the simulation of a trigger efficiency (see sec. 4.3.3) at this stage. Considering all reconstructed muon tracks present in the event a flag is set reflecting which muon trigger would have (theoretically) been passed by the event. This does however not imply that the muon used for the reconstruction is the same muon that would have passed a trigger. All the relevant information including the trigger flags is finally written into a ROOT TTree for further processing.

- In the final stage of the analysis the candidates are cast into the $p_T(\mu D^0)$ binning using the information from the TTree. The only additional cut introduced for the final analysis was the simulation of a trigger on single muons with transverse momenta greater than 5 GeV, see section 4.3.3 for a description of this procedure, and figure 38, upper left, for the isolated efficiency of this cut.

4.3.2 Overall Efficiency

The most straightforward way to find a precise estimate for the overall efficiency is to generate an independent sample, run the reconstruction tools on it and observe how many signal events 'survived'. This was accomplished using the signal sample mentioned in sec. 2.3. The exact same analysis procedures used on the 'data' sample were performed on this signal sample resulting in a number of reconstructed events. Then the generator level information was used to find the number of events initially generated in each $p_T(\mu D^0)$ bin. The overall efficiency is then the ratio of reconstructed events to generated events. For each of the fitting algorithms described above this procedure was executed, and the different results were processed separately. The resulting overall efficiencies per $p_T(\mu D^0)$ bin for each algorithm are shown in figures 35, 36 and 37.

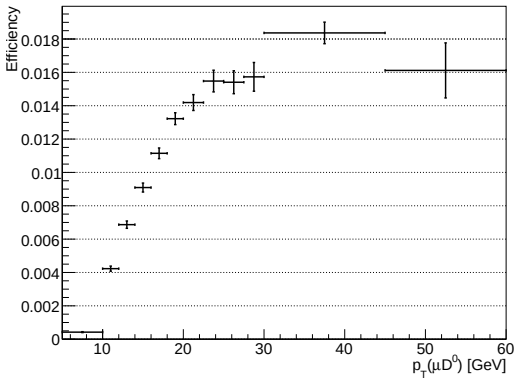

Figure 35: Overall efficiency using single gaussian fitting, in bins of $p_T(\mu D^0)$. Branching fraction $\mathcal{B}(D^0 \to K\pi)$ is not included.

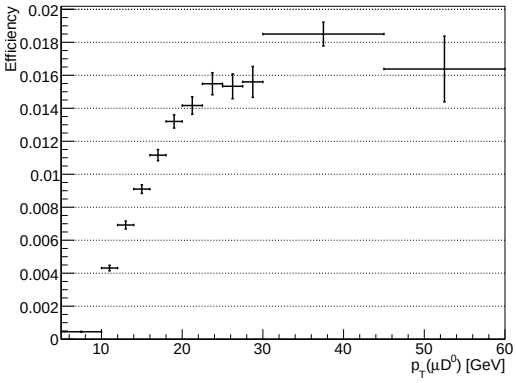

Figure 36: Overall efficiency using triple gaussian fitting, in bins of $p_T(\mu D^0)$. Branching fraction $\mathcal{B}(D^0 \to K\pi)$ is not included.

4.3.3 Trigger Efficiency

As mentioned in the introduction, the CMS detector uses a variety of selection criteria – 'triggers' – to decide which events are to be recorded for permanent storage. For this analysis the most convenient trigger is one that acts on muon transverse momenta, i.e. it fires the trigger when an event contains a muon track with p_T above a certain threshold. The exact value of the threshold is adjusted according to the instantaneous luminosity. For the modest luminosities in the startup phase of the LHC, for which this analysis is designed, the cutoff can be lowered to allow for the recording of relatively low p_T muon events. At high luminosities approaching the design specifications, these events would become so abundant that the data acquisition system could no longer handle the incoming data rate. In addition to single muon transverse momenta, one can also trigger on the occurrence of multiple muons in a single event, thereby enforcing a far more restrictive selection. Triggering on two muons each with relatively low p_T allows to probe lower transverse momenta regions at higher luminosities.

The CMSSW version used to develop this analysis (1.6.x) unfortunately does not include trigger information in the generated data. The muon transverse momenta triggers were therefore simulated using completely reconstructed muon tracks. Each event was tagged according to the presence of a muon track that would in principle have fired a certain muon trigger. I.e. if there existed a reconstructed muon track with a transverse momentum greater than 5 GeV, or two individual muon tracks each with p_T greater than 3 GeV, a flag signifying the passing of a single muon trigger at 5 GeV or a dimuon trigger at 3 GeV respectively, was set. The possible flags and therefore the simulated triggers are shown in table 4.

Table 4: The different single muon and dimuon trigger thresholds simulated in the analysis.

Single Muon:	5 GeV	11 GeV	16 GeV	
Di Muon:	$3+3$ GeV	$3+5$ GeV	$3+11$ GeV	$3+16$ GeV

The corresponding efficiencies were then determined by counting how many signal events ful-

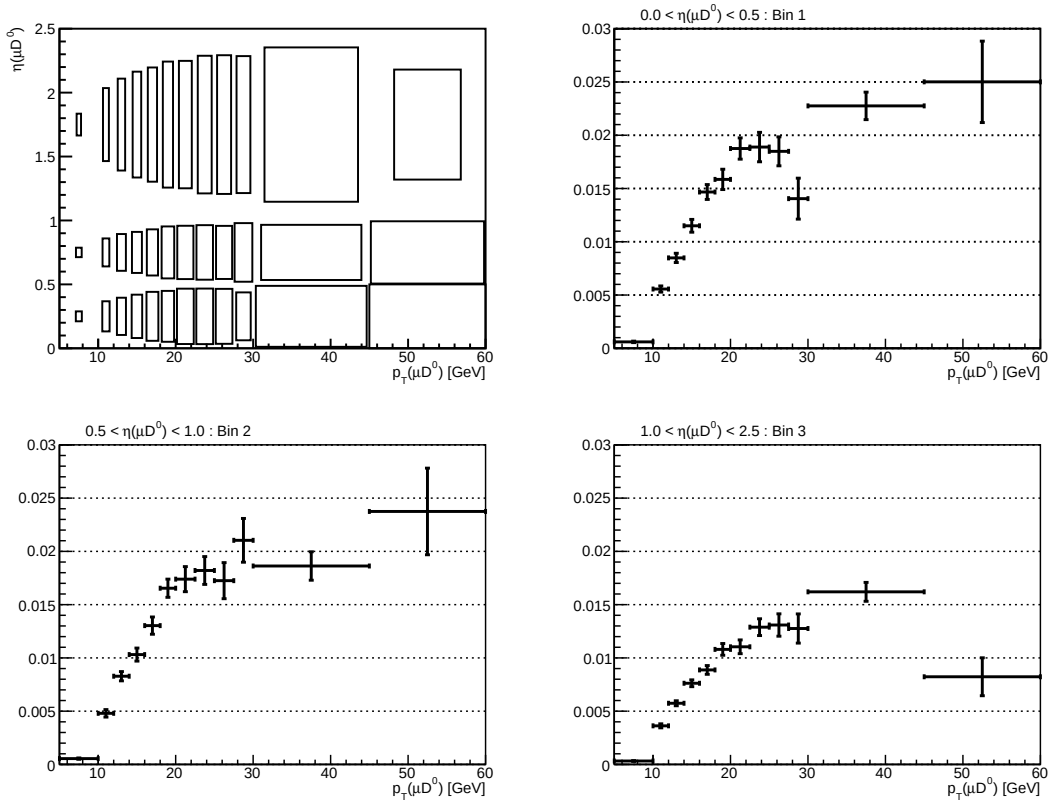

Figure 37: Overall efficiency using single gaussian fitting in bins of $\eta(\mu D^0)$ and $p_T(\mu D^0)$. Branching fraction $\mathcal{B}(D^0 \to K\pi)$ is not included.

filled each separate trigger, the signal events being tagged using generator information. The problem with this reenacted trigger is the fact that it utilizes 'too much' data: the actual muon trigger does not access data from the tracker or pixel system, but only acts on hits in the muon system. The trigger efficiencies so reconstructed should therefore be systematically lower than those of the real muon triggers. The so simulated signal efficiencies for each of the threshold values mentioned above are displayed in figures 38 and 39.

For the measurement of the differential cross-section the chosen trigger was the lowest and least restrictive of the above, the single muon trigger at 5 GeV. The trigger efficiency however is included in the previously described overall efficiency. The results from this and the next section are therefore to be understood as a study of two individual contributions to the overall efficiency.

4.3.4 Muon Identification Efficiency

A muon in the detector is identified by matching a succession of hits (or even a single hit) in the muon chambers to a corresponding sequence of hits in the tracker and pixel detector. This is a straightforward and highly effective method, as muons are the only particles expected to reach the muon system without being absorbed in ECAL, HCAL or any other material in-between. A

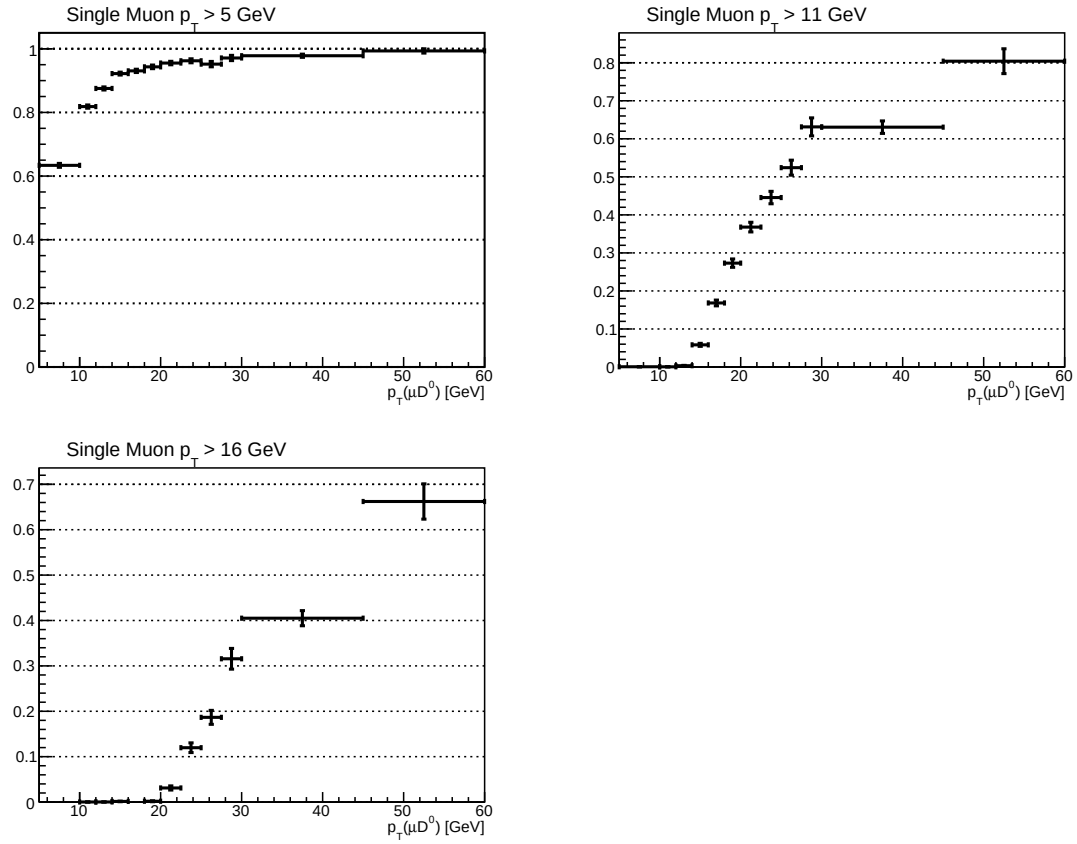

Figure 38: Simulated single muon trigger efficiencies for the three trigger thresholds indicated in the text.

track extending from the interaction point to the muon system is therefore highly likely to indeed originate from a muon – hence the name 'muon system'. However, the mean distance a muon can traverse inside the detector material depends linearly on the muon momentum, assuming it only loses energy by ionization and atomic excitation. For low p_T muons it is therefore primarily a question of reaching the first muon chamber without being absorbed in the intermediate material. For high p_T tracks there is also the issue of punch-through pions and kaons producing hits in the muon chambers, and thereby faking a muonic signal. The overall probability to correctly identify a muon in the detector is therefore a function of the muon transverse momenta and pseudo-rapidity. This issue is investigated in a fellow diploma thesis [6], and we just quote the results obtained there, see figure 40. However, since we are working in bins of $p_T(\mu D^0)$ in this analysis, we need some means of implementing this information to determine a muon identification efficiency for our binning. In order to do so, we average the identification probability over all muons in a given $p_T(\mu D^0)$ bin. For muons outside the range of the probability histogram ($2\,\text{GeV} < p_T(\mu) < 25\,\text{GeV}$, $0 < |\eta(\mu)| < 0-2.5$), we assume a constant efficiency of $(95\pm1)\,\%$ for $p_T(\mu) > 25\,\text{GeV}$ and $(5\pm1)\,\%$ for $p_T(\mu) < 2\,\text{GeV}$. The result is shown in figure 41.

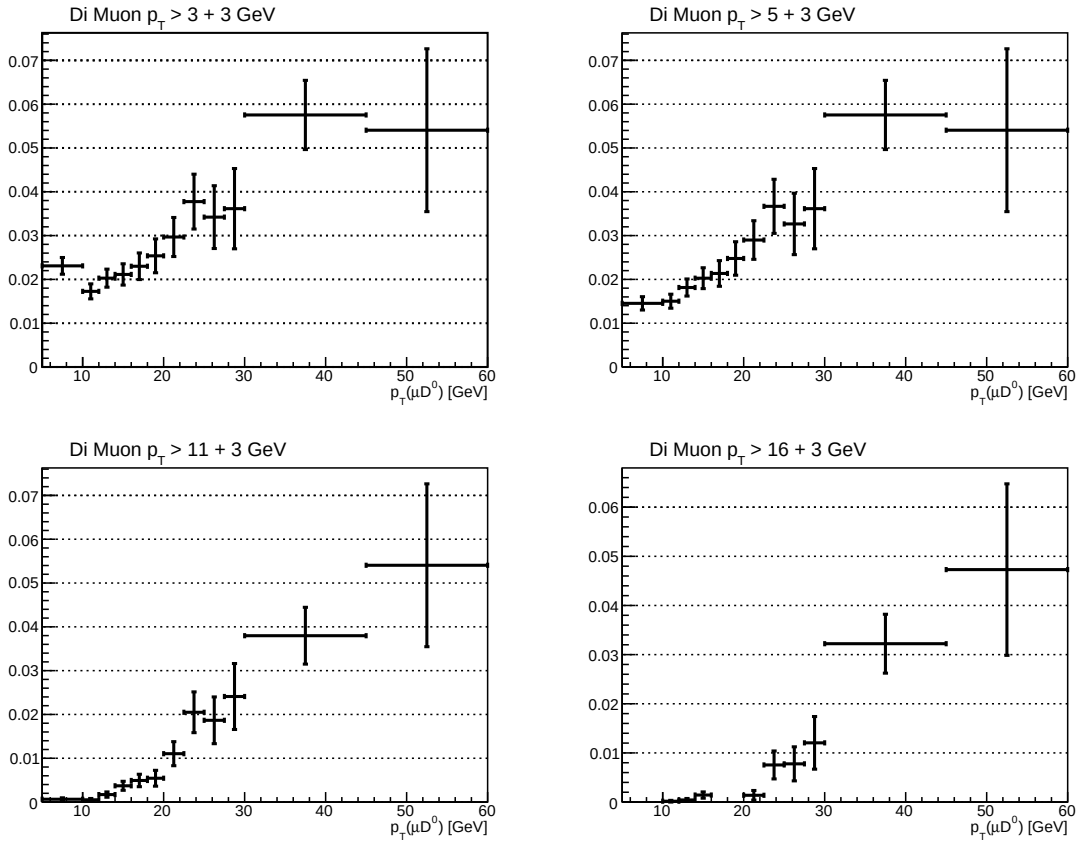

Figure 39: Simulated dimuon trigger efficiencies for the four trigger thresholds indicated in the text. Note that the available statistics for the lower bins of the highest triggers are severely limited.

4.4 Background Estimation

In light of the fact that we do not apply any additional cuts other than the single muon trigger at 5 GeV, we are forced to deal with background contributions. The substantial, non-peaking background in our mass-window of 1.7 - 2.0 GeV is taken care of by the fitting algorithm(s). There is however a significant D^0-background, producing a peak in the mass histograms that cannot be distinguished from the $B \to \mu D^0$ signal by means of a fitting technique. To classify and quantify this peaking background, we tag all candidates according to the following criteria, using generator information:

- To separate peaking from non peaking candidates we first investigate whether the three tracks used in the D^0 reconstruction were correctly identified, i.e. the μ track was indeed caused by a muon, and so forth. In the same step, we check if the K and π generator candidates were indeed the only two daughter objects of a D^0 parent object. If all that is satisfied the D^0 candidate is classified as contributing to the peak.

- To distinguish between peaking background and signal events, we inquire on a possible common heritage of the muon generator candidate and the D^0-meson. I.e. we check whether the objects from the $D^0 \to K\pi$ decay originate in some generation from the mother of the μ. If

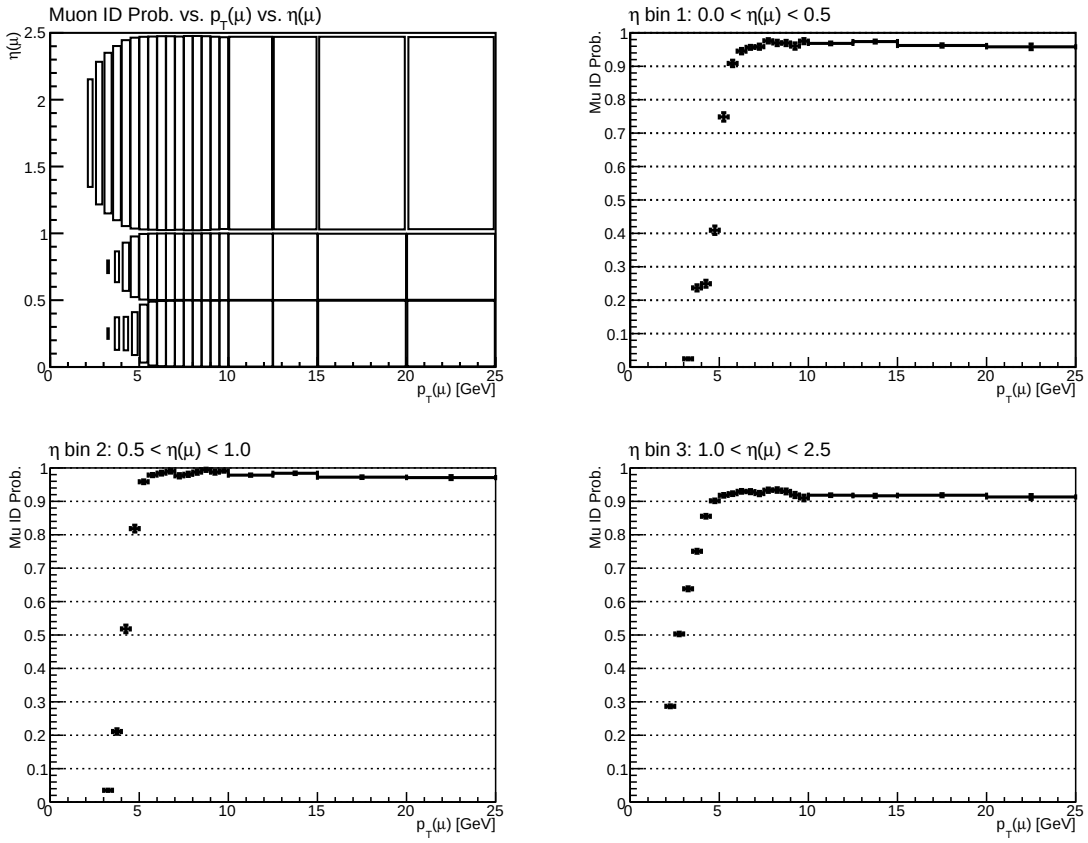

Figure 40: Muon identification probability from [6], as used for the muon identification efficiency determination, calculated for positive muons.

this check was positive, three different cases are investigated:

- The mother particle of the muon was a B^+-meson. In this case there was an additional distinction between direct and indirect decays. I.e. whether the D^0 originated directly from the B^+ or had an intermittent state. These two cases correspond to the decays $B^+ \to \bar{D}^0 \mu^+ \nu_\mu$ and $B^+ \to \bar{D}^*(2007)^0 \mu^+ \nu_\mu$ respectively.
- The mother particle of the muon was a B^0-meson. This case corresponds to the decay $B^0 \to D^*(2010)^- \mu^+ \nu_\mu$.
- Neither of the above. These few events (roughly one percent of the peak area) are a result of a $B_s \to D_{s1}(2536)^- \mu^+ \nu_\mu$ decay implemented in EvtGen. The $D_{s1}(2536)^-$ then decays into a $D^*(2010)^-$ or a $D^*(2007)^0$, providing a source for a $\bar{D}^0$, resulting in a correct $\mu K \pi$ sign combination. We will neglect these events in the further discussion, and include them in the systematic error (see sec. 4.7).

- For the cases in which the mother of the muon was not an ancestor of the D^0-meson, two different distinctions were made:
  - The muon came from a B^+ or B^0 meson, which case is most likely caused by $B^0/\bar{B}^0$ mixing on the 'other' side of the event, producing an otherwise Cabibbo-suppressed

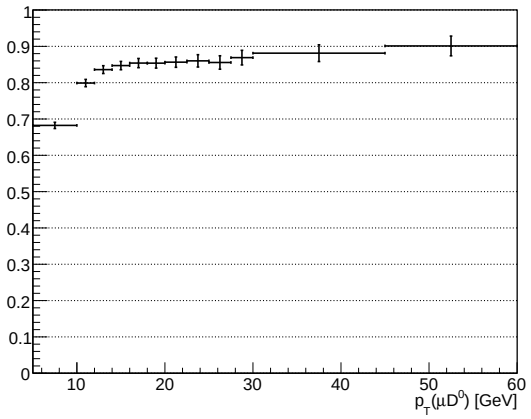

Figure 41: Efficiency from muon identification, averaged over $p_T(\mu)$ and $\eta(\mu)$ for each $p_T(\mu D^0)$ bin.

channel for $B^0 \to \mu D^0 X$ with the right sign combination. The tracks used for the D^0-reconstruction can still survive the cut on ΔR when produced by gluon-splitting rather than gluon-gluon-fusion, thereby receiving little angular separation.

– The muon did not come from a B^+ or B^0 meson, in which case we assume the D^0 to be produced in a $c\bar{c}$ event, neglecting other contributions to this case like the decay $B^+ \to \bar{D}^0 \tau^+$ with a subsequent $\tau^+ \to \mu^+ \nu_\mu \bar{\nu}_\tau$, which rarely fulfill the requirements on the building of a D^0 candidate due to a substantially softened μ from the τ-decay. The muon in this case could stem from a semileptonic decay of the associated c-quark. Again the angular separation cut can be circumvented when produced in gluon-splitting.

The so obtained composition of the signal peak is depicted in figure 42. Using this information we estimate the background component of the peak as the ratio of events tagged as coming from $c\bar{c}$-production or $B\bar{B}$-mixing to the total number of events in the peak. The values obtained from the above histogram were a contribution of $(7.75 \pm 0.19)\,\%$ from charm production, and one of $(5.37 \pm 0.16)\,\%$ from $B\bar{B}$-mixing. Note that these designations are not to be understood as precise, but rather as a nomenclature derived from the main contributors of the two background channels. The quoted errors are statistical only.

For want of a more sophisticated background analysis, these contributions are simply subtracted from the reconstructed number of signal events. Background suppression methods less dependent on Monte Carlo simulations could include requirements on kinematic properties of the reconstructed objects,[18] or independent measurements of high energy charm-production. However, since the first requires more statistics and the latter is not available yet, we stand by the simple Monte Carlo estimation explained above.

[18] See sec 4.6 for a first application of this using vertex separations.

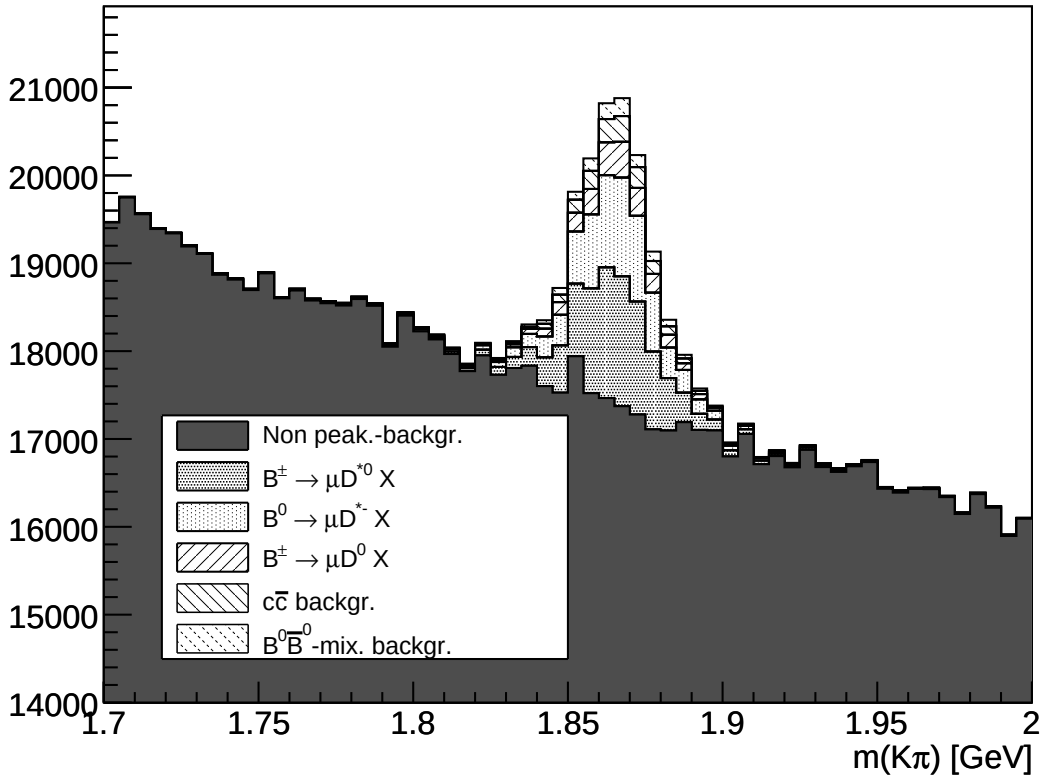

Figure 42: Different contributions to the μD^0 signal peak in the 'data'-sample, distinguished using generator information. The classifications are made as described in the text. The contribution from B_s decays is neglected. Note that the figure is zero-suppressed.

4.5 Cross-Section Results

Having obtained a value for the number of signal events in each bin, and an efficiency, one can now calculate the differential cross-section using the following formula:

$$\frac{d\sigma_B}{dp_T(\mu D^0)} = \frac{N_B}{\mathcal{L} \times \Delta p_T(\mu D^0)} = \frac{N_{\mu K\pi}}{2 \times \mathcal{B}(B \to \mu K\pi X) \times \epsilon_{tot.} \times \mathcal{L} \times \Delta p_T(\mu D^0)}.$$

Note that $N_{\mu K\pi}$ is assumed to be number of $\mu K\pi$ events originating from B-mesons, hence the subtraction of peaking background is implied. A factor of 2 is included to indicate that the production cross-section of a single B-meson ($B^\pm$ or B^0) is quoted. The integrated Luminosity $\mathcal{L}$ can be estimated from the number of generated events and the total pp cross-section:

$$\mathcal{L} = \frac{N_{gen.pp}}{\sigma_{pp}} = \ \simeq 0.4489\,\mathrm{pb}^{-1}.$$

The branching fraction of $B \to \mu K\pi X$ has a value of $(0.47 \pm 0.02)\,\%$ as described in the introduction. We now have all the ingredients necessary to build the differential cross-section. Table 5 gives all cross section results for all bins, with the generated cross-section for comparison; figures 43,

44 and 45 show the values depicted in histograms. Note that at this point the histograms were normalized by the bin-width.

Table 5: Results of the measurement of differential cross-section for the three different fitting techniques. Values are given in nb/ GeV. The generated cross-section, determined as described in the next subsection is printed for comparison. The quoted errors describe the statistical and systematic uncertainties in that order. The errors on the efficiency, resulting from the fitting procedure are counted as statistical errors. The error bars in the depicted histograms represent the sum in quadrature of the statistical and systematic uncertainties.

Bin	1	2	3	4
	5-10 GeV	10-12 GeV	12-14 GeV	14-16 GeV
Generated	10556 ± 482	3695 ± 169	2057 ± 94	1176 ± 54
Sin. Gauss.	$11114 \pm 1255 \pm 1977$	$3896 \pm 255 \pm 693$	$1898 \pm 115 \pm 338$	$1151 \pm 59 \pm 205$
Tr. Gauss.	$10818 \pm 1243 \pm 1924$	$3825 \pm 264 \pm 680$	$1893 \pm 117 \pm 337$	$1182 \pm 60 \pm 210$
$\eta(\mu D^0)$-sum.	$10512 \pm 2215 \pm 1164$	$3841 \pm 368 \pm 441$	$1884 \pm 167 \pm 211$	$1159 \pm 98 \pm 132$
η bin 1	$2663 \pm 662 \pm 474$	$705 \pm 137 \pm 125$	$462 \pm 70 \pm 82$	$262 \pm 41 \pm 47$
η bin 2	$2350 \pm 903 \pm 418$	$961 \pm 164 \pm 171$	$410 \pm 72 \pm 73$	$248 \pm 43 \pm 44$
η bin 3	$5499 \pm 1911 \pm 978$	$2176 \pm 300 \pm 387$	$1012 \pm 133 \pm 180$	$649 \pm 78 \pm 115$
Bin	5	6	7	8
	16-18 GeV	18-20 GeV	20-22.5 GeV	22.5-25 GeV
Generated	688.9 ± 31.5	420.9 ± 19.3	251.9 ± 11.5	149.5 ± 6.9
Sin. Gauss.	$689 \pm 38 \pm 123$	$370 \pm 21 \pm 66$	$250 \pm 15 \pm 44$	$156 \pm 12 \pm 28$
Tr. Gauss.	$692 \pm 40 \pm 123$	$372 \pm 22 \pm 66$	$253 \pm 16 \pm 45$	$157 \pm 13 \pm 28$
$\eta(\mu D^0)$-sum.	$706 \pm 66 \pm 82$	$365 \pm 41 \pm 40$	$258 \pm 29 \pm 31$	$158 \pm 21 \pm 18$
η bin 1	$146 \pm 25 \pm 26$	$95.8 \pm 18.0 \pm 17.0$	$42.9 \pm 10.3 \pm 7.6$	$32.8 \pm 8.2 \pm 5.8$
η bin 2	$145 \pm 26 \pm 26$	$82.6 \pm 17.6 \pm 14.7$	$54.6 \pm 11.6 \pm 9.7$	$33.3 \pm 10.3 \pm 5.9$
η bin 3	$415 \pm 55 \pm 74$	$187 \pm 33 \pm 33$	$161 \pm 24 \pm 29$	$92.1 \pm 15.9 \pm 16.4$
Bin	9	10	11	12
	25-27.5 GeV	27.5-30 GeV	30-45 GeV	45-60 GeV
Generated	92.31 ± 4.25	59.34 ± 2.74	19.27 ± 0.88	3.15 ± 0.15
Sin. Gauss.	$105 \pm 9 \pm 19$	$56.7 \pm 6.7 \pm 10.1$	$19.0 \pm 1.4 \pm 3.4$	$3.5 \pm 0.7 \pm 0.6$
Tr. Gauss.	$104 \pm 9 \pm 19$	$58.2 \pm 7.5 \pm 10.4$	$18.7 \pm 1.5 \pm 3.3$	$3.5 \pm 0.7 \pm 0.6$
$\eta(\mu D^0)$-sum.	$102 \pm 16 \pm 11$	$60.2 \pm 13.3 \pm 6.3$	$19.0 \pm 2.2 \pm 2.1$	$4.0 \pm 2.4 \pm 0.5$
η bin 1	$31.0 \pm 8.1 \pm 5.5$	$19.9 \pm 8.7 \pm 3.5$	$4.7 \pm 1.0 \pm 0.8$	$0.72 \pm 0.32 \pm 0.13$
η bin 2	$26.1 \pm 7.3 \pm 4.6$	$14.5 \pm 4.7 \pm 2.6$	$4.8 \pm 1.1 \pm 0.9$	$0.72 \pm 0.31 \pm 0.13$
η bin 3	$45.0 \pm 11.2 \pm 8.0$	$25.8 \pm 8.9 \pm 4.6$	$9.5 \pm 1.6 \pm 1.7$	$2.6 \pm 2.3 \pm 0.5$

The generated cross-section was determined using the signal sample. This data set has only events in which a process $B \to \mu D^0 X$ was generated, and therefore contains no background events. The signal decay of the D^0, namely $D^0 \to K\pi$ was however not enforced such that the sample still contains background in the form of different decay channels of the D^0. This fact has already been taken advantage of when determining the overall efficiency by performing the reconstruction on this sample. We now extract a $p_T(\mu D^0)$ spectrum from this sample, using the

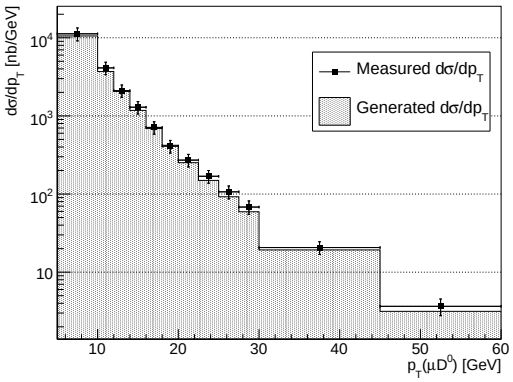

Figure 43: Result of the differential cross-section for production of a single B-meson, in bins of $p_T(\mu D^0)$. Measurement performed as described in the text, using single gaussian fitting. The filled background is the generated cross-section. The included systematical uncertainties are correlated over the bins, see sec. 4.7.

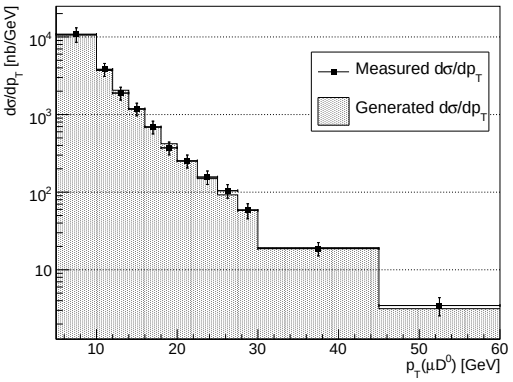

Figure 44: Result of the differential cross-section for production of a single B-meson, in bins of $p_T(\mu D^0)$. Measurement performed as described in the text, using triple gaussian fitting. Generated cross-section is indicated. The included systematical uncertainties are correlated over the bins, see sec. 4.7.

generated B-meson decays. A cut on $|\eta(\mu D^0)|$ to be smaller than 2.5 is introduced at this point, to exclude the generated η-overflow. This ensures that the same number of events enter into this histogram as into the $\eta(\mu D^0)$-binned one.[19] The result is then scaled by the integrated luminosity of the signal sample, estimated once more using the total pp cross-section, and by the branching fraction of $B \to \mu D^0 X$. The same factor of two as mentioned above is also included here, and the histogram was again normalized by the bin-width before comparing it to the values received from the measurement procedure performed on the 'data'-sample. Using the same method we also produce a generator-level cross-section in bins of $p_T(B)$, to be used later on.

4.6 Background Suppression using Distance to Primary Vertex

As an alternative to estimating the peaking background contribution using our produced Monte Carlo samples, we present an approach that utilizes a cut on the distance of primary vertex (i.e. pp interaction point) and reconstructed μD^0 vertex, which is an approximation to the point of decay of the B-meson. The advantage of this method is of course the gained independence from input parameters in the simulations, however more data is needed to sustain the cut. The here presented cut on the primary vertex distance is a rather simple first attempt, and could easily be improved, employing e.g. a comparison of the flight-direction of the B (using the reconstructed secondary and primary vertices) and the direction of the reconstructed decay products.

The distance of primary and secondary vertex is correlated to the lifetime of a particle; the longer it lives, the further it gets away from its production point. One of our main contributions to the peaking background was suspected to come from $c\bar{c}$-production. Due to Cabibbo-suppression c-hadrons generally live much shorter than b-hadrons, making the primary vertex distance a good indicator for a distinction between the two. We therefore try to take advantage of this fact by

[19] Recall that the $\eta(\mu D^0)$ binning ranges from 0 to 2.5

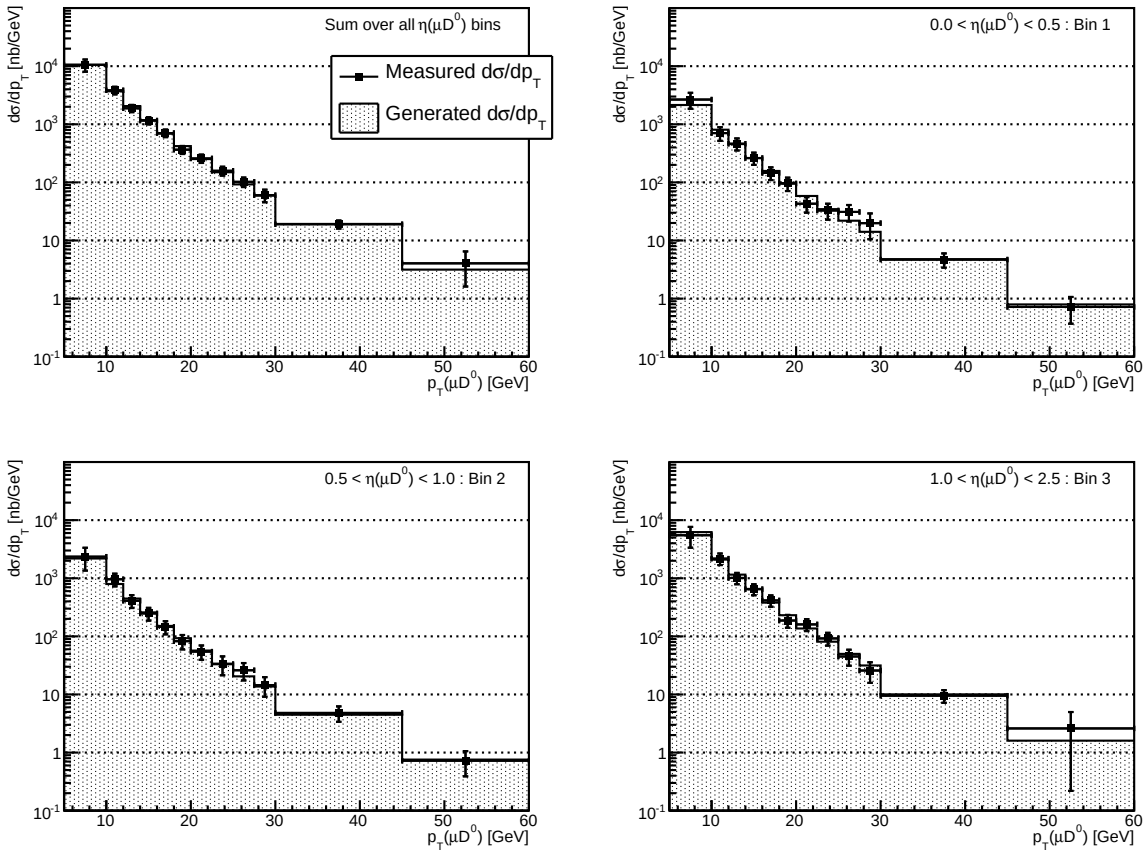

Figure 45: Result of the differential-cross section measurement using single gaussian fitting in bins of $\eta(\mu D^0)$ and $p_T(\mu D^0)$, compared with the generated cross-section. The included systematical uncertainties are correlated over the bins, see sec. 4.7.

requiring a distance to the primary vertex above a certain threshold value. Note however that peaking backgrounds caused by $B\bar{B}$-mixing are unaffected by this measure.

We first show a comparison of the distribution of the primary vertex distance for signal events and peaking background events, i.e. we only show events that contribute to the peak (see fig. 46 and 47).

From these figures a threshold value of 0.1 cm was chosen as a compromise between cut-efficiency and background-rejection. With more data available, this value can be increased to further suppress $c\bar{c}$ background. The selection efficiency of this additional cut is shown in figure 48. We once again determine the contributions from both background channels using generator-information, and find a correspondingly reduced value of $(2.04 \pm 0.12)\,\%$ for $c\bar{c}$ background contributions, and a virtually unchanged contribution of $(5.69 \pm 0.19)\,\%$ from $B\bar{B}$-mixing. Using this global cut value, we again determine the overall efficiency and the differential cross-section as described above,[20] without subtracting a background component however. We therefore expect – and observe – a

[20] We only perform the single gaussian fitting procedure here.

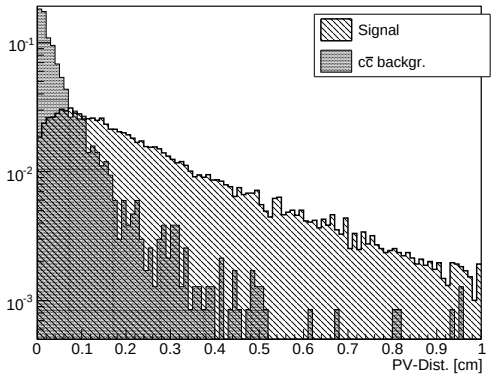

Figure 46: Distribution of primary vertex distance for signal events and peaking background events from $c\bar{c}$-production (as described in sec. 4.4). Histograms are normalized to unity for comparison.

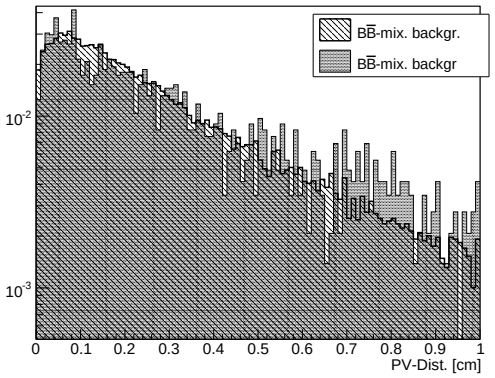

Figure 47: Distribution of primary vertex distance for signal events and peaking background events from $B\bar{B}$-mixing (as described in sec. 4.4). Histograms are normalized to unity for comparison.

systematically higher result for the cross-section, compared to the generated one, which, as you recall, does not contain peaking background events. The results so obtained are plotted in figures 49 and 50. The systematical shift towards higher values is visualized by plotting the ratio of measured to generated cross-sections in figure 51.

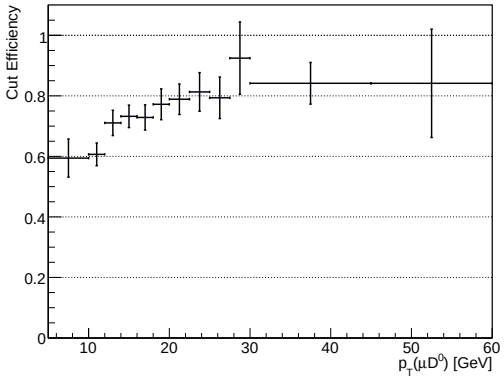

Figure 48: Efficiency of additional primary vertex distance cut at 0.1 cm.

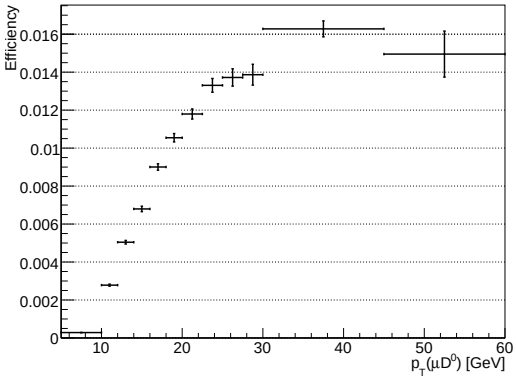

Figure 49: Overall efficiency for the reconstruction of $B \to \mu K \pi X$, using a cut on primary vertex distance greater than 0.1 cm. Branching fraction $\mathcal{B}(D^0 \to K\pi)$ is not included.

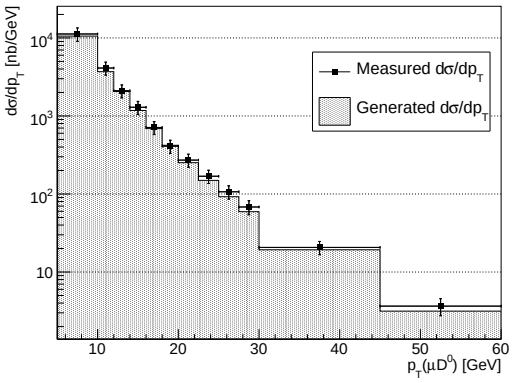

Figure 50: Measured differential cross-section of the production of a single B-meson, compared to generated cross-section. Background from $c\bar{c}$-production is suppressed using a cut on primary vertex separation. As that cut has no effect on the peaking backgrounds from $B\bar{B}$-mixing and other sources, the resulting cross-section is systematically too large.

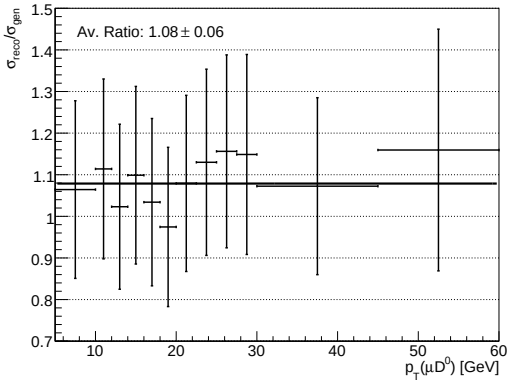

Figure 51: Ratio of measured to generated cross-section for a cut on primary vertex distance to be greater than 0.1 cm. Note that peaking background has not been subtracted, resulting in a shift towards higher cross-section values in the measurement. Also note that the included systematic errors are correlated from bin to bin.

4.7 Systematic Uncertainties

The precision of the measurement of a cross-section depends on various factors. The statistical uncertainties on the number of reconstructed signal events, entering into the number of events itself and the efficiency determination, was already mentioned. It is estimated using the toy fittings described in sec. 3.4.4, and decreases with the inverse square root of the numbers of events taken into consideration. The part of the uncertainty that cannot be reduced by repeating the measurement, the systematic errors, can be divided into various different contributions. We attempt to give an overview over all the different sources of uncertainties in this subsection. Since this analysis was developed entirely on simulated data, most systematic uncertainties arise because we cannot assume that the input parameters of the simulations coincide with the parameters experienced in the real experimental environment.

The luminosity enters prominently into the calculation of the cross-section from the number of signal events. Measuring this quantity at a high precision is not an easy task, and we assume an uncertainty of 5 % on the integrated luminosity of our 'data' sample.

Our measurement method relies heavily on the efficiency of the tracking system: failing to reconstruct a relevant track is equivalent to losing a signal event. The tracking efficiency depends on various environmental issues, e.g. the amount of material present inside the detector, the strength of the magnetic field at each point and so forth. For these parameters certain values have been assumed and put into the Monte Carlo simulations. Due to many unforeseeable circumstances these parameters most likely differ from the actual conditions present when the experiment is running. We assume a systematic uncertainty on the tracking efficiency (and thereby on the number of reconstructed events) of 4 % per track, i.e., since we reconstruct three tracks in this analysis, we assume a systematic uncertainty of 12 % from tracking issues. Note that we neglect any p_T dependence of this value. Effects caused by misalignment of tracker, pixel detector and

muon systems are to be understood as included in this number.

The muon identification probability is determined by using a control sample, e.g. $J/\psi \to \mu^+\mu^-$, which can introduce a bias. To account for such effects we include a systematic uncertainty of 4 % in our measurement.

In our fitting procedures we observed a slight tendency to reconstruct too few events. As previously mentioned this inadequacy cancels exactly when calculating the differential cross-section, and therefore does not show in the cross-section results. It does however add an uncertainty on both the reconstructed number of signal events and the efficiency. The average deviation from the number obtained in truth counting is roughly 3 %. We therefore include this value as a – rather conservative – estimate for the systematic uncertainty caused by the determination of signal events. Again we neglect any dependence on transverse momenta of this value.

As mentioned in the description of the peaking backgrounds (sec. 4.4) there is a contribution from a $B_s \to D_{s1}(2536)^-\mu^+\nu_\mu$ decay of about one percent that we neglected in the background assessment. We therefore formally add an additional systematic uncertainty of one percent, and understand it to include additional sources of μD^0 combinations that were overlooked or neglected hitherto.

The underlying motivation for this analysis is to measure the p_T spectrum of b-mesons to be used in background estimations at previously not reached center-of-mass energies. However, this is also the main input parameter relevant to the development of this study. The ability to reproduce the generated cross-section is therefore not equivalent to a successful measurement of the real cross-section.

Other effects included in the systematical uncertainty are the errors on the branching fractions on the order of 5 %. Adding all these contributions in quadrature we find an estimate for the overall systematic uncertainty of about 15 %. This is included in the error bars on each separate bin, hence the bin errors are correlated.

5 Unfolding

A measured differential cross-section in bins of $p_T(\mu D^0)$ has limited use for the estimation of backgrounds spectra, as naturally most analyses are not concerned with only this one decay channel of the b-quarks. The desired result is therefore a B-production cross-section in bins of $p_T(B)$, which can then be used to calculate a variety of beauty-backgrounds as a function of transverse momenta. A simple method how to transform from $p_T(\mu D^0)$ binning to $p_T(B)$ bins is presented in this section.

The momentum carried by the initial B-meson is of course distributed to its daughter particles upon decay. Since we only partially reconstruct the decay products, we will generally end up with a p_T value lower than the initial $p_T(B)$.[21] To find a $p_T(B)$ spectrum from its corresponding $p_T(\mu D^0)$ spectrum, we generally need the probability density function for the two correlated transverse momenta, i.e. we want to determine the probability for a certain value of $p_T(B)$ to result in a value of $p_T(\mu D^0)$ or vice versa. The fact that we work in bins of p_T, instead of continuous spectra, simplifies the problem to the calculation of a matrix of discrete probabilities rather than a continuous 2-dimensional probability density function.

5.1 Unfolding Matrix Building

We define the unfolding matrix M as the transformation matrix from $p_T(\mu D^0)$ to $p_T(B)$, as that is its intended use: we wish to transform our measured cross-section in bins of $p_T(\mu D^0)$ to one binned

[21] Even in the direct decay channel we ignore the neutrino, which of course carries part of the transverse momentum. In the case where the B had a very low p_T to begin with, and the neutrino was emitted in opposite direction, the μD^0 can in principle carry away more transverse momentum than the initial B had.

in $p_T(B)$. A vector containing the entries of each bin of $p_T(B)$ is then the matrix-multiplication of the unfolding matrix with the vector containing the measured entries in $p_T(\mu D^0)$ bins:

$$N_i(p_T(B)) = \sum_{j=0}^{N_{bins}} M_{ij} N_j(p_T(\mu D^0)).$$

The meaning of the matrix element M_{ij} is therefore the probability for an event in $p_T(\mu D^0)$ bin j to result from an event in $p_T(B)$ bin i. To estimate these probabilities we use our signal-sample to produce a 2-dimensional histogram of $p_T(B)$ vs $p_T(\mu D^0)$, depicted in figure 52, where the $|\eta(\mu D^0)| > 2.5$ overflow is once again rejected. For reasons of simplicity the binning in $p_T(B)$ was

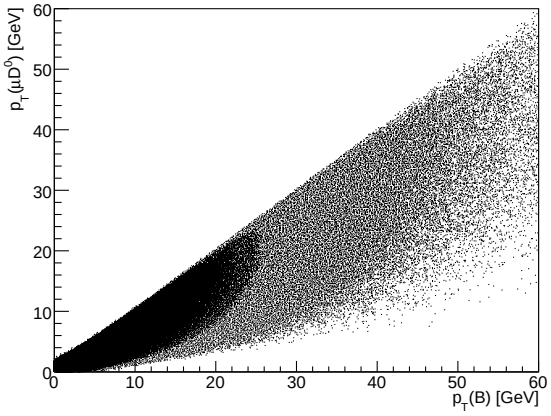

Figure 52: Distribution of transverse momenta of the initial B-meson and its daughter particles μD^0.

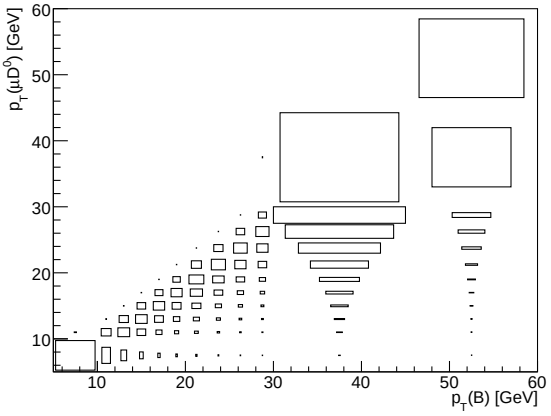

Figure 53: The same histogram filled into the used p_T binning and normalized as described in the text.

chosen to be equal to that in $p_T(\mu D^0)$. The entries of our unfolding matrix are proportional to the number of entries in the corresponding bin of the 2D histogram. We normalize the matrix such that it preserves the entries of the input histogram, i.e. we divide each entry by the sum over the whole column, including an additional line corresponding to a $p_T(B)$-overflow. The entries of each bin of the input histogram are then distributed such that their overall number is conserved, when again including the $p_T(B)$-overflow bin. The resulting 12×12 matrix is the following:

$$\begin{bmatrix}
0.6528 & 0.0020 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.1931 & 0.2557 & 0.0013 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0860 & 0.3639 & 0.2237 & 0.0009 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0354 & 0.1972 & 0.3537 & 0.1938 & 0.0007 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0144 & 0.0953 & 0.2032 & 0.3378 & 0.1660 & 0.0005 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0059 & 0.0444 & 0.1060 & 0.2083 & 0.3218 & 0.1442 & 0.0004 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0029 & 0.0241 & 0.0625 & 0.1362 & 0.2503 & 0.3699 & 0.1688 & 0.0001 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0010 & 0.0099 & 0.0271 & 0.0642 & 0.1272 & 0.2218 & 0.3332 & 0.1436 & 0.0002 & 0.0000 & 0.0000 & 0.0000 \\
0.0004 & 0.0040 & 0.0119 & 0.0298 & 0.0658 & 0.1212 & 0.2110 & 0.3106 & 0.1226 & 0.0000 & 0.0000 & 0.0000 \\
0.0002 & 0.0018 & 0.0056 & 0.0143 & 0.0326 & 0.0649 & 0.1221 & 0.2125 & 0.2927 & 0.1104 & 0.0001 & 0.0000 \\
0.0001 & 0.0015 & 0.0049 & 0.0143 & 0.0347 & 0.0744 & 0.1570 & 0.3138 & 0.5469 & 0.8120 & 0.6571 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0004 & 0.0010 & 0.0029 & 0.0068 & 0.0177 & 0.0334 & 0.0696 & 0.2906 & 0.5120
\end{bmatrix}$$

Note that, in particular for higher p_T matrix-elements, a significant amount of entries are 'lost' to the $p_T(B)$-overflow bin. Also note that this matrix describes the bin-to-bin correspondence

between $p_T(\mu D^0)$ and $p_T(B)$ histograms for our chosen p_T-binning. We must therefore apply it to an unnormalized histogram in bins of $p_T(\mu D^0)$, and separately normalize the resulting histogram to the bin-widths for a binning-independent differential result. The statistical errors on these matrix-elements are virtually negligible, with over 2×10^7 entries in the 2D histogram.

5.2 Underflow Estimation

Owing to the fact that our $p_T(\mu D^0)$ binning starts at a value of 5 GeV, we need to take into account an additional effect here. The lower bins of a $p_T(B)$ histogram, also starting at 5 GeV, contain events with a $p_T(\mu D^0)$-value below 5 GeV, as the neutrino or other decay products can carry away enough momentum to push the transverse momentum of the remaining μD^0 below that threshold. These additional events are not implemented in our matrix so far. In fact they cannot be, since we do not have and do not intend to have, a measurement for the $p_T(\mu D^0)$ underflow. We can however estimate the content of such an underflow bin using our Monte Carlo simulation. To do so, we scale the underflow of the generator-level cross-section such that the integrals of the visible $p_T(\mu D^0)$ range coincide. I.e. we multiply the content of the generated underflow bin by the ratio of measured to generated integrated visible cross-section. This estimation yields a value of 0.0 ± 0.0 nb. Once obtained, this value can be distributed to the $p_T(B)$ bins using an equivalent technique as described above. We cast the entries of our 2D histogram with a $p_T(\mu D^0)$ value below 5 GeV into a 1D histogram of $p_T(B)$ bins, corresponding to an additional column of our unfolding matrix. Again the entries are normalized by the overall sum, including $p_T(B)$ under- and overflows. The contributions to the $p_T(B)$ underflow and first 5 bins are displayed in table 5.2. The uncertainties on these percentages are – in the relevant bins – below 0.1 %.

Table 6: Distribution of estimated $p_T(\mu D^0)$ underflow bin to $p_T(B)$ bins of unfolded cross-section.

$p_T(B)$/ GeV	$0-5$	$5-10$	$10-12$	$12-14$	$14-16$	$16-18$
Underfl. matrix element	63.8 %	34.2 %	1.40 %	0.400 %	0.114 %	0.0334 %
Contr. from underfl.	0.00 nb	0.00 nb	0.00 nb	0.00 nb	0.00 nb	0.00 nb

5.3 Unfolded Cross-Section

We now apply the unfolding matrix to the cross-section measurement (from single gaussian fitting only), and compare the result with the generator-level cross-section in bins of $p_T(B)$. The errors are propagated through the matrix-multiplication and include the uncertainties on the matrix-elements and underflow estimations. We include an additional systematic uncertainty of 10 % to account for a probable discrepancy between the simulated and physical cross-sections.[22] The results of our unfolding are displayed in figures 54 and 55, the bin-for-bin values are printed in table 7.

[22] When performing the unfolding on real data, one would attempt to account for this by iterating the matrix building process such that generated cross-section and measured (unfolded) cross-section converge towards one another.

Table 7: Unfolded cross-section result in bins of $p_T(B)$, from single gaussian fitting result. The quoted errors include systematic and statistical uncertainties. Contributions from $p_T(\mu D^0)$ underflow are estimated as described in the text. Values in units of nb/ GeV.

Bin	1	2	3	4
	5-10 GeV	10-12 GeV	12-14 GeV	14-16 GeV
Generated	12509 ± 571	6618 ± 302	4239 ± 193	2666 ± 122
Unfolded	7259 ± 1734	6364 ± 1341	4232 ± 730	2648 ± 397
Bin	5	6	7	8
	16-18 GeV	18-20 GeV	20-22.5 GeV	22.5-25 GeV
Generated	1676 ± 77	1070 ± 49	668.8 ± 30.6	406.4 ± 18.6
Unfolded	1661 ± 228	1054 ± 138	649 ± 82	396 ± 50
Bin	9	10	11	12
	25-27.5 GeV	27.5-30 GeV	30-45 GeV	45-60 GeV
Generated	255.7 ± 11.7	167.3 ± 7.7	55.59 ± 2.54	9.51 ± 0.44
Unfolded	253 ± 32	169 ± 21	55.9 ± 7.0	9.6 ± 1.6

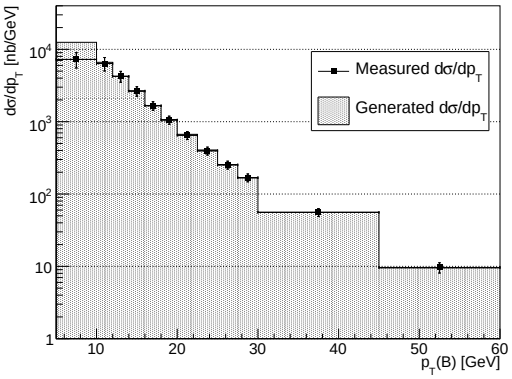

Figure 54: Unfolded differential cross-section for production of a strange-less B-meson in bins of $p_T(B)$. Measurement performed using single gaussian fitting. Contribution from $p_T(\mu D^0)$ underflow is not included. Generator-level cross-section is shown for comparison.

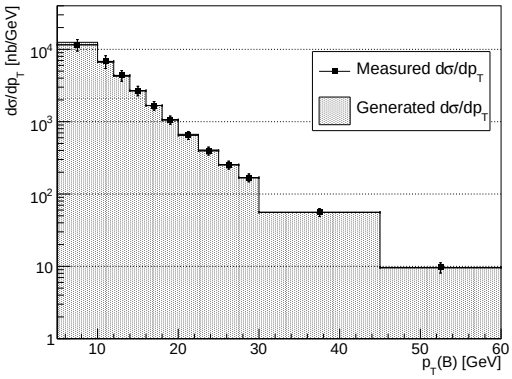

Figure 55: The same histogram with the contribution from a simulated $p_T(\mu D^0)$ underflow bin included, as described in the text.

A Additional Figures

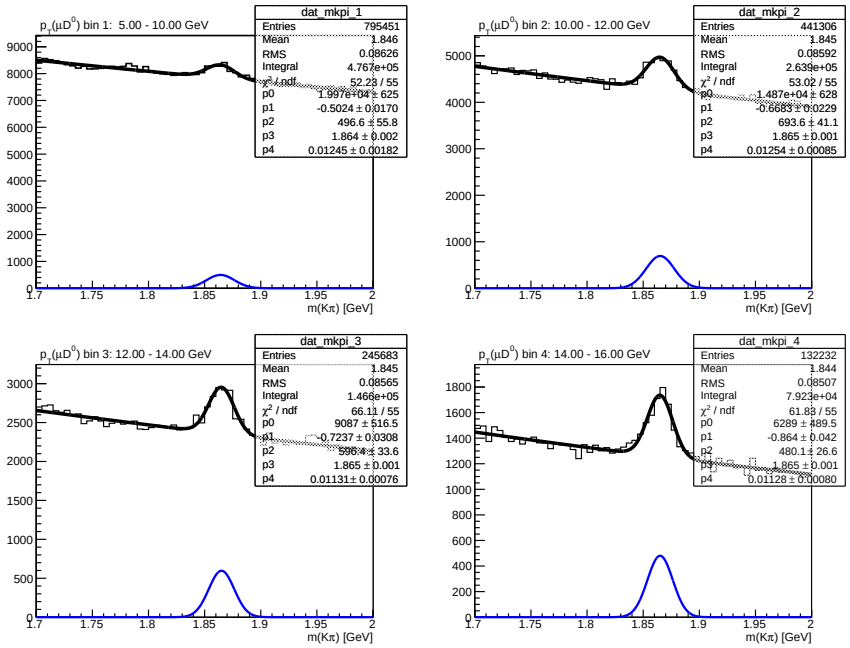

Figure 56: Single Gaussian fitting results for bins 1-4 ('data' sample)

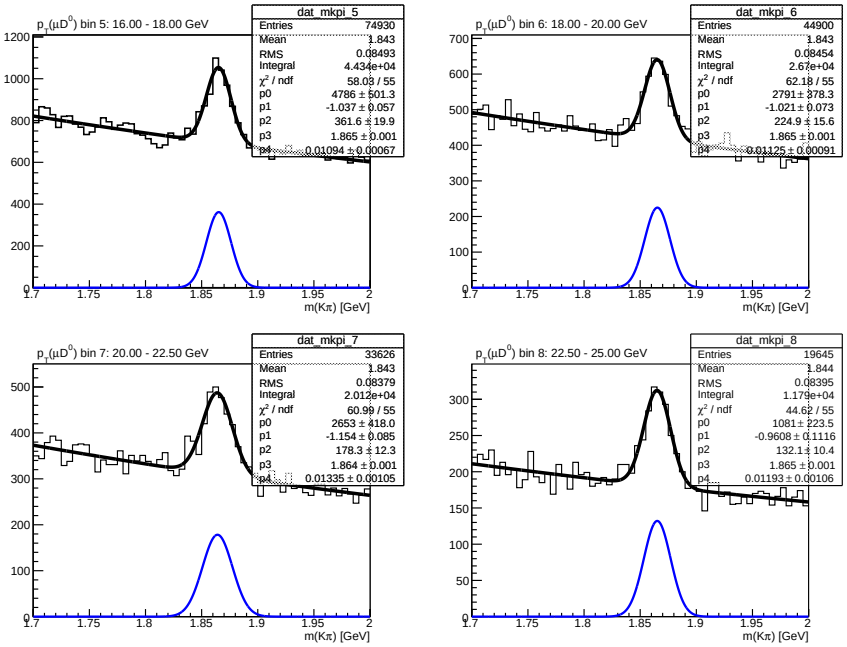

Figure 57: Single Gaussian fitting results for bins 5-8 ('data' sample)

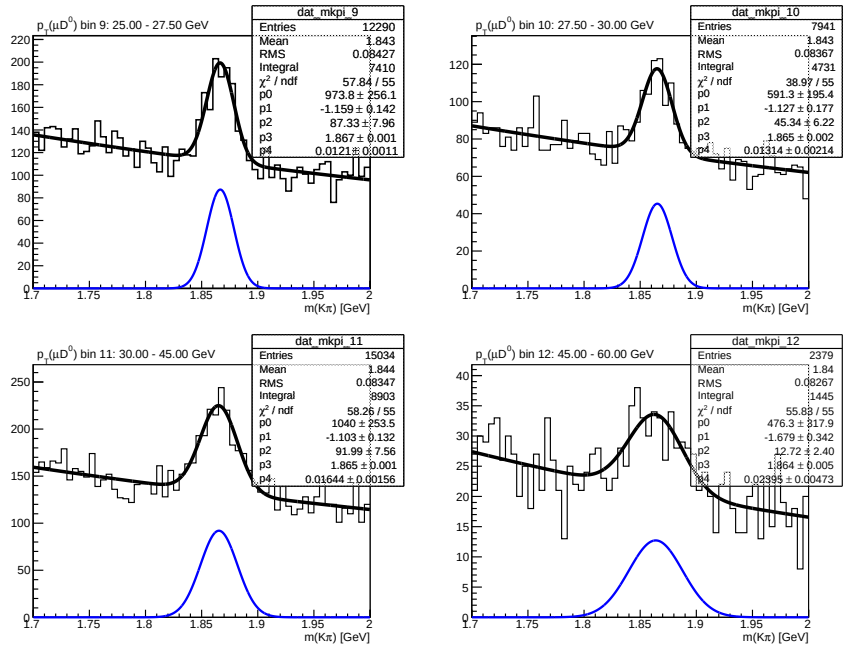

Figure 58: Single Gaussian fitting results for bins 9-12 ('data' sample)

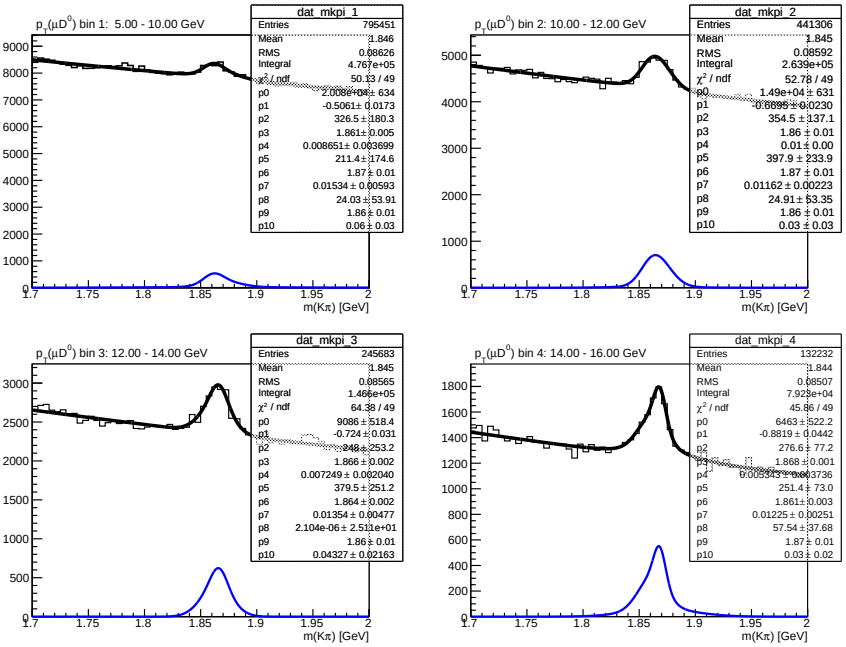

Figure 59: Triple Gaussian fitting results for bins 1-4 ('data' sample)

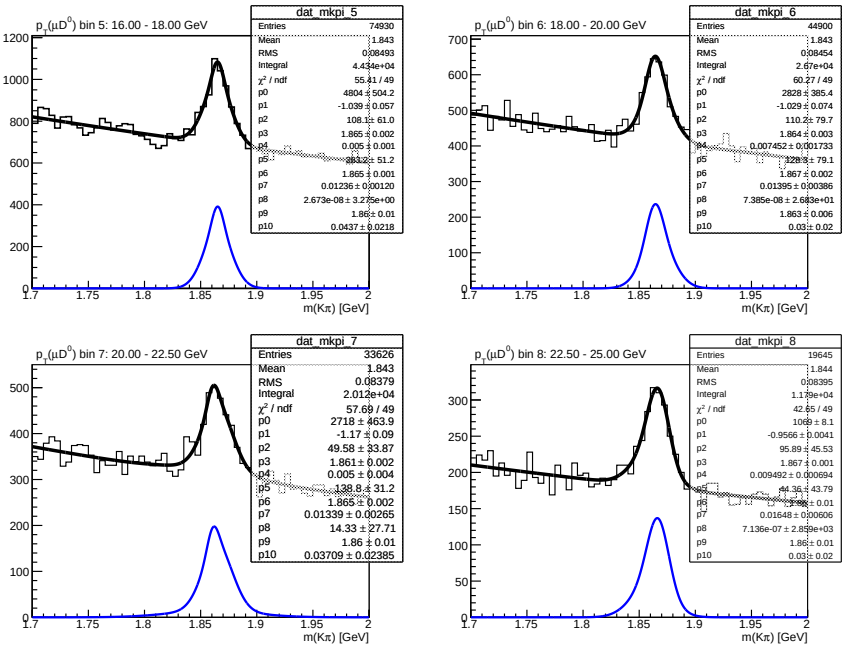

Figure 60: Triple Gaussian fitting results for bins 5-8 ('data' sample)

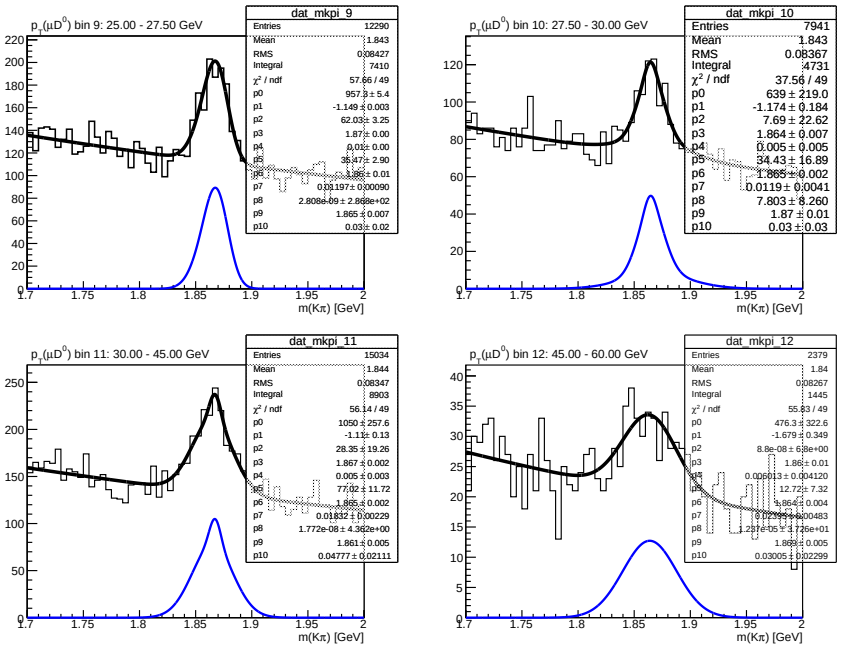

Figure 61: Triple Gaussian fitting results for bins 9-12 ('data' sample)

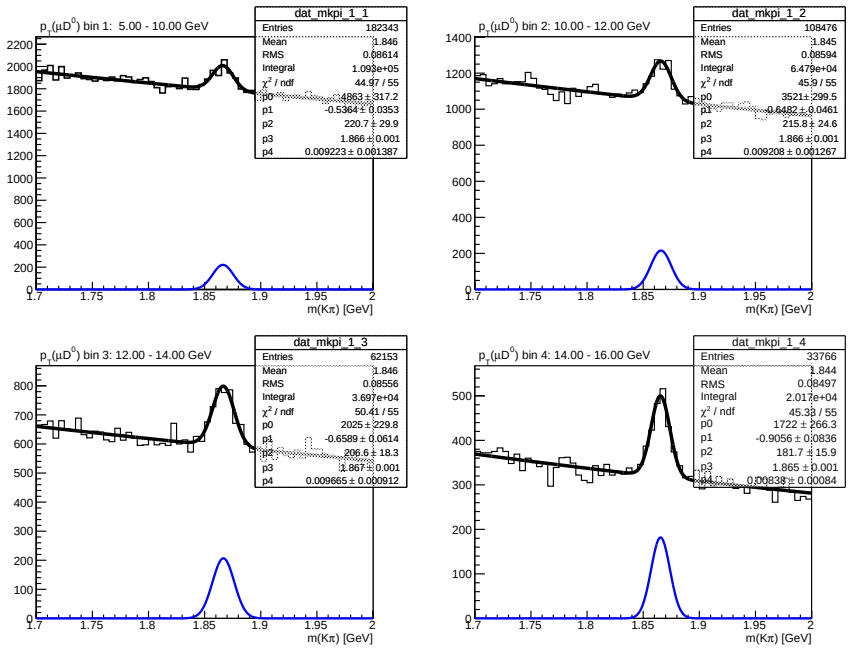

Figure 62: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 1 ($0.0 < |\eta| < 0.5$) and $p_T(\mu D^0)$ bins 1-4 ('data' sample)

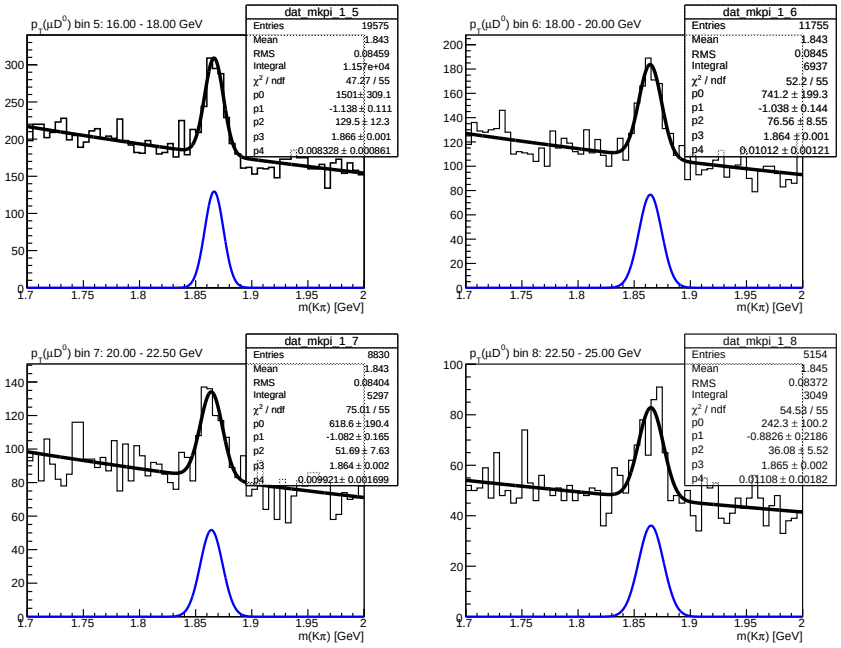

Figure 63: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 1 ($0.0 < |\eta| < 0.5$) and $p_T(\mu D^0)$ bins 5-8 ('data' sample)

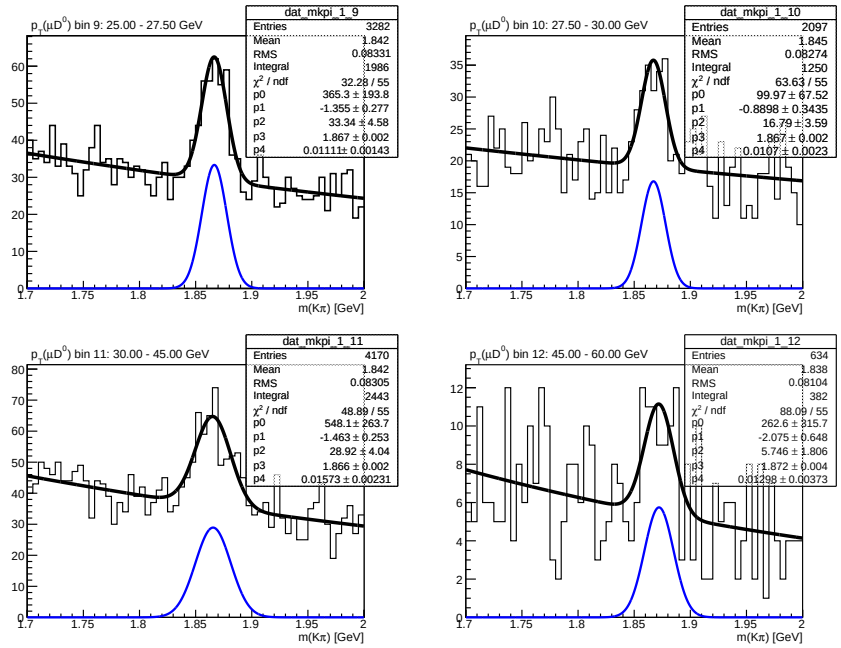

Figure 64: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 1 ($0.0 < |\eta| < 0.5$) and $p_T(\mu D^0)$ bins 9-12 ('data' sample)

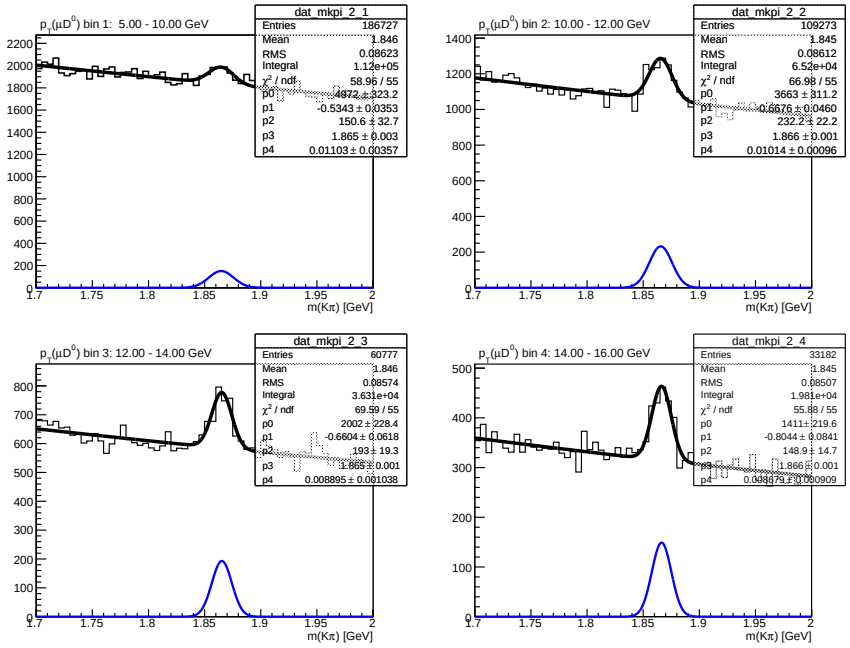

Figure 65: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 2 ($0.5 < |\eta| < 1.0$) and $p_T(\mu D^0)$ bins 1-4 ('data' sample)

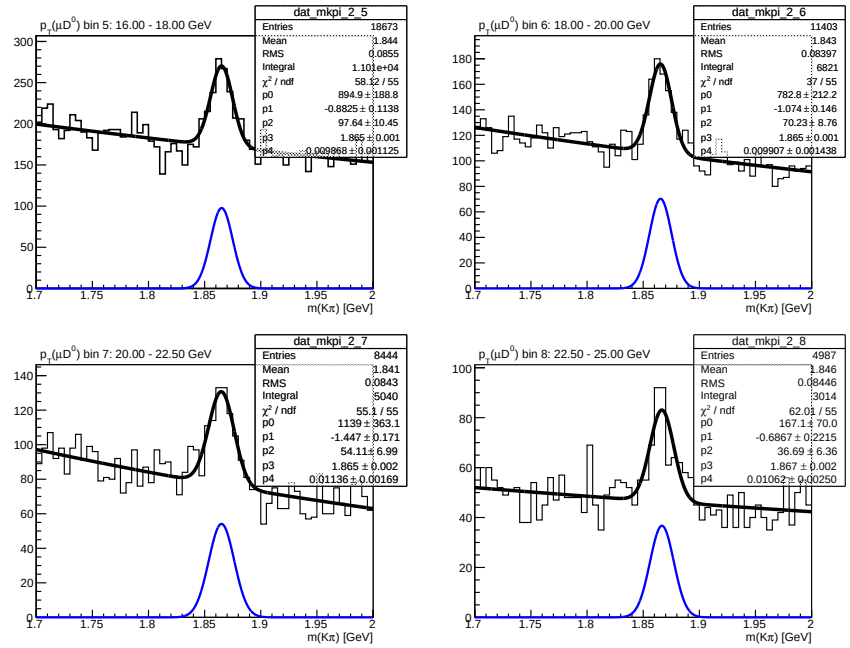

Figure 66: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 2 ($0.5 < |\eta| < 1.0$) and $p_T(\mu D^0)$ bins 5-8 ('data' sample)

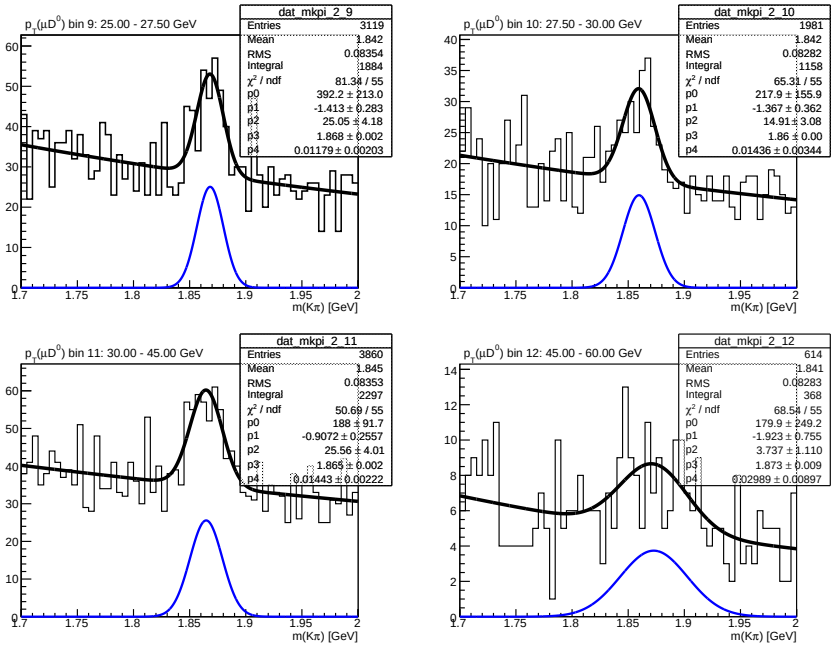

Figure 67: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 2 ($0.5 < |\eta| < 1.0$) and $p_T(\mu D^0)$ bins 9-12 ('data' sample)

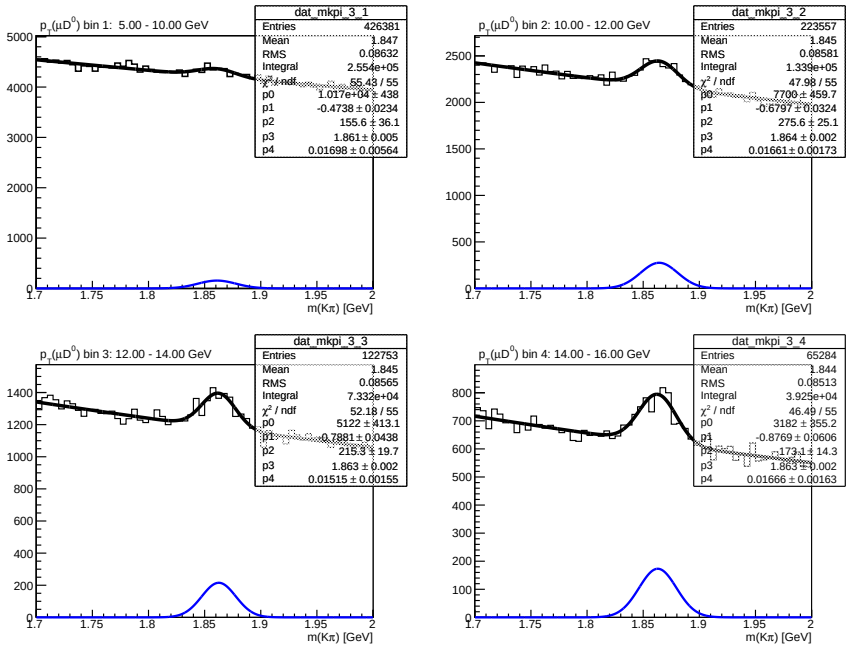

Figure 68: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 3 ($1.0 < |\eta| < 2.5$) and $p_T(\mu D^0)$ bins 1-4 ('data' sample)

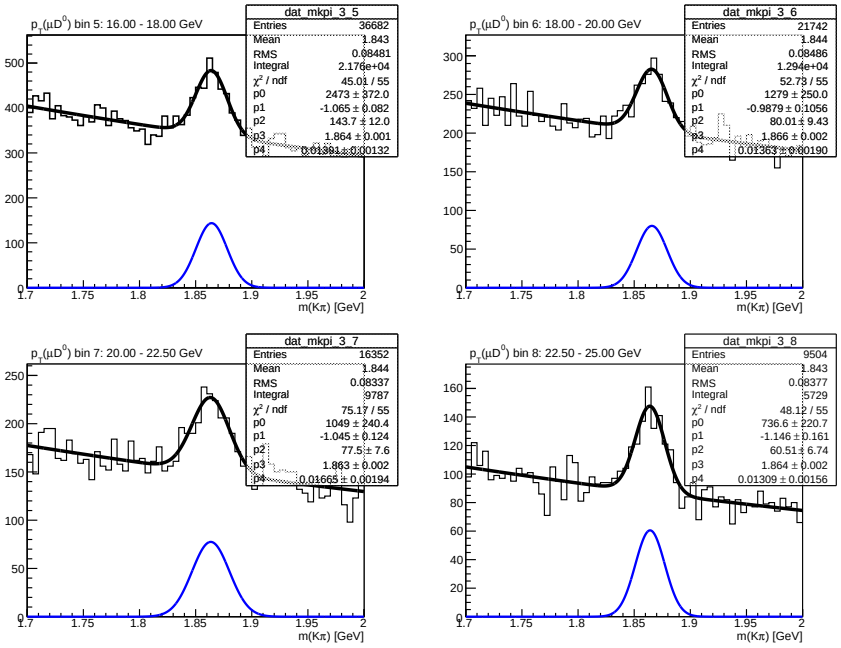

Figure 69: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 3 ($1.0 < |\eta| < 2.5$) and $p_T(\mu D^0)$ bins 5-8 ('data' sample)

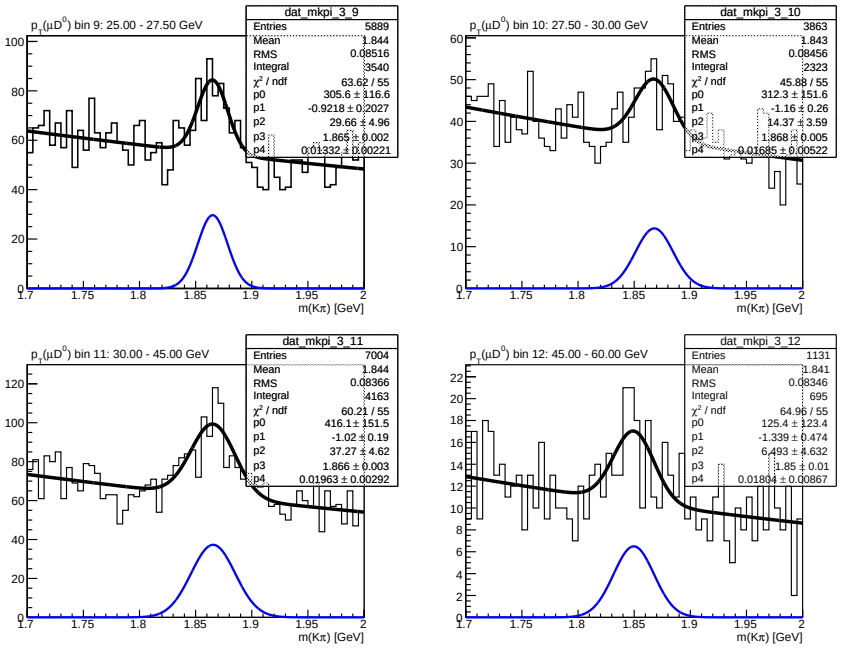

Figure 70: Single Gaussian fitting results for $\eta(\mu D^0)$ bin 3 ($1.0 < |\eta| < 2.5$) and $p_T(\mu D^0)$ bins 9-12 ('data' sample)

B Acknowledgements

I wish to express my deep gratitude towards Urs Langenegger for the many hours of his precious time he invested in me, as well as the generally pleasant working environment he maintained for us. I also want to thank my co-students Michel De Cian and Marco Reinhard, and the members of Urs' group, in particular Christina Eggel and Sarah Dambach for many enlightening discussions, delightful conversations and answers to stupid questions. Last but not least I wish to thank my parents without whose moral and material support none of this would have been possible.

References

[1] E. Norrbin and T. Sjostrand. Production and hadronization of heavy quarks. *Eur. Phys. J.*, C17:137–161, 2000.

[2] W. M. Yao et al. Review of particle physics. *J. Phys.*, G33:1–1232, 2006.

[3] Michel Della Negra et al. CMS Physics TDR: Volume I (PTDR1), Detector Performace and Software. *CERN-LHCC*.

[4] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

[5] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[6] M. Reinhard, C. Eggel, and U. Langenegger. Determination of Muon Reconstruction and Identification Efficiencies in the CMS Experiment. 2008.