Invariant mass determination using the output from a specialized electron track reconstruction algorithm

HÅVARD GJERSDAL

Abstract

The Large Hadron Collider at CERN will produce proton-proton collisions with a center of mass energy of 14 TeV, making it possible to examine physics at the TeV scale. Four detectors are located along the LHC, one of them is ATLAS, a general-purpose detector designed to identify a wide range of physics processes. In order to discover new physics, presice measurements of the four-momentum of the particles traversing the detector is important. For electrons this is done in two detector sub-systems. The inner detector reconstructs the trajectories of all charged particles, providing estimates of the momentum. The electromagnetic calorimeter measures the energy of the electrons.

The Kalman filter (KF) is the optimal track reconstruction algorithm if all stochastic processes encountered during the reconstruction procedure follow a normal distribution, and the track model is truly linear. Electrons lose energy predominantly through bremsstrahlung, which is a highly non-Gaussian stochastic process, and in this situation it is plausible that a non-linear estimator better taking the actual shape of the energy loss distribution into account can do better. The Gaussian-sum filter (GSF) models the bremsstrahlung energy loss as a Gaussian mixture instead of a single Gaussian, and the estimate of the track parameters provided by the GSF also becomes a Gaussian mixture.

It has been shown in a simulation experiment in the CMS tracker that the GSF can improve the momentum resolution of electrons as compared to the KF. In a subsequent simulation experiment in the ATLAS Inner Detector, it has been shown that also the invariant mass resolution can be improved compared to the standard KF by "collapsing" the Gaussian mixture describing each reconstructed electron into a single Gaussian, and then using the single Gaussian to reconstruct the invariant mass.

This thesis gives a summary of some aspects of particle physics and the ATLAS experiment. The process of track reconstruction is then discussed in more detail. Finally, a method of invariant mass determination using the full shape of the GSF track estimates is discussed and tested with simulated data.

Acknowledgements

- First of all, I wish to thank my supervisors Are Strandlie an Lars Bugge for a very interesting thesis subject, and for all the theoretical, technical and moral support you have given me.

- Thank you, Anthony Morley, for guiding me through the jungle that is Athena, and providing the tools to extract the information I needed.

- Thank you, Tom Atkinson, for letting me look at your ntuples and for writing a thesis that helped me understand what the GSF is all about.

- Thank you, Ole Røhne and Erlend Bolle, for introducing me to the wonderful world of 3DSi pixels, and giving me the opportunity to spend the summer at CERN.

- Thanks to everybody working with the BCM for letting me play with diamonds.

- Thanks to the entire EPF group for being a very smart and very helpful group of people.

Contents

Chapter 1

Introduction

Thanks to remarkable experimental and theoretical advances in the last decades, a model describing what is assumed to be the fundamental building blocks of matter, and how these particles interact, has emerged. This is the Standard Model of particle physics [3, 5].

The model is a relativistic quantum mechanical description of nature, and accounts for all known fundamental interactions, sans gravity. The matter particles, called fermions, interact by exchanging force mediating particles, called gauge-bosons. There are four known fundamental forces. The Standard Model contains quantum field theories describing strong, weak and electromagnetic interactions.

1.1 Fundamental particles and forces

There are two main groups of particles, fermions and bosons. All directly observed matter consist of fermions, particles with half integer spin. Fermions are split into two groups, quarks and leptons. The six quarks in the Standard Model are listed in table 1.2. In addition to carrying an electric charge, the quarks carry one of three color charges linked to the strong interaction. The other group of fermions, the leptons, are listed in table 1.2. Lepton do not carry a color charge.

The leptons and quarks are further divided into three generations. Each generation consists of two quarks and two leptons. The quarks and leptons in the different generations have the same quantum numbers, like electric charge, spin and color charge, but the mass of the particles differ. The first generation contain the lightest particles, the third has the heaviest. The heavier particles in the second and third generation are not stable, and quickly decay to a lighter generation. Because of this, all stable matter in the universe is made up of the quarks and leptons in the first generation.

All the fermions in the Standard Model have anti-particles, particles with opposite quantum numbers, but the same mass.

With the discovery of the tau neutrino at the DONUT(Direct Observation of the NU Tau) experiment at Fermi lab [18] in 2000, all the fermions in the Standard

1

Quarks spin= 1/2 $\hbar$			
Generation	Flavor	Mass	Electric charge
First	Up(u)	1.5 - 3.0 MeV	2/3
	Down(d)	3 - 7 MeV	-1/3
Second	Charm(c)	1.25 GeV	2/3
	Strange(s)	95 MeV	-1/3
Third	Top(t)	172.5 GeV	2/3
	Bottom(b)	4.2 GeV	-1/3

Table 1.1: Standard Model quarks, with mass and electric charge. All the quarks carry a color charge of either red, green or blue.

Leptons spin= 1/2 $\hbar$			
Generation	Flavor	Mass	Electric charge
First	Electron neutrino(ν_e)	<2 eV	0
	Electron(e^-)	0.511 MeV	-1
Second	Muon neutrino(ν_μ)	< 0.19 MeV	0
	Muon(μ^-)	106 MeV	-1
Third	Tau neutrino(ν_τ)	< 18.2 MeV	0
	Tau(τ^-)	178 GeV	-1

Table 1.2: Standard Model leptons, with mass and electric charge.

Model have been observed.

Bosons are particles with integer spin. The standard model contain 12 force mediating gauge-bosons with spin 1, and one spin 0 particle, called the Higgs boson. The gauge-bosons are the mediators of the strong, weak and electro magnetic force, as shown in figure 1.3.

Fundamental forces			
	Weak	Electro magnetic	Strong
Acts on	flavor	electric charge	color charge
Particles experiencing	fermions	electrically charged	quarks, gluons
gauge-boson	Z, W$^\pm$	photon	8 gluons
gauge-boson mass	91GeV,80GeV	0	0
strength	10^{-5}	10^{-2}	1
range	10^{-18}	∞	10^{-15}

Table 1.3: The forces described by the Standard model. All the gauge-bosons have spin=1$\hbar$.

All the fundamental interactions in the model are derived from a single general principle, the requirement of local gauge-invariance. Quantum Chrommo Dynamics(QCD) are based on the symmetry group $SU(3)_C$. This group describes the

strong force, mediated by 8 gluons. The gluons couple only to particles carrying a color charge, this means that only quarks and gluons can interact through the strong force.

The electro magnetic force, mediated by the photon, and the weak force, mediated by the massive Z and $W^\pm$ bosons are unified in a single theory. This is the Glashow - Weinberg - Salam theory of electro-weak processes, based on the $SU(2)_L \times U(1)_Y$ symmetry group. Under $SU(2)_L$, the fermions are ordered in left-handed isodoublets and right-handed isosinglets, anti-particles in right-handed doublets and left-handed singlets. A left-handed particle is a particle where the spin and momentum point in opposite directions, a right-handed particle has a spin pointing in the same direction as the momentum.

The Z and $W^\pm$ bosons only couple to left-handed fermions and right-handed anti-fermions. Photons couple to all particles with an electric charge.

The standard model contains a mechanism that breaks the symmetry of the $SU(2)_L \times U(1)_Y$ group. Without this spontaneous symmetry breaking, including fermion masses or massive Z and $W^\pm$ bosons break the gauge invariance of the theory. This mechanism is called the Higgs mechanism, after the British physicist Peter Higgs. The mechanism uses a scalar field with a non-vanishing vacuum expectation value, i.e. a field that is present throughout the universe, to break the electro-weak symmetry. The symmetry breaking leads to massive Z and $W^\pm$ bosons, and fermions obtain mass by interacting with the scalar field.

The scalar field is called the Higgs field, and a neutral, spin 0 particle, called the Higgs boson, is associated with it. The Higgs boson is the only particle predicted by the Standard Model that has not yet been observed.

1.2 Problems with the Standard Model

The Standard Model has gone through extensive experimental testing, and so far all the observed QCD and electroweak processes observed are in agreement with the theory. Because of this, the theory is a tremendous success. There are, however, several areas where the theory does not offer an explanation to observed phenomenon.

The Standard Model is unable to account for the matter - anti-matter asymmetry in the universe. A small preference for matter over anti-matter is introduced due to CP-violation in the CKM-matrix. The CKM-matrix describes the transition likelihood between up-type and down-type quarks in flavor-changing weak interactions. This CP-violation is several orders of magnitude to low to account for the observed matter - anti-matter asymmetry.

Observations of the cosmos suggest vast amounts of weakly interacting massive particles, called dark matter. Observation of neutrino oscillation suggests that neutrinos are massive, but this is not enough to account for all the "missing" matter in the universe. The existence of dark matter points to the existence of a new type of fundamental particle that only undergo weak interactions.

Super Symmetry(SUSY) models predict that all fermions has a partner boson, and vice versa. This symmetry must be broken, so that the SUSY-partners are heavier than their Standard Model counterparts. If this was not the case, SUSY-particles would already have been observed in experiments. Some SUSY-models predict stable, neutral particles, which would account for the dark matter.

The Higgs boson has not yet been observed. The mass of the Higgs boson is not predicted by theory. A very heavy Higgs, $m_H \gtrsim 1 TeV$, would lead certain scattering processes to violate unitarity. The mass of the Higgs boson is very sensitive to radiative corrections. Without an extreme fine-tuning of the parameters, the radiative corrections to the mass by Standard Model particles could push the mass close to the Planck-scale($10^{19} GeV$). If SUSY-particles exist with a mass below 1 TeV, radiative corrections from these particles would cancel out the corrections from Standard Model particles.

The Standard Model contains a number of free parameters. The fermion masses, the matrix elements of the CKM-matrix, the coupling constant of the three forces, the Higgs mass and the vacuum expectation value of the Higgs field are not predicted by theory, and must be experimentally determined. The fact that so many parameters must be put in by hand, and that no explanation for gravity is included in the theory, makes it hard to consider the theory fundamental.

1.3   The Large Hadron Collider

The Standard Model predicts the existence of the Higgs boson with a mass of less than 1 TeV. Several SUSY-models predict SUSY- particles with masses of less than 1 TeV. In order to probe for these particles, a particle accelerator able to produce collisions with unprecedented center of mass energies is needed.

The Large Hadron Collider(LHC) built at CERN is a proton-proton collider able to accelerate protons to an energy of 7TeV. In the resulting collisions, the quarks and gluons that make up the protons, called the partons, are able to collide at high enough energies to probe the 1 TeV scale.

A cross section is a value describing the likelihood of a physics process taking place. Colliding protons with anti-protons gives higher cross-sections, but at very high energies the difference is small. The difference in cross sections is more than made up for by the higher luminosity that can be achieved with proton-proton collisions.

$$\mathscr{L} = \frac{n_b^2 f_b}{\sigma_b^2 4\pi} \tag{1.1}$$

The LHC is designed to collide bunches of protons with $n_b = 10^{11}$ protons, with a frequency of $f_b = 40 MHz$ and a bunch width $\sigma_b = 15.9 \mu m$, bringing the design luminosity to approximately $10^{34} cm^{-2} s^1$.

The protons are extracted from a hydrogen plasma before being accelerated to 50MeV by a linear accelerator. Then they are further accelerated by the Proton

Synchrotron Booster(PSB) to an energy of 1.4GeV. They are then injected to the Proton Synchrotron, where the protons are ordered in bunches, further accelerated and injected to the Super Proton Synchrotron(SPS). In the SPS they are accelerated to an energy of 450GeV before being injected to the LHC, where the bunches are accelerated to 7TeV.
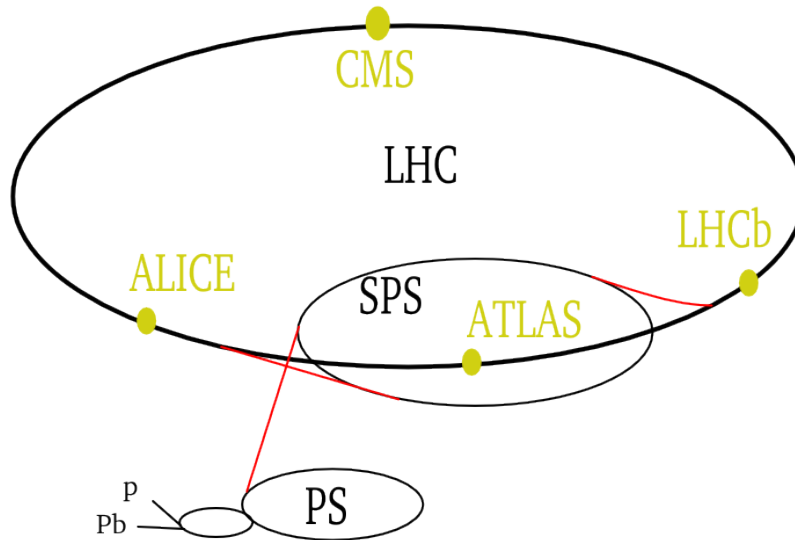

Figure 1.1: The accelerator chain at CERN.

The LHC bunches are made to cross at four points along the ring. At the beam crossing points, four detectors will examine the collisions produced in the LHC. Two general purpose detectors, CMS(Compact Muon Solenoid) and ATLAS(A Toroidal LHC AparatuS), and one dedicated B-physics experiment, LHCb are designed for proton-proton collisions.

The LHC will also produce lead-lead ion collisions. ALICE(A Large Ion Collider Experiment) is designed to investigate the heavy ion collisions.

Chapter 2

The ATLAS detector

The ATLAS detector(figure 2.1) has been designed to examine a wide range of processes, in the full energy range produced by the 14 TeV center of mass proton-proton collisions. The high luminosity and energy environment at the LHC places strong requirements on the detectors.

Most interesting processes involve the creation of unstable particles, decaying instantly after they have been created. These particles must be identified from the decay products. This process requires precise four-momentum measurement for a wide energy range, and good particle identification. Neutrinos and the possible SUSY dark matter candidates can not be directly observed in the detector. For this reason it is important that all the particles from an interesting decay are detected. Non-interacting particles can then be indirectly observe by missing transverse energy. In order to prevent particles from escaping detection, the interaction point must be enclosed by the detector.

Because of the high particle flux close to the interaction point, the detector must be able to withstand huge amounts of radiation.

The ATLAS detector is cylindrical in shape, consisting of a central barrel region, with an end-cap on each side. Cylindrical coordinates are used to describe the position of the detector elements, as well as particle momentum. The direction from the interaction point is give by an azimuthal angle, ϕ, and a polar angle, θ. Another way of describing the polar angle is to use the pseudo rapidity, $\eta = -ln(tan(\theta/2))$. The reason for this, is that particle production is almost constant as a function of η. The detector is also described by two planes, the $R\phi$ plane and the Rz plane. The z-axis runs along the center of the detector in the same direction as the beam line, R is the distance of a point to the z-axis.

The detector consists of three major subsystems:

- The inner tracker, located closest to the interaction point. A 2 Tesla solenoidal magnetic field bends the trajectories of charged particles. From the measurements in the inner detector, these trajectories are calculated.

- Electro magnetic and a hadronic calorimeters, located outside the inner detector. The calorimeters measure the energies of electrons, photons and
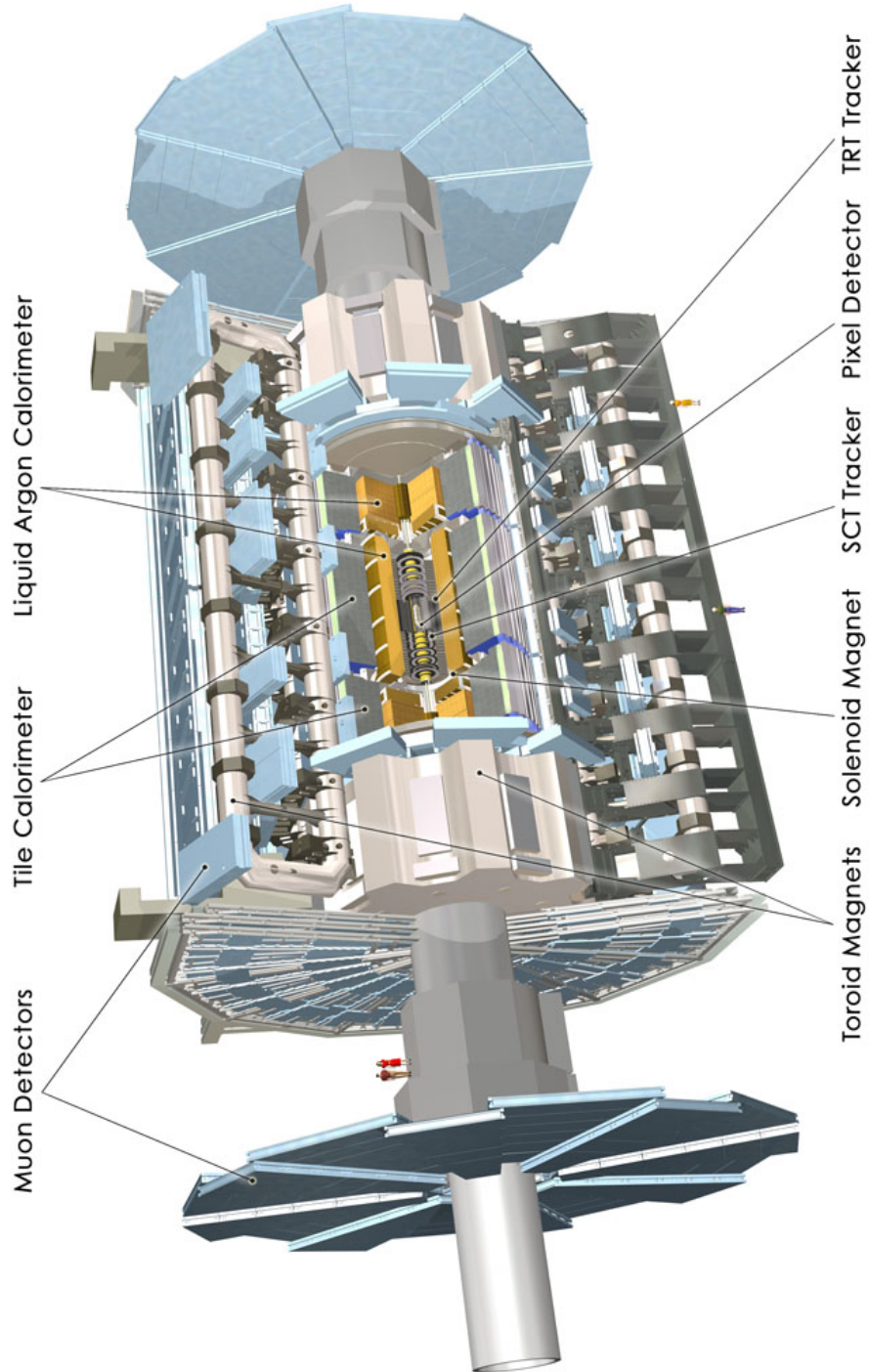
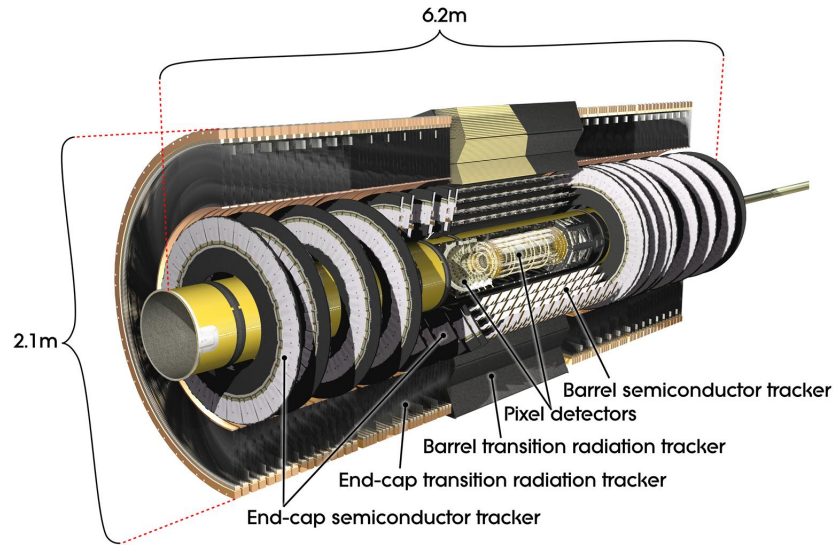Figure 2.1: The ATLAS detector(from the CERN web pages)

Figure 2.2: The ATLAS inner detector detector(from the CERN web pages)

hadrons.

- The muon spectrometer, located farthest away from the interaction point. The muon spectrometer measures the trajectories of muons.

2.1   The inner detector

The ATLAS inner detector [15–17] serves the purpose of measuring the trajectories of charged particles passing through it. A super conducting solenoid magnet produces a 2T magnetic field that bends the trajectories of charged particles in the $R\phi$ plane. From the amount of bending a particle trajectory undergoes, the momentum of the particle can be determined. The direction of the bending is used to determine the sign of the electric charge of the particle.

Precise measurements of the direction and momentum of the particles can be used to determine the location of the production vertex of the particle. Most particles are created at the interaction point, or the primary vertex. However, some particles, like taus and B-mesons, live long enough to create secondary vertices. Recognizing secondary vertices is an important part of identifying b-jets and tau leptons.

The inner detector is made up of three sub systems(figure 2.2), the pixel detector, the semi-conductor tracker(SCT) and the transition radiation tracker(TRT).

The pixel detector

The pixel detector is the detector located closest to the interaction point. It consists of three barrel layers and three discs in each end-cap. The barrel layers are located 4, 10 and 13 cm away from the beam-line. Because of its proximity to the interaction point, the measurements in the detector are very important in estimating secondary vertices.

The pixels are significaly smaller in the bending plane, this is to provide accurate measurements of the transverse momentum(p_T). The pixel detector covers the range $|\eta| < 2.5$, producing 3 measurements per track.

The pixel detector consists of 1456 silicon modules in the barrel region, and 288 modules in the end-cap. The silicon pixels gives a two dimentional read out, with a total of approximately 140 million channels, with a resolution of $\sim 10\mu m$ in the $R\phi$ plane and $\sim 115\mu m$ in the z direction(barrel layers) and R direction (end-cap layers).

The SCT

The SCT lies outside the pixel detector, in the radial range from 30 to 56 cm. The barrel region has four layers of silicon micro-strip sendsors, the end-caps have nine layers. The SCT cover the range $|\eta| < 2.5$. The barrel modules are rectangular, with two silicon wafers, one mounted on each side. Each silicon layer produce a one dimentional measurement, but measurements from the front and back side of the module can be combined to a two dimentional space point. The silicon strips provide less granularity than the pixels, but because they are located further away from the interaction point, less resolution is required. The average number of hits per track is eight, leading to four space points. The resolution of the SCT is $\sim 17\mu m$ in $R\phi$, and $\sim 580\mu m$ in R and z.

The TRT

The bulk volume of the inner detector is made up of the transition radiation tracker(TRT) It occupies the outermost part of the inner detector, with a radius from 56 to 107 cm. Like the other parts of the inner detector, it consists of a barrel and two end-cap regions. The TRT coverage is limited to $|\eta| < 2.0$.

The TRT consists of straw tube detectors. The inside of each tube is covered with a aluminum foil acting as a cathode. The tubes are filled with a xenon based gas mixture, and have a central anode wire. When a charge particle passes through a straw tube, the gas is ionized, the charge carriers drift in the electric field, and a current is read out at the end of the straw tube. The TRT contributes with an average of 35 one dimentional measurements, with a resolution of 130 μm, in the $R\phi$ plane per track.

In addition to these presision measure ments, the TRT can be used to identify electrons. Between the straws there is embedded ploypropylene foils and fibers. When a highly relativistic charged particle traverses a boundary between materials

with different dielectric constants, it can emit soft photons, called transition radia-
tion. Because electrons are light particles, they move closer to the light speed than
heavier particles with the same momentum, and emits more transition radiation.
Electrons can be seperated from charged pions for momentum up to around 100
GeV. The transition radiation interacts with the gas in the straw tubes, and produce
a higher read out than charged particles.

The inner detector porvides a momentum estimates with uncertaionties of $\Delta p_t / p_t =$
$0.04\% \oplus 2\%$, and approximately 15 μm for the impact parameters. The impact pa-
rameters are z and R estimates of the trajectory at the point where the trajectory is
closest to the z-axis. The uncertainties in the momentum estimates increases with
higher transverse momentum.

2.2 Calorimetry

The calorimeters measure the energy of electrons, photons and jets. The ATLAS
detector uses two types of calorimeters, an electromagnetic calorimeter(ECAL)
and a hadronic calorimeter(HCAL) [17].

Both calorimeters are sampling calorimeters. Thick absorbers are used to cre-
ate electromagnetic cascades in ECAL, and hadronic showers i HCAL. Between
the layers of absorbers, active detector material is used to sample the particle show-
ers.

The mean free path of a hadronic jet is about one order of magnitude greater
than the mean free path of an electron, the radiation length(x_0). For this reason,
ECAL is located closer to the interaction point than HCAL. Electrons and photons
then create showers that are absorbed before they enter HCAL.

ECAL

The electromagnetic calorimeter uses lead absorbers to create cascades, and liq-
uid argon as sampling material. The cascades are created by electrons emitting
bremsstrahlung photons($e \rightarrow e\gamma$), and photons converting to electron/positron
pairs($\gamma \rightarrow e^+ e^-$).

The thickness of the calorimeter is approximately 24 x_0 in the in the barrel
region, and 35 x_0 in the end caps. Before the electrons and photons reach the
ECAL, they pass through several radiation lengths(figure 3.4). To correft for the
energy loss in the inner detector, a liquid argon pre sampler is installed.

The ECAL covers the regiona up to $|\eta| < 3.2$, with uncertainties in the energy
measurement about $\sigma(E)/E = 10\%/\sqrt{E} \oplus 0.2\%$ [17].

HCAL

The HCAL uses iron plates as absorbers, and plastic scintillator tiles to sample the
showers. The hadronic showers are created through inelastic interactions between
the hadrons and the nuclei in the absorbers. This produces wide cascades of the

	a	b
barrel	56%	5%
end-cap	70%	6%
FCAL	70%	3%

Table 2.1: Parameters describing uncertainties in energy measurements [17].

secondary leptons and hadrons. At large η there is a very high level of hadronic radiation, that would damage the scintillator tiles. For this reason the end-caps use the radiation hard liquid argon as active sensor material.

In order to keep particles escaping detection to a minimum, a forward calorimeter(FCAL) is installed to detect particles at high $|\eta|$. The HCAL and FCAL cover the range $|\eta| < 4.8$.

The uncertainties in the energy measurements goes as $\sigma(E)/E = a/\sqrt{E} \oplus b$. The values for a and b differs in the different part of the calorimeter. Approximate values for the parameters are listed in table 2.1

2.3 The muon spectrometer

The outermost detector system in ATLAS is the muon spectrometer. Three barrel layers an four end-cap wheels are in place to measure the momentunm of muons. A air-core torroidal magnet produces a field that bends the particles in the Rz plane.

Two different detector types are used to register hits. The range $|\eta| < 2.0$ is covered by muon drift tubes. In the region $2.0 < |\eta| < 2.7$, there is a higher neutron-induced background. In this range cathode strip chambers are used, because they have a shorter responce time and higher granularity.

Resistiv plate chambers ($|\eta| < 1.05$) and thin gap chambers ($1.05|\eta| < 2.4$) are used aswell, mostly to provide information to the event trigger.

2.4 Trigger

The LHC produces proton-proton collisions at a rate of 40MHz. This produces an amount of data several orders of magnitude greater than what the ATLAS data aquisition system can handle. A maximum frequnzy of ~ 100 events per second can be stored for offline analysis.

A three level trigger trigger is used to select events that show signs of interesting physics processes. The first level is a hardware based trigger selecting events with large energy deposits in the calorimeters, hits in the muon chambers or missing p_T. The level-2 trigger is software based, and selects events based on a rudimentary analysys of regions of interest identified in level-1. The level-3 trigger does a preliminary reconstruction of the entire event, events that are selected by this trigger is stored for offline analysis.
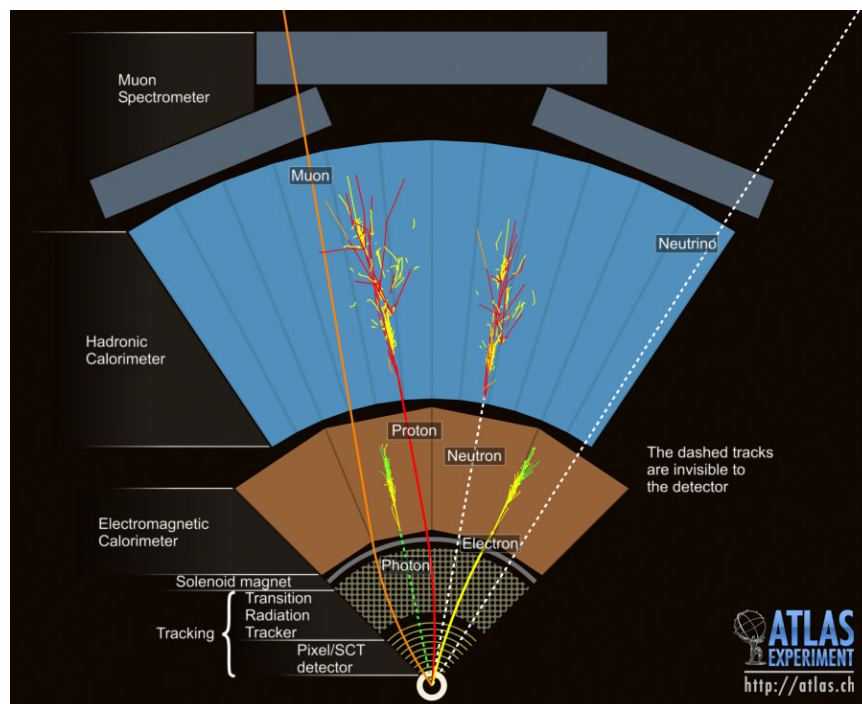
Figure 2.3: Examples ofr particls passing through the detctor(from the CERN web pages)

Chapter 3

Charged particles passing through matter

Understanding the trajectories of charged particles in the detector, and how these trajectories are impacted by the particle interacting with the detector material is important for tracking to be successful. In this chapter the most important aspects of how particles interact with matter, and how this relates to the ATLAS inner detector are discussed.

3.1 Particles passing through matter

3.1.1 Ionizing energy loss

Except for electrons, moderately relativistic particles passing through matter lose energy predominantly by ionization and atomic excitation. The mean ionization energy loss is described by the Bethe-Block equation [4]:

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{\beta^2}[\frac{1}{2}ln\frac{2m_ec^2\beta^2\gamma^2T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}], \qquad (3.1)$$

where T_{max} is the maximum kinetic energy the particle can lose in a single collision.

$$T_{max} = \frac{2m_ec^2\beta^2\gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2} \qquad (3.2)$$

The other constants and variables in the equation are described in table 3.1.

Ionizing energy loss is a stochastic process, but the standard deviation is small compared to the mean, so a deterministic approximation is normally used [6].

3.1.2 Radiative energy loss

When a charge particle with high momentum is deflected by another charged particle, like an atomic nucleus, it can decelerate by emitting a photon. This process
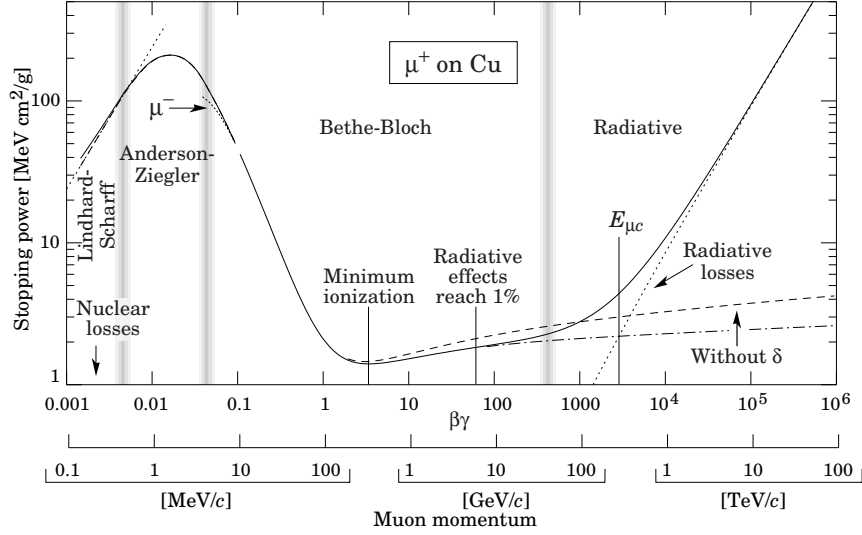
Figure 3.1: Stopping power for muons in copper. The central region is well described by the Bethe-Bloch equation (From [4]).

Symbol	Definition	Units or value
K/A	$4\pi N_A r_e^2 m_e c^2/A$	$0.307075 MeV g^{-1} cm^2$
N_A	Avogadro's number	$6.0221415 \times 10^2 3 mol^{-1}$
r_e	Classical electron radius	$2.817940325(28) fm$
$m_e c^2$	Electron mass $\times c^2$	$0.510998918(44) MeV$
M	Incident particle mass	MeV/c^2
Z	Atomic number of absorber	
A	Atomic mass of absorber	$g\ mol^{-1}$
ze	Charge of incident particle	
β		v^2/c^2
γ	Lorentz factor	$\frac{1}{\sqrt{1-v^2/c^2}}$
$\delta(\beta\gamma)$	Density effect correction to ionization energy loss	
T	Kinetic energy	MeV
I	Mean excitation energy	eV

Table 3.1: Constants and variables used in the Bethe-Bloch equation.

is known as bremsstrahlung. The energy emitted is proportional to $1/m^2$, where m is the mass of the incident particle [13]. Because of this dependence on mass, the only particle substantially affected by bremsstrahlung is the electron. Above a few tens of MeV the electron lose energy predominantly through bremsstrahlung, as seen in figure 3.2.
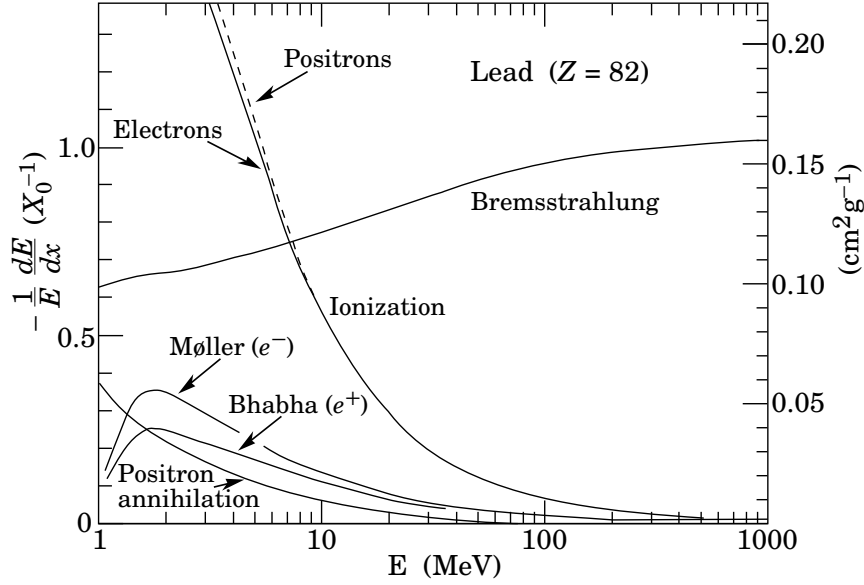

Figure 3.2: Fractional energy loss per radiation length for electron in lead as a function of energy (From [4]).

Bremsstrahlung is a stochastic process with large variance relative to the mean. In this case a deterministic approximation can not be used. This highly non-Gaussian process is modelled by the Bethe-Heitler probability density function [2].

$$f(z) = \frac{(-ln(z))^{c-1}}{\Gamma(c)}, \tag{3.3}$$

where

$$c = \frac{t}{ln(2)}, \tag{3.4}$$

and $\Gamma(c)$ is the gamma function.

The probability density function models the fractional energy loss, final energy(E_f) over initial energy(E_i).

$$z = \frac{E_f}{E_i} \tag{3.5}$$

t is the amount of material traversed measured in radiation lengths(x_0).

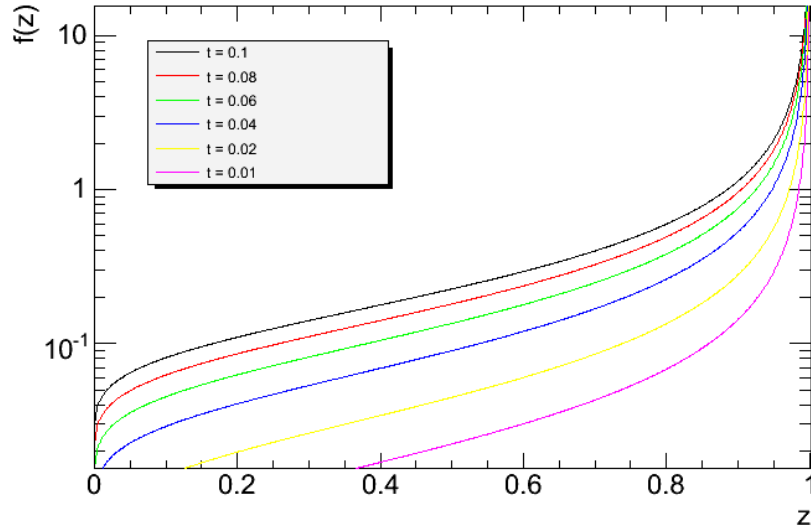$$t = \frac{x}{x_0} \tag{3.6}$$

Figure 3.3: The Bethe-Heitler distribution for different path lengths,t.

A radiation length is is a constant of a material connected with the probability of radiative energy loss for electrons, and the probability of photons converting to e^+e^- pairs. x_0 is the mean distance for a high-energy electron to lose all but $1/e$ of its energy by bremsstrahlung, and $7/9$ of the mean free path of high-energy photons converting to an electron/positron pair.

For muons and pions radiative energy loss can not be ignored for energies $\gtrsim$ $100\ GeV$, but in this case, similar to the ionizing energy loss case, a deterministic approximation is normally used.

3.1.3 Multiple scattering

Multiple scattering is the sum of many small deflections caused by Coulomb scattering as the particle traverses the detector material. This contributes very little to energy loss, but impacts the angle of the trajectory. In a given plane this process is roughly Gaussian, with a mean of zero and the width:

$$\theta_0 = \frac{13.6(MeV)}{\beta cp} z \sqrt{\frac{x}{x_0}} \left[1 + 0.038 ln \left(\frac{x}{x_0} \right) \right], \tag{3.7}$$

where βc is the speed of the particle, p the momentum, z the charge and $\frac{x}{x_0}$ the length of material traversed measured in radiation lengths(x_0) [4]. Large angles are underestimated by this Gaussian approximation, around 98% of the central distribution is covered.

3.1.4 Material in the ATLAS inner detector

Because of the high track density and frequency in the inner detector, quick read-out, radiation hardness and high granularity are important. The pixels ,SCT and TRT offer these features, but at the cost of a high amount of material in the detector. Compared to gaseous detectors, semi-conducting materials contribute more to the amount of material in terms of radiation lengths. On top of active sensor material, services like front-end electronics, power distribution and cooling are needed. Because of the size of the detector, support structures are needed. All of this add to the total amount of material a charged particle must traverse to pass through the tracking detector.
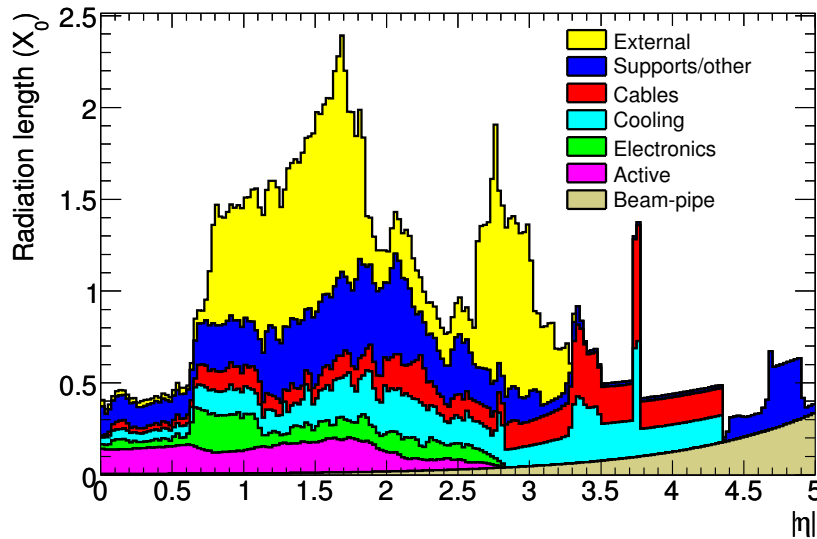

Figure 3.4: Accumulated material distribution in radiation lengths as function of $|\eta|$(From [17]).

Figure 3.4 shows the accumulated material distribution in radiation lengths as a function of the pseudo-rapidity, η.

$$\eta = -ln\left[tan(\theta/2)\right] \tag{3.8}$$

The material in figure 3.4 called external lies between the sensors in the inner detector and ECAL, so it does not affect the tracking performance.

The amount of material in the inner detector leads to a high probability for high energy electrons to lose a substantial amount of energy through bremsstrahlung. The stochastic nature of radiative energy loss leads to problems when the trajectories of electrons are to be calculated.

Chapter 4

Track reconstruction

Track reconstruction is the process of going from measurements in the detector to a description of the trajectories of particles. This task can be split into pattern recognition and track finding. The job of the pattern recognition algorithms is to associate a set of measurements with a track candidate. The track fitter tries to obtain the optimal estimate of the trajectory with uncertainties from the measurements provided by pattern recognition.

4.1 Track recognition

In the ATLAS inner detector track recognition is done in three steps: pre-processing, track finding and post-processing.

4.1.1 Pre-processing

In the pre-processing stage, the raw detector read-out is prepared for further analysis. For the pixels and the SCT this consists of identifying clusters. A cluster is the set of pixels or silicon strips where there have been an above threshold readout.

A pixel sensor gives the local two dimensional position for each pixel above threshold. Combining this two-dimential measurement with the known location of the module, the spatial position of the hit is calculated.

Each SCT layer gives a one dimensional position of a measurement in a silicon strip. From a cluster in a SCT layer, the most probable one dimensional position of a hit is calculated. Each SCT module consists of two layers of silicon micro-strips, the direction of the strips are orthogonally aligned. Combining the information from both layers give a two dimensional local coordinates for the hits.

From a read-out in a TRTstraw tube, the local radius describing the distance from a track to the anode in the center of the straw is measured. Hits in the straw tubes only produce a one dimential measurement, so a space-point can not be calculated.

4.1.2   Track finding

Track finding is the process identifying a subset of the detector hits as belonging to a track. The main track reconstruction method used ATLAS is to start with a few hits close to the beam line, called a track seed. The track seed consists of hits in the three pixel layers and the innermost SCT layer, taking advantage of the high granularity in the innermost scilicon tracker layers. The trajectory suggested by the track seed is then extrapolated through the SCT using a Kalman Filter. If there are enough hits within this narrow "road" through the SCT, this set of hits is considered a track candidate.

This method produces a high number of track candidates. A track fit is applied to the track candidates. The candidates are kept or discarded depending on a number of factors including the goodness of fit, the number of hits shared between candidates and the number of tracks with "holes". A hole is in this case a track passing through a sensor without creating a cluster.

The remaining track candidates are then extended through the TRT and refitted. If the refit with the TRT hits improves the track fit, all hits are used in the final track fit. If the fit with TRT hits does not improve the track fit, the TRT hits are associated with the track, but only the pixel and SCT hits are included in the final track

4.1.3   Post-processing

The post-processing stage uses the information from the track finding stage to calculate the position of the primary vertex, points of photon conversion and position of second vertices.

4.2   Track fitting

In a physics analysis, the position of the individual measurements of a track is not very interesting. However, with the help of a track fitter these measurements can be used to obtain information about the momentum, charge and trajectory of a particle.

Six parameters are needed to fully describe the position, momentum and charge of a particle. However, if a track is described at an arbitrary surface, the state of the track can be described by five parameters and a 5×5 covariance matrix describing the uncertainties. The surface often corresponds to a detector element, but it can also be a completely abstract construction.

For this to be successful a detailed description of how charged particles behave in the detector is crucial. This includes descriptions of the magnetic field, material effects and the geometry of the detector.

4.2.1   Track model

The track model is a mathematical model describing the trajectory of a charged particle. If the state of a track is described at a surface, the track model can be used to obtain an estimate of the state on a different surface.

In simple cases, the track model can be analytical. Examples of this is a vacuum without a magnetic field, where the tracks are straight lines, or a vacuum with a constant magnetic field, where the trajectories become helices.

The ATLAS inner detector does not have a homogeneous magnetic field, and it has a complex geometry. The ATLAS track model is therefor not analytical, but employs approximations and numerical methods.

4.2.2   The track parameters in the ATLAS inner detector

When a track is fitted, the five track parameters used are $(locX, locY, \phi, \theta, q/p)$. The track fitters examine the predicted state at the same surface as the detector element containing a measurement. $locX$ and $locY$ are the local Cartesian coordinates on the surface. ϕ and θ describe the direction of the particle, ϕ in the bending plane and θ in the non bending plane. q/p is the charge of the particle divided by the momentum. The transverse momentum of a particle determines the amount of bending the particle undergoes in the magnetic field and the charge determines the direction of the bending.

In the study of track fitter performance and invariant mass estimates in the following chapters, the tracks are evaluated at the perigee. The perigee is an abstract surface where the track is closest to the z-axis. In this case the track parameters are $(d_0, z_0, \phi, \theta, q/p)$. d_0 and z_0 are called the impact parameters. d_0 is the transverse distance to the z-axis, z_0 the position along the z-axis.

The parameters ϕ and d_0 are sensitive to information only in the bending plane($R\phi$), θ and z_0 only the non-bending plane(Rz). q/p depends on information in both planes. p_T measurements depend only on the $R\phi$ plane, the total momentum also requires an estimate of the angle θ.

$$p_t = p \, sin(\theta) \tag{4.1}$$

4.3   The Least squares method

The state of the trajectory at a surface, k, can not be measured directly, a measurement only contains information about the position of the trajectory at a specific surface. The measurements are, however, a function of the true state of the track [1, 6].

$$\mathbf{m}_k = h_k(\mathbf{x_{true}}) + \epsilon_k \tag{4.2}$$

$\mathbf{m}_k$ is here a vector containing the measurement at surface k, $\mathbf{x_{true}}$ is the true state of the track, and $\epsilon_{\mathbf{k}}$ is the noise vector corrupting the measurements. The noise

vector depends on the detector modules, it should have an expectation value of zero, and the uncertainties should be well known from calibration of the modules.

$$\langle \epsilon_k \rangle = 0, \ cov(\epsilon_k) \equiv \mathbf{V}_k \tag{4.3}$$

The function h_k simply maps the state vector of the track to the local coordinates of the measurement at the surface.

The process of track fitting is to find an optimal estimate of the true state vector, $\mathbf{x_{true}}$. In the simplest case only the initial state of the track, $\mathbf{x_0}$, is estimated.

In this case, the measurements are a function of the initial state of the track.

$$\mathbf{m}_k = f_k(\mathbf{x_0}) + \gamma_k \tag{4.4}$$

The function $f_k(\mathbf{x})$ uses the track model to predict the state vector at surface k, and maps this vector to local coordinates. γ_k is a combination of ϵ_k and uncertainties arising from material effects.

All the local measurement vectors, noise vectors and functions are collected in global vectors

$$\mathbf{m} = \begin{pmatrix} m_1 \\ . \\ . \\ m_n \end{pmatrix}, \ \mathbf{f} = \begin{pmatrix} f_1 \\ . \\ . \\ f_n \end{pmatrix}, \ \gamma = \begin{pmatrix} \gamma_1 \\ . \\ . \\ \gamma_n \end{pmatrix}, \tag{4.5}$$

the full set of measurements are now described as

$$\mathbf{m} = f(\mathbf{x_0}) + \gamma. \tag{4.6}$$

If the track model is approximately linear,

$$f(\mathbf{x_0}) = \mathbf{F}x_0 + c, \tag{4.7}$$

the expression becomes

$$\mathbf{m} = \mathbf{F}x_0 + c + \gamma, \tag{4.8}$$

where $\mathbf{F}$ is the Jacobian of the function f.

From the covariance matrix $\mathbf{V}$ of γ, a weight matrix, $\mathbf{W}$, can be found, describing the "quality" of the individual measurements.

$$\mathbf{W} \equiv \mathbf{V}^{-1} \tag{4.9}$$

The track fitter is a function that estimates the state vector($\mathbf{x}_0$), from the measurement vector($\mathbf{m}$), without introducing bias and with minimum variance,

$$\mathbf{\tilde{x}_0} = t(\mathbf{m}), \tag{4.10}$$

where $\tilde{x}_0$ is the fitted state vector.

The least squares method finds the state vector that minimizes the sum of squares of the weighted residuals $\mathbf{m} - f(\mathbf{x})$.

$$M(\mathbf{x}) = (\mathbf{m} - f(\mathbf{x}))^T W (\mathbf{m} - f(\mathbf{x})) \tag{4.11}$$

The weight matrix, $\mathbf{W}$, makes measurements with small variance contribute more than measurements with a large variance.

The estimate of the least squares method is

$$\tilde{\mathbf{x}} = (F^T W F)^{-1} F^T W (m - c), \tag{4.12}$$

with a covariance matrix

$$C(\tilde{\mathbf{x}}) = (F^T W F)^{-1}. \tag{4.13}$$

4.3.1   Properties of the least squares method

If $\mathbf{m}$ is unbiased, i.e. $\langle \epsilon \rangle = 0$, the least squares method is an unbiased estimator.

The Gauss-Markov theorem states that the least squares method is the unbiased estimator with the minimum variance given that:

- The measurements are unbiased.

- The experimental errors follow a Gaussian distribution.

- The track model is linear.

The least squares method is a consistent estimator. This means that the estimate converges to the true value as the number of measurements increases.

4.4   The Kalman filter

If the requirements listed above hold, the least squares method is the optimal estimator. When implemented as in the previous section it requires inversion of large matrices, which is a time consuming operation. The Kalman filter [1, 6] is a recursive formulation of the least squares method that improves the estimate sequentially. It reads in one measurement at the time, avoiding large matrices. When all the measurements are included, the estimate is equivalent to the one obtained by the global least square estimator.

For each measurement that is included the Kalman filter goes through two steps. The first step is to predict the state at a surface k, $\mathbf{x}_k$, from the state of the track at surface k-1, $\mathbf{x}_{k-1}$. To do this it uses the system equation

$$\mathbf{x}(s_k) \equiv \mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k. \tag{4.14}$$

The function f_k is the track model between the two layers, $\mathbf{w}_k$ is the process noise arising from material effects like multiple scattering and energy loss. Bias from material interactions are accounted for in the track model, so $\mathbf{w}_k$ is unbiased.

$$\langle \mathbf{w}_k \rangle = 0, \ cov(\mathbf{w}_k) \equiv \mathbf{Q}_k \tag{4.15}$$

If the track model is linear, the predicted state at layer k, given the state at k-1 is

$$\mathbf{x}_{k|k-1} = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_k, \tag{4.16}$$

where $\mathbf{F_k}$ again is the Jacobian of the track model. The predicted state depends on all the previous measurements.

Similar to the case of the least square method, the measurement can now be expressed as a function of the predicted state at the surface.

$$\mathbf{m}_k = \mathbf{h}_k(\mathbf{x}_{k|k-1}) + \epsilon_k \tag{4.17}$$

The function $\mathbf{h}_k$ maps $\mathbf{x}_k$ to local coordinates. For the ATLAS track parameters this is simply

$$\mathbf{h}_k \rightarrow \mathbf{H}_k = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.18}$$

Again ϵ is the measurement errors.

$$\langle \epsilon_k \rangle = 0, \ cov(\epsilon_k) \equiv V_k \equiv W_k^{-1} \tag{4.19}$$

The predicted covariance matrix is calculated using linear error propagation on the covariance matrix at surface k-1. Linear error propagation is a method of calculating how a transformation function changes the uncertainties of an estimate.

$$\mathbf{C}_{k|k-1} = \mathbf{F}_k \mathbf{C}_{k-1} \mathbf{F}_k^T + \mathbf{Q}_k \tag{4.20}$$

The term $\mathbf{Q}_k$, which is due to material effects, inflates the covariance matrix.

The next step is to update the estimate state with the measurement at surface k. This is called filtering. The state vector is combined with the measurement and becomes

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k(\mathbf{m}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}), \tag{4.21}$$

where $\mathbf{K}_k$ is the Kalman gain matrix

$$\mathbf{K}_k = \mathbf{C}_{k|k-1} H_k^T (\mathbf{V}_k + \mathbf{H}_k \mathbf{C}_{k|k-1} \mathbf{H}_k^T)^{-1}. \tag{4.22}$$

The updated covariance matrix becomes

$$\mathbf{C}_{k|k} = (\mathbf{I} - \mathbf{K_k H_k}) \mathbf{C}_{k|k-1}, \tag{4.23}$$

where I is the unity matrix. In this step the covariance matrix shrinks because of the additional information gained by the measurement.

If the Kalman filter starts far away from the beam pipe, reading in points as it works its way to the interaction point, the optimal estimate is reached at the surface of the innermost measurement. To obtain the optimal estimate on all surfaces, a second Kalman filter runs in the opposite direction. Information from both Kalman filters are used in the estimate at all surfaces, this way all the measurements are a part of the estimate.

4.5   Electron track fitting

One of the requirements for the linear least squares method to be the optimal estimator, is that all the stochastic processes encountered should follow a Gaussian distribution. This is often not the case in particle detectors. Energy loss by ionization and excitation follow a Landau distribution in thin layers and multiple scattering has non-Gaussian tails, the same can be true for the errors in the detector modules. In these cases the use of a linear estimator can be defended because the uncertainties in energy loss adds little to the covariance matrix of the estimate, and can be treated as a deterministic process. The distributions of multiple scattering and measurement errors are close to being Gaussian, and can be approximated as such.

This can not be said for the energy loss of electrons. Electrons lose energy predominately through bremsstrahlung. This stochastic, highly non-Gaussian process follow the Bethe-Heitler distribution(eq.3.3, fig.3.3).

The only way for bremsstrahlung energy loss to be included in a Kalman filter is to model it as a Gaussian. This can be done by using a Gaussian with the same mean and variance as the Bethe-Heitler distribution, but this is a crude approximation. In this case it is plausible that a non linear track fitter that takes the actual shape of the energy loss distribution into consideration can perform better than the Kalman filter.

4.5.1   The Gaussian sum filter

Athena, the ATLAS software framework, contains a track fitter specialized for electron tracks called the Gaussian-Sum Filter(GSF) [1, 8, 10, 11]. The GSF is a non-linear generalization of the Kalman filter.

The GSF deals with the energy loss of electrons by modelling the Bethe-Heitler distribution as a weighted sum of Gaussians with a fixed number of components, N_{BH}.

$$g(z) = \sum_{i=0}^{N_{BH}} g^i \varphi(z; z^i, var(z^i)) \tag{4.24}$$

z is the fractional energy loss, E_f/E_i and φ is the probability density function describing a Gaussian with mean z^i and variance $var(z^i)$. Each component also has a weight, g_i.

$$\sum_{i=0}^{N_{BH}} g_i = 1 \tag{4.25}$$

The Gaussian mixture is calculated by minimizing this distance between the cumulative distribution functions of the Bethe-Heitler distribution , $F(z)$, and the Gaussian sum modeling it, $G(z)$ [9].

$$D_{CDF} = \int_{-\infty}^{\infty} |F(z) - G(z)|dz \tag{4.26}$$

For a given number of components, the mean, weight and variance of the components are calculated as a function of the amount of material traversed in radiation lengths, $t = x/x_0$. In ATLAS, the default number of components used in the model is $N_{BH} = 6$.

In order to use this model in track fitting, the mean and variance of the components must be transformed from z to the track parameter q/p.

$$g(z) \to g(\Delta(q/p)) = \sum_{i=0}^{N_{BH}} g^i \varphi(\Delta(q/p); \Delta(q/p)^i, var(\Delta(q/p)^i)) \tag{4.27}$$

Like the Kalman filter, the state on a surface is predicted using all the previous measurements, and then updated with the measurement at the surface. In the case of the Kalman filter, the state on the surface is a multivariate Gaussian probability density function, described by a state vector and covariance matrix.

$$p(\mathbf{x_k}|\mathbf{M_{k-1}}) = \varphi(\mathbf{x}; \mathbf{x}_{k|k-1}, \mathbf{C}_{k|k-1}), \ M_{k-1} = (m_1.....m_{k_1}) \tag{4.28}$$

For the GSF, the state of the track is generally non-Gaussian, described by a sum of multivariate Gaussian components

$$p(\mathbf{x}_{k|M_k}) = \sum_{i=0}^{N_{SoS}} \pi_k^i \varphi(\mathbf{x}; \mathbf{x}_{k|k-1}^i, \mathbf{C}_{k|k-1}^i), \tag{4.29}$$

where N_{SoS} is the number of components describing the state on the surface. Each component in the Gaussian mixture has its own state vector with a corresponding covariance matrix and a weight.

In the prediction step, each component in the Gaussian mixture is updated with several independent Kalman filters. Each Kalman filter uses one of the components in the Gaussian mixture modeling the Bethe-Heitler distribution, $g(\Delta(q/p))$, to describe the energy loss.

$$p(\mathbf{x}_{k|k-1}) = \sum_{j=1}^{N_{BH}} \sum_{i=1}^{N_{SoS}^{(k-1)}} g^i \pi^i$$

$$\varphi(\mathbf{x}; \mathbf{x}_i + \mathbf{H}_{el}[\Delta(q/p)^j], \mathbf{C}^i + \mathbf{H}_{el}^T[var(\Delta(q/p)^j)]\mathbf{H}_{el}) \tag{4.30}$$

$\mathbf{H}_e l$ is a matrix that maps $\Delta(q/p)$ to the state vector and $var(\Delta(q/p))$ to the covariance matrix of the trajectory estimate.

$$\mathbf{H}_{el} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.31}$$

The result of this is that for each prediction step the number of components describing the state on the surface increases.

$$N_{SoS}^k = N_{SoS}^{k-1} N_{BH} \tag{4.32}$$

This quickly leads to an unmanageable number of components, so component reduction is needed. If the number of components exceeds a fixed maximum, the two closest components are collapsed into a single equivalent Gaussian. This process of merging the two closest components continues until the total number of components becomes less than a preset maximum. The default maximum number of components in ATLAS is $N_{SoS} = 12$.

When the state on a surface is predicted, the state vector and covariance matrix of each component is updated with the measurement at the surface in the same way as for the Kalman filter (eq 4.21 and eq. 4.23).

The implementation of the GSF in ATLAS is made to deal with non-Gaussian energy loss for electrons, but a GSF can be made to deal with any stochastic non-Gaussian process encountered in the fitting process, including measurements modeled as Gaussian mixtures. If the measurements are modeled as Gaussian mixtures, the number of components increase in the update step, not the prediction step.

Chapter 5

Performance of the Gaussian Sum Filter

In this chapter three different track fitters are used on Monte Carlo simulated data. The tracks are estimated at the perigee, the point where the trajectory is closest to the center of beam pipe. The estimates of single electron tracks for the different track fitters are compared to each other, as well as the estimated invariant mass for physics processes for events with pure electron final states. For a more in depth analysis of single electron tracks, and analysis of the same simulated data sample as used here, see [1].

5.1   Track fitters and final estimates

The Gaussian sum filter is compared with two Kalman filters.

- Kalman filter - Bethe-Bloch(KF-BB): Energy loss is treated as a deterministic process. The Bethe-Bloch equation is used to calculate the energy loss between two surfaces. Radiative energy loss is disregarded, witch means that energy loss is under-estimated when the fitter is used on electron tracks.

- Kalman filter - Bethe-Heitler(KF-BH): Energy loss is treated as a stochastic process modeled by a single Gaussian. The model uses the first and second moment of the Bethe-Heitler distribution as mean and variance [14].

$$\langle z \rangle = \langle \frac{E_f}{E_i} \rangle = e^{-t} \tag{5.1}$$

$$var(z) = e^{-t(ln3/ln2)} - e^{-2t} \tag{5.2}$$

In analysis, the variance and shape of an estimate is very often disregarded. A final estimate is needed, a numerical value to be used for cuts and histogramming. For a Gaussian estimate the probability amplitude reaches its maximum at the mean

27

value. The distribution is also symmetric around this value, so the mean value is clearly the best estimate.

The estimates from the GSF are non-Gaussian. The mean and the value with maximum probability amplitude is generally not the same. In the following, two values are used as final estimates of the GSF, the mean and the mode, the most probable value, of the distribution.

The mean of the distribution can be obtained simply as a weighted sum of the means of the components. However, similar to the component reduction step in the track fitting process, the Gaussian components can be merged into a single Gaussian [8].

$$\mathbf{x} = \sum_{i=1}^{N_c} q^i \mathbf{x}^i \tag{5.3}$$

$$\mathbf{C} = \sum_{i=1}^{N_c} q^i \mathbf{C}^i + \sum_{i=1}^{N_c} \sum_{k>i} q^i q^k (\mathbf{x}^i - \mathbf{x}^k)(\mathbf{x}^i - \mathbf{x}^k)^T \tag{5.4}$$

$\mathbf{x}^n$, $\mathbf{C}^n$ and q^n is the mean vector, the covariance matrix and the weight of component n, N_c is the total number of components. Although information about the shape of the estimate is lost, a single covariance matrix describing the uncertainties greatly simplifies the analysis.

The mode estimate uses the Newton-Raphson method to obtain the mode of the q/p track parameter. The track model takes energy loss into account by adding a $\Delta(q/p)$ to the q/p parameter, so q/p is generally the track parameter least resembling a Gaussian. The mean of the other parameters are used.

There is no trivial way of getting an accurate variance describing the distribution around the mode, so the estimate is collapsed into a single point. Unless specifically mentioned, the estimate used in this chapter is the mode of the probability distribution function produced by th GSF.

5.2 $J/\psi \rightarrow e^+ e^-$

The J/ψ is a meson consisting of a charm and an anti-charm quark. It has a mass of 3096.916 ± 0.011 MeV and a width of 93.4 ± 2.1 keV [4]. It can decay to an electron and a positron, with a branching ratio of $BR(J/\psi \rightarrow e^+ e^-) = (5.94 \pm 0.06) \times 10^{-2}$.

5000 J/ψ decaying to an electron and a positron have been generated, and ATLAS full simulation has been used to determine the response of the inner detector. The Monte Carlo truth momentum distributions of the leading and trailing momentum of the J/ψ decays are shown in figure 5.1. The leading lepton is the lepton with the highest momentum in a decay, the trailing the one with the lowest momentum.
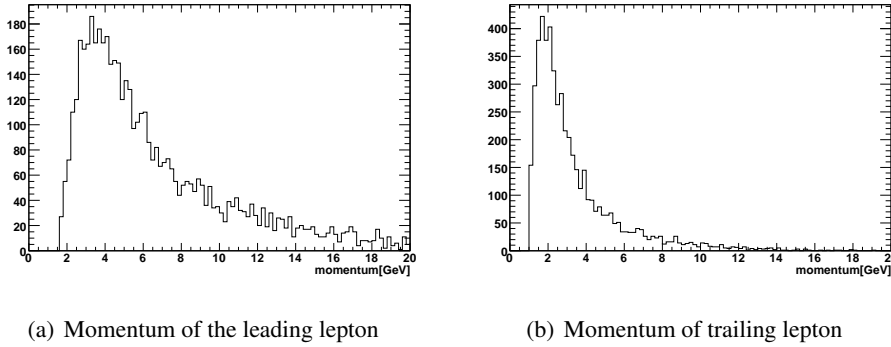
(a) Momentum of the leading lepton   (b) Momentum of trailing lepton

Figure 5.1: True simulated momentum of the electrons and positrons in the J/ψ decay.

5.2.1 Single electron/positron tracks

To determine the resolution of the collapsed estimates, the normalized residuals,

$$\frac{\mathbf{x}_{Rec} - \mathbf{x}_{truth}}{\mathbf{x}_{truth}}, \tag{5.5}$$

are calculated for each event. To get a resolution with sensitivity to the true value, the symmetric half width around the origin of the distribution of normalized residuals is used. The half width containing 68% of the events is the 1σ resolution. This resolution describes the width of the core of the distribution. The 2σ half width, containing 95% of the entries, is also sensitive to tails in the distribution.

The simulated J/ψ decays occur at the ideal interaction point, $(d0, z0) = (0,0)$. For this reason, the residuals are not normalized and half widths are calculated from standard residuals. Figure 5.2 shows the normalized residuals for the track parameters, except $d0$ and $z0$, where standard residuals are shown.

The q/p resolutions have been histogrammed separately to show how the momentum resolution decreases with momentum. Figure 5.3 shows the difference between the two estimates from the GSF.

Table 5.1 lists the 1σ and 2σ resolutions of the track parameter q/p and the momentum. The residual momentum distributions are shown in figure 5.4. It is interesting to note that the Kalman filter disregarding radiative energy loss has the second best q/p resolution for both the leading and trailing lepton. However, the resolution of the estimated momentum is more important in a physics analysis, and in this regard, the Bethe-Bloch Kalman filter does not compare well to the other track fitters.

The GSF mode gives the best 1σ resolution for both q/p and the momentum, but the 2σ resolution shows that it has a long tail. For the 2σ momentum resolution, the Bethe-Heitler Kalman filter performs best.

Table 5.2 shows the half width of the residuals of $d0$ and $z0$, and the half widths of the normalized residuals of ϕ and θ. The GSF improves the 1σ resolution of the
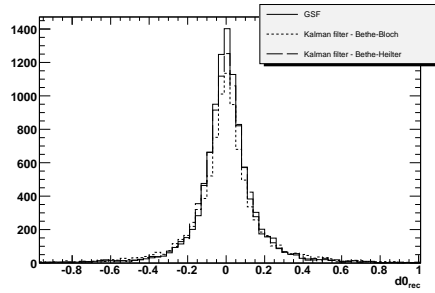
(a) *d0* distribution
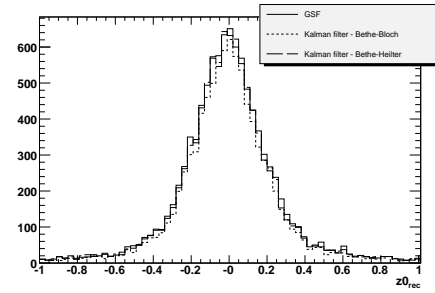
(b) *z0* distribution
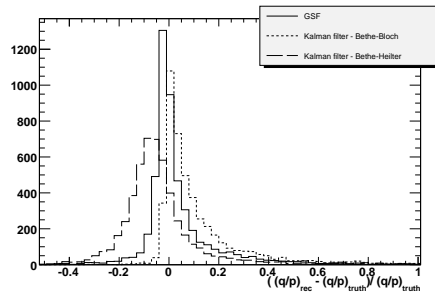
(c) Normalized residuals for q/p for the leading lepton
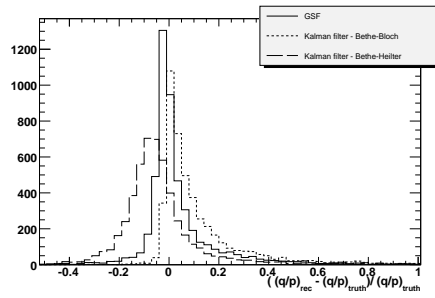
(d) Normalized residuals for q/p for the trailing lepton
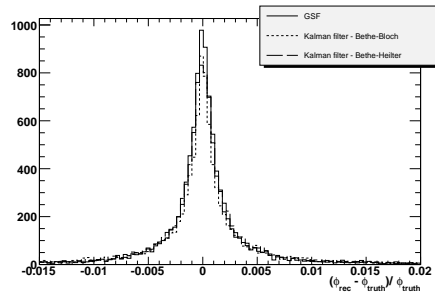
(e) Normalized residuals for ϕ
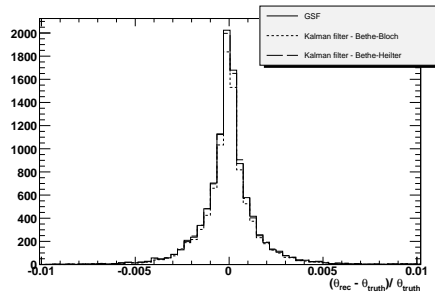
(f) Normalized residuals for θ

Figure 5.2: The track parameters for electrons and positrons from a J/ψ-decay.
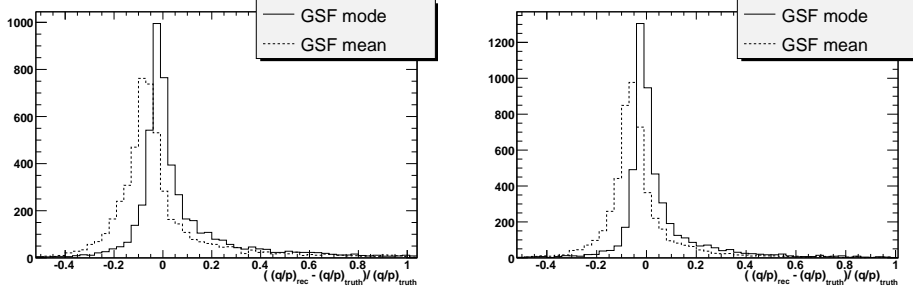
(a) Normalized Residuals of the leading lepton

(b) Normalized residuals of the trailing lepton

Figure 5.3: Normalized residuals for q/p for the GSF mode and mean.
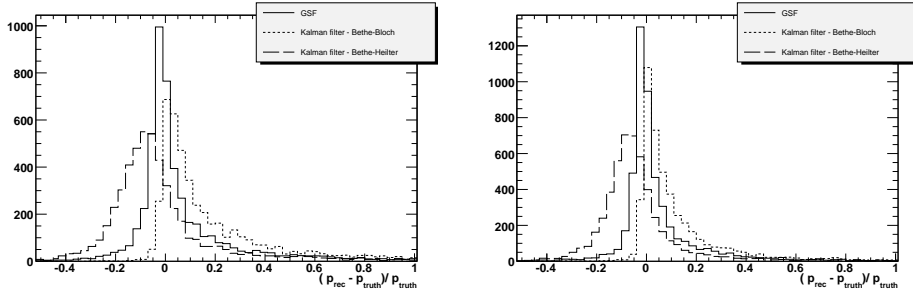
(a) Normalized Residuals of the leading lepton

(b) Normalized residuals of the trailing lepton

Figure 5.4: Distribution of momentum residuals

	$(q/p)_l$	$(q/p)_t$	p_l	p_t
GSF - Mode	$1\sigma = 0.0342$	$1\sigma = 0.0266$	$1\sigma = 0.119$	$1\sigma = 0.0738$
	$2\sigma = 0.589$	$2\sigma = 0.328$	$2\sigma = 1.00$	$2\sigma = 0.592$
GSF - Mean	$1\sigma = 0.0696$	$1\sigma = 0.0553$	$1\sigma = 0.149$	$1\sigma = 0.106$
	$2\sigma = 0.367$	$2\sigma = 0.233$	$2\sigma = 0.687$	$2\sigma = 0.357$
KF - BH	$1\sigma = 0.0786$	$1\sigma = 0.0625$	$1\sigma = 0.169$	$1\sigma = 0.127$
	$2\sigma = 0.373$	$2\sigma = 0.247$	$2\sigma = 0.649$	$2\sigma = 0.349$
KF - BB	$1\sigma = 0.0542$	$1\sigma = 0.0322$	$1\sigma = 0.222$	$1\sigma = 0.122$
	$2\sigma = 0.766$	$2\sigma = 0.370$	$2\sigma = 1.156$	$2\sigma = 0.604$

Table 5.1: Half widths of the normalized residuals of momentum and (q/p). p_l and $(q/p)_l$ refer to the leading lepton, p_t and $(q/p)_t$ to the trailing.

	d_0	z_0	ϕ_0	θ
GSF	$1\sigma = 0.111$	$1\sigma = 0.222$	$1\sigma = 0.000840$	$1\sigma = 0.00106$
	$2\sigma = 0.500$	$2\sigma = 0.930$	$2\sigma = 0.0113$	$2\sigma = 0.00356$
KF - BH	$1\sigma = 0.121$	$1\sigma = 0.221$	$1\sigma = 0.000928$	$1\sigma = 0.00105$
	$2\sigma = 0.495$	$2\sigma = 0.903$	$2\sigma = 0.0113$	$2\sigma = 0.00350$
KF - BB	$1\sigma = 0.139$	$1\sigma = 0.214$	$1\sigma = 0.000962$	$1\sigma = 0.00104$
	$2\sigma = 0.605$	$2\sigma = 0.830$	$2\sigma = 0.0142$	$2\sigma = 0.00332$

Table 5.2: (Normalized) residuals for track parameters.

parameters that rely on information in the bending plane, ϕ, $z0$ and q/p. The track parameters using information only in the non bending plane, θ and $d0$ have a similar resolution for all the track fitters.

5.2.2   Invariant mass reconstruction

The GSF produces better estimates of the momentum than the Kalman filter. In this section the invariant mass of the J/ψ is calculated from tracks reconstructed in the inner detector.

The energy of the decaying particle is

$$E^2 = m^2 + \mathbf{p}^2. \tag{5.6}$$

The energy and momentum is conserved, and can be reconstructed from the electron and positron produced in the decay,

$$E^2 = (\sum E_e^i)^2 = (\sum m_e)^2 + |\sum \mathbf{p}_e|^2, \ \mathbf{p} = \sum \mathbf{p}_e^i, \tag{5.7}$$

where E_e, m_e and $\mathbf{p}_e$ are the energy, mass and momentum of the particle. In this case $m_e \ll E_e$, and can be disregarded.

$$E^2 \approx |\sum \mathbf{p}_e|^2 = p_{ex}^2 + p_{ey}^2 + p_{ez}^2 \tag{5.8}$$

The reconstructed mass of the decayed particle then becomes

$$m = \left[\left(\sum \sqrt{p_{ex}^2 + p_{ey}^2 + p_{ez}^2} \right)^2 - \left(\sum p_{ex} \right)^2 - \left(\sum p_{ey} \right)^2 - \left(\sum p_{ez} \right)^2 \right]^{1/2} \tag{5.9}$$

Figure 5.5 shows the reconstructed invariant mass of the J/ψ. The resolution of the invariant mass is, like earlier, determined by the symmetric half width of the normalized residuals of the mass estimates. The normalized residuals are shown in figure 5.6 for the two Kalman filters and the GSF mode estimate, and figure 5.7 for the two GSF estimates.

The resolutions for the different estimators are listed in table 5.3. The GSF estimate coming from the mode of the momentum distribution is a clear improvement on the 1σ resolution. The GSF mean estimate produces the best 2σ resolution.
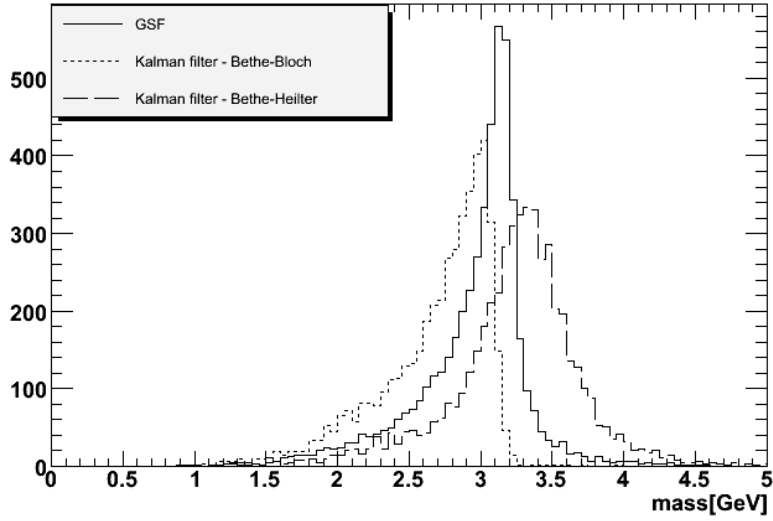
Figure 5.5: Reconstructed invariant mass of J/ψ.
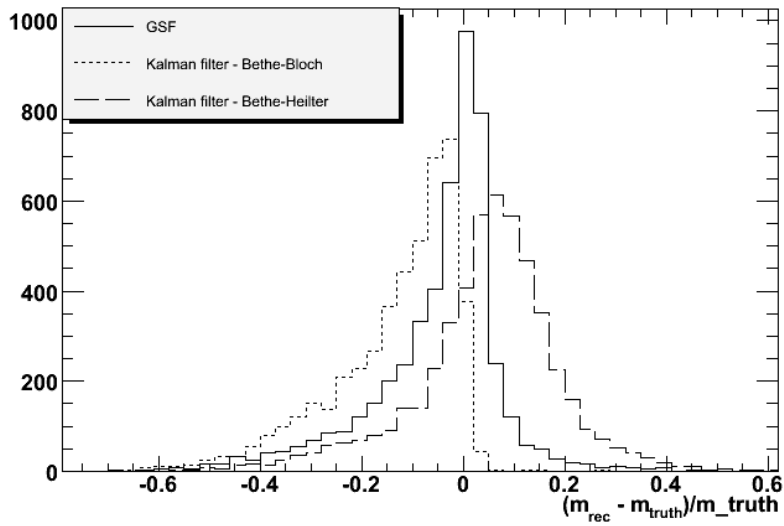
Figure 5.6: Normalized residuals of J/ψ invariant mass.
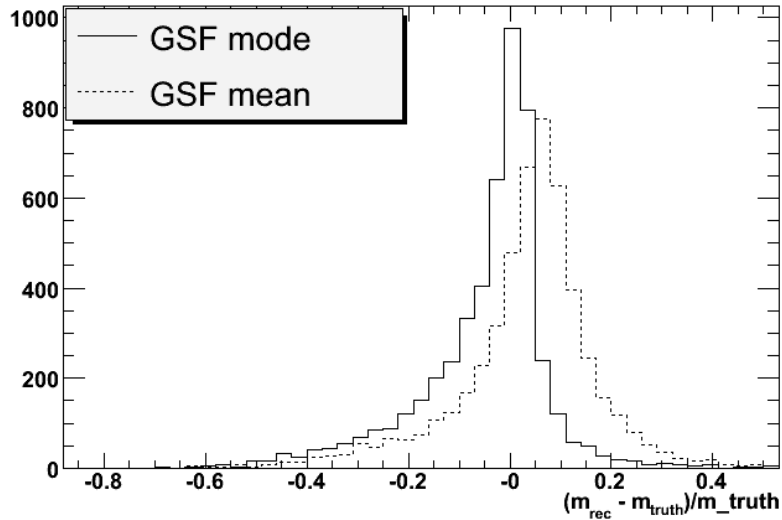
Figure 5.7: Normalized residuals of J/ψ invariant mass.

Half width	1σ	2σ
GSF mode	0.092	0.343
GSF mean	0.120	0.317
KF(BH)	0.138	0.329
KF(BB)	0.160	0.384

Table 5.3: $1\sigma(68\%)$ and $2\sigma(95\%)$ resolution of J/ψ invariant mass

5.3 $H \to ZZ^* \to e^+e^+e^-e^-$

In order to add boson and fermion masses to the standard model without breaking gauge invariance and the renormalizability of the Lagrangian, a mechanism that breaks the electro weak symmetry is needed. This mechanism is the Higgs mechanism [5]. This mechanism predicts the existence of the Higgs boson, the only particle in the standard model not yet discovered.

Experiments at LEP have put a lower bound on the Higgs mass of $m_H = 114$ GeV. The ATLAS detector will search for the Higgs boson in the mass range $114 GeV < m_H < 1000 GeV$.

The branching ratios of the decay modes of the Higgs boson are shown in figure 5.8. The Higgs only couples directly to massive particles. The coupling to fermions is proportional to m_f, the coupling to the massive vector bosons is proportional to m_V^2. For this reason the WW and ZZ decay modes are dominant if the Higgs mass is large enough to allow these decays.

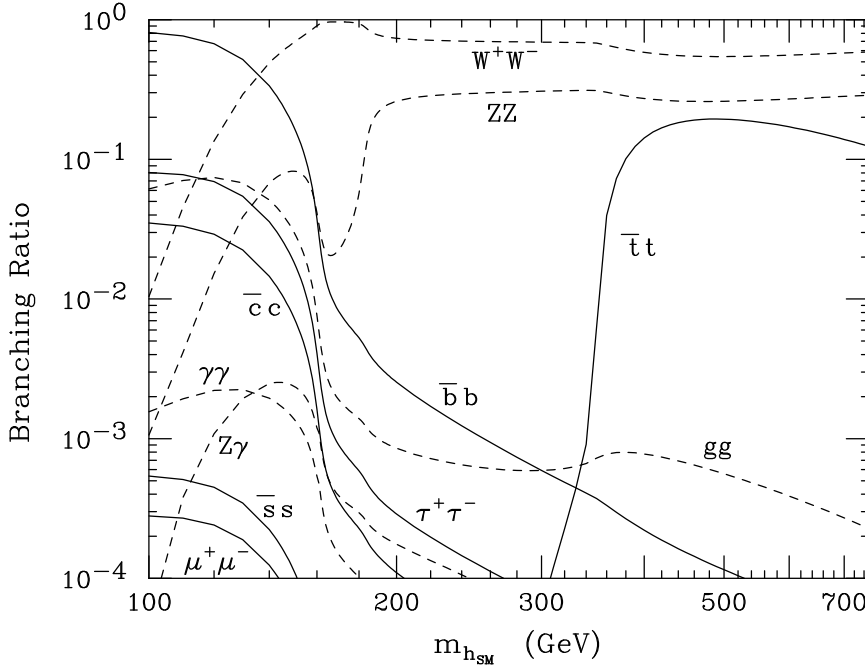Due to the clean final state and low background, the $H \to ZZ \to 4l^\pm$ decay is

Figure 5.8: Branching ratios of the Higgs boson as function of the Higgs mass, from [4]

.

called the "gold plated mode". The Higgs boson can decay to ZZ^*, where Z^* is an off shell Z boson, so this decay is also allowed below the $m_H > 2m_Z$ threshold. For m_H above $\sim$ 130 GeV, this is the most promising discovery mode for the Higgs boson.

The width of the Higgs boson increases rapidly with m_h [7], and for a heavy Higgs boson, the experimental mass resolution is dominated by the intrinsic width of the Higgs, not the detector resolution.

5.3.1   Invariant mass reconstruction

20,000 $H \to ZZ^* \to e^+e^-e^+e^-$ events have been generated to test the performance of the GSF, with $m_H = 130 GeV$. At this mass, the resolution of m_H is dominated by the uncertainties in the tracker. The tracks are reconstructed with the three track fitters and the true simulated momentum distributions for the four leptons are shown in figure 7.10.

Figure 5.10 shows the normalized residuals of the reconstructed momentum of the four leptons. Table 5.4 shows the resulting half widths. As in the case of J/ψ decays, the decreasing resolution with rising momentum is clear.

The best 1σ resolution is provided by the mode of the GSF estimate, but the Bethe-Heitler Kalman filter still has a better 2σ half width. However, in this case
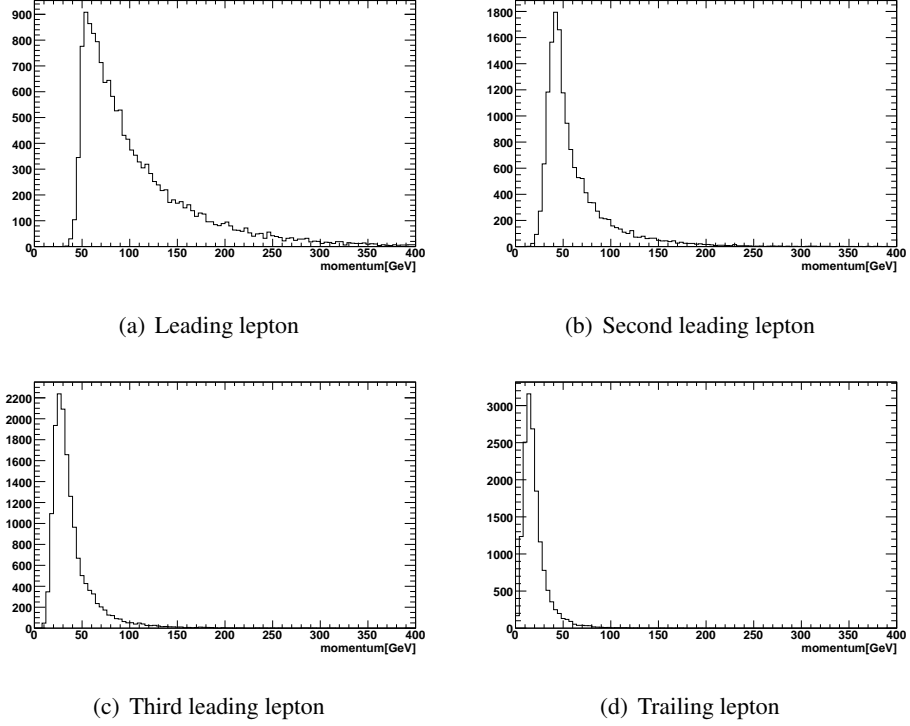
(a) Leading lepton

(b) Second leading lepton

(c) Third leading lepton

(d) Trailing lepton

Figure 5.9: Monte Carlo truth distribution of the momentum of the four leptons

	p_1	p_2	p_3	p_4
GSF - Mode	$1\sigma = 0.144$	$1\sigma = 0.106$	$1\sigma = 0.0882$	$1\sigma = 0.0709$
	$2\sigma = 2.10$	$2\sigma = 1.74$	$2\sigma = 1.48$	$2\sigma = 1.16$
GSF - Mean	$1\sigma = 0.178$	$1\sigma = 0.144$	$1\sigma = 0.127$	$1\sigma = 0.108$
	$2\sigma = 1.79$	$2\sigma = 1.54$	$2\sigma = 1.26$	$2\sigma = 1.00$
KF - BH	$1\sigma = 0.189$	$1\sigma = 0.155$	$1\sigma = 0.140$	$1\sigma = 0.123$
	$2\sigma = 1.87$	$2\sigma = 1.61$	$2\sigma = 1.36$	$2\sigma = 1.00$
KF - BB	$1\sigma = 0.273$	$1\sigma = 0.207$	$1\sigma = 0.173$	$1\sigma = 0.135$
	$2\sigma = 3.25$	$2\sigma = 2.71$	$2\sigma = 2.10$	$2\sigma = 1.66$

Table 5.4: Normalized residuals for the momentum. p_1 is the momentum of the hardest lepton, p_4 of the softest.

(a) Leading lepton

(b) Second leading lepton
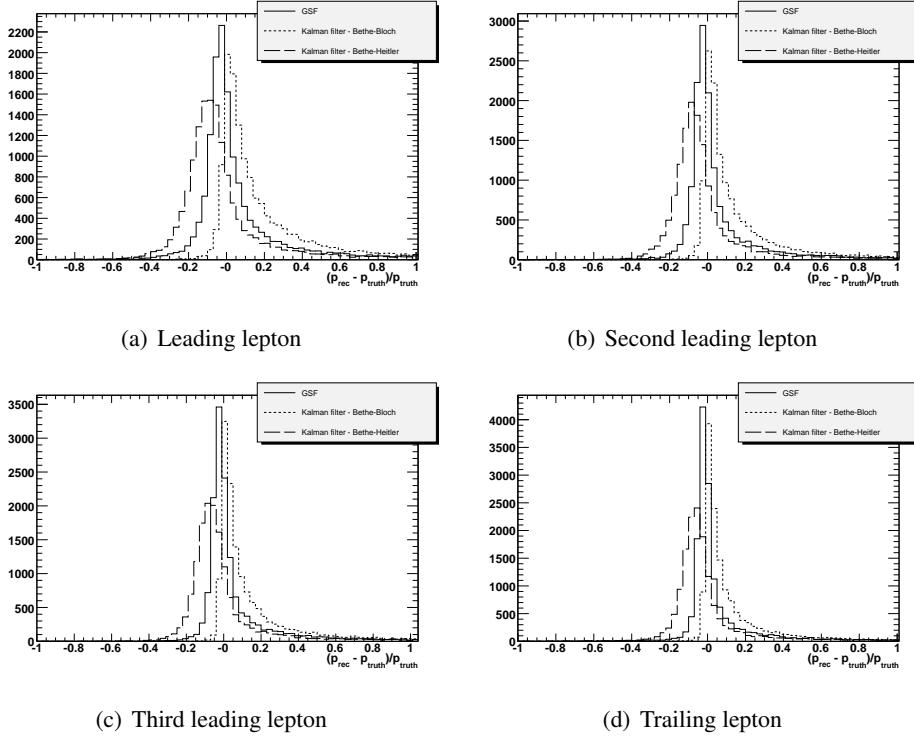

(c) Third leading lepton

(d) Trailing lepton

Figure 5.10: Normalized residuals of the momentum of the four leptons

the GSF mean estimate produce a better 2σ resolution than all the other estimates.

The reconstructed invariant mass distributions are shown in figure 7.11, normalized residuals in figure 5.12. The peak of the distribution of the GSF mode estimate is close the true value, and table 5.5 shows it has the narrowest 1σ half width.

	m_h
GSF - Mode	$1\sigma = 0.118$
	$2\sigma = 0.402$
GSF - Mean	$1\sigma = 0.126$
	$2\sigma = 0.364$
KF - BH	$1\sigma = 0.131$
	$2\sigma = 0.360$
KF - BB	$1\sigma = 0.211$
	$2\sigma = 0.473$

Table 5.5: Normalized residuals of the reconstructed Higgs boson mass.
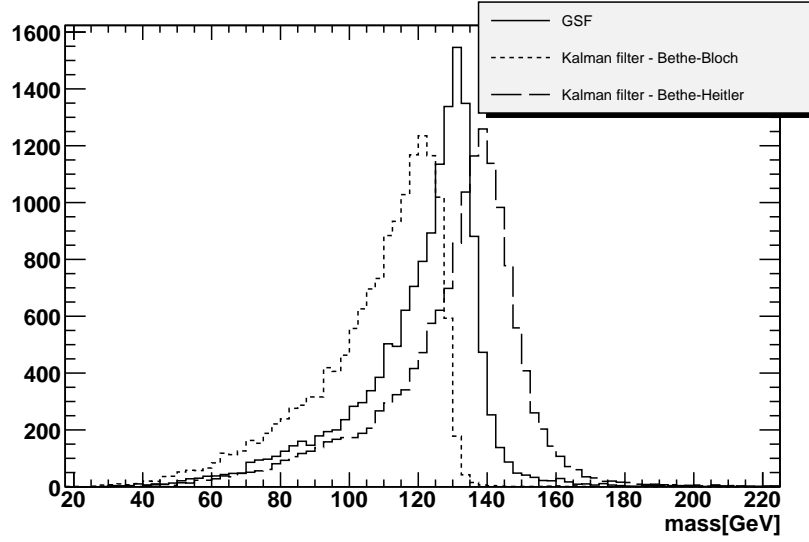
Figure 5.11: Reconstructed invariant mass of the Higgs boson
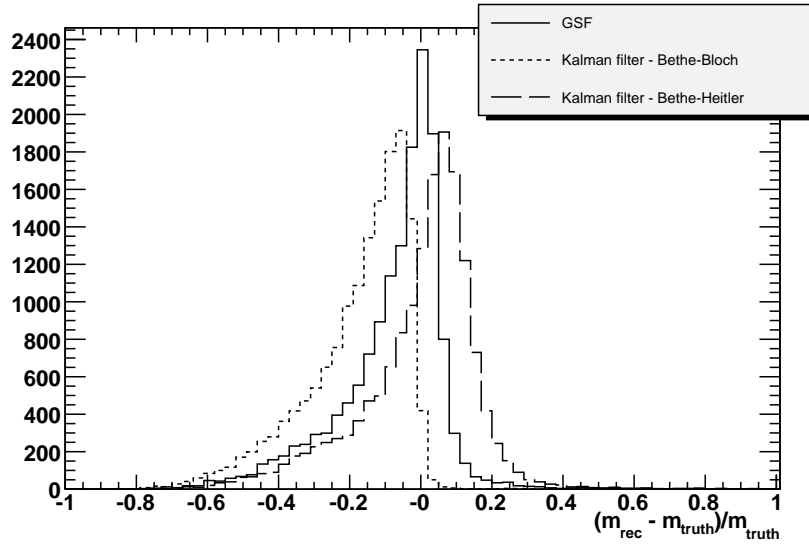
Figure 5.12: Normalized residuals of the reconstructed Higgs boson mass

Chapter 6

Calculating a probability distribution function of the invariant mass

6.1 Invariant mass of tracks described by Gaussian mixtures

Taking the mean or mode of the probability distribution describing the tracks discards a lot of the available information. It is plausible that a method for calculating the invariant mass keeping all information in the probability distribution functions describing the tracks can improve the invariant mass estimates. The method discussed here calculates a probability distribution function for the invariant mass, from tracks described as Gaussian mixtures.

6.1.1 Calculating a probability distribution function of the invariant mass

The method discussed here assumes that Gaussian track estimates transform to Gaussian invariant mass distributions. This is not strictly the case, but it leads to a straight forward way of calculating the invariant mass distribution.

This section will describe a way of transforming a set of n track estimates into a probability distribution function describing the invariant mass. The invariant mass distribution becomes a sum of single variate Gaussians, where each component is fully described by a mean, variance and a weight.

In order to calculate a probability distribution function from tracks described as Gaussian mixtures, a method of transforming Gaussian tracks to a Gaussian estimate of the invariant mass is needed. A Gaussian tracks is described in full by a vector containing the estimated mean of the track parameters($\vec{\mu}_i$) and the corresponding covariance matrix($\mathbf{C}_i$). When calculating the invariant mass, only the parameters containing information about the momentum and direction of the track

are relevant. For the ATLAS track parameters, only ϕ, θ and (q/p), with the corresponding 3×3 covariance matrix, are needed in the calculation.

$$\vec{\mu}_i = [\phi_i, \theta_i, (q/p)_i] \tag{6.1}$$

$$\mathbf{C}_i = \begin{pmatrix} \sigma^2_{\phi_i} & cov(\phi_i, \theta_i) & cov(\phi_i, (q/p)_i) \\ cov(\theta_i, \phi_i) & \sigma^2_{\theta_i} & cov(\theta_i, (q/p)_i) \\ cov((q/p)_i, \phi_i) & cov((q/p)_i, \theta_i) & \sigma^2_{(q/p)_i} \end{pmatrix} \tag{6.2}$$

The set of n tracks can be described by a single Gaussian with a $n \times 3$ mean vector, $\vec{\mu}$, and a $(n \times 3) \times (n \times 3)$ covariance matrix, $\mathbf{C}$.

$$\vec{\mu} = [\vec{\mu}_1, \vec{\mu}_2 ... \vec{\mu}_n] \tag{6.3}$$

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_{12} & . & . & & . \\ \mathbf{C}_{21} & \mathbf{C}_2 & . & . & & . \\ . & . & . & . & & . \\ . & & . & \mathbf{C}_{(n-1)} & \mathbf{C}_{(n-1)n} \\ . & & . & \mathbf{C}_{n(n-1)} & \mathbf{C}_n \end{pmatrix} \tag{6.4}$$

The matrices $\mathbf{C}_{ij}$ in $\mathbf{C}$ are matrices describing the covariance between two tracks. This information, as well as better track estimates, can be found using a vertex fitter. If a set of tracks originate from a common vertex, the position of this vertex is estimated. The tracks are then refitted with the estimated vertex used as a measurement. The fact that all tracks sharing a common vertex are used in estimating this fake measurement, leads to covariance between the tracks. A GSF vertex fitter has not yet been implemented in ATLAS, so the covariance matrix of the n tracks becomes block diagonal, with no covariance between tracks.

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & 0 & . & . & & . \\ 0 & \mathbf{C}_2 & . & . & & . \\ . & . & . & . & & . \\ . & & . & \mathbf{C}_{(n-1)} & 0 \\ . & & . & 0 & \mathbf{C}_n \end{pmatrix} \tag{6.5}$$

Using the assumption that Gaussian tracks transform to a Gaussian invariant mass estimate, all that is needed is to calculate a mean and variance of this estimate. The mean of the invariant mass is calculated from the estimated energy and momentum of the reconstructed particle.

$$W = h(\vec{\mu}) = \sqrt{E^2 - \left(p_x^2 + p_y^2 + p_z^2\right)} \tag{6.6}$$

The energy and momentum is calculated from the estimated track parameters. In the relativistic approximation, the electron mass can be disregarded.

$$p_x = \sum_{i=1}^{n} p_{xi} = \sum_{i=1}^{n} q_i \left(q/p\right)_i^{-1} cos(\phi_i) sin(\theta_i) \tag{6.7}$$

$$p_y = \sum_{i=1}^{n} p_{yi} = \sum_{i=1}^{n} q_i \left(q/p\right)_i^{-1} sin(\phi_i) sin(\theta_i) \tag{6.8}$$

$$p_z = \sum_{i=1}^{n} p_{zi} = \sum_{i=1}^{n} q_i \left(q/p\right)_i^{-1} cos(\theta_i) \tag{6.9}$$

$$E = \sum_{i=1}^{n} E_i = \sum_{i=1}^{n} \sqrt{p_{xi}^2 + p_{yi}^2 + p_{zi}^2} \tag{6.10}$$

The variance of the invariant mass is calculated using linear error propagation.

$$var(W) = J\mathbf{C}J^T, \tag{6.11}$$

where J is the $1 \times (n \times 3)$ Jacobian of the transformation function.

$$J = \nabla W \tag{6.12}$$

Using the chain rule multiple times leads to analytic expressions for the partial derivatives.

$$\frac{dW}{d\phi_i} = \frac{1}{2W} \left[2E \frac{1}{2E_i} \left(2p_{xi} \frac{dp_{xi}}{d\phi_i} + 2p_{yi} \frac{dp_{yi}}{d\phi_i} + 2p_{zi} \frac{dp_{zi}}{d\phi_i} \right) \right.$$
$$\left. - 2p_x \frac{dp_{xi}}{d\phi_i} - 2p_y \frac{dp_{yi}}{d\phi_i} - 2p_z \frac{dp_{zi}}{d\phi_i} \right] \tag{6.13}$$

$$\frac{dW}{d\theta_i} = \frac{1}{2W} \left[2E \frac{1}{2E_i} \left(2p_{xi} \frac{dp_{xi}}{d\theta_i} + 2p_{yi} \frac{dp_{yi}}{d\theta_i} + 2p_{zi} \frac{dp_{zi}}{d\theta_i} \right) \right.$$
$$\left. - 2p_x \frac{dp_{xi}}{d\theta_i} - 2p_y \frac{dp_{yi}}{d\theta_i} - 2p_z \frac{dp_{zi}}{d\theta_i} \right] \tag{6.14}$$

$$\frac{dW}{d\left(q/p\right)_i} = \frac{1}{2W} \left[2E \frac{1}{2E_i} \left(2p_{xi} \frac{dp_{xi}}{d\left(q/p\right)_i} + 2p_{yi} \frac{dp_{yi}}{d\left(q/p\right)_i} + 2p_{zi} \frac{dp_{zi}}{d\left(q/p\right)_i} \right) \right.$$
$$\left. - 2p_x \frac{dp_{xi}}{d\left(q/p\right)_i} - 2p_y \frac{dp_{yi}}{d\left(q/p\right)_i} - 2p_z \frac{dp_{zi}}{d\left(q/p\right)_i} \right], \tag{6.15}$$

where $i = 1, ..., n$.

The transformation function, $h(\vec{\mu})$, and its Jacobian are dependent on the number of tracks involved in the transformation. However, using this formulation the method can be implemented to work with any number of tracks.

As discussed earlier, the Gaussian sum filter describes the track as a weighted sum of Gaussian components, where each component is described by a state vector, a covariance matrix and a weight.

To transform a set of tracks described by Gaussian mixtures, the method above is repeated several times. Each iteration produces a Gaussian estimate of the invariant mass using one component from each of the tracks involved. This estimate then gets a weight, calculated from the track components used

$$g_i = \prod_{j=1}^{n} \delta_j,$$

(6.16)

where g_i is the weight of the invariant mass component, and δ_j the weight of the track component.

One Gaussian mass component is calculated for every possible combination of track components. If each track has n_{SoS} components, the resulting invariant mass estimate gets

$$n_W = \prod_{j=1}^{n} n_{SoS}$$

(6.17)

components.

The full probability distribution of the invariant mass is the sum of the components, each described by a mean, W, variance and weight.

$$f(x) = \sum_{i}^{n_W} \frac{g_i}{\sqrt{2var(W_i)\pi}} exp\left[-(x - W_i)/(2var(W_i))\right]$$

(6.18)

6.1.2   Other estimates

In addition to calculating the probability distribution function of the invariant mass, a probability distribution function describing the square of the invariant mass is calculated. The reason for this is to test if the transformation from track parameters to the square of the invariant mass is more linear than the transformation to the invariant mass. This is accomplished simply by changing the transformation function and its Jacobian in the method described above.

Using equations 5.3 and 5.4, the Gaussian components in the probability distribution functions describing the tracks can be merged into a single Gaussian around the mean. From these Gaussian track estimates, a single Gaussian probability distribution function of the invariant mass can be calculated. The mean of this estimate is the same as the GSF - Mean invariant mass estimate used in the previous chapter, but in this case the variance of the estimate is known.

6.2   Simulating a probability distribution of the invariant mass

To test the calculated probability distribution of the invariant mass, a way to simulate the distribution is needed.

To simulate the invariant mass distribution from Gaussian tracks, all that is needed is to pull out covariant, Gaussianly distributed random numbers from the tracks involved and calculate the invariant mass. The method used for getting covariant random numbers, is to fill a vector $\vec{x_g}$ with Gaussianly distributed random numbers, multiply this vector with the Cholesky decomposed covariance matrix and then add the vector containing the mean of the track parameters, $\vec{\mu}$ [4].

$$\vec{x} = \vec{\mu} + \mathbf{L}\vec{x_g} \tag{6.19}$$

Cholesky decomposition is a method of obtaining a lower triangular matrix, $\mathbf{L}$, from a symmetric and positive definite matrix,

$$\mathbf{C} = \mathbf{L}\mathbf{L}^T. \tag{6.20}$$

The covariance matrix, $\mathbf{C}$, satisfies these criteria. A matrix element, l, of the matrix $\mathbf{L}$ can be calculated from the matrix elements, c, in $\mathbf{C}$ [12],

$$l_{ii} = \sqrt{\left(c_{ii} - \sum_{k=1}^{i-1} l_{ik}^2\right)} \tag{6.21}$$

$$l_{ij} = \left(c_{ji} - \sum_{k=1}^{i-1} l_{jk}l_{ik}\right)/l_{ii}, \tag{6.22}$$

where $i = 1, ..., n$ and $j = i + 1, ...n$.

Because the full track mixtures do not have one covariance matrix, but a set of matrices, getting covariant random numbers is more tricky. The method used is to choose one component each cycle, and then get covariant random numbers from this component. The component is chosen based on its weight, using a random number from a uniform distribution. For Gaussian mixtures, this method is not optimal, but it appears to be a good approximation.

Figures 6.1, 6.2 and 6.3 show the simulated and calculated invariant mass distributions. Each simulation consists of 500,000 Monte Carlo simulated invariant masses.

From figure 6.2, where the invariant masses are calculated from Gaussian tracks, it is clear that the simulated distributions differ from the calculated ones. The simulated distributions are not Gaussians, generally they have a positive skewness, i.e. the tails of the distributions are heavier to the right. The mode of the simulated distribution also generally differ from the calculated distributions.

In the cases where the distributions come from tracks described by Gaussian mixtures, figure 6.1 and 6.3,the difference between the simulated and calculated distributions is generally smaller. The reason for this is that the invariant masses are calculated from the components of the track mixtures, who generally have a smaller variance than the approximated single Gaussian around the mean.
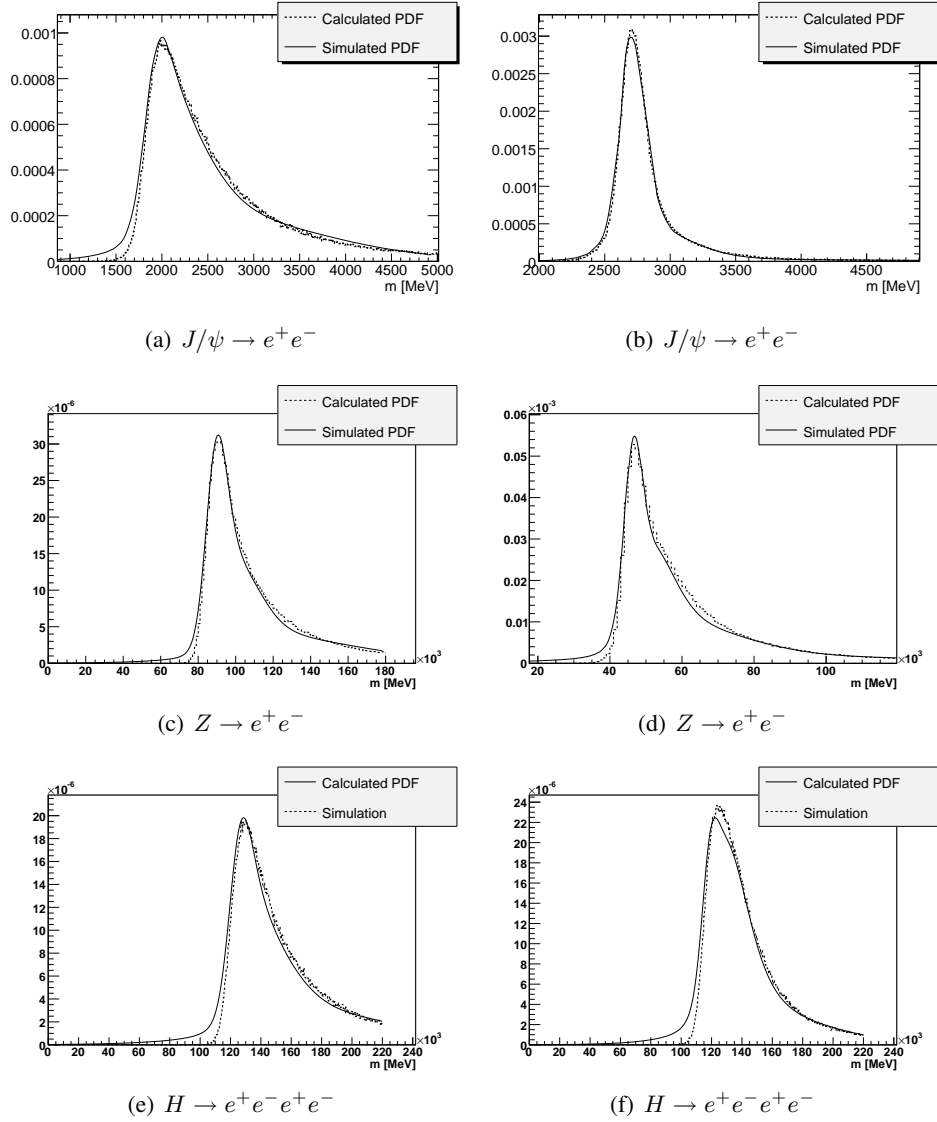
(a) $J/\psi \rightarrow e^+ e^-$

(b) $J/\psi \rightarrow e^+ e^-$

(c) $Z \rightarrow e^+ e^-$

(d) $Z \rightarrow e^+ e^-$

(e) $H \rightarrow e^+ e^- e^+ e^-$

(f) $H \rightarrow e^+ e^- e^+ e^-$
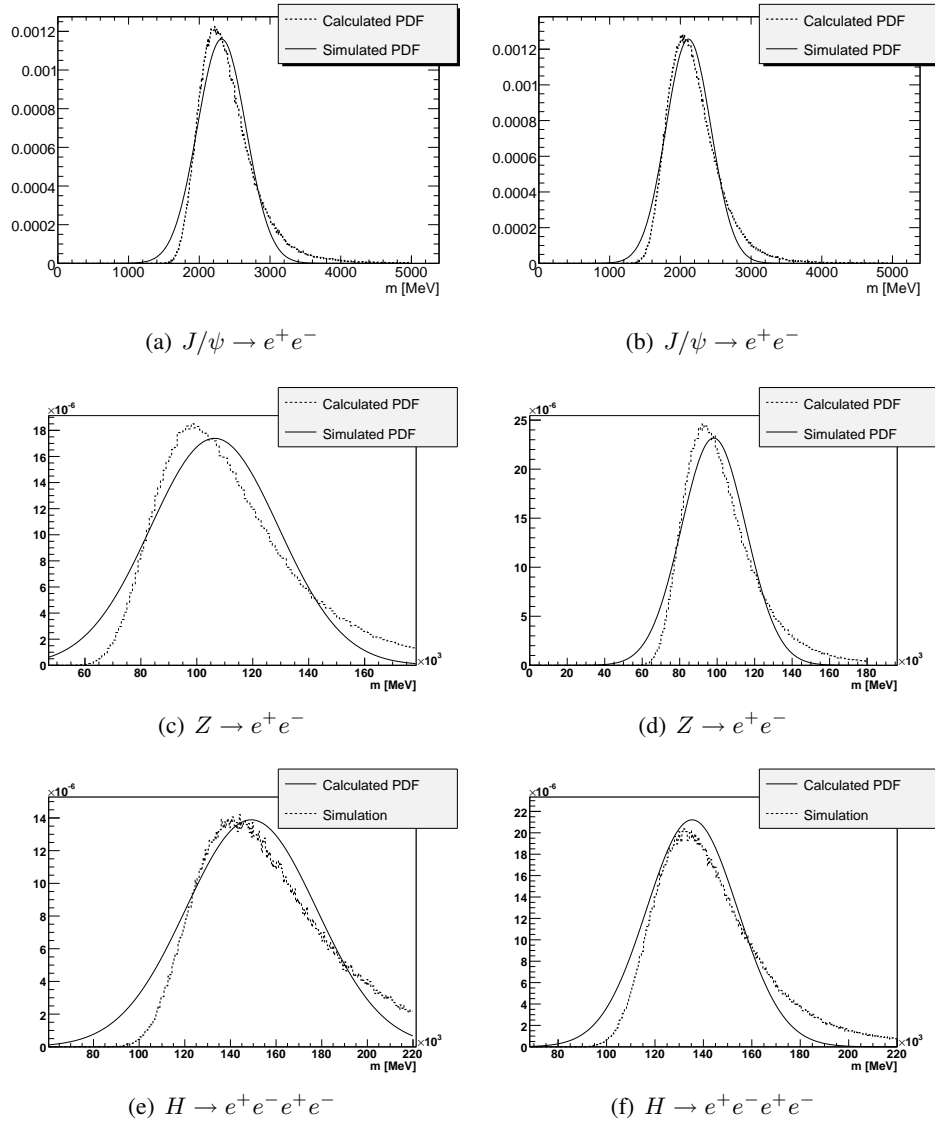
Figure 6.1: Example of calculated and simulated invariant mass probability distributions calculated from GSF track estimates.
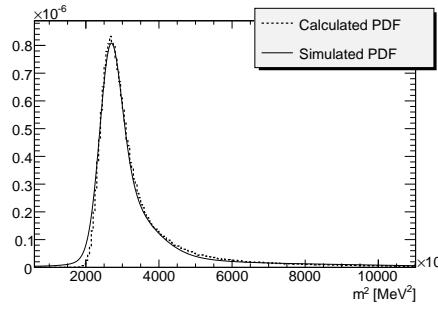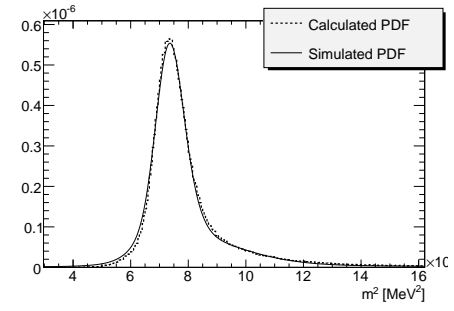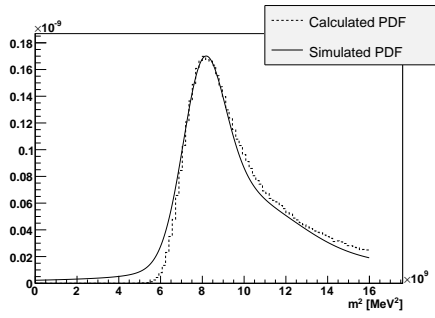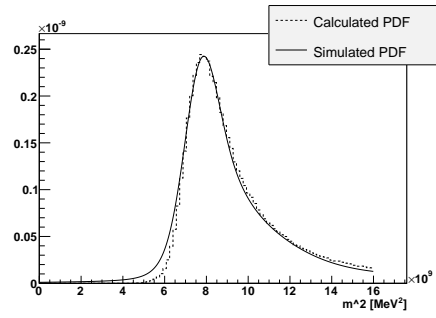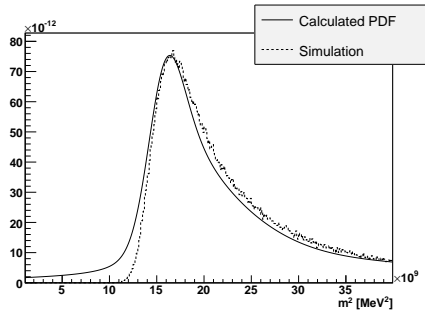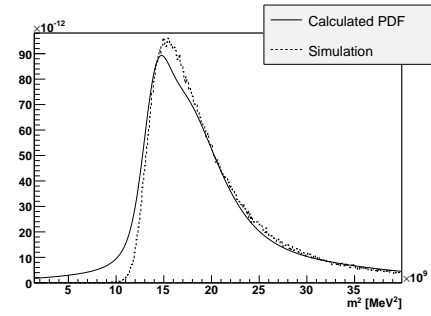
(a) $J/\psi \to e^+e^-$

(b) $J/\psi \to e^+e^-$

(c) $Z \to e^+e^-$

(d) $Z \to e^+e^-$

(e) $H \to e^+e^-e^+e^-$

(f) $H \to e^+e^-e^+e^-$

Figure 6.2: Example of calculated and simulated invariant mass probability distributions calculated from Gaussian track estimates.

(a) $J/\psi \rightarrow e^+e^-$

(b) $J/\psi \rightarrow e^+e^-$

(c) $Z \rightarrow e^+e^-$

(d) $Z \rightarrow e^+e^-$

(e) $H \rightarrow e^+e^-e^+e^-$

(f) $H \rightarrow e^+e^-e^+e^-$

Figure 6.3: Example of calculated and simulated probability distributions of the squared invariant mass, calculated from GSF track estimates.

Chapter 7

Invariant mass reconstruction

To test the new invariant mass estimate, two physics processes are considered, $J/\psi \rightarrow e^+e^-$ and $H \rightarrow ZZ^* \rightarrow e^+e^-e^+e^-$.

The physics events have been generated from proton-proton bunch crossings using PYTHIA. Geant has then been used to simulate material effects and detector response. The digital detector response and the Monte Carlo truth is stored in RDO(Raw Data Object) files. The RDO's are normally converted to ESD'a(Event Summary Data) and AOD's(Analysis Object Data) files, that are smaller in size and uses an abstract description of the particles in the event instead of raw detector read out. Most physics analysis is done on ntuples produced by processing AOD's or ESD's, and replicas of the datasets are generally available at several grid-sites.

The curent implementation of th GSF, however, works directly on RDO's. Because of the large size of the RDO files, they are normally not replicated, and very often stored on tape. This makes the RDO datasets hard to access by the distributed analysis tools. The Higgs data sample used in this chapter does not contain as many events as desiered.

The package InDetRecExample has then been used through Athena to fit tracks using the GSF on the inner detector response. The track fitter is configured to treat all tracks as electrons, leading to a bias on all non-electron tracks.

The standard approach of the GSF is to collapse the track mixtures to a single Gaussian after the track is fitted, so an algorithm, GSFPerigeeValidation, has been developed by Anthony Morley to extract the full track mixture at the perigee. Full mixture information from the tracks, the collapsed estimate of the track and Monte Carlo truth obtained from the RDOs are stored in ntuples.

In the analysis, electron and positron tracks are identified from Monte Carlo truth information. From the Monte Carlo truth, the invariant mass is reconstructed, and if the reconstructed mass is sufficiently close to the true mass of the decaying particle, the mass is estimated using the reconstructed tracks.

The fact that the events evaluated in this chapter are generated from proton-proton bunch crossings, unlike the case in chapter 5, leads to large amounts of tracks per event. An average LHC bunch crossing produces ~ 20 inelastic inter-

actions per event. Selecting tracks based on the invariant mass reconstructed from Monte Carlo truth alone is not a stringent enough cut. The Monte Carlo truth of the $z0$ track parameter of the involved particles are used as a crude estimate of the vertex. If the particles involved in the invariant mass reconstruction appear to originate from different vertices, the event is cut.

From events that pass these cuts, the invariant mass is reconstructed from the fitted tracks. Probability distribution functions of the invariant mass and the square of the invariant mass are calculated from the track mixtures. The track mixtures are collapsed, and the probability distribution function arising from these Gaussian track estimates are calculated. These track estimates are compared to the estimated invariant mass arising from the mode of the track mixtures, the estimate with consistently the best 1σ resolution in chapter 5.

7.1   Final estimates

To obtain resolution quantities from the invariant mass estimates, the calculated probability distributions must be collapsed into a single point. Similar to the case of the (q/p) track parameter, the mode and mean of the invariant mass distribution is calculated and used as the final estimate.

The mean of a Gaussian mixture can be obtained simply by taking the weighted average of the mean of the components.

$$m_{mean} = \frac{\sum_i g_i \mu_i}{\sum_i g_i} \tag{7.1}$$

To calculate the mode, the Newton-Raphson method has been implemented to obtain the root of the first derivative of the probability distribution function. The Newton-Raphson method takes an initial guess of the mode(x_0) as input, and uses the first and second derivative of the probability distribution to improve on this initial guess iteratively.

$$x_{i+1} = x_i - \frac{f'(x)}{f''(x)} \tag{7.2}$$

Figure 7.1 shows the first and second derivatives of an invariant mass distribution. If the mode finder is to be successful, the initial guess has to be within a region where the first iteration moves the initial guess in the direction towards the mode. The mean of the Gaussian mixture is often outside this region, and using it as the initial guess has a low success rate.

The maximum probability amplitude of a Gaussian component can be calculated from the weight and the variance of the component.

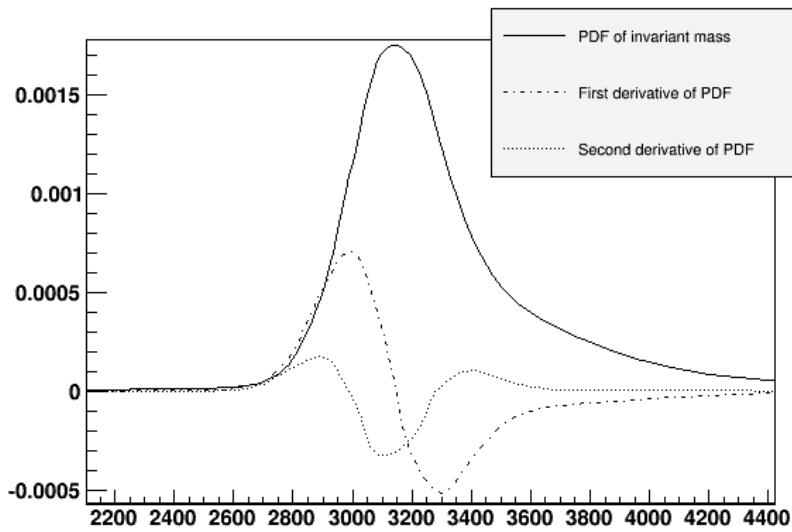$$A_{max} = \frac{g}{\sigma\sqrt{2\pi}} \tag{7.3}$$

Figure 7.1: PDF of invariant mass with first and second derivatives. The first and second derivatives have been scaled up.
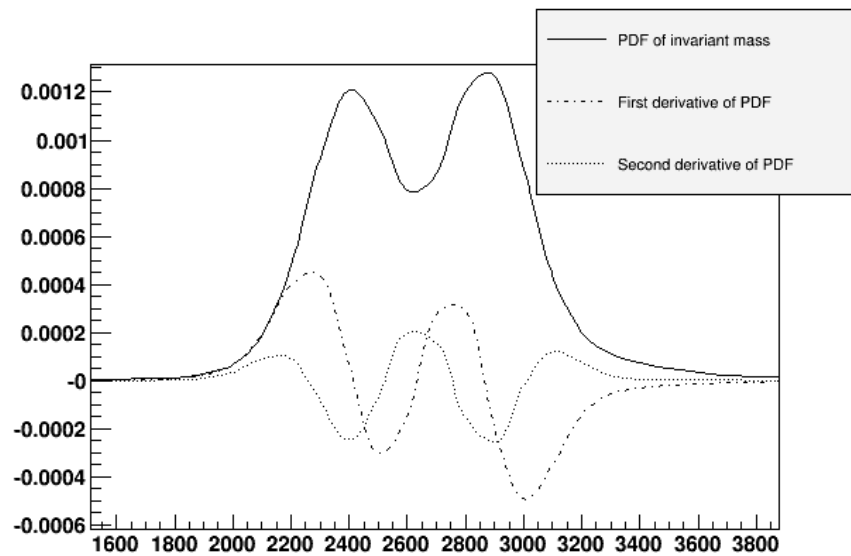

Figure 7.2: Event with multiple modes. The first and second derivatives have been scaled up.

Using the mean of the component with the highest maximum probability amplitude as initial guess has a success rate of $\sim 95\%$ in the case of invariant masses calculated from $J/\psi \to e^+ e^-$.

Some events get mass distributions with more than one local maximum(Fig 7.2). These events can be identified by using the the Newton-Raphson several times, using the mean of the n components with highest probability amplitudes as initial guesses. If more than one peak is identified using the different initial guesses, the event is discarded. Using $n = 5$ in the case where the invariant mass is calculated from two tracks, each described by 12 Gaussian components, all events with multiple local maxima are detected. Even though some events are discarded, the overall statistics is increased. This is because the success rate for events with a single maximum is very close to 100% when multiple initial guesses are used.

7.2　$J/\psi \rightarrow e^+ e^-$

~ 8500 J/ψ invariant masses have been reconstructed from electron and positron tracks. The J/ψ events are simulated from direct proton-proton collisions in the bunch crossings. The resulting electron and positron both have energy above 3 GeV. Figure 7.3 shows th momentum distribution of the electrons and positrons.

For each event, combinations of electron and positron tracks are used to calculate invariant masses. If the invariant mass of a combination is sufficiently close to the J/ψ mass of 3097 MeV, the combination is kept.

The impact parameters of electron/positron combinations close to the J/ψ mass are shown in figure 7.4. Figure 7.5 shows the difference between the $z0$ parameters from the two tracks. The distribution has long tails. Events in the core of the distribution are assumed to come from the same vertex, and these events are kept.
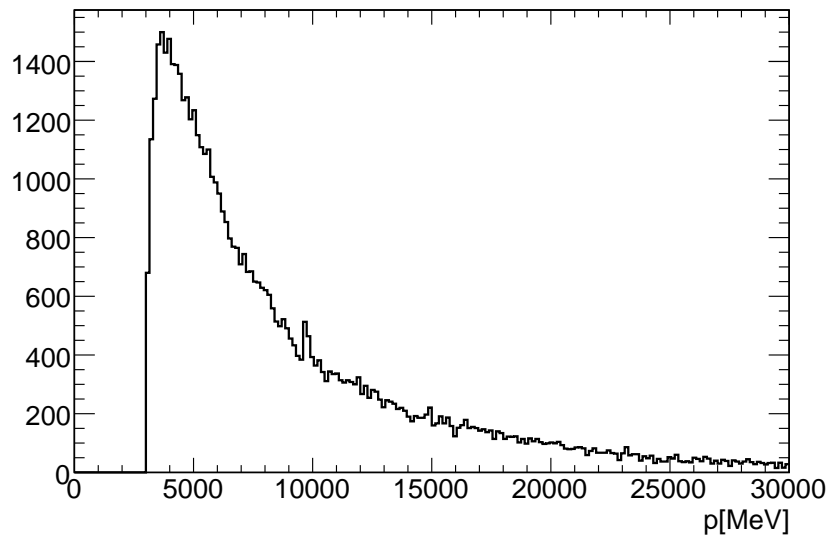

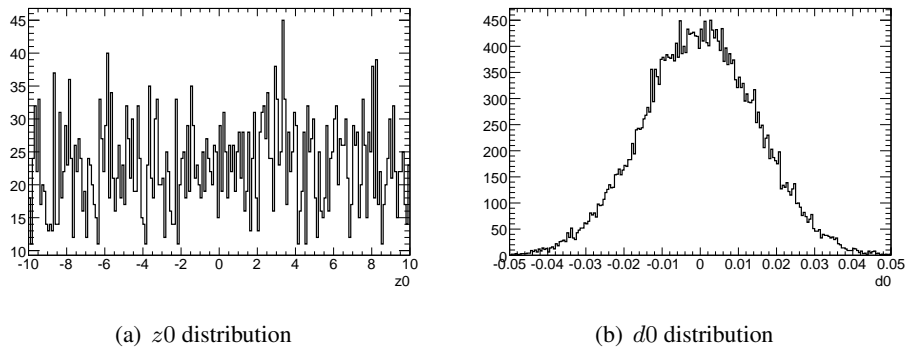Figure 7.3: Monte Carlo truth momentum distribution

(a) $z0$ distribution

(b) $d0$ distribution

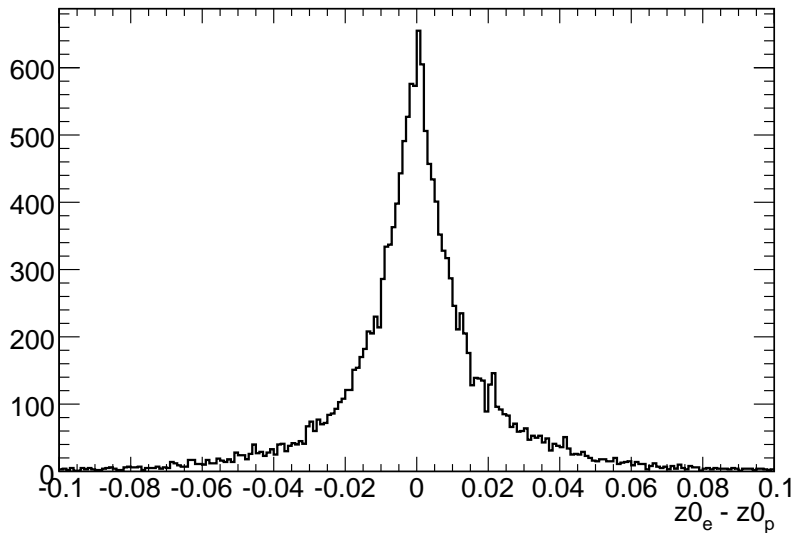Figure 7.4: Distribution of Monte Carlo truth for the impact parameters[μm].

Figure 7.5: The difference of the $z0$ true track parameter for the electron and positron.

The reconstructed invariant mass is shown in figure 7.6. The three estimates used, are the mode of the calculated probability distribution function of the invariant mass, the mode of the calculated probability distribution function for the square of the mass and the estimate using the mode of the track mixtures. The three invariant mass distributions are very similar.

Figure 7.7 shows the normalized residuals of the same estimates, the $1\sigma(68\%)$ and $2\sigma(95\%)$ symmetric half widths of the residuals are listed in table 7.1. The performance of the estimator using the mode of the calculated probability distribution is close to identical to the estimator using the mode of the track mixtures. The mode of the probability distribution of the squared mass performs slightly worse.

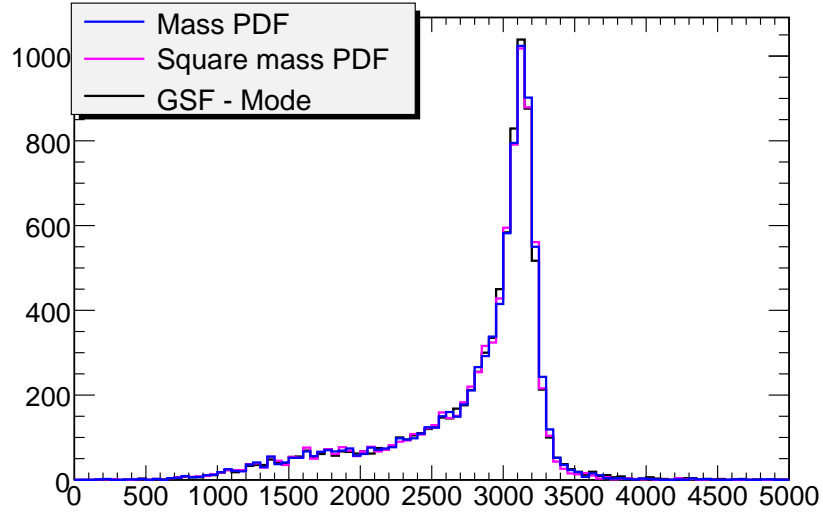The half widths of the normalized invariant mass residuals calculated from the

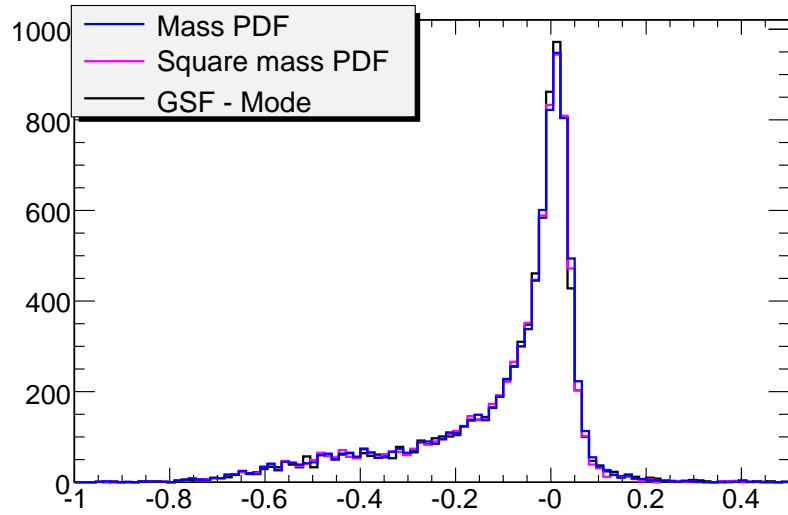Figure 7.6: Reconstructed J/ψ invariant mass.

Figure 7.7: Normalized residuals of the reconstructed J/ψ mass.

	Mass resolution
Mass PDF - Mode	$1\sigma = 0.114$
	$2\sigma = 0.505$
Square mass PDF - Mean	$1\sigma = 0.118$
	$2\sigma = 0.524$
GSF - Mode	$1\sigma = 0.115$
	$2\sigma = 0.506$

Table 7.1: Halfwidths of the normalized residuals for the reconstructed J/ψ mass.

mode of the track mixtures is larger than in the case discussed in section 5.2.2. Part of the difference is explained by the higher momentum of the leptons involved in the invariant mass resonstruction.

The resolution of the invariant mass residual from the calculated probability distributions show a strong dependance on track selection. The results presented here accepts events where the mass reconstructed from Monte Carlo truth is 3096.88 ± 0.1 MeV, and the Monte Carlo truth z0 parameters are seperated by less than $0.015\mu m$. Several different cuts have been explored. The resolution of the estimate stemming from the calculated probability distribution and the mode of the track estimates only differ by a few percent for all cuts. It is very plausible that some of the reconstructed tracks do not truely originate from a J/ψ decay. The relativly low momentum electrons that dominate the final states of the decays, generally have a small variance compared to the mean. Tracks like this transform to invariant mass distributions with relativly narrow peaks, which give little room for improvement.

Figure 7.8 is a scatter plot of the normalized residuals for the mode of the calculated distribution versus the estimate using the mode of the track mixtures. Events where the two estimates are very different are generally above the diagonal. This implies that for events where the two methods disagree, the new method generally return lower estimates. The vast majority of the estimates are located close to the diagonal, where the two estimators return similar results.

7.2.1 Probability transform

A way of testing the accuracy of the probability distribution functions is to calculate the p-value of the distributions. This is done by integrating the probability distribution functions from $-\infty$ to the true value.

$$p = \int_{-\infty}^{m} f(x)dx = 0.5 + \sum_{i=1}^{n} 0.5 \, g_i erf\left(\frac{m - W_i}{\sqrt{2var(W_i)}}\right), \qquad (7.4)$$

where W_i, $var(W_i)$ and g_i are the mean, variance and weight of a component, n the total number of components , m the true value, and erf is the error function.
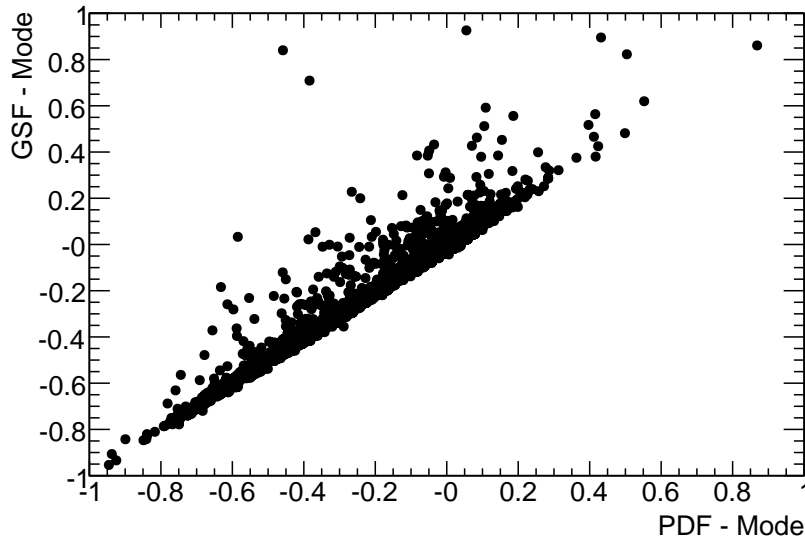
Figure 7.8: Normalized residuals of the mode of the calculated probability distribution function vs the mode of the track mixtures.

If the calculated probability distribution function accurately describes the uncertainty in the invariant mass, the p-values should become a flat distribution between 0 and 1.

The probability distribution functions of the mass and squared mass are compared to the probability distribution calculated from tracks collapsed into a single Gaussian around the mean. Figure 7.9 indicates that the probability distributions calculated from the full track mixtures give a better description of the uncertainties than the probability distribution from the collapsed tracks.

7.3 $H \rightarrow ZZ^* \rightarrow e^+e^-e^+e^-$

A sample of 50,000 130GeV Higgs bosons decaying to two Z bosons have been studied. The Z bosons furter decay to electrons/positrons or muons. The probability of a Z boson decaying to electrons is equal to the probability of it decaying to muons, $BR(Z \rightarrow e^+e^-) = BR(Z \rightarrow \mu^+\mu^-) = 1.12\%$. This means that 3/4 of the events have muons in the final state. The muons are biased because they are fitted with the GSF, so only events with four electrons in the final state are kept.

The momentum of the final state electrons/positrons are shown in figure 7.10. The electrons in this case have significantly higher momenta than in the J/ψ case. The momentum resolution decreases as the momentum rises, leading to larger variance in the q/p estimates, which in turn leads to larger uncertainties in the invariant mass estimate. The fact that there now are four particles in the final states instead of two, also lead to larger variance in the invariant mass estimates. Because of this,
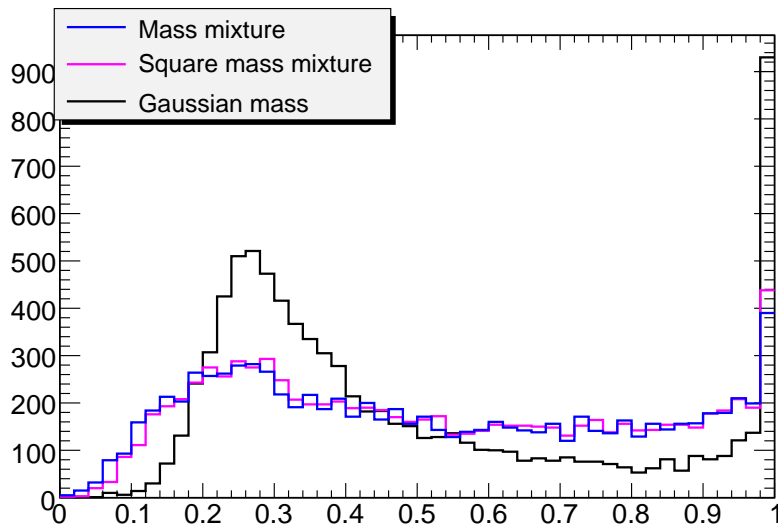
Figure 7.9: p-values of the calculated probability distribution functions.

the invariant mass resolutio of of Higgs decays are worse than in the J/psi case. The peaks in the invariant mass probability distribution functions are therefore expected to be wider, which means there is more room for improvement.
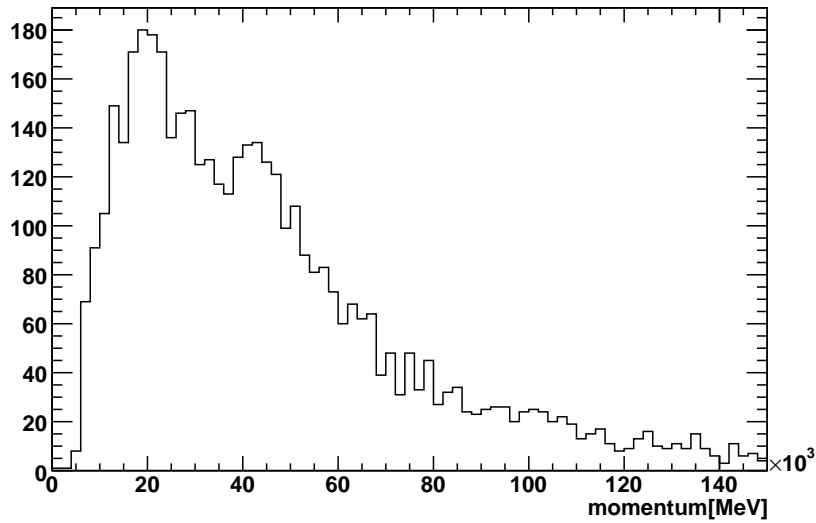

Figure 7.10: Monte Carlo truth momentum distribution.

The invariant mass is again calculated from Monte Carlo truth, and cuts on

the z0 parameters are applied to see if the track originate from approximately the same vertex. ~ 1200 events pass these cuts, and the invariant masses are estimated from the reconstructed tracks. The tracks reconstructed by the GSF are described by 12 components each, and all 12^4 possible combinations of track components are needed to calculate the invariant mass probability distribution function. This is obviously a slow process.

The reconstructed invariant masses of events where the Monte Carlo truth return a Higgs mass of (130 ± 0.1) Gev, and a spread in the z0 parameters of less than $0.015 \mu m$ are shown in figure 7.11. The three estimates used are the mode of the calulated probability distribution function describing the invariant mass, the mode of the calulated probability distribution function describing the squared invariant mass and the estimate using the mode of the track mixtures.
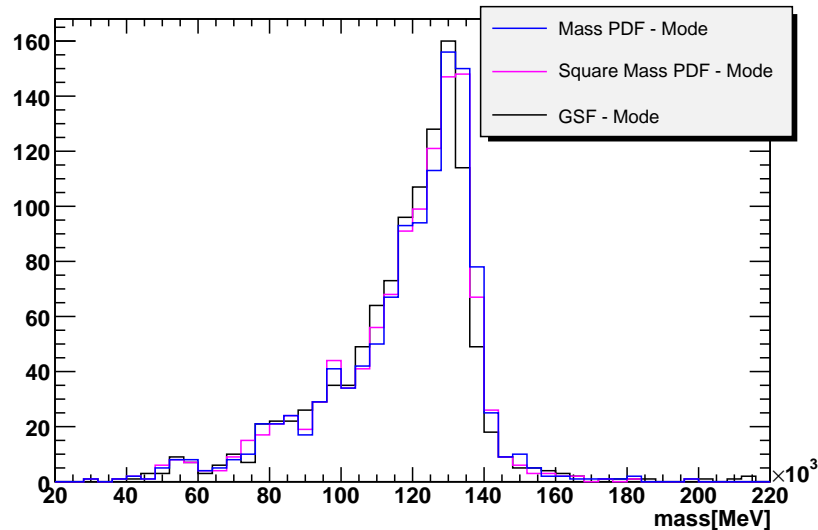

Figure 7.11: Reconstructed Higgs invariant mass.

Figure 7.12 shows the normalized residuals of the invariant mass estimates, the 1σ and 2σ halfwidths are listed in table 7.2. Both the calculated probabilty distribution functions produce better resolution than the estimate discarding the shape of the traks. Different cuts have been explored, and the new method always improves the mass resolution with between 7% and 10% compared to the estimate using the collapsed tracks. The probability distribution function describing the mass consistently outpreforms the estimate of the square mass.

Figure 7.13 shows the normalized residuals of the estimate from the mass probability distribution vs the estimate from the collapsed tracks. Similar to the J/ψ case, events where the estimates are very different are generally located above the diagonal.

The p-values for the two probability distribution functions describing the mass
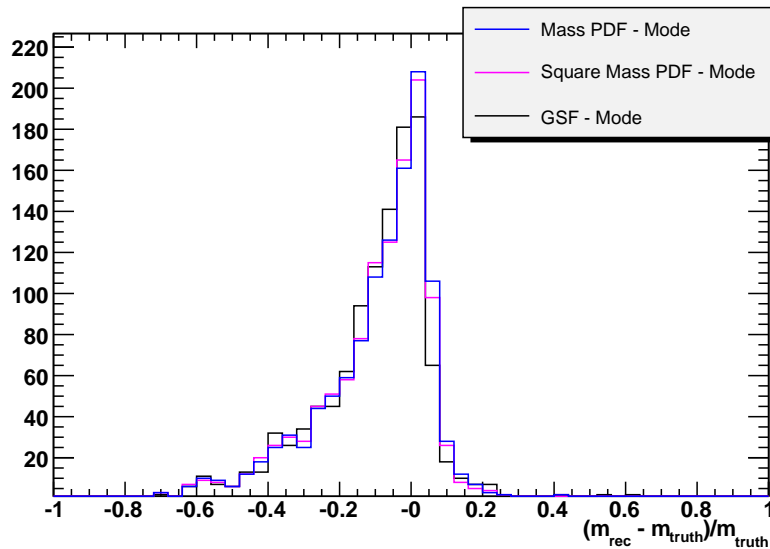
Figure 7.12: Normalized residuals of the reconstructed Higgs invariant mass.

	Mass resolution
Mass PDF - Mode	$1\sigma = 0.135$
	$2\sigma = 0.409$
Square mass PDF - Mode	$1\sigma = 0.138$
	$2\sigma = 0.417$
GSF - Mode	$1\sigma = 0.147$
	$2\sigma = 0.442$

Table 7.2: Halfwidths of the normalized residuals of the reconstructed Higgs mass.

and the square mass are compared with a probability distribution calculated from the tracks collapsed to a single Gaussian around the mean. The estimates described by Gaussian mixtures give p-value distributions that closer resembles a uniform distribution than the single Gaussian estimate, implying a better description of the uncertainties. The p-value distribution of the Gaussian estimate is, however, considerably closer to the multi-Gaussian p-values than in the J/ψ case.
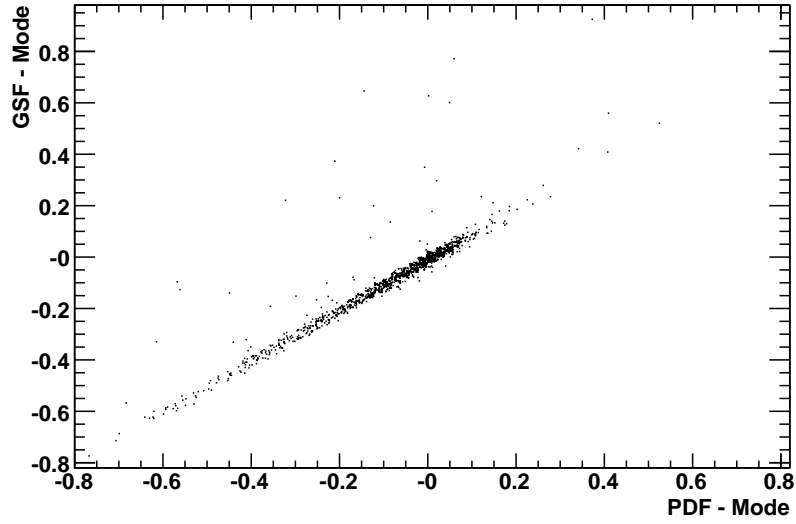
Figure 7.13: Normalized residuals of the mode of the calculated probability distribution function vs the mode of the track mixtures.
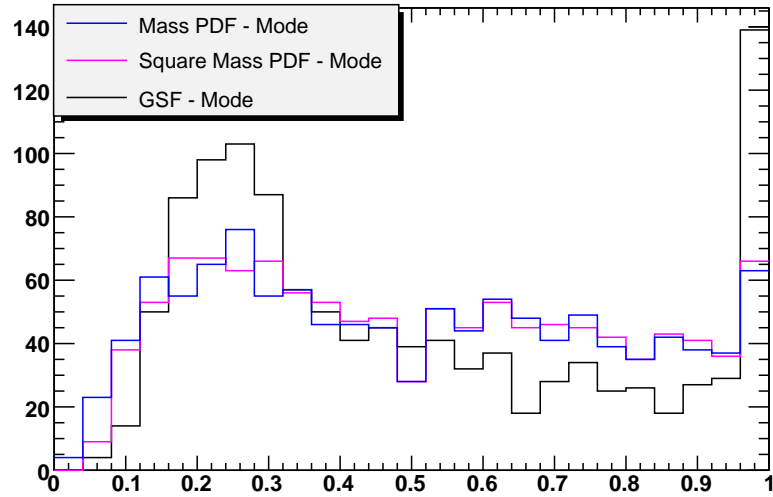

Figure 7.14: p-values of the calculated probability distribution functions.

Chapter 8

Conclusions

In this thesis, a method of calculating a probability distribution function describing the invariant mass using track estimates produced by the Gaussian sum filter(GSF) has been presented. The GSF track estimates are given by a weighted sum of Gaussian components, describing the tracks as a non-Gaussian probability distribution function.

The method has been tested on Monte Carlo simulated data of $J/\psi \to e^+e^-$ and $H[130GeV] \to ZZ^* \to e^+e^-e^+e^-$ and compared with the invariant mass estimated from the mode of the GSF track estimates.

In the J/ψ decays, no significant improvement in the invariant mass resolution has been found. However, the distribution of the p-values of the calculated invariant mass imply that the calculated probability distribution function gives a more accurate description of the invariant mass than estimates using the collapsed track mixtures.

In the analysis of the sample of Higgs decays, an improvement in the invariant mass resolution is found. This is a very promising result, but the sample used does not contain sufficient amount of data. In both cases the simulated events contain large numbers of tracks, and the track selection used is not optimal. A large sample of clean Higgs decays should be produced to further test the method. The p-value distribution again implies that the new method gives a better description of the uncertainties of the invariant mass.

The analysis in this thesis is based solely on tracks reconstructed in the inner detector. Combining the track estimates with information from ECAL will affect the invariant mass estimates. A GSF vertex fitter is currently not available, if this is implemented, it will also affect the invariant mass estimates. How this will influence the performance of the new method compared to using the collapsed track mixtures is hard to tell, and further testing is needed when methods for doing this become available.

Bibliography

[1] Thomas Martin Atkinson. *Electron Reconstruction with the ATLAS Inner Detector*. PhD thesis, The University of Melbourne, 2006.

[2] H. Bethe and W. Heitler. On the Stopping of fast particles and on the creation of positive electrons. *Proc. Roy. Soc. Lond.*, A146:83–112, 1934.

[3] C. Burgess, G. Moore. *The Standard Model, A Primer*. Cambridge, 2007.

[4] W.-M. Yao et. al. Review of Particle Physics. *Journal of Physics G*, 33:1+, 2006.

[5] G. Shaw F. Mandl. *Quantum Field Theory*. John Wiley and Sons Ltd, revised edition, 1993.

[6] Frühwirth, Regler, Bock, Grote, Notz. *Data Analysis Techniques for High-Energy Physics*. Cambridge, second edition, 2000.

[7] D. Green. *High Pt Physics at Hadron Colliders*. Cambridge University Press, 2005.

[8] R. Frühwirt. Track fitting with non-gaussian noise. *Comp. Phys. Comm.*, 100:1–16, 1997.

[9] R. Frühwirt. A gaussian-mixture approximation of the bethe-heitler model of electron energy loss by bremsstrahlung. *Comp. Phys. Comm.*, 154:131, 2003.

[10] R. Frühwirt, S. Frühwirt-Schnatter . On the Treatment of Energy Loss in Track Fitting. *Comp. Phys. Comm.*, 110, 1998.

[11] R. Frühwirt, W. Adam, A. Strandlie, T. Todorov. Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc. *Journal of Physics G: Nuclear and Particle Physics*, 31:N9–N20(1), September 2005.

[12] R.K. Bock, W. Krisher. *The Data Analysis BriefBook*. 1998.

[13] R.K. Bock,A.Vasilescu. *The Particle Detector BriefBook*. 1998.

[14] Stampfer et al. Track fitting with energy loss. *Comp. Phys. Comm.*, 79:157, 1994.

[15] The ATLAS collaboration. ATLAS Inner Detector Technical Design Report. *Technical Report, CERN*, 1997.

[16] The ATLAS collaboration. ATLAS Detector and Physics Performance Technical Design Report. *Technical Report, CERN*, 1999.

[17] The ATLAS collaboration. The ATLAS Experiment at the CERN Large Hardon Collider. *Technical Report, CERN*, 2008.

[18] The DONUT collaboration. Observation of tay-neutrino interactions. *Phys. Lett.*, B504, 2001.