Centro de Investigación y de Estudios Avanzados
del Instituto Politécnico Nacional

Physics Department

Observation of the doubly strange $b-$Baryon Ω_b^-

A dissertation submitted by

José de Jesús Hernández Orduña

in partial fulfillment for the award of the Degree of

Doctor in Physics

under the guidance of

Dr. Heriberto Castilla Valdez
&
Dr. Eduard De La Cruz Burelo

Mexico City February 2011

Abstract

This thesis reports the first experimental evidence of the doubly strange b–baryon $\Omega_b^-(ssb)$ following the decay channel

$$\Omega_b^- \longrightarrow \overbrace{J/\psi(1S)}^{\mu^+\;\mu^-}\quad \underbrace{\Omega^-}_{\underbrace{\Lambda}_{p\;\pi^-}\quad K^-}$$

in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV.

Using approximately 1.3 fb^{-1} of data collected with the DØ detector at the Fermilab Tevatron Collider, we observe 17.8 $\pm$ 4.9 (stat) $\pm$ 0.8 (syst) Ω_b^- signal events at 6.165 $\pm$ 0.010 (stat) $\pm$ 0.013 (syst) GeV/c^2 with a corresponding significance of 5.4 σ, meaning that the probability of the signal coming from a fluctuation in the background is 6.7×10^{-8}.

Resumen

Esta tesis reporta la primera evidencia experimental del barión pesado $\Omega_b^-(ssb)$ en el siguiente canal de decaimiento

$$\Omega_b^- \longrightarrow \overbrace{J/\psi(1S)}^{\mu^+\ \mu^-}\ \ \underbrace{\Omega^-}_{\underbrace{\Lambda}_{p\ \pi^-}\ \ K^-}$$

en colisiones $p\bar{p}$ a $\sqrt{s} = 1.96$ TeV.

Usando datos tomados con el detector de la colaboración DØ situado en el colisionador *Tevatron* de Fermilab, y que suman aproximadamente 1.3 fb^{-1} de luminosidad integrada, observamos 17.8 ± 4.9 (est.) ± 0.8 (sist.) eventos de Ω_b^- a una masa de 6.165 ± 0.010 (est.) ± 0.013 (sist.) GeV/c^2 que corresponden a una señal de 5.4 σ de significancia, es decir, la probabilidad de que la señal observada provenga de una fluctuación del ruido es 6.7×10^{-8}.

Dedication

My life was happily changed the
morning of December 30, 2005.

My work since then have been
dedicated so far to the origin
of that change.

Even if I won't see him again.

To my beloved son... Bogdan.

Acknowledgements

Along the last few years I have met the people that I'm likely gonna deal with for the rest of my career. To my fortune, they are people that I admire and respect and from whom I learn new things on a daily basis.

They deserve more than this words, but it would be absolutely unfair not to mention them here, where I intend to present the work that I think belongs also to them.

First, I want to express my gratitude for the enormous support I have received from my advisor **Dr. Heriberto Castilla** who have done much more than just to show me the HEP world.

Next, I want to thank **Dr. Eduard De La Cruz Burelo** –who happens to be my advisor, as well as one of the key stones on the analysis presented here–, and **Dr. Alberto Sánchez Hernández**. Both CINVESTAV professors who have been there for me, sharing their knowledge and giving me support and advice whenever I need them.

The evaluation committee integrated by **Dr. Salvador Carrillo Moreno** from Universidad Iberoamericana, **Dr. Omar Miranda Romagnoli**, **Dr. Luis Fernando Rojas Ochoa** and **Dr. Alberto Sánchez Hernández** from CINVESTAV, who reviewed and contributed to a better work on this thesis. Also **Dr. Abdel Pérez** who as Academic Coordinator, gave me all the support that I needed at my home institution.

My friends. From the CINVESTAV–HEP group at Fermilab, who I consider both younger and smarter: **Ivan Heredia** and his wife **Liliana Hernández**, **Ricardo Magaña** and his wife **Laura Robles**, **Jorge Martínez** and **Roger Hernández**. My comrades, **Luis A. López**, **Pablo Paniagua** and **Marco Carrasco** from CINVESTAV in Mexico City, and **Aaron Higuera** from DCI, University of Guanajuato, –who collaborates with MINERνA at Fermilab–, also **Omar Molina**, **Mario Corona** and **Moises Vargas**; we share lots of great memories. All of them have been there for me in really difficult times as well.

The women who inspired me and made me enhance the respect that I've always had for them. Alphabetically: **Dalia Cervantes**, **Lizbeth M. Fernández**, **Estela Garcés**, **Flor Ibañez** and **Mayra J. Reyes**.

I have spent roughly half of my Ph. D. based at Fermilab on the DØ Collaboration and it would be unforgivable not to mention the people I met there.

Spokesperson. **Дмитрий Сергеевич Денисов**, **Darien R. Wood**, **Stefan Söldner Rembold** and **Terry Wyatt**.

Run Coordinators. **William Lee**, **George Ginther**, **Stefan Grünendahl** and **Marc Buehler**.

SMT group. **Lynn Bagby**, **Satish Desai**, **Masato Aoki**, **Zhenyu Ye**, **Michael Kirby**, **Nirmalya Parua**, **Peter Svoisky**, **Дмитрий Цыбышев**, **Sung Woo Youn**, **Derek Strom**, **Ron Lipton** and **Marvin Johnson**.

B–Physics group. **Penelope Kasper**, **Marj Corcoran**, **Gene Fisk**, **Геннадий Борисов**, **Aurelio Juste** and **Brendan Casey**.

System experts, DØ collaboratos and crew members. **Jadwiga Warchoł**, **Michelle Prewitt**, **Mandy Rominsky**, **Carrie McGivern**, **Albert Ito**, **Geoff Savage**, **Takahiro Yasuda**, **Дмитрий Смирнов** and **Rolando Flores**.

Of course, the foundation in which I am trying to build myself as a human being was given since the very beginning in my family. They have taught me to be prudent, just, temperate, strong, as well as to have faith and hope, but the most important I have learned from them is to feel and express love. To **Marta**, **Paola**, **Juan Pablo** y **José** all my love and gratitude.

To finish I want to rephrase a fragment of a speech given by Steve Jobs at Stanford in 2005 to say: *my work is gonna fill a large part of my life and the only way to be truly satisfied is to do what I believe is great work, and the only way to do great work is* ***to love what I do***. While is true that I'm still on the way to find what I love, I deeply thank all of these people, they have helped me to find it and now I know that every step I take, I get closer and closer.

This work was supported in part by **CONACyT**.

Contents

List of Figures

List of Tables

Introduction

The theoretical model we have to describe what we believe are the building blocks of nature and the interactions between them, is known as Standard Model.

The Standard Model is the combination of Electroweak Theory and Quantum Chromodynamics into a single core in the attempt to include all interactions of subatomic particles except those due to gravity in a simple framework.

This model has proved highly accurate in predicting certain interactions, but it does not explain all aspects of subatomic particles. For example, it cannot say how many particles there should be or what their masses are. The search goes on for a more complete theory, and in particular an unified field theory describing the strong, weak, and electromagnetic forces.

Twelve elementary particles are known in the Standard Model: the Fermions. They have spin –1/2 and obey the Pauli Exclusion Principle. Fermions are divided into six Quarks: up u, down d, charm c, strange s, top t and, bottom b; and six Leptons: electron e, muon μ, tau τ, electron neutrino ν_e, muon neutrino ν_μ and, tau neutrino ν_τ.

Quarks interact via the strong force because they carry color charge, electromagnetically because of their electric charge and via the weak nuclear interaction because of the weak isospin.

Quarks form color–neutral composite particles known as Hadrons which are divided in Mesons, containing a quark and an antiquark and Baryons, made up three quarks.

Leptons have no color charge and can not interact via the strong force. Only three of them have electric charge, hence interact electromagnetically. The motion of non–electrically charged leptons, the neutrinos, is influenced only by the weak nuclear interaction.

Every fermion have an associated antiparticle. For quarks, the antiparticle carry opposite electric charge, color charge and baryon number. For leptons, the antiparticle carry opposite electric charge and lepton number.

Fermions are suitably grouped together considering their properties and three generations of them are defined. A higher generation fermion have greater mass than those in lower generations. Charged members of the first generation do not decay and form the ultimate building blocks for all the baryonic matter we know about. Charged members of higher generations have very short half lives and are found normally in high–energy environments. Non–electrically charged fermions do not decay and rarely interact with baryonic matter.

The way particles interact and influence each other in the Standard Model is result from matter particles exchanging other particles, known as Force Mediating Particles. They are believed to be the reason of the existence of the forces and interactions between particles observed in the laboratory and the universe.

Force mediating particles have spin 1, *i. e.*, they are Bosons, and do not follow the Pauli Exclusion Principle. The types of force mediating particles are: the photon γ, three gauge bosons $W^{\pm}$ and Z and, eight gluons g.

Photons have no mass, the theory of Quantum Electrodynamics describe them very well and are responsible for mediation of the electromagnetic force between electrically charged particles.

Gauge bosons are massive, being Z heavier than $W^{\pm}$. They are responsible for the mediation of the weak interactions between particles of different flavors but $W^{\pm}$ act only on left–handed particles and right–handed antiparticles while Z with both left–handed particles and antiparticles. Due to the electric charge of $W^{\pm}$, they couple also to electromagnetic interactions.

Photons and the three gauge bosons are grouped together and collectively mediate the electroweak interactions.

Finally, gluons have no mass, the theory of Quantum Chromodynamics describe them and are responsible for the mediation of the strong interactions between particles with color charge. Having an effective color charge, gluons can interact among themselves.

The Higgs Boson is the only particle in the SM without direct experimental evidence. Its detection would help in the explanation of the difference between massive bosons mediating the weak force and the massless photon mediating the electromagnetism.

Even when Tevatron has already set some important limits on the existence of the Higgs boson, the conditions in which the Large Hadron Collider (LHC) will operate, are believed to be such that they can provide new insights into the Higgs boson's existence.

Figure 1 shows the summary of properties of the Standard Model fermions and gauge bosons while Fig. 2 shows the review of the interactions between all the particles described in the Standard Model.

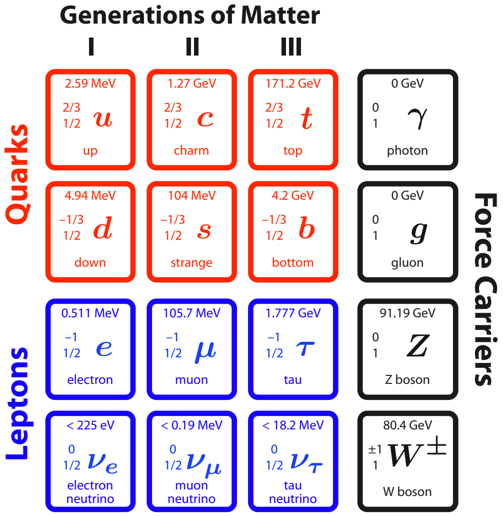


Figure 1: Building blocks of matter and force–carrying particles in the Standard Model.

The hadron classification in terms of their valence quarks is known as Quark Model and has been particularly successful describing the spectroscopy of hadrons.

The Quark Model was originally just a very good classification scheme to organize the depressingly large number of hadrons that were being discovered starting in the 1950's and continuing through the 1960's but has received good experimental verification beginning in the late 1960's with the observation of the Ω^- with quark content $|sss\rangle$, and continuing to the present.

However, at some point the observation of anomalous particles, those which do not fit in the Quark Model because of its quantum numbers, like the X(3872) observed first by the Belle Experiment and confirmed by other collaborations, together with the large datasets collected at Fermilab at the required energies for heavy baryon observation, motivated the continuation for the search of the particles predicted by the Quark Model.

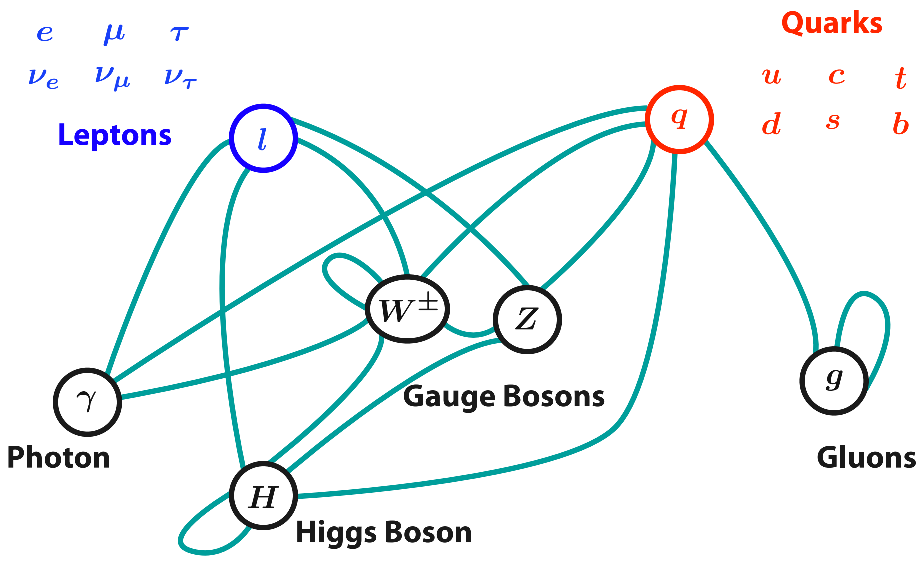


Figure 2: Interactions between all the particles described by the Standard Model.

Observation of Bottom Baryons

Until 2006, the only one bottom baryon directly observed was the Λ_b with quark content $|udb\rangle$. Then, in 2007 some other $b-$baryons appeared published after their observation: the Ξ_b^- with quark content $|bsd\rangle$ –the first one observed having a quark from each family,– and the $\Sigma_b^{(*)}$ states, the $\Sigma_b^{(*)+}$ ($|buu\rangle$) and $\Sigma_b^{(*)-}$ ($|bdd\rangle$). Table 1. The current prediction of bottom baryons with $J^{+1/2}$ is shown in Fig. 3.

This thesis presents the results for the search of the bottom baryon with quark content $|bss\rangle$ –which according to the rules for classification of baryons defined by the Particle Data Group, is named Ω_b^-–, within the

$b-$baryon	Quark content	J^P	SU(3)	(I,I_3)	Strangeness	Mass (MeV$/c^2$)
Λ_b^0	bud	+1/2	3*	(0,0)	0	5,620.2±1.6
Ξ_b^-	bsd	+1/2	3*	(1/2,–1/2)	–1	5,792.4±3.0
Σ_b^+	buu	+1/2	6	(1,1)	0	5,807.8±2.7
Σ_b^-	bdd	+1/2	6	(1,–1)	0	5,815.2±2.0
$\Sigma_b^{(*)+}$	buu	+3/2	6	(1,1)	0	5,829.0±3.4
$\Sigma_b^{(*)-}$	bdd	+3/2	6	(1,–1)	0	5,836.4±2.8

Table 1: Observed bottom baryons until 2007.

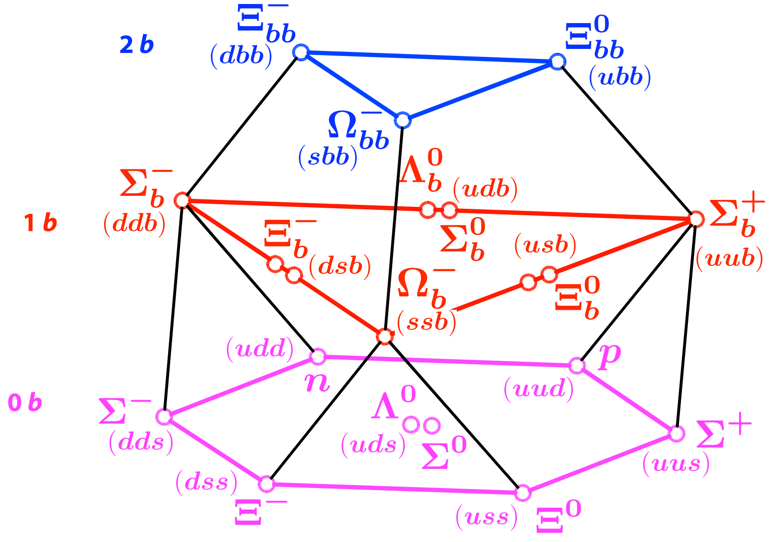


Figure 3: Bottom baryons with $J^{+1/2}$.

data collected by the DØ Detector between April 2002 and February 2006 which led to its first observation in real data.

The first part of the thesis is dedicated to the overview of the experimental apparatus, the accelerator which provides the collisions detected in order to look for experimental evidence of some theoretical predictions. Next, a brief description of the DØ Detector is presented referring the characteristics of all the main components of the detector, its construction and relationships, as well as the description of the triggering and data acquisition systems used in the detector. Then the description of the analysis itself. Beginning with the presentation of the data in which the search was made, the need of a special processing of this information due to the long–lived particles involved in the particular decay channel used, the step–by–step reconstruction of the decay by means of a blind selection of candidates and the application of the method of selection to control samples to discard artificial production of resonances. Then the actual application over pre–selected candidates and the set of crosschecks made to give further support of the resonance found. Also, the calculation of the production ratio of the b–baryon with respect to another previously observed, the Ξ_b^-, which is itself very similar in topology and kinematics to the Ω_b^-. The conclusion of this thesis shows the summary of the work done and present a little treatment of further work that can be done on this subject.

Fermilab Tevatron Collider

The **Fermi National Accelerator Laboratory**, known as Fermilab, is located thirty miles west Chicago in the city of Batavia, Illinois USA. Its mission is to help scientist from all over the world to advance in the understanding of the fundamental nature of matter and energy.

Fermilab is the home of the **Fermilab Tevatron Collider**.

The Fermilab Tevatron Collider was until 2009 the highest–energy particle collider in the world. Produces collisions between protons and antiprotons at a center–of–mass energy of $\sqrt{s} = 1.96$ TeV.

Each bunch of protons and antiprotons is accelerated to 980 GeV in a circular ring to cross their paths at two interaction points, located in the middle of two detectors: CDF and DØ.

In order to accelerate massive particles to this energy, a number of individual acceleration stages is needed.

The Fermilab Tevatron Collider is shown in Fig. 4. In the following sections, an overview of the main components is described. Beginning with negative hydrogen ions until bunches of protons and antiprotons collide at the interaction point of both detectors.

The Proton Source and the Linear Accelerator

Gaseous hydrogen is injected into the ion source to produce negatively–charged hydrogen ions. The ions are extracted from the source at 18 keV and the Preaccelerator produces a beam at 750 keV to be injected into a linear accelerator called Linac.

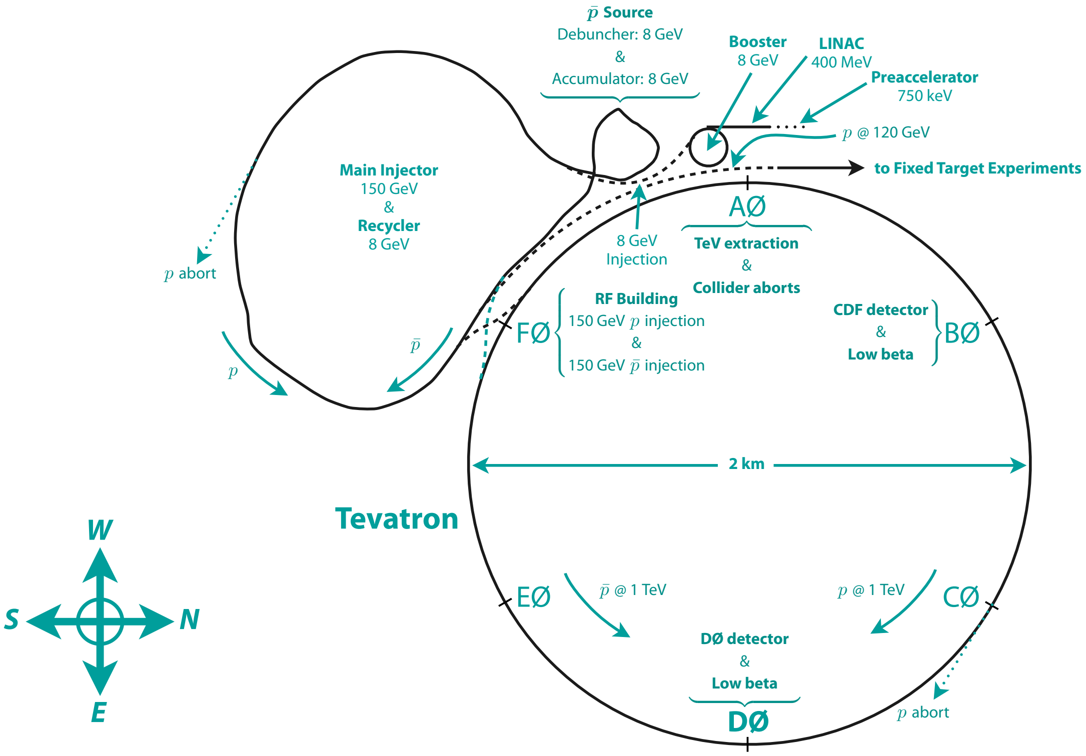
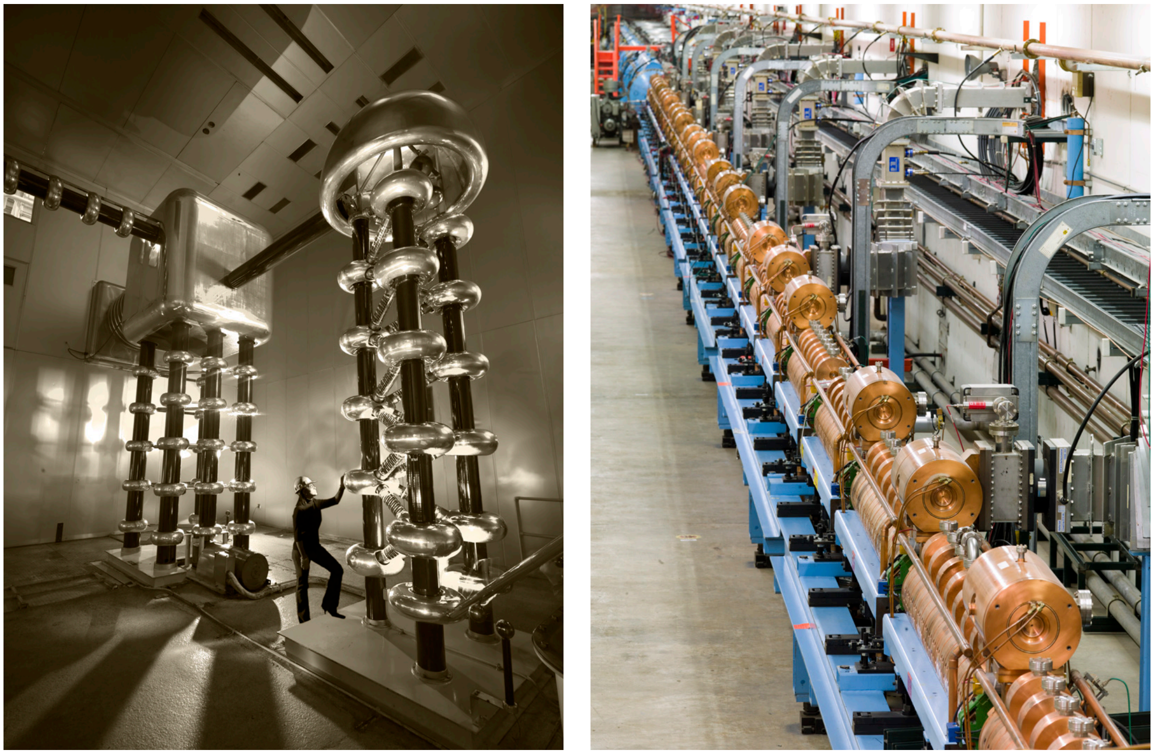


Figure 4: Main components of the Fermilab Tevatron Collider.

The Linac provides two more acceleration stages to raise the energy up to 400 MeV. The first stage consist of five Alvarez Drift–tube tanks with a pulsed RF resonance at 201.25 MHz to raise the energy from 750 keV to 116.5 MeV. The second stage is a side–coupled linear accelerator with seven cavities at four times the Alvarez frequency, such that only one in four RF cycles contains beam. At the end of this stage, the beam energy is 400 MeV with a pulse length of 40 μs, which at this point goes into the Booster.

Throughout the Linac, alternating focusing and defocusing magnets ensure that the transverse beam size remains within the acceptable cavity aperture of around 3 cm, compensating for certain RF effects which can broaden the beam profile.

Figure 5a shows the Preaccelerator and Fig. 5b the Linear Accelerator.



(a) Preaccelerator (b) Linear Accelerator

Figure 5: H^- ions are produced in the Proton Source and the Preaccelerator (left) raise their energy to 750 keV. The Linear accelerator increases their energy to 400 MeV

Booster

The Booster is a synchrotron. It takes the H^- pulsed ion beam from the Linac at 400 MeV, removes the electrons, and accelerates the resulting protons to 8 GeV. The beam of protons passes then to the Main Injector.

Eighteen RF cavities distributed along the Booster are responsible for the acceleration. In total, there are 84 RF regions in which the particles are captured and accelerated. Suitable located magnets are used to focus and bend the beam.

The Booster has three purposes:

Inject. An electrostatic deflector located after the Linac controls the injection of the 400 MeV proton beam. The momentum spread of the particles is controlled by a debuncher; it focuses the ions in the momentum space preventing the particles from having undesirable radii of curvature. In this stage, the H^- ions from the Linac are injected through the carbon foil which removes the electrons. Holding the injection for enough time as to complete approximately six revolutions in the Booster, assures the 84 RF buckets to be filled. This stage does not produce any acceleration.

Accelerate. After the injection, each RF cavity experience a sinusoidal field oscillation. Looking at the beam energy, the phase difference between adjacent cavities is modified to make every bunch to feel a stable accelerating electric field when it crosses the cavities. At the end of this stage, the beam have 8 GeV and is transferred to the Main Injector.

Extract. Four pulsed magnets deflect each bunch out of the synchrotron towards the transfer line in a single Booster revolution. Typically, the first bunch is lost while turning the magnets on. The amount of bunches sent to the Main Injector depends on the operation mode. Locking the Booster RF to match that of the Main Injector is what makes possible to maintain the bucket structure with almost no loss in the transition.

Figure 6 shows the Booster.

Main Injector

The Main Injector is also a synchrotron. It is made up of 18 RF cavities, 208 quadrupole and 344 dipole magnets, which accelerate, focus and bend the beam along the ring. A total of 588 RF buckets are found in the Main Injector.

The Main Injector also serve to three main purposes:

Antiproton Production. In this mode, the Main Injector receives single groups of 8 GeV protons from the Booster with approximately 5×10^{12} protons over all 83 bunches. They are accelerated to 120 GeV and sent to the Antiproton Source using the transfer lines P1 and P2. This process is repeated as many times as needed to collect the necessary amount of antiprotons and is known as Stacking. On average, 5 million protons are needed to extract a single antiproton. This is the reason

Figure 6: The Booster takes H^- ions from the Linac at 400 MeV and produces a proton beam of 8 GeV which is transferred to the Main Injector.

why the Main Injector is dominated by the Antiproton Production, which nowadays takes place in parallel with collisions occurring in the Tevatron.

Proton Injection to the Tevatron. In this case, the Main Injector receives twelve groups of 8 GeV protons from the Booster. Each group have a few bunches of around 4×10^{11} protons. This time, the bunch structure is removed and the resulting beam accelerated to 150 GeV which is then transferred to the Tevatron in a single revolution. Three of this injection processes are needed to provide the required 36 proton bunch configuration in which Tevatron operates.

Antiproton Injection to the Tevatron. Now, the Accumulator sends a single group of 8 GeV antiprotons to the Main Injector[1]. Each group have four sets of eleven bunches

---

[1] Antiprotons can also be transferred from the Recycler

with around 3×10^{10} antiprotons. Like in the proton case, the bunch structure is removed and the resulting beam accelerated to 150 GeV to be transferred to the Tevatron in a single revolution. Nine of this processes are needed to provide the required 36 antiproton bunch configuration in which Tevatron operates.

The Main Injector also provides 120 GeV proton beams to other fixed target experiments located at Fermilab, typically during the Antiproton Production stage.

Figure 7: The Main Injector and the Recycler share the same tunnel, the later stores protons and antiprotons at the same time that provides additional cooling.

Recycler

The Recycler is located in the same tunnel that the Main Injector. It is used to store antiprotons coming from the Accumulator and provides them with additional cooling. To control the beam position, the Recycler is composed of a series of a Focusing quadrupole followed by a dipole together with a Defocusing quadrupole followed by another dipole. The four components are known as FODO lattice. The Main Injector has the same configuration.

The Accumulator transfers the antiprotons to the Recycler using the Main Injector. The Recycler then uses a four–cavity RF system to capture, stabilize and produce discrete buckets of antiprotons. The Recycler do not produce any acceleration.

When the amount of antiprotons collected is enough to produce a quality beam and the 36 proton bunches are already in the Tevatron, the antiprotons from the Recycler are moved into the Main Injector for acceleration to 150 GeV and then transferred to the Tevatron.

Figure 7 show the aerial view of the Main Injector in which the Recycler is also located.

Antiproton Source

Protons at 120 GeV are directed to a fixed target to produce antiprotons. The antiprotons are captured, stabilized and accumulated until enough of them have been collected to be sent to the Main Injector and the Recycler.

The group of protons coming from the Main Injector, are focused using quadrupole magnets to a spot of around 0.15 mm of radius. This beam impact a fixed target made of Nickel producing secondary particles focused by a lithium lens. Magnetic bending is used to select negative particles of momentum 8.9 GeV/c. The resulting particles, which remain in bunches as the original incident protons, are transferred to the debuncher. The debuncher have 90 RF buckets and manipulates the particles received in such a way that stabilizes its orbit and transforms the spatial bunch distribution to an almost continuous strip of antiprotons.

The random motion of the antiproton beam is reduced by stochastic cooling and the beam is now transferred to the Accumulator just before a new set of protons from the Main Injector enters the Antiproton Source.

The Accumulator a is smaller synchrotron located in the same tunnel as the Debuncher. Contains 84 RF buckets to produce the antiproton bunches. Each group of antiprotons sent to the Accumulator is gradually merged to the existing amount of antiprotons by a small deceleration followed by an additional stage of stochastic cooling. The antiprotons are grouped into a single core. Once there, they are concentrated in both spatial and momentum distributions in a continuous way in order to collect the required 10^{12} antiprotons.

The transference to the Main Injector is possible because the Accumulator RF is locked to match that of the Main Injector while part of the beam is accelerated to achieve the larger radius needed for the extraction. This part of the beam is resized to provide eleven bunches and repeated four times.

Fig. 8 show the aerial view of the site in which the $\bar{p}$'s are produced.

Figure 8: Aerial view of the $\bar{p}$ source.

Tevatron

The Tevatron is a synchrotron of radius 1 km and is the final stage in the acceleration process. It has 216 quadrupole and 774 dipole magnets to focus and bend the beam with a total of 1,113 buckets. Eight RF cavities provide the acceleration to the 36 bunches composing each beam. The bunches are grouped in sets of 12 bunches, each one separated by 139 buckets by the others. The total spread of a particular set of 12 bunches is 232 buckets.

Tevatron operates at low temperature, around 4.6 K, because of the superconductor characteristics needed to reach the final high–energy. Takes the protons and antiprotons from the Main Injector at 150 GeV and accelerate them in opposite directions up to the collision energy of 980 GeV.

Both beams at the final energy are focused and their halos (particles far from the center of the beam) are removed by collimators.

Beams under this condition have enough quality to produce collisions, which occur at an average rate of 1.7 MHz. The intensity of the beams fall over time because of the interaction between them and with the surrounding material. Typically, once the amount of accumulated antiprotons is large enough to be transferred to the Tevatron, the beams in the Tevatron are dumped and a new cycle of acceleration takes place.

The period of time between the first collisions of quality beams until the beams are dumped is known as Store. In normal condition, the Store is the time when both detectors are in recording configuration.

The next chapter contains the global description of the DØ Detector, located in one of the two interaction points along the Main Ring, and used to collect the data used in this analysis.

Figure 9 shows the aerial view of the actual Tevatron.

Figure 9: The Tevatron or Main Ring is the actual synchrotron where both beams reach their final energy. Both detectors: CDF and DØ, are located on the Main Ring to record the products of the collisions.

The DØ Detector

Note. For the scope of this analysis and thesis, the DØ Detector corresponds to the configuration prior to the shutdown that took place in 2006. The DØ Detector suffered major upgrades during this shutdown. The most important being the introduction of the innermost layer of silicon detectors, known as LØ, and the removal of the two farthermost disks of silicon detectors.

The DØ detector, Fig. 10, is made up the following components:

- the Central Tracking System: the Silicon Microstrip Tracker and the Central Fiber Tracker,
- a Solenoidal Magnet,
- the Central and Forward Preshower Detectors,
- the Uranium/Liquid–Argon Calorimeters,
- a Toroid with the Central and Forward Muon Detectors,
- a Forward Proton Detector and
- the Luminosity Monitor.

All the systems are described by a right–handed coordinate system in which the z–axis is along the proton direction and the y–axis points upward, Fig. 11. The angles ϕ and θ are the azimuthal and polar angles, respectively. The r coordinate denotes the perpendicular distance from the z–axis.

Another quantity used in analysis is the pseudorapidity η, defined as

$$\eta = -\ln\left\{\tan\left(\frac{\theta}{2}\right)\right\}$$

which approximates the true rapidity,

$$y = \frac{1}{2}\ln\left\{\frac{E + p_z c}{E - p_z c}\right\}$$

for finite angles in the limit that $mc^2/E \rightarrow 0$. Based on $|\eta|$, the detector is divided in two regions. The *central region* which corresponds to $|\eta| < 2.5$. The remaining sections in the detector are referred as *forward regions*.

The first part of this chapter is dedicated to the overview of the different subsystems of the DØ Detector while the second covers aspects of the operation of the detector itself.

Central Tracking System

The Central Tracking System, Fig. 12, consists of the Silicon Microstrip Tracker and the Central Fiber Tracker. These two systems are immersed in a 2 T solenoidal magnet. The information from the Central Tracking System is used to locate the primary interaction vertex with a resolution of about 35 μm along the beamline.

Silicon Microstrip Tracker: SMT

The SMT covers almost all the η region and is composed by single silicon detectors arranged in two kind of arrays known as Barrels and Disks. Some of the Disks are interspersed between the Barrels in the central region and two assemblies of Disks are located in the forward regions.

Barrels are groups of rectangular–shaped detectors parallel to the beam–line. In the other hand, the trapezoidal–shaped detectors composing the Disks remain perpendicular to the beam direction.

The SMT design is such that ideally most of the tracks cross detectors perpendicular to their surfaces for all the coverage region. The actual design and construction of the SMT is itself an extremely complex subject, specially for the difficult mechanical challenges in arranging the detector components and minimizing dead areas while providing enough space for the required cooling and cabling.

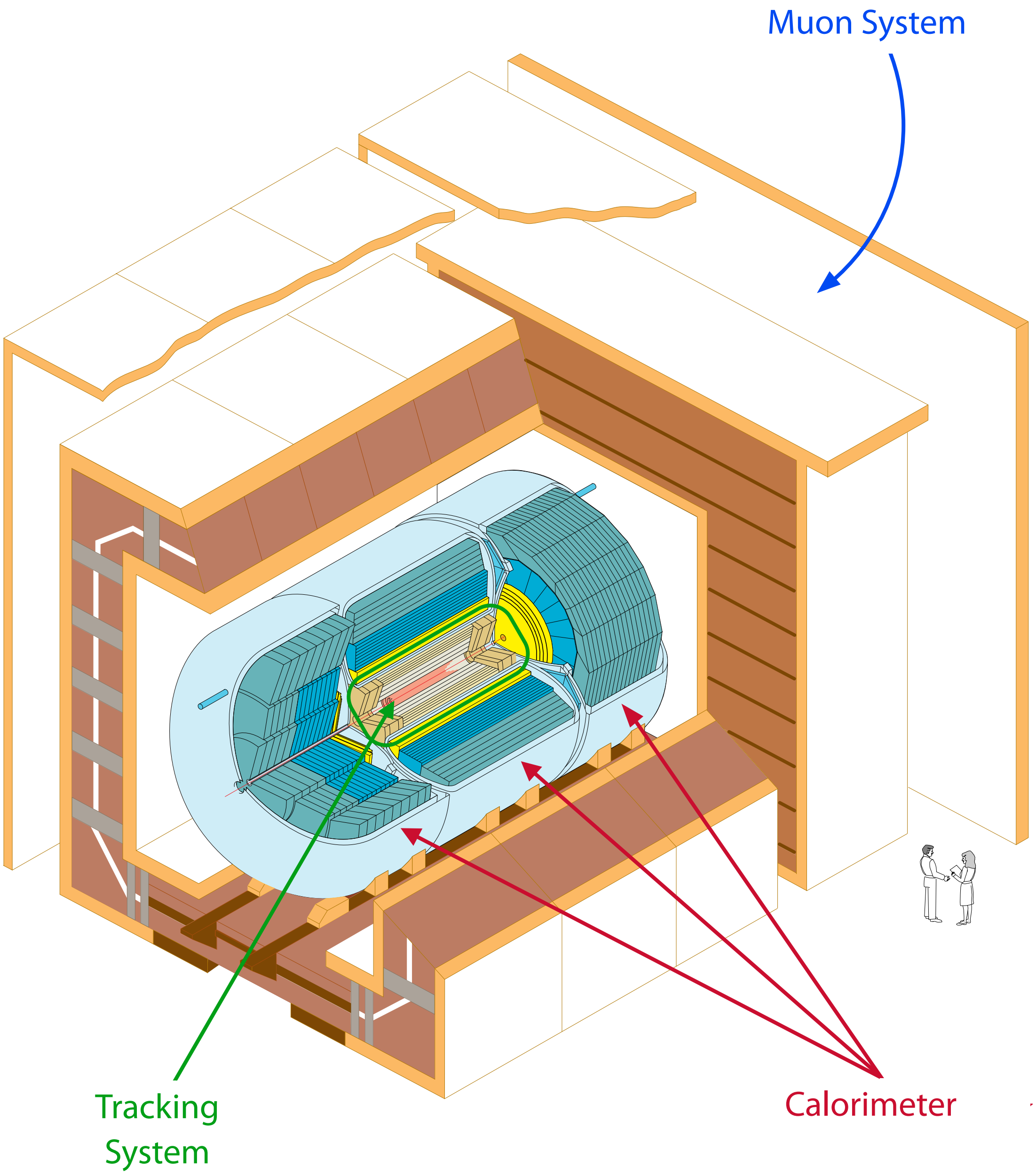


Figure 10: DØ Detector.

The SMT has six Barrels, all in the central region. Each Barrel has four silicon readout layers numbered 1 to 4 (the last the farthest from the center). The centers of the Barrels are located at $|z| = 6.2, 19.0, 31.8$ cm.

The silicon modules installed in the Barrels are called Ladders. Layers 1 and 2 have twelve Ladders each;

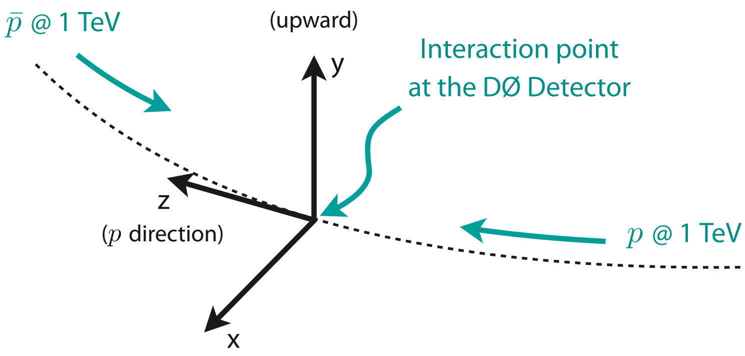


Figure 11: Coordinate system in the DØ Detector.

layers 3 and 4 have twenty–four to count 432 Ladders in total. Barrels are constructed out of three types of sensors.

The first and third layers of the outer two barrels use Single–Sided sensors (SS) with axial strips. Two of these sensors are bonded together to form a 12 cm long ladder. All the second and fourth layers use Double–Sided stereo sensors (DS) with the n–side implants at a 2° angle with respect to the p–side axial strips, and, again, two sensors are bonded together to make one 12 cm ladder. The first and third layers of the inner four barrels use single 12 cm long Double–Side–Double–Metal 90° stereo sensors (DSDM). The term Double–Metal is attributed due to the fact that to achieve the ninety degree readout, it uses a second metal layer on the n–side insulated from the first metal.

Barrels measure primarily the r–ϕ coordinate.

Each Barrel is capped at high $|z|$ with a Disk of twelve double–sided wedge detectors. They are known as F–Disks. After the three Barrel/F–Disk assemblies on each side, is a unit consisting of three more F–Disks to count a total of twelve F–Disks with 144 double–sided F–wedges. Readout strips in all Disk sensors are arranged parallel to the long edge of the devices. For the F–Disks, this means an effective stereo angle of 30°.

The F–Disks are located at $|z| = 12.5, 25.3, 38.2, 43.1, 48.1$ and 53.1 cm.

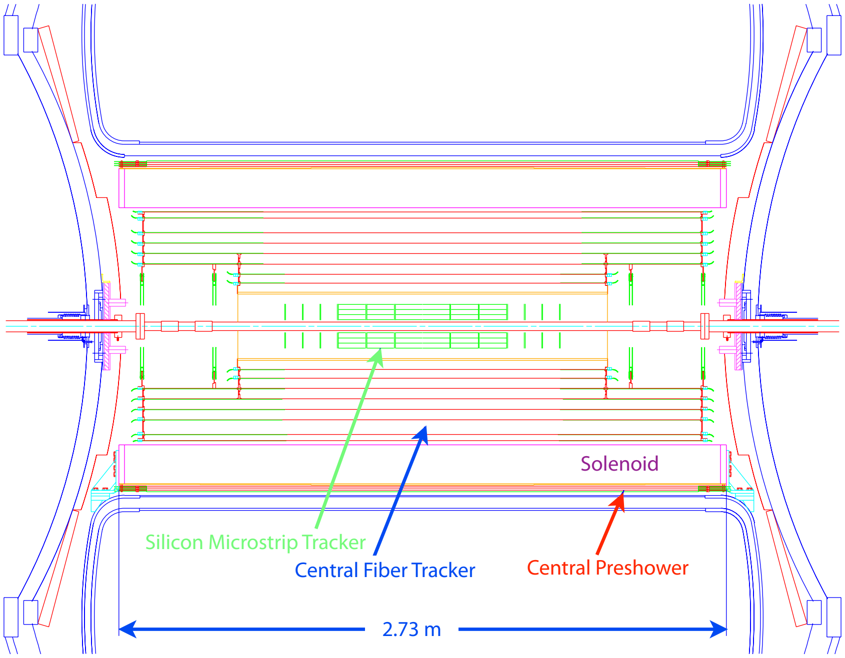


Figure 12: The Central Tracking System is composed of the Silicon Microstrip Tracker and the Central Fiber Tracker.

In both forward regions, there is a set of two larger disks providing tracking at high $|\eta|$. They are known as H–Disks and have twenty–four full wedges. Each wedge consist of two back–to–back single–sided half wedges and each half is read out independently. There are 96 full H–wedges in the SMT with an effective stereo angle of 15°.

The centers of the H–Disks are located at $|z| = 100.4$ and 121.0 cm.

Both kind of Disks (F and H), measure r–z as well as r–ϕ.

The total number of readout modules in the SMT is 912, with 792,576 channels which are readout by custom–made 128–channel readout chips. Table 2, shows the main characteristics of the different silicon modules used in the SMT.

Module	Type	Layer	Pitch (μm) p/n	Length (cm)	Inner radius (cm)	Outer radius (cm)
Central Barrels (4)	DSDM	1, 3	50 / 153.5	12.00	2.715	7.582
	DS	2, 4	50 / 62.5	6.00	4.550	10.510
Outer Barrels (2)	SS	1, 3	50 / –	6.00	2.715	7.582
	DS	2, 4	50 / 62.5	6.00	4.550	10.510
F–Disks	DS	–	50 / 62.5	7.93	2.570	9.960
H–Disks	SS	–	40 / –	7.63^{i}	9.500	26.000
	SS	–	80 / –	6.33^{o}	9.500	26.000

Table 2: Characteristics of the SMT different sensors. ^{i} Indicates the length of the inner H–Disk sensor; ^{o} is the length of the outer H–Disk sensor.

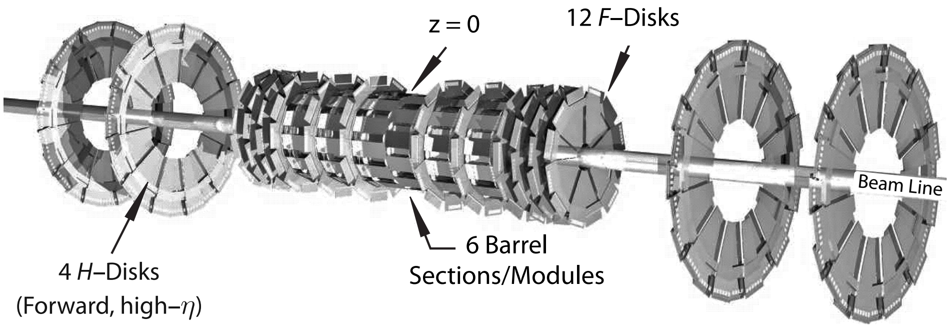


Figure 13: The Silicon Microstrip Tracker is composed by sets of detectors grouped in Barrels, F–Disks and H–Disks.

Central Fiber Tracker: CFT

The CFT consists of eight concentric cylinders, 1 to 8 beginning with the innermost and outward. Their radii go from 20 cm to 52 cm. Cylinders 1 and 2 are 1.66 m and the remaining are 2.52 m long. The reason of this difference is that the SMT H–Disks are supported by the cylinder 3. The CFT outermost cylinder covers $|\eta| < 1.7$.

Each CFT cylinder have two layers of scintillating fibers mounted on its surface. The first layer is oriented along the beam direction and is referred as Axial Layer. The second layer is referred as Stereo Layer because its fibers are oriented with an stereo angle. Odd number Stereo Layers are *u*–Layers and have an stereo

Layer	Radius (cm)	Fibers × Layers	Fiber Separation (μm)	Active Length (m)
1	20.04	1,280 × 2	982.4	1.66
1_u	20.22	1,280 × 2	990.3	1.66
2	24.93	1,600 × 2	978.3	1.66
2_v	25.13	1,600 × 2	985.1	1.66
3	29.87	1,920 × 2	976.1	2.52
3_u	30.05	1,920 × 2	980.9	2.52
4	34.77	2,240 × 2	974.4	2.52
4_v	34.95	2,240 × 2	979.3	2.52
5	39.66	2,560 × 2	971.7	2.52
5_u	39.86	2,560 × 2	976.3	2.52
6	44.56	2,880 × 2	970.0	2.52
6_v	44.74	2,880 × 2	974.3	2.52
7	49.49	3,200 × 2	969.8	2.52
7_u	49.67	3,200 × 2	973.3	2.52
8	51.97	3,520 × 2	926.1	2.52
8_v	52.15	3,520 × 2	927.8	2.52

Table 3: Design parameters of the eight axial and stereo Layers comprising the CFT; indexes refer to the stereo angle of the corresponding layer according to $u_{angle} = +3°$, $v_{angle} = -3°$.

angle of $\phi = +3°$. Even number Stereo Layers are *v*–Layers and have an stereo angle of $\phi = -3°$.

The scintillating fibers, are 835 μm in diameter and match the length of the cylinder they are mounted in. The core of the fiber is made from Polystyrene whose refractive index is $n = 1.59$ and is surrounded by two coats of approximately 25 μm each. The inner is a layer of Polymethylmethacrylate with $n = 1.49$, and an outer layer of fluoro–acrylic with $n = 1.42$.

The scintillating fibers are optically connected to clear fiber waveguides of identical diameter which are 7.8 m to 11.9 m long and carry the scintillation light to Visible Light Photon Counters, VLPC's, for readout. The scintillating fibers are structural and chemically similar to the clear fibers, but contains fluorescent dyes.

The light is observed by only one end of each scintillating fiber. The opposite end of each of the scintillating fibers is mirrored with a sputtered aluminum coating that provides a reflectivity of about 90%.

The total 76,800 readout channels of the CFT use around 184 km of scintillating fiber and 800 km of clear fiber. Table 3 shows the summary of the design parameters of the CFT.

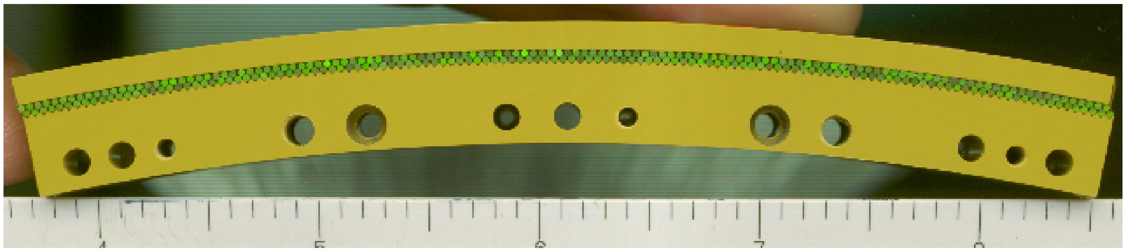


Figure 14: The Central Fiber Tracker is made by eight cylinders covered by two layers of scintillating fibers each, like the one shown here. The first layer is axially–oriented and is located right in the surface of the cylinder, the second layer follows and is stereo–oriented.

Cryostat Warm Bore Diameter	1.067 m
Cryostat Length	2.729 m
Thickness	0.87 X_0
Central Field	2.0 T
Operating Current	4,749 A
Cold Mass	1,460 kg

Table 4: Main parameters of the Solenoidal Magnet

Solenoidal magnet

The superconducting Solenoidal Magnet help to optimize the momentum resolution, $\delta p_T/p_T$, and tracking pattern recognition. Its size is constrained by the available space in the hollow inside the Central Calorimeter. It is 2.73 m in length and 1.42 m in diameter and it provides a central field of 2 T.

The main parameters of the Solenoidal Magnet are listed in Table 4. A perspective view of the Solenoidal Magnet inside the Central Calorimeter with its chimney and control dewar is shown in Fig. 15.

Central and Forward Preshower Detectors

The Preshower Detectors have two purposes: to work as calorimeters as well as tracking detectors; enhancing the spatial matching between tracks and calorimeter showers. In accordance with the η region they cover, there are two different Preshower Detectors.

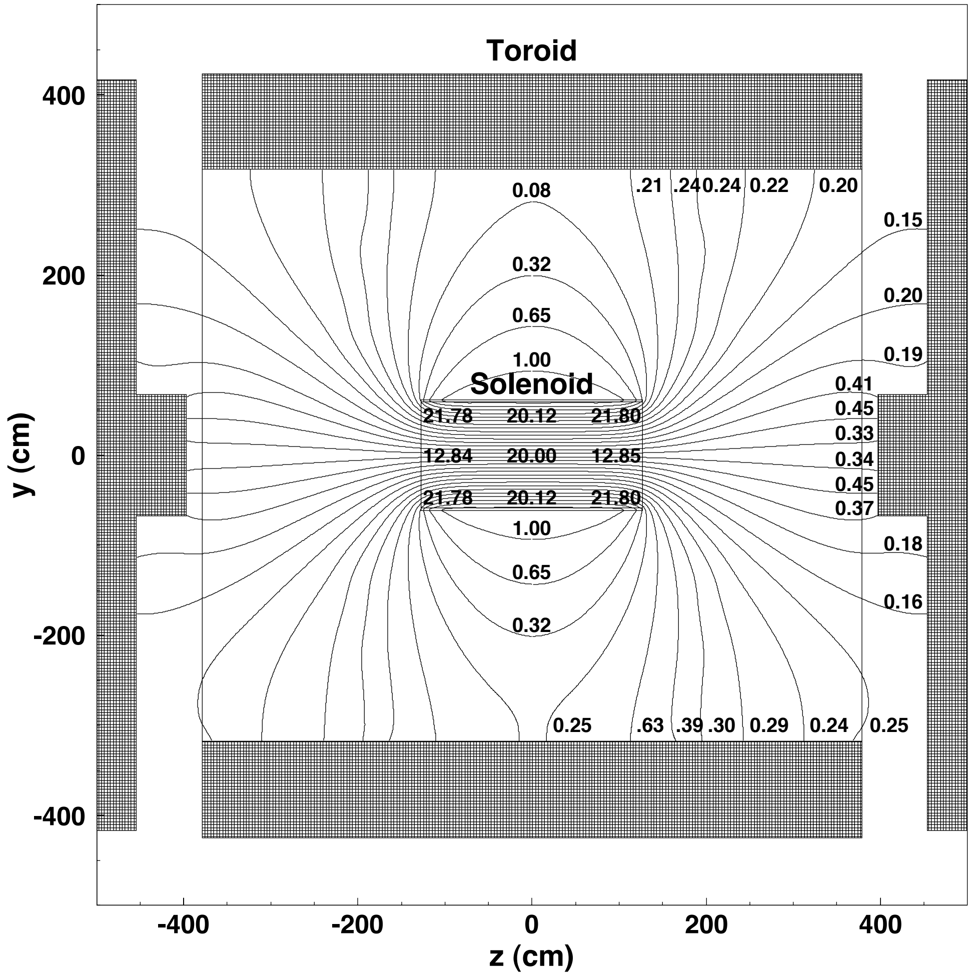


Figure 15: The 2 T Solenoidal Magnet resides in the space between the Central Fiber Tracker and the Central Calorimeter.

Central Preshower (CPS). Covers the region $|\eta| < 1.3$ and is located between the Solenoidal Magnet and the Central Calorimeter with an inner radius of 71.75 cm and an outer radius of 74.19 cm.

Forward Preshower (FPS). The two Forward Preshower Detectors cover $1.5 < |\eta| < 2.5$ and are attached to the faces of the End Calorimeters.

Both are made from triangular strips of scintillator. Since the triangles are interleaved, there is no dead space between strips and most tracks traverse more than one strip, allowing for strip–to–strip interpolations and improved position measurement.

For optical isolation, each scintillator strip is machine–wrapped in aluminized Mylar, which is a film of bi-axially–oriented Polyethylene Terephthalate (boPET). The ends are painted white to enhance reflectivity.

Embedded at the center of each triangular strip is a wavelength–shifting (WLS) fiber that collects and carries the light to the end of the detector.

The non–readout ends of the WLS fibers are diamond–polished and silvered. In the other end, the readout end, fibers are grouped into bunches of sixteen and potted into connectors for transition to clear light–guide fibers. Light is transmitted via the clear fibers to VLPC's for readout. Both the WLS and clear fibers are the same as in the CFT.

Central Preshower Detector

The CPS consists of three concentric cylindrical layers of triangular scintillator strips and is located in the nominal 5 cm gap between the Solenoidal Magnet and the Central Calorimeter.

Between the solenoid and the CPS is a lead radiator 5.56 mm thick (approximately 1 radiation length X_0) and 243.84 cm long, covering $|\eta| < 1.31$. The lead is covered by stainless steel skins 0.8 mm thick and 261.62 cm long. The solenoid itself is 0.9 X_0 thick, providing a total of about two radiation lengths of material for particles at normal incidence, increasing to about four radiation lengths at the largest angles.

The three layers of scintillator are arranged in an axial–u–v geometry. The u–stereo–angle is 23.774° and the v–stereo–angle is 24.016°. Each layer contains 1,280 strips. The WLS fibers are split at $z = 0$ and read out from each end resulting in 2,560 readout channels/layer.

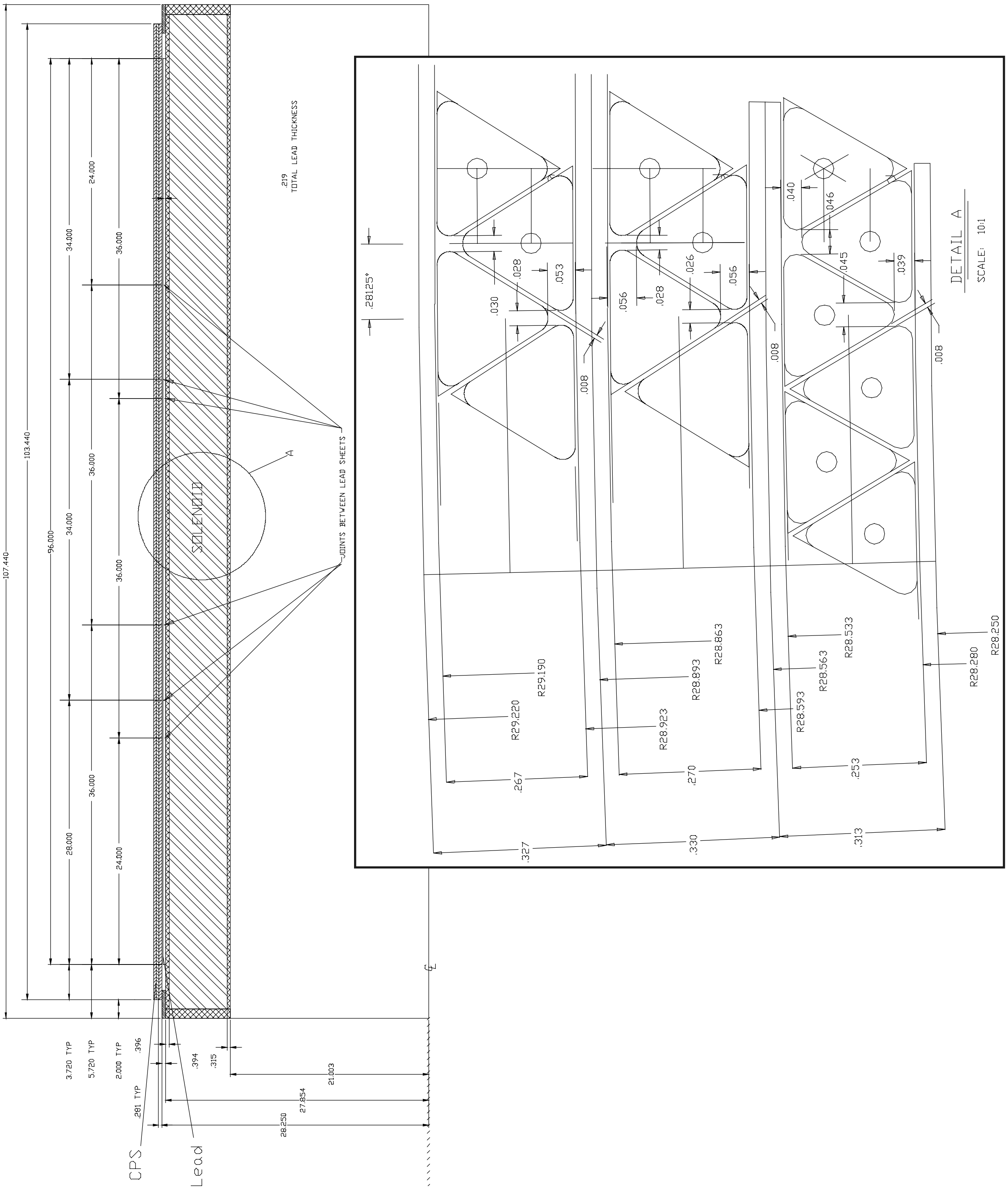


Figure 16: The Central Preshower Detector is located in the gap between the Solenoidal Magnet and the Central Calorimeter. On the frame: Layers are axial, *u*, and *v* stereo in increasing radius. Note the change in strip to strip spacing in the three layers to accommodate the similar channel count of each layer. All dimensions in inches.

Forward Preshower Detector

The two FPS Detectors are mounted on the spherical heads of the End Calorimeter Cryostats, between the Luminosity Monitor and the Intercryostat Detectors.

Measuring planes in the FPS contain eight 22.5° wedges of active material which have two sub–layers of nested scintillator strips with a stereo angle of 22.5° with respect to one another, separated by eight wedges of inactive material. Two measuring planes which are named u and v, respectively, are grouped together to form a layer.

Each detector is made from two of these layers, at different z, with a 2 X_0–thick lead–stainless–steel absorber between them. The layer nearest the interaction region is known as the Minimum Ionizing Particle Layer, or MIP Layer, has 206 scintillator strips –which are perpendicular to an edge at constant ϕ edge– and covers the region 1.65 $<|\eta|<$ 2.5. The farthest –behind the absorber– is called Shower Layer, has 288 strips –also perpendicular to a constant ϕ edge– and covers the region 1.5 $<|\eta|<$ 2.5. The absorber covers the same η region that the MIP–Layer and is divided into forty–eight wedge–shaped segments subtending 7.5° in ϕ. The extra region of the Shower Layer not covered by the MIP Layer, lies in the shadow of the Solenoidal Magnet Coil, which provides up to 3 X_0 of material in front of the FPS. This amount of material readily induces showers that can be identified in the Shower Layer.

In the vicinity of the Solenoid Cryogenics service pipes, the wedges are notched to allow these to enter and have only 142 strips per wedge in both the MIP and Shower Layers. The presence of these special wedges reduces the coverage to 1.8 $<|\eta|<$ 2.5 in this area.

Charged particles passing through the detector will register minimum ionizing signals in the MIP Layer, allowing measurement of the location in η, ϕ, and z of the track.

Electrons will readily shower in the absorber, leading to a cluster of energy, typically on the order of three strips wide, in the shower layer that is spatially matched with the MIP Layer signal. Heavier charged particles are less likely to shower, typically producing a second MIP signal in the Shower Layer. Photons will not generally interact in the MIP Layer, but will produce a shower signal in the Shower Layer.

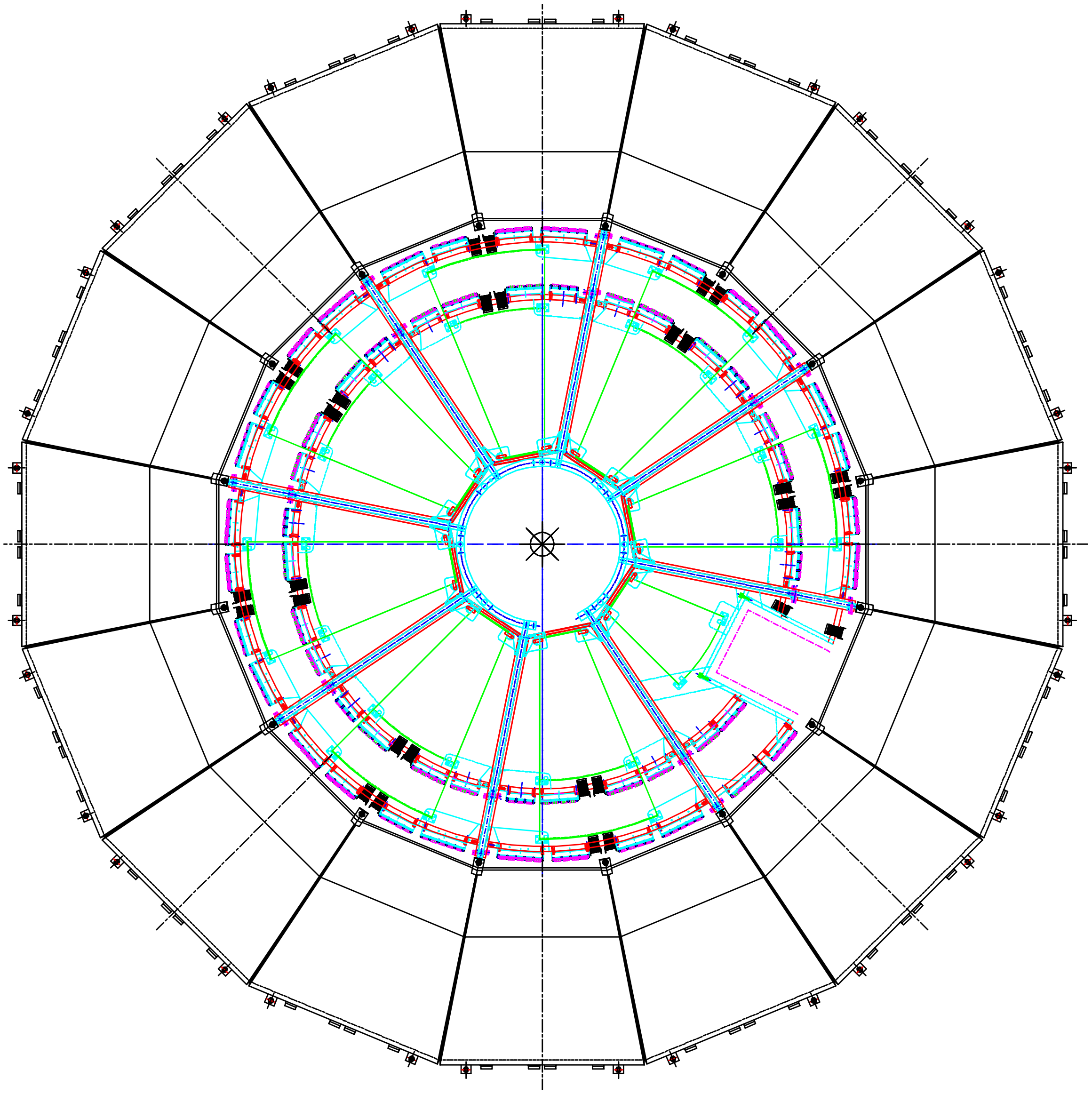


Figure 17: The Forward Preshower Detectors are mounted on the spherical heads of the End Calorimeter Cryostats.

Calorimeter

The liquid–argon calorimeter provides energy measurements for electrons, photons and jets. It remains unchanged from Run I, but the reduced bunch crossing interval of 396 ns in Run II (from 3.5 μs in Run I) has required an upgrade of its front–end electronics. The former peak sampling time of 2.2 μs was reduced to 400 ns, mainly by decreasing the effective integration time and reducing the intrinsic noise of the pre–amplifier.

The calorimeter, as shown in Fig. 18, is divided into a central calorimeter covering and two end calorimeters to cover the region $1 < |\eta| < 4$. Both calorimeters contain three basic types of modules: An electromagnetic section with thin nearly pure depleted uranium plates (thickness 3 or 4 mm), fine hadronic modules with 6 mm thick uranium–niobium (2%) alloy plates and coarse hadronic sections. These latter sections consist of 46.5 mm thick plates, made of copper for the central calorimeter and stainless steel in the end calorimeters. There are four separate layers for the electromagnetic modules, three for the fine hadronic and one or three for the coarse hadronic modules. These layers are used to measure the longitudinal shower shape to distinguish between electrons and hadrons.

The smallest unit of the calorimeter is a read–out cell. The read–out cells are organized in semi–projective towers, pointing toward the interaction region. Semi–projective means that a straight line can be drawn through the interaction point and the cells of one tower, but that the sides of the cells are not aligned toward the interaction point. The segmentation of the read-out cells is (0.1×0.1) in $(\eta,\ \phi)$, except for the third layer in the electromagnetic modules, where cells cover (0.05×0.05) in $(\eta,\ \phi)$.

Intercryostat Detector

The region between the central and end calorimeters ($0.8 < \eta < 1.4$) contains large amounts of inactive material, mostly the walls of the calorimeter cryostats, light guides coming from the CFT and read–out electronics connected to the SMT and the magnetic field monitors for the solenoid. To supplement the energy measurement for particles in this region, an inter–cryostat detector has been installed. This detector consists of a single layer of 384 scintillating tiles, covering (0.1×0.1) in $(\eta,\ \phi)$, to match the calorimeter cells. The tiles are mounted on the front surface of the end calorimeters. The light signals are picked up by wavelength shifting fibers in the tiles and transported via clear fiber ribbon cables toward photo–multiplier tubes outside the magnetic field.

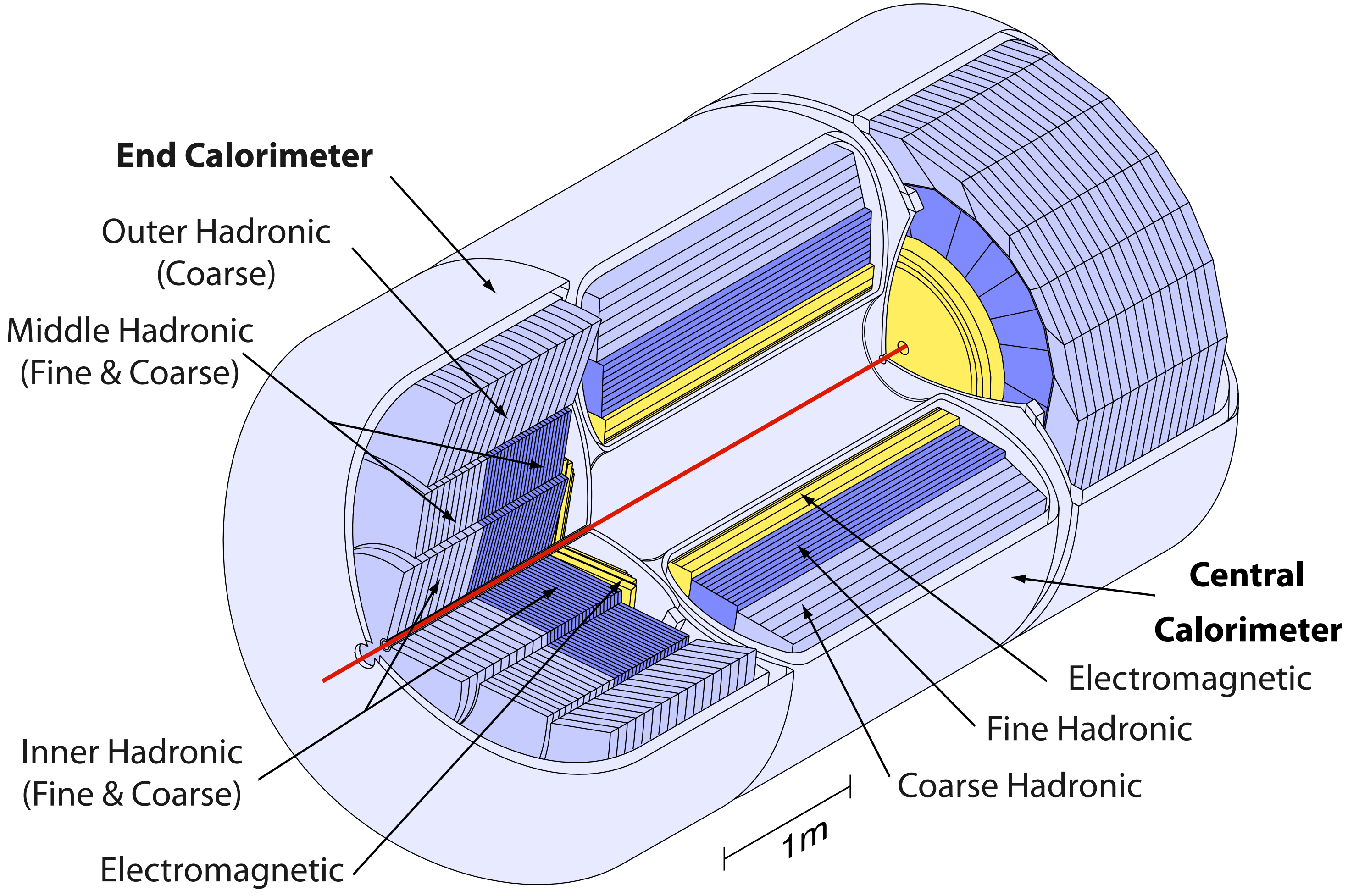


Figure 18: The calorimeter in the DØ Detector.

Muon system

The muon detectors are based on two technologies: Drift chambers and scintillation counters. Both technologies are low–cost relative to the area that can be covered. This is important, because the muon system is located at the outer radius of the detector and therefore large area coverage is needed to obtain the necessary angular coverage. A drawback of drift chambers is their poor measurement of the time when a particle passes through. Therefore the drift chambers are supplemented with scintillation counters, that do not provide accurate tracking information, but have very good timing resolution. The timing resolution of the scintillator counters is of the order of 1 ns. Finally, a magnet system is provided, to enable momentum measurement in the muon system. The DØ muon system consists of three layers of drift–chambers and scintillation counters, called the *A*–, textitB– and textitC–layer, labeled starting at the inner layer, and a toroid magnet system. Fig. 19 shows the schematic view of the muon detector at DØ.

It is divided into central and forward systems, covering the range $|\eta| < 1$ and the range $1 < |\eta| < 2$, respectively. The central muon system consists of the WAMUS drift chambers, the Cosmic cap and bottom scintillation counters, and the *A*–ϕ scintillation counters. The forward muon system consists of the FAMUS drift chambers, as well as the pixel scintillation counters. The design of these detectors will be described in the following paragraphs.

Muon Magnets

The magnet for the muon system is an iron–core toroid. The iron core helps reaching a high field value, at modest current (and operational cost). As for the muon system itself, it consists of a central and a forward section. The central toroid magnet, is a square annulus, with a thickness of 109 cm and a weight of 1973 metric tons. The forward toroid core is made of single large weldments of plate steel. The central and forward toroids are currently operated in series at 1,250 A, resulting in a magnetic field of approximately 1.8 T.

WAMUS Drift Chambers

WAMUS is an acronym for Wide Angle Muon System. It consists of proportional drift tube chambers (PDT's). Its purpose is to provide muon identification and an independent momentum measurement. The PDT's are typically 250×560 cm. They are made of rectangular aluminum tubes, 10.1 cm wide. These aluminum

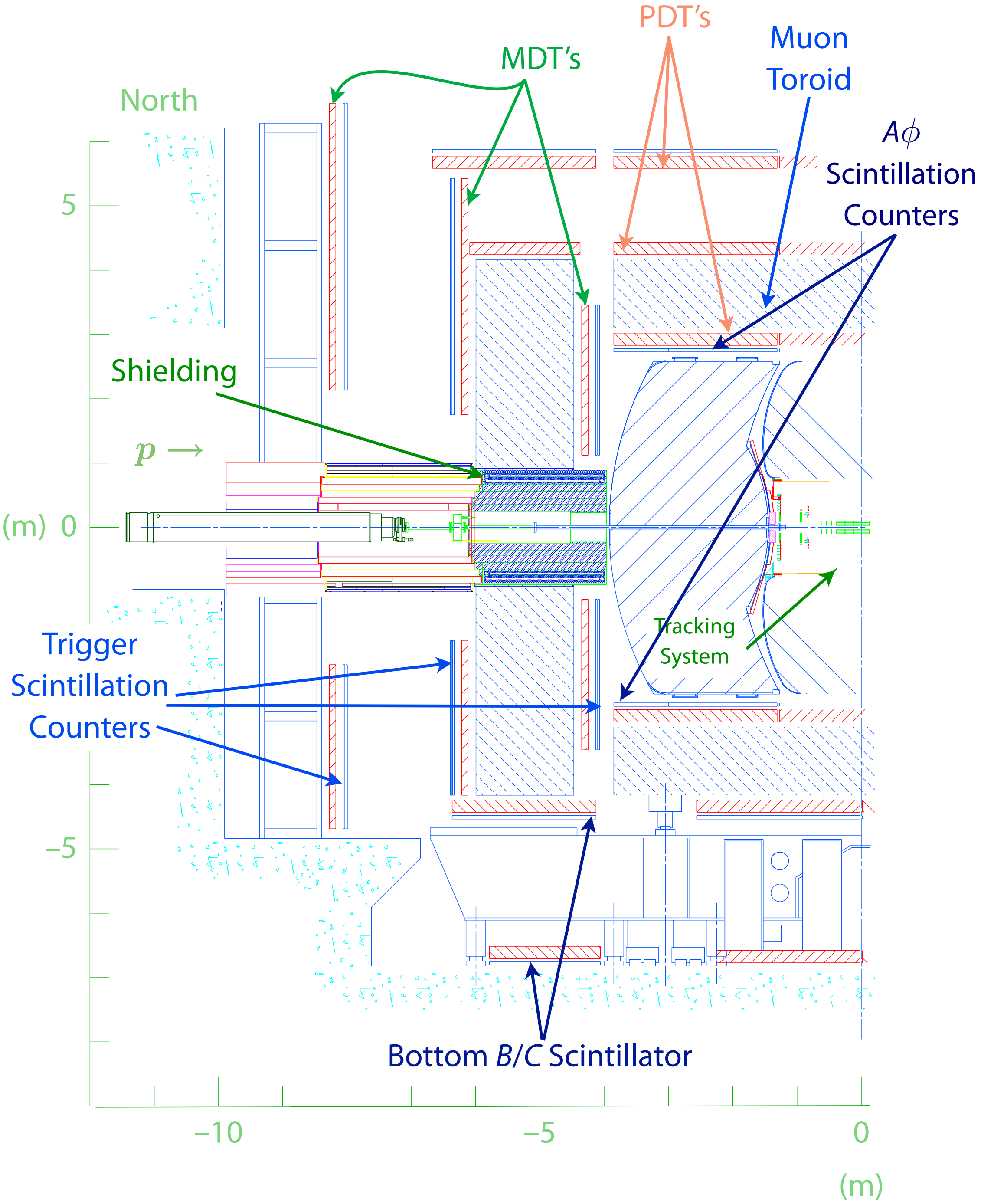


Figure 19: Muon System in the DØ Detector.

tubes are the drift–cells. The PDT's outside the magnet have three decks of drift–cells. The PDT's inside the magnet (the *A*–layer) have four decks of drift cells, with the exception of the bottom PDT's inside the magnet, which have three decks. Each drift cell is equipped with Vernier pads. The Vernier pads provide an additional measurement of the hit–coordinate along the signal wire. The pad read–out however is only instrumented in the layer inside the toroid and in about 10% of the other two layers, because with the addition of scintillators in Run II, the Vernier pad measurement only represents a minor resolution improvement.

The signal from the PDT wires is the time measured on the anode wire in the center of the cell. The two wires of neighboring cells are connected with a delay, so that two times are measured for each hit. Knowing the propagation times of the signal along the wire and the delay–time, it is possible to calculate the drift–time and axial time (time of propagation along the wire). After calibration, these times can be converted into distances. The resolution on the drift distance is of the order of 500 microns. The resolution on the axial distance varies with the location of the hit along the wire. If the hit occurs far from the electronics, the resolution is approximately 10 cm. If the hit occurs near the read–out, the signal has to propagate two wire lengths, and due to dispersion of the signal the resolution on the axial distance will be approximately 50 cm.

Cosmic cap and bottom scintillation counters The "Cosmic Cap" scintillation counters cover the top, sides and bottom of the PDT *C*–layer. They are used to determine the timing of muon tracks in the PDT chambers relative to the bunch crossing and thereby determine from which bunch crossing they originate. Similarly they are used to reject cosmic rays, by determining that the timing of the track is inconsistent with any bunch crossing. The limited coverage of the counters underneath the bottom *C*–layer, due to the detector support structure, is supplemented with counters underneath the PDT *B*–layer.

A–ϕ Scintillation Counters

The *A*–ϕ counters have been mounted on the inside of the *A*–layer PDT's. Similarly to the Cosmic Cap counters, the *A*–ϕ counters measure the timing of muon tracks in the PDT's they are mounted on, to determine the bunch crossing of the track. They are also used to reject out–of–time particles. For the *A*–ϕ counters, however, these are mostly not cosmic particles, but backscatter from the forward direction. The *A*–ϕ counters are 4.5° wide, matching the expected multiple scattering for high–p_T muons. One of the arguments for the construction and installation of the *A*–ϕ counters, was to use them in a level 1 $J/\psi(1S) \to \mu^+\mu^-$ trigger. The threshold for muon identification in Run I was around 4 GeV, corresponding to the energy needed to penetrate both the calorimeter and the muon toroid. Tracks that were stopped in the muon toroid were not usable (so–called *A*–stubs), due to the high background of out–of–time hits in the PDT *A*–layer. With the addition of the *A*–ϕ counters, these background hits have been reduced, thereby improving the signal to noise ratio of the *A*–stubs to a useful level. The threshold for muon identification could therefore be lowered to 1.1–1.4 GeV.

FAMUS Drift Chambers

The Forward Angle Muon System, or FAMUS, is made of 3 layers of mini drift tube detectors (MDT's), and covers the range $|\eta| < 1$ and the range $1 < |\eta| < 2$, on both sides of the interaction point. An MDT is rectangular in shape and contains a stack of 8 drift–cells. "Iarocci tubes" have been chosen as the technology for the drift–cell, because this technology is known to be robust, reliable, and efficient. To increase efficiency, the FAMUS *A*–layer contains four layers of MDT tubes, and the *B*– and *C*–layers each contain three layers. The Iarocci drift–cells have an efficiency close to 100%. However, each 9.4 mm wide cell is separated from the next by a 0.6 mm thick wall, which introduces an inefficiency of approximately 6%. Also, the efficiency is reduced near the wire support in each cell. Furthermore, there are dead areas between octants, for gas, high voltage and signal connectors. Overall, the efficiency for finding two hits on a muon track in one layer of the MDT detector (out of three expected) is around 90%. The momentum resolution of the MDT detector is limited by the coordinate resolution of the detector, as well as multiple scattering in the toroid. The drift distance measurement by the Iarocci tubes has an accuracy of around 0.7 mm, and after digitization in 18.8 ns bins, the accuracy is about 0.9 mm. With 3 or 4 planes, the resolution is effectively about 0.7 mm. Taking into account this resolution, the multiple scattering in the toroid and the layout of the planes, the forward muon system has a momentum resolution of around 20% for low momentum muons.

Pixel Scintillation Counters

The purpose of the forward trigger counters, or pixel scintillation counters, is to provide accurate time information to supplement the measurements from the MDT's. This allows substantial reduction of the background rates in the MDT's. The ϕ segmentation of the counters is 4.5°, and the segmentation in η is 0.1. This segmentation was chosen to reasonably limit the total number of counters over the total area to be covered. The area of the *C*–layer is 12×10 m^2, and there is a total of almost 5,000 counters in this system.

Luminosity Monitor

The luminosity is determined from the total event rate for inelastic $p\bar{p}$ interactions. This rate is measured with scintillator tiles mounted close to the beam pipe on the front surfaces of the end calorimeters. From the event rate, the luminosity can be calculated using the total cross–section for $p\bar{p}$ inelastic interactions and the known acceptance of the luminosity monitor.

Trigger

Must of the $p\bar{p}$ collisions lead to events of little interest. To accumulate a large sample of events of interest without having to store and reconstruct a staggering number of uninteresting events, DØ employs an event trigger which decides whether to store an event or to disregard it. The trigger system is organized in three major levels, the hardware based level 1 and level 2, and the software based level 3. These levels are referred to as L1, L2 and L3.

Each level examines the event in more detail than lower tiers and restricts the rate of events to higher tiers. This is reflected in the amount of decision time per event that each level has available.

The hardware L1 trigger takes input from the calorimeter, the muon system and the luminosity monitor, with an event rate of 2.5 MHz. As output, it produces a list of so–called trigger terms. The L1 framework, employing a series of field programmable gate arrays (FPGA's), examines these trigger terms and issues an L1 accept when certain criteria are met.

The level 1 muon subsystem takes input from the drift chambers and the muon scintillators. A line through the two scintillators should approximately point back to the interaction point. This is a very efficient trigger for high–p_T muons, but low–p_T muons can be stopped either in the calorimeter or in the iron magnet core between the *A*– and *B*–layers of the muon system. On average an energy of around 4 GeV is required to pass through the calorimeter and muon toroid.

The L2 trigger is the first trigger to match information from different sub–detectors. It is comprised of two stages, a preprocessor stage and a global processor stage. At the preprocessor stage, each individual sub–detector forms physics objects. At the global processor stage, physics objects from the different preprocessors can be combined to make the final L2 trigger decision.

Once an L2 accept has occurred, digitized data is loaded onto the so–called single board computers (SBC's). Data from the SBC's are transferred to the L3 farm. Each processor of the farm is called a "level 3 node", runs an independent instance of the L3 filtering software and processes a complete event and makes the trigger decision. Events that are passed, are transmitted to the data logger, from where they are copied to tape.

The basic elements of the L3 filtering software are the tools, which contain the L3 reconstruction algorithms. These algorithms mimic the reconstruction algorithms used offline, but have improved processing speed at the cost of some precision. The reconstructed objects provided by the tools (e.g. tracks, muons, jets) are passed to filters, which apply quality criteria. The filter scripts are easily modifiable lists defining the trigger parameters.

Analysis

On this chapter, the actual analysis and its details is described. It contains the main core of this thesis, and presents the extensive set of tests that this analysis had to pass internally at the DØ Collaboration before it was approved to see the public light.

DØ Data

On April 2002, DØ started to take Physics Data from the Tevatron period known as RunII. By the end of 2010, Tevatron had delivered 10.06 fb^{-1} from which DØ had recorded 8.99 fb^{-1} [14].
From February to June 2006, a long shutdown took place. It was during this time that the DØ Detector was overhauled and significant changes were made in its configuration with the addition and remotion of a number of detector components, as well as other modifications in order to improve its performance. Because of the different configurations of the DØ Detector and/or data availability periods, the DØ Data is suitable divided. RunIIa is the name given to all data before the 2006 shutdown. RunIIb follows after this shutdown.

This analysis was made with all RunIIa data, *i. e.*, approximately 1.3 fb^{-1}. Fig. 20 shows the DØ Data used for this analysis.

Decay chain and search procedure

The decay chain of interest is:

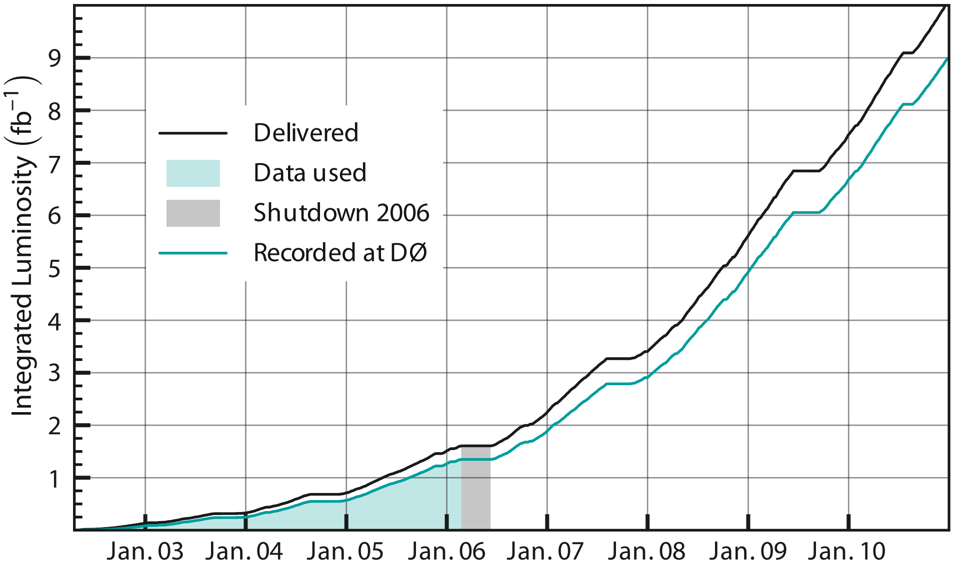


Figure 20: Tevatron had delivered 10.06 fb^{-1} since April 2002 until the end of 2010. The DØ Detector had recorded 8.99 fb^{-1} [14]. The shaded areas represent: in green, the DØ Data used in this analysis and in grey the long shutdown during 2006.

$$\Omega_b^- \to \overbrace{J/\psi(1S)}^{\mu^+\,\mu^-}\ \underbrace{\Omega^-}_{\underbrace{\Lambda}_{p\,\pi^-}\ K^-}$$

Charge conjugation is assumed from now on.

The procedure for the search of this specific decay is based on the individual reconstruction of all the intermediate particles, starting with the final products available for us, *i. e.*, charged particles like μ's, p's, π's and K's, and going backwards until de whole decay chain is reconstructed.

This way, two opposite–charge tracks identified as muons are used to form a $J/\psi(1S)$, then two more, this time supposing they are a proton and a pion are joined to construct a Λ. Another charged–track is added to the constructed Λ in order to produce a Ω^-, which combined with the $J/\psi(1S)$ will produce an Ω_b^-. Fig. 21 shows schematically the decay chain.

In order to construct the particle of interest, the whole dataset is divided isolating samples containing the

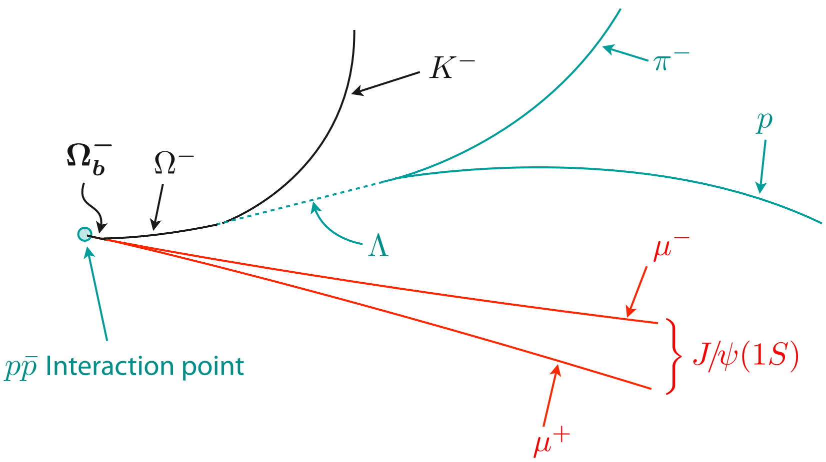


Figure 21: Schematic description of the decay chain.

sequence of intermediate particles needed. The events in the final subset go trough a set of previously validated steps to make sure no artificial generation of any kind is taking place. The events that survive all this process are those in which a relevant resonance is expected to be found.

Even when the implementation of the global reconstruction of the decay chain is relatively complex, it can be easily summarized by the following pseudocode:

1. Isolate events with at least a $J/\psi(1S)$.
2. Find those who contain a Λ.
3. Leave only the ones who reconstruct Ω^- candidates.
4. Proceed to a final combination between the $J/\psi(1S)$ and the Ω^-.
5. Validate the combination.

Reprocessing

As all the information recorded with the detector is useful, ideally we should reconstruct as many tracks as we have available information to be used. In practice, to process the recorded information is a relatively time–consuming operation. This is particularly true with the increasing luminosity because of the large multiplicity of tracks. To alleviate this situation, a series of offline cuts are applied during the process of reconstruction of the tracks from the information recorded with the detector.

We define the Primary Vertex, PV, as the point at which the initial interaction $p\bar{p}$ took place, and the Track Impact Parameter, IP, as the perpendicular distance to the closest approach of the track referred to the PV. Fig. 22.

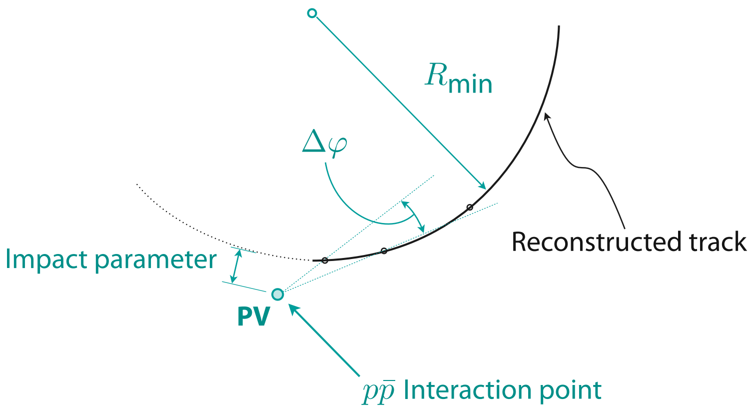


Figure 22: Primary Vertex (PV) and Impact parameter.

The standard reconstruction of DØ data takes into account all tracks with an Impact Parameter up to 2.5 cm, Fig. 23. This requirement provides a good compromise between the CPU–time needed to reconstruct the tracks and the physics we are able to extract from the result of the reconstructed information.

However, some difficulties emerge when long–lived particles are involved. The Maximum Impact Parameter described, decreases the efficiency in the reconstruction of long–lived particles. This kind of particles travel a few centimeters in the detector before decaying, and their resulting products may have impact parameters

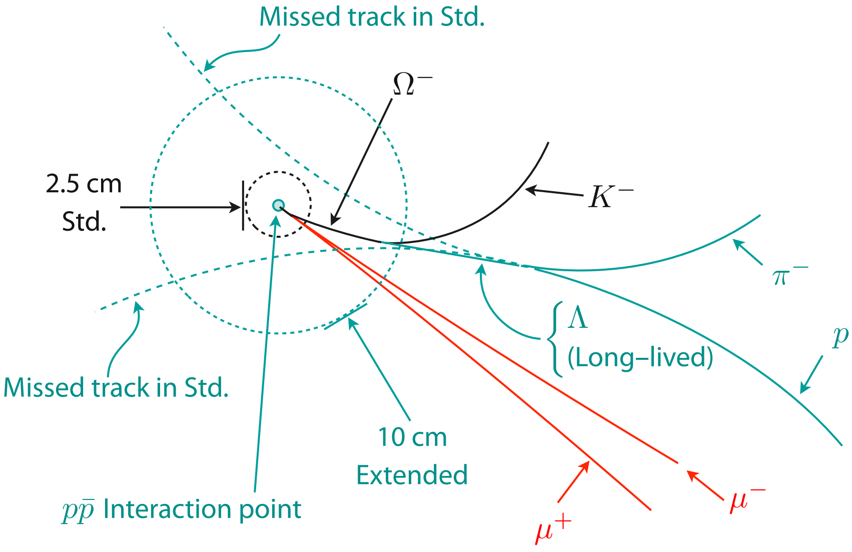


Figure 23: Long–lived particles may be missed by the standard reconstruction due to its large impact parameter.

greater than 2.5 cm, for this reason, even when the information to process these tracks is there, they are not reconstructed. Typical examples of decay chains which are affected by this characteristic are: K_S^0, Ξ^-, and Ω^-.

An extended reconstruction of particle trajectories was used in this analysis in order to increase the reconstruction efficiency of long–lived particles. Events with dimuon candidates chosen in accordance with the DØ Common Samples Group, were selected to be reprocessed with an official extended version of the DØ tracking algorithm.

In the extended reconstruction, in order to process a particle trajectory:

- its minimum radius, R_{min}, may be as short as 20 cm (reduced from the standard 30 cm), Fig. 22,
- its impact parameter –relative to the XY beam position– can be as large as 10 cm (increased from the standard 2.5 cm), Fig. 23.
- the maximum acceptance angle, $\Delta\varphi$, is allowed to be as large as 0.20 rad (increased from the standard 0.08 rad), Fig. 22.

In this reprocessed RunIIa data, the Ξ_b^- was discovered in the summer of 2007 [15].

This process starts on one of the stages of processed information, called TMB++ format, and preforms only particle trajectory reconstruction. The remaining information in the event is copied from the original standard reconstruction.

Monte Carlo

In Experimental High Energy Physics, the collision of particles in accelerators lead to the creation of unstable particles, these particles decay quickly. The event in which the collision took place may be reconstructed with the help of the measurements made by a series of detectors.

The role of a Monte Carlo simulation is to mimic what happens when a collision of some kind takes place and what happens when the products of the collision interact with the detector. Two steps are needed to reach a suitable sample of Monte Carlo events. The first one is the generation of events using an event generator and some specialized decay package. Then, in order to include the behavior of the detector itself, the generated events undergo a detector simulation.

For this analysis, two Monte Carlo samples with different lifetimes were generated containing the decay chain previously described. A total of 77,750 events were generated with $c\tau = 0$ μm and 127,500 events with $c\tau = 462$ μm.

The input mass used was 6.052 GeV$/c^2$. PYTHIA [16], was used to generate the Ω_b^- particles and EVTGEN [17] to produce the decays to the intermediate and final products. Then, GEANT [18], is used to include the DØ detector simulation. The complete simulation was performed with the DØ release p17.09.06 and the same extended reprocessing as in real data was applied.

At generation level (`d0_mess`), the following cuts are required[1]:

```
string Cut1 = "PdgId == 443 && AbsParentId == 5332"
string Cut2 = "PdgId == 13 && Pt > 1.5 && AbsEta < 2.5 && ParentId == 443"
string Cut3 = "PdgId == -13 && Pt > 1.5 && AbsEta < 2.5 && ParentId == 443"
string Cut4 = "AbsPdgId == 3334 && AbsParentId == 5332"
```

---

[1] Particle identification numbers from [21].

$J/\psi(1S)$ selection

In the reprocessed RunIIa data, a sample of events with at least one dimuon candidate is isolated. These events satisfy the following requirements [20]:

- Two certified muons of opposite charge,
- $p_T > 1.0$ GeV for each particle,
- both muons have at least one hit on the CFT (N(CFT) > 1),
- at least one of the muons leaves signatures on the three muon–detector layers (nseg $= 3$).
- if nseg > 0, then $p_T > 1.5$ GeV (*i. e.* allow $p_T < 1.5$ GeV only for muons with nseg $= 0$).
- for muons with nseg $= 0$, extra requirements are imposed
  - $p_T\left(\mu^+\mu^-\right) > 4.0$ GeV,
  - $p_{tot} < 7.0$ GeV,
  - the second muon must have a $p_T > 2.5$ GeV, and
  - the global muon fit must have a $\chi^2 < 25$ for both particles.
- M $\left(\mu^+\mu^-\right) > 2.5$ GeV$/c^2$ for their masses[2].

which are implemented into the BAna package [19].

The $J/\psi(1S)$ sample for this analysis, is composed by all dimuon events in the mass range from 2.75 to 3.4 GeV$/c^2$, whose vertex have a χ^2 probability greater than 1% and the distance between the primary and dimuon vertices is 10 cm at most.

All the events passing the previous requirements are taken into account to produce a histogram of the $J/\psi(1S)$ resonance. The model used to fit the mass distribution is composed by a gaussian peak and a linear background as follows

$$h_j(m) = j_0 \exp\left\{-\frac{1}{2}\left(\frac{m-j_1}{j_2}\right)^2\right\} + j_3 + j_4 \cdot m, \tag{1}$$

---

[2] The conditions for high dimuon–mass candidates inherent to the code, M $\left(\mu^+\mu^-\right) > 3.6$ GeV$/c^2$, are not described because they do not apply due to the mass window selected.

Parameter	Mean Value	Error	Meaning
j_0	6,747.29	26.3118	Gaussian height
j_1	3.083	0.000255643	Gaussian mean value (GeV$/c^2$)
j_2	67.5332	0.283236	Gaussian width (MeV$/c^2$)
j_3	4,534.95	86.4156	Linear background h–intercept
j_4	–679.978	27.9218	Linear background slope (c^2/GeV)

Table 5: $J/\psi(1S)$ signal fit results. Fig. 24.

where m is the mass and $j_i \ \ (\forall \ i = 0, \ \ldots, \ 4)$ are the parameters extracted from the fit.

The mass distribution as well as the resulting fit is shown in Fig. 24 and the summary of the fit parameters is given in Table 5. Resulting in 175,721 $J/\psi(1S)$ events[3].

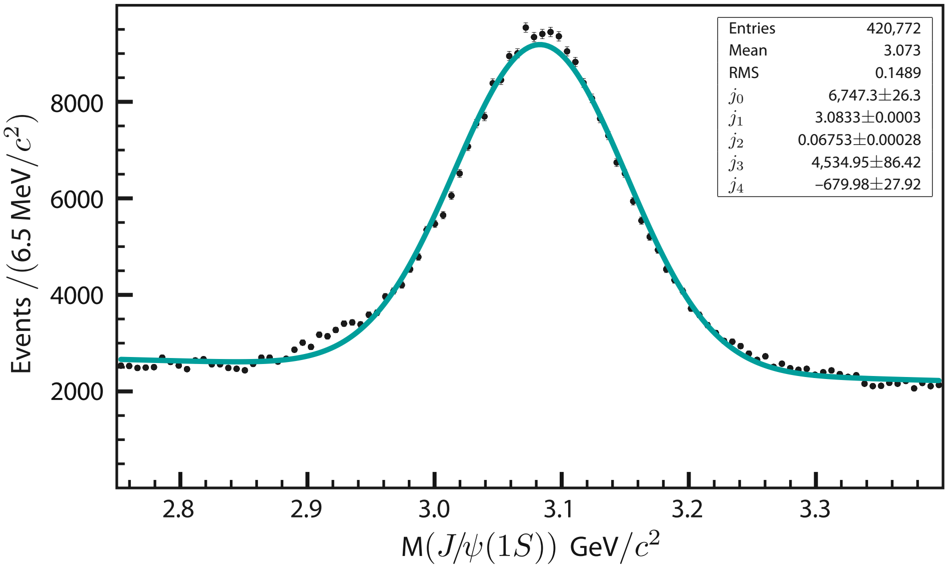


Figure 24: $J/\psi(1S)$ sample.

---

[3] Events under the Gaussian peak between the 6 $\cdot\,\sigma$ range centered in the Gaussian mean value.

Λ sample

The events containing a $J/\psi(1S)$ are the departing point for the search of the next intermediate particle of interest: the Λ decaying to $p\ \pi^{-}$.

On every event, the rest of the tracks –not belonging to the products of the $J/\psi(1S)$,– are used to look for opposite charged pairs, that assuming some conditions on their masses match the reconstruction of a Λ. This two tracks must have a common vertex whose χ^2 probability is greater than 1%, and have no more than two hits in the tracking system before the vertex.

MC simulations show that in almost all $\Lambda \longrightarrow p\ \pi^{-}$ events generated –as well as the corresponding anti–decay,– the track with the higher p_T is the one corresponding to the proton/antiproton. For this reason, in the search for Λ, we assign to the leading track[4] the proton mass and give the pion mass to the other track.

With this mass assignation, the charge of the leading track will give the information about whether it corresponds to a particle or the antiparticle.

A minimum p_T of 0.2 GeV is required for both products of the Λ decay and in order to reduce combinatorial backgrounds, the Λ decay distance from the primary vertex is required to be greater than 4 times its estimated uncertainty when this uncertainty is less than 0.5 cm.

$\Lambda\ K^{\pm}$ sample

At this point, another charged track is selected from the remaining tracks on the event, to be added to the events on the identified Λ sample, in order to reproduce the next particle: the Ω^{-}.

The mass assumption given to this extra track is that of a kaon, but this time two cases are distinguished:

Wrong Sign. If the charge of the kaon is opposite to the charge of the pion coming from the Λ decay, the baryonic number is not conserved and this sample is considered as background.

Right Sign. The charge of the kaon matches the charge of the pion coming from the Λ decay.

---

[4] The track with the greater p_T.

The Wrong Sign sample is used to study the background and optimization, while the Right Sign is the one to be used for the search of the Ω^-.

In both cases, the Λ and the K must have a common vertex whose χ^2 probability is greater than 1% and, in order to reduce combinatorial backgrounds, the Ω^- decay distance from the primary vertex is required to be greater than 4 times its estimated uncertainty when this uncertainty is less than 0.5 cm. In addition, the uncertainty of the transverse decay distance, *i. e.*, the uncertainty on the projection of its decay path over the plane perpendicular to the beam direction, is required to be less than 0.2 cm.

Even when some conditions have been applied to reduce the level of background for the Λ signal first, and then for its combination with a K, the signal may be cleaned a little more using known conditions from the particles involved.

The Λ travels a long distance in the detector before it decays ($c\tau =$ 7.89 cm), then it is reasonable to think that a cut on the significance of its proper decay length may be useful to clean the signal even more before the next step is taken.

The proper decay length is defined as a function of the projection over the plane perpendicular to the beam of the distance that the Λ travels before it decays, L_{XY}, the momentum M, and the transverse momentum p_T as:

$$\lambda = \frac{L_{XY} \cdot M}{p_T},$$

and is divided by its error, σ_λ, to compute the significance of the proper decay length. Then, this quantity is used to analyze the statistical significance of the signal, defined by the ratio between the signal events, S, and the square root of the total number of events ($S + B$), being B the number of background events.

The Wrong Sign sample is employed here to avoid the use of the Right Sign sample in which the actual Ω^- events are supposed to be included.

The Λ events corresponding to the Wrong Sign sample in the mass window between 1.6 and 1.8 GeV$/c^2$, are shown in Fig. 25, with a superimposed fit using a model similar to the one on Eq. 1, *i. e.*, with a gaussian peak and a linear background. Table 6 gives the summary of the results from the fit.

The signal significance reaches its maximum value around $\lambda/\sigma_\lambda =$ 10, as can be seen in the Fig. 26, for this reason we cut on this value in order to reduce the background.

Parameter	Mean Value	Error	Meaning
l_0	797.093	19.7593	Gaussian height
l_1	1.11627	0.000084	Gaussian mean value (GeV$/c^2$)
l_2	3.1471	0.0963	Gaussian width (MeV$/c^2$)
l_3	–37,096.2	561.188	Linear background h–intercept
l_4	34,335.4	503.988	Linear background slope ($c^2/$GeV)

Table 6: Fit parameters for the Λ signal corresponding to Wrong Sign. Fig. 25.

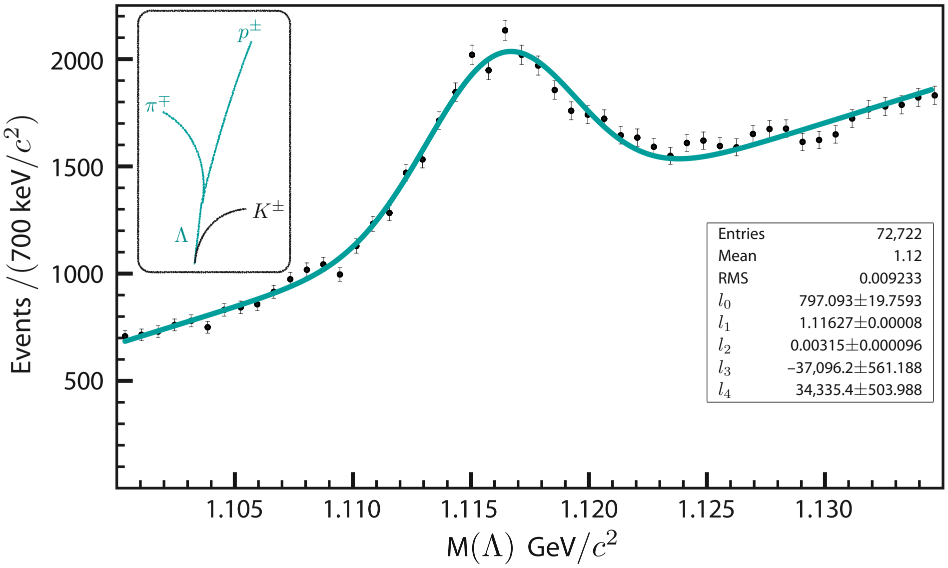


Figure 25: Λ signal corresponding to Wrong Sign.

Once the cut is applied, the resulting Λ signal for both cases are shown in Fig. 27 and Fig. 28.

From the fit results reported in Table 6, the chosen mass window for the Λ signal is set between 1.107 and 1.126 GeV$/c^2$, the 6$\cdot\sigma$ window centered on the Gaussian mean value.

The combination between the optimized Λ and $K^{\pm}$ is shown in Fig. 29. It is evident from these two distributions that there is an apparent excess of events around 1.67 GeV$/c^2$ for the Right Sign while, on the other hand, the Wrong Sign seems to follow quite well the shape of the background. Still, the level of background is too high and the distinction between signal and background events at this point, is done by means of multivariate analysis techniques.

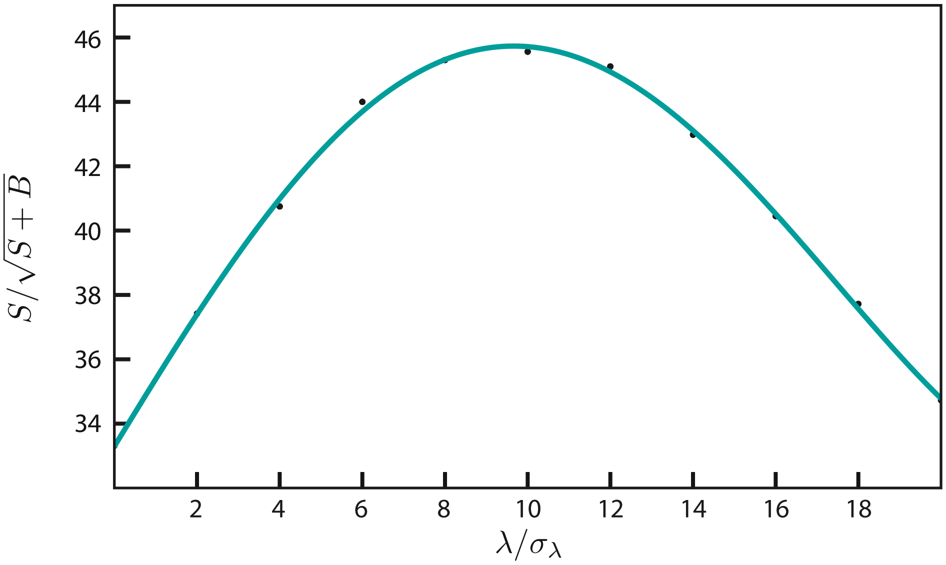


Figure 26: Signal significance as a function of the proper decay length significance for the Λ.

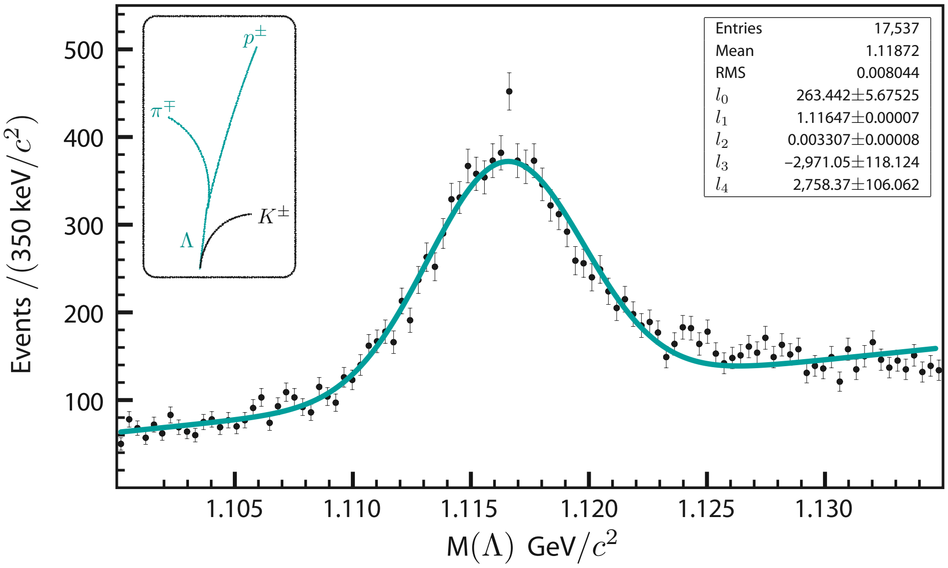


Figure 27: Wrong sign Λ sample.

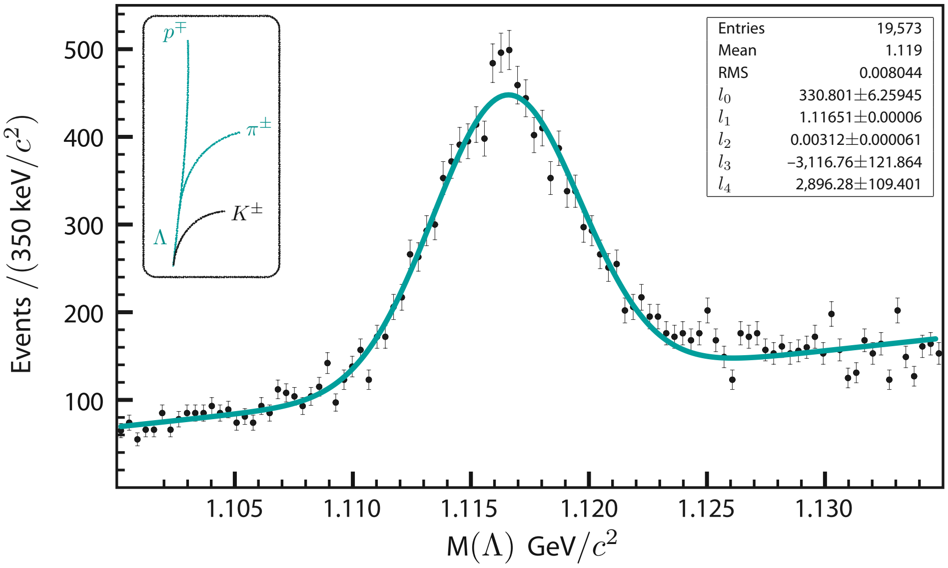


Figure 28: Right sign Λ sample.

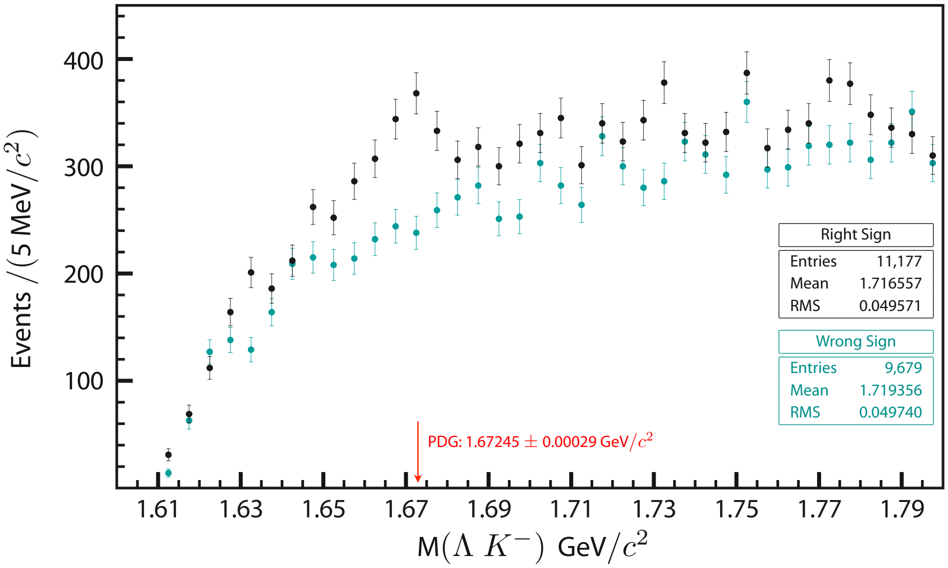


Figure 29: Λ K^- sample.

Multivariate Analysis. BDT

Multivariate analysis is based on the statistical principle of multivariate statistics, which involves observation and analysis of more than one statistical variable at a time. Its techniques are used to perform studies across multiple dimensions while taking into account the effects of all variables on the responses of interest. In this case, variables containing kinematic and quality information for the particles, tracks and vertices conforming the Ω^- are used to train a classifier able to distinguish in a smarter way between signal and background events. One implementation of Multivariate analysis techniques commonly used in high energy physics is knows as TMVA [24]. Various techniques were tested to look for the one doing the best classification, being Boosted Decision Trees the most accurate for our purposes. The first appendix of this thesis is dedicated to the coverage of boosted decision trees.

The training of this classifier was made using 20 variables, Table 7, from two samples:

- 500 MC $\Omega^- \to \Lambda\ K^-$ decays from the original $\Omega_b^- \to J/\psi(1S)\ \Omega^-$ events to describe signal, and
- 800 Wrong Sign events, Fig. 29, to describe background.

The number of trees in the forest was 300, the separation criterion applied for the node splitting was the Gini Index and the trees were build using Adaptive Boost as the boosting type.

The purpose of the classifier is to help in the isolation of $\Omega^- \to \Lambda\ K^-$, specifically coming from $\Omega_b^- \to J/\psi(1S)\ \Omega^-$ decays.

Once the training is complete, the result is used over a test sample in order to evaluate the classifier performance. The output given by the classifier is a number between −1 and +1, being the first the corresponding to background and the later to signal. The application of the classifier over both, training and test samples is shown in Fig. 30. From here, it can be seen that a BDT output of 0.0 gives a good separation between events recognized as signal from those who seem not to be. In the actual application of the classifier, all events with a BDT output greater than zero are kept while the rest are rejected as background.

Fig. 31, show the $\Lambda\ K^-$ distribution after the BDT output criteria is applied. The visible peak is identified as being the result of Ω^- events.

Contamination from Ξ^-

The topology of the $\Xi_b^- \to J/\psi(1S)\ \Xi^-$ decay [15], is essentially the same that in the case of the Ω_b^-, Fig. 32. The mass of a kaon was given to the charged track added to Λ in order to look for Ω^-. If instead, the

Variable name	Variable description	BDT Ranking (1 for the best)
v_0	Λ vertex χ^2	4
v_1	Λ collinearity	15
v_2	Λ lifetime significance	2
v_3	χ^2 for the p track coming from the Λ	11
v_4	Combined impact parameter significance for the p coming from the Λ	18
v_5	p SMT hits	12
v_6	p CFT hits	14
v_7	χ^2 for the π track coming from the Λ	10
v_8	Combined impact parameter significance for the π coming from the Λ	19
v_9	π SMT hits	17
v_{10}	π CFT hits	8
v_{11}	p_T for the p coming from the Λ	6
v_{12}	p_T for the π coming from the Λ	5
v_{13}	Λ transverse decay length	13
v_{14}	Error of the Λ transverse decay length	20
v_{15}	Error of the $\Omega-$ transverse decay length	16
v_{16}	$\Omega-$ transverse decay length	3
v_{17}	$\Omega-$ collinearity	9
v_{18}	p_T for the K coming from the Ω^-	1
v_{19}	Combined impact parameter significance for the K coming from the Ω^-	7

Table 7: BDT input variables and their corresponding ranking.

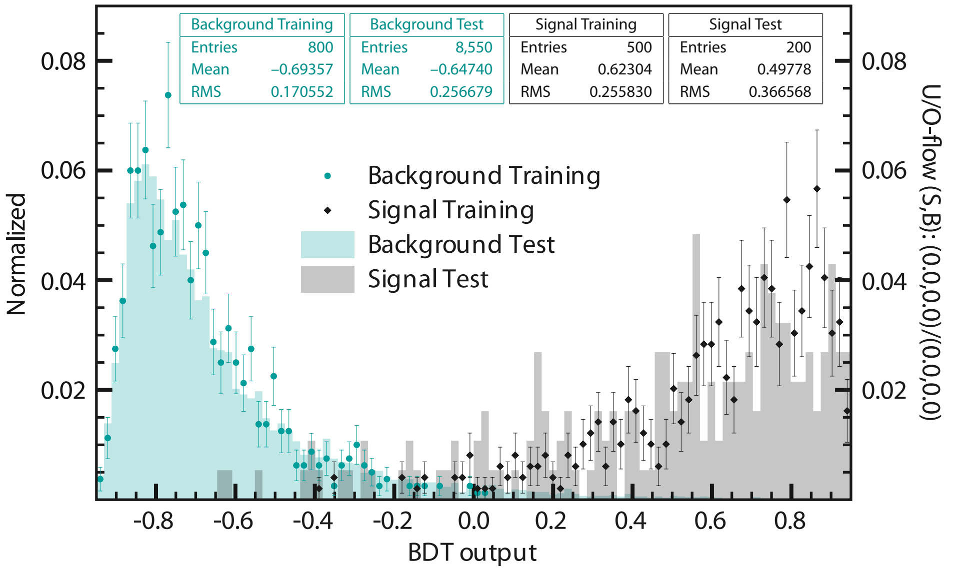


Figure 30: BDT output.

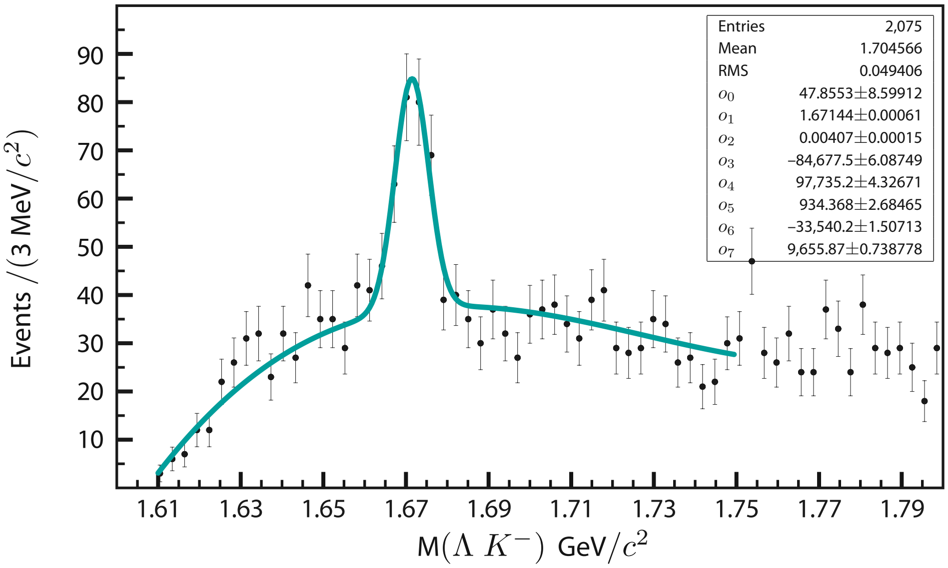


Figure 31: Λ K^- sample for events with BDT output greater than zero.

mass of this track is assumed to be that of a pion, the resulting reconstructed particle is a Ξ^-. Then, a test of this nature is needed in order to discard all events that due to this condition, may be misidentified as Ω^- events while actually coming from Ξ^- decays.

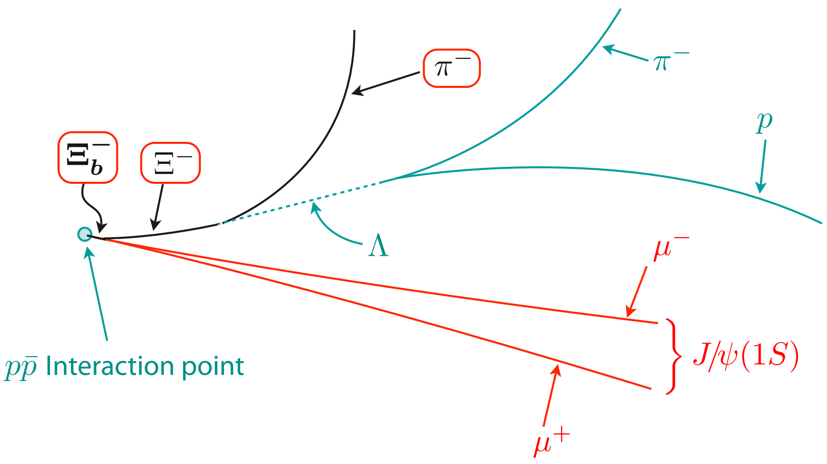


Figure 32: Schematic description of the Ξ_b^- decay chain. The red squares show the differences with the Ω_b^- decay.

The signal obtained from the $\pi-$track assumption is shown in Fig. 33 in the mass window corresponding to the Ξ^-, from 1.26 to 1.4 GeV$/c^2$. A clear signal with almost null background is found there, and all events with a mass 1.34 GeV$/c^2$ or less are removed from the $\Lambda\ K^-$ sample to avoid contamination from this signal.

Furthermore, the same events identified above as Ξ^-, on the $\Lambda\ K^-$ distribution appear distributed all over the histogram without a particular accumulation of any kind, as shown in Fig. 34.

$\boldsymbol{\Omega^-}$ signal

The distribution of $\Lambda\ K^-$ events after the Ξ^- contamination is removed is shown in Fig. 35, with an estimation of around 146 $\Omega^- \rightarrow \Lambda\ K^-$ events. Even here, the Wrong Sign events still show background behavior.

The mass window for the Ω^- signal is set between 1.660 and 1.683 GeV$/c^2$, the $6\cdot\sigma$ window centered on the Gaussian mean value.

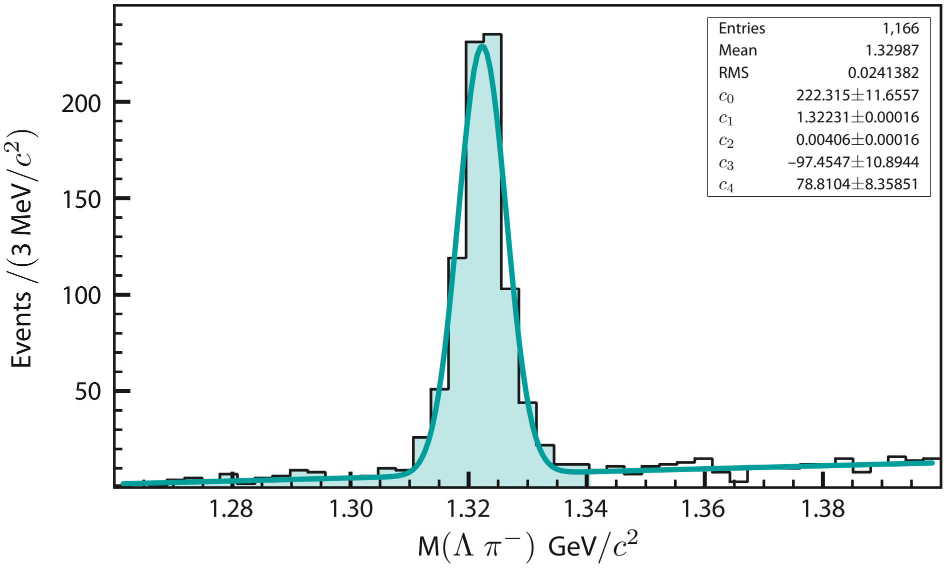


Figure 33: $\Lambda\ \pi^-$ signal. The shaded area show the events to be removed from the $\Lambda\ K^-$.

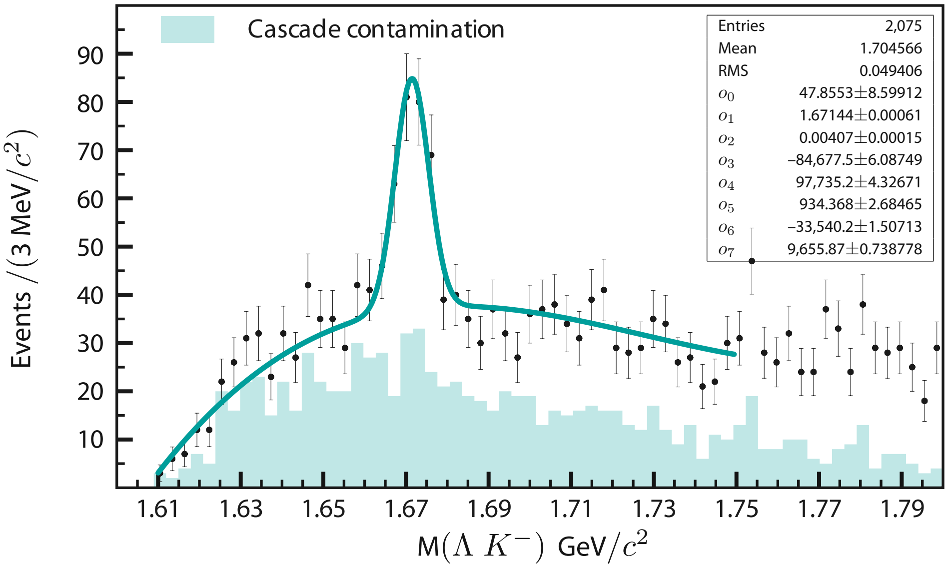


Figure 34: $\Lambda\ K^-$ showing the distributed $\Lambda\ \pi^-$ events.

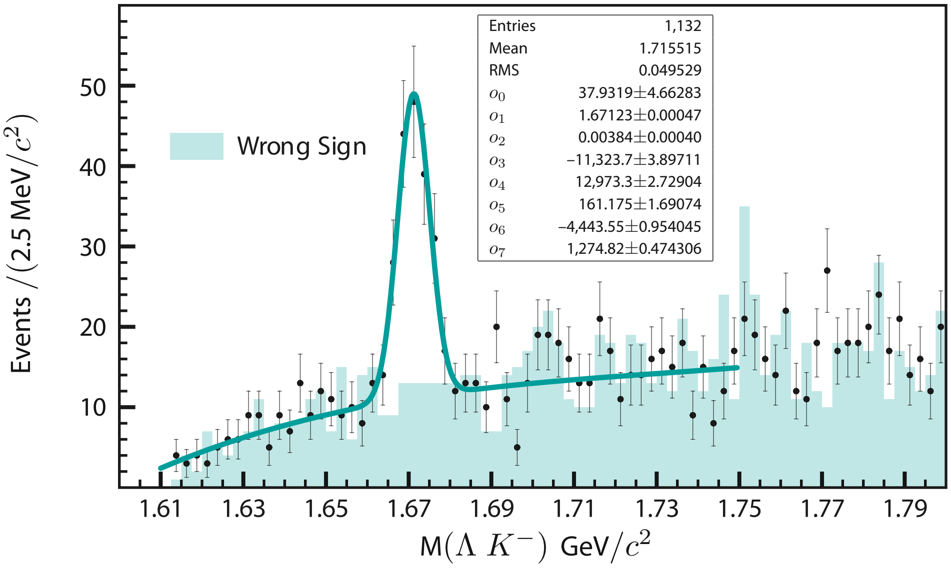


Figure 35: $\Lambda\ K^-$ after removing $\Lambda\ \pi^-$ contamination.

$J/\psi(1S)\ \Omega^-$ selection

With an optimized $\Omega^- \ \rightarrow\ \Lambda\ \ K^-$ signal, Fig. 35, the reconstruction of Ω_b^- events continue by selecting events with a common vertex between the already found $J/\psi(1S)$ and Ω^-, the χ^2 probability for this vertex should also be greater than 1%.

The mass window chosen for the search of the Ω_b^- decays is between 5.6 and 7.0 GeV$/c^2$, since 5.62 GeV$/c^2$ is the mass of the lightest $b-$baryon, the Λ_b, and 7.0 GeV$/c^2$ is almost 1 GeV$/c^2$ above the predicted mass of the Ω_b^- [3]. In addition, as in the the case of the Ξ_b^- observation [15], to reduce detector effects on the scale of the mass of any possible signal, the mass definition used here correspond to the mass difference:

$$\mathrm{M}\left(\Omega_b^-\right) = \mathrm{M}\left(J/\psi(1S)\ \Omega^-\right) - \mathrm{M}\left(J/\psi(1S)\right) - \mathrm{M}\left(\Omega^-\right) + \mathrm{M}_{\mathrm{PDG}}\left(J/\psi(1S)\right) + \mathrm{M}_{\mathrm{PDG}}\left(\Omega^-\right) \tag{2}$$

Without the use of this mass definition, the results of the unbinned fit over the mass distribution for MC events are 6.047 $\pm$ 0.008 GeV$/c^2$ for the central value and 0.080 $\pm$ 0.004 GeV$/c^2$ for the width, Fig. 36.

When Eq. 2 is used to compute the mass of MC events, the unbinned fit to the resulting distribution gives 6.052 $\pm$ 0.002 GeV$/c^2$ as the result for the mass and 0.034 $\pm$ 0.002 GeV$/c^2$ for its width, Fig. 37. Note that the mass central value reproduces precisely the input mass value used to generate the MC itself, *i. e.*,

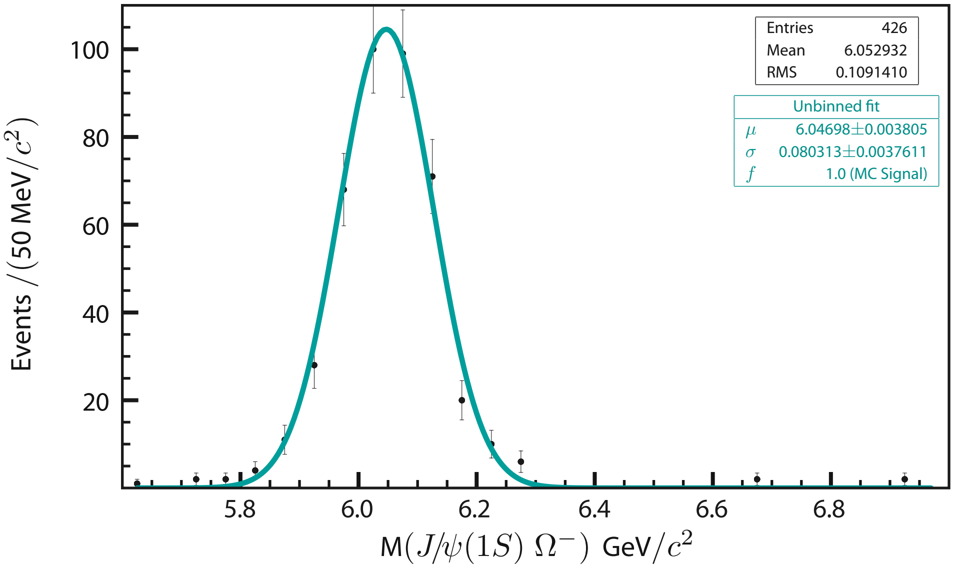


Figure 36: Distribution of mass of MC events without using the mass definition in Eq. 2.

6.052 GeV$/c^2$. Also, the width of the signal is significantly reduced in this way.

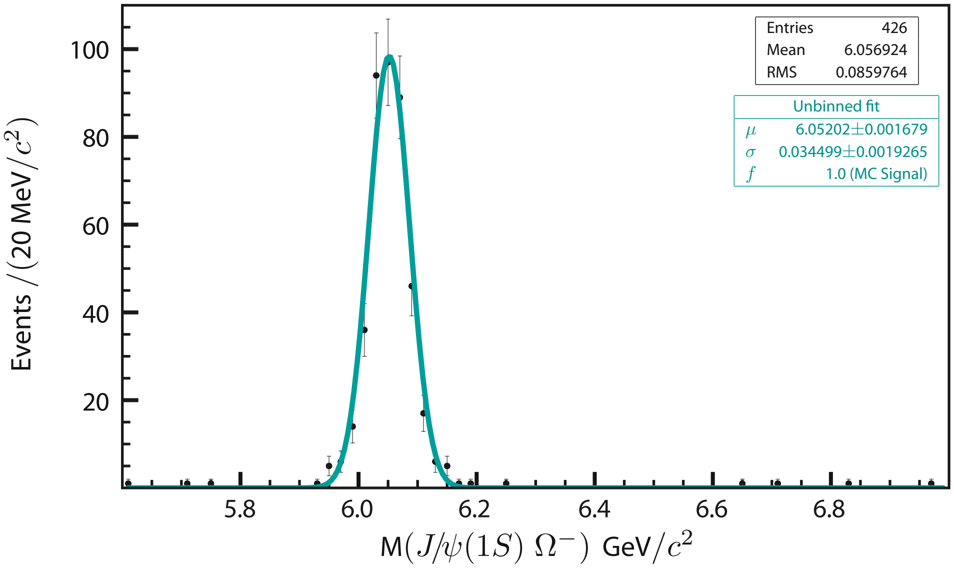


Figure 37: Distribution of mass of MC events using the mass definition in Eq. 2.

Now, a set of straight cuts is to be found in order to identify and remove background events while leaving as much signal events as possible. The optimization of these cuts takes place looking again at samples independent from the actual set of final $\Lambda \;\; K^-$ events. For this reason, MC is used as signal events and compared against Wrong Sign, serving as background.

In total, 41 Wrong Sign events from the Ω^- signal window survive all the previous selection criteria. These events are used as background sample for the cut–optimization purposes.

The first variable explored is the error on the proper decay length for $J/\psi(1S) \;\; \Omega^-$ candidates. Previous B–hadron analyses show that this variable gives a good discrimination power between signal and background events and it combines uncertainties from vertices and momentum reconstruction. It should be noted that lifetime–related analyses should not use this kind of variable for background rejection purposes, however, as the goal of this analysis is not the lifetime measurement but the experimental observation, this variable may be used without restriction.

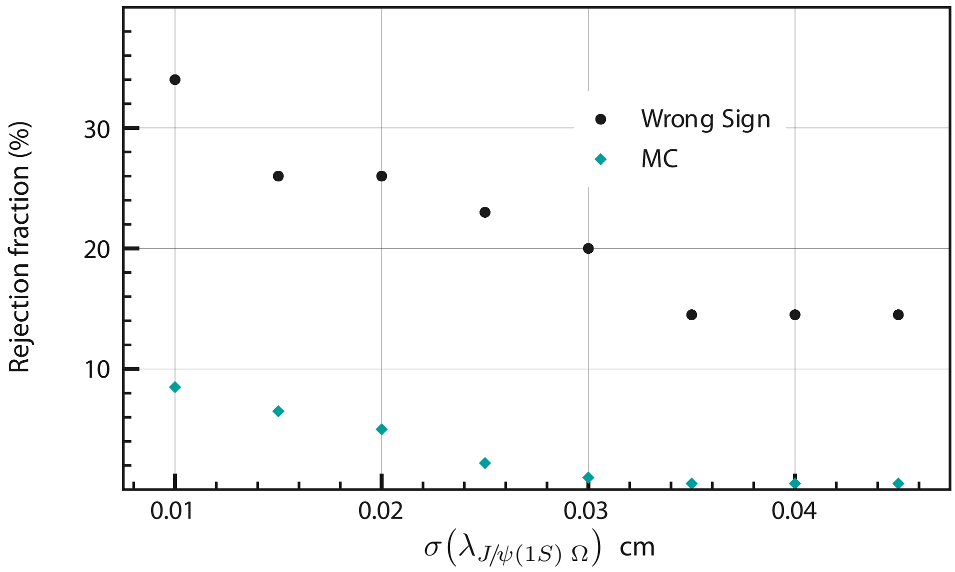


Figure 38: Rejection fraction comparison between MC and Wrong Sign events for different cuts on $\sigma\left(\lambda_{J/\psi(1S)\;\Omega^-}\right)$.

Fig. 38, shows the variation of rejected MC and Wrong Sign events for different cuts on the error of the proper decay length to be less than a given value. Requiring a cut of 0.03 cm, less than 0.5% of signal is removed while the background is reduced by 20%. This cut will be used as a default selection cut for $\lambda_{J/\psi(1S)\;\Omega^-}$ candidates.

Another requirement explored, this one inherited from the observation of the Ξ_b^-, is the angle between the $b-$baryon daughters in the plane perpendicular to the beam, in this case the $J/\psi(1S)$ and Ω^-, should be less than $\pi/2$. We found that while removing no MC events, it removes around 17% of the Wrong Sign events. Different values were used: $3\pi/8$, $\pi/4$, and $\pi/8$, finding that requiring this angle to be less than $3\pi/8$ yields to a rejection of 28% of the Wrong Sign and less than 1% of MC events. However, to be consistent with previous studies, the requirement for this analysis is $\pi/2$.

Finally, another commonly used cut to remove background while keeping a good amount of signal, like in the case of $\Lambda_b \rightarrow J/\psi(1S)\ \Lambda$ lifetime measurements previously studied by the CINVESTAV Group at DØ, is the p_T of the $b-$baryon. However, it is also known that the p_T spectrum in MC tends always to be softer than in real data. Then, an optimization in MC *vs.* Wrong Sign could not be totally correct for the signal, but it could give us at least an idea of how the background responds to this cut. Fig. 39, shows the behavior of MC and Wrong Sign events after different cuts in the p_T of the $b-$baryon. Almost any cut as low as 6 GeV/c removes background events while not removing signal events from MC. This specific value is set as a nominal selection cut.

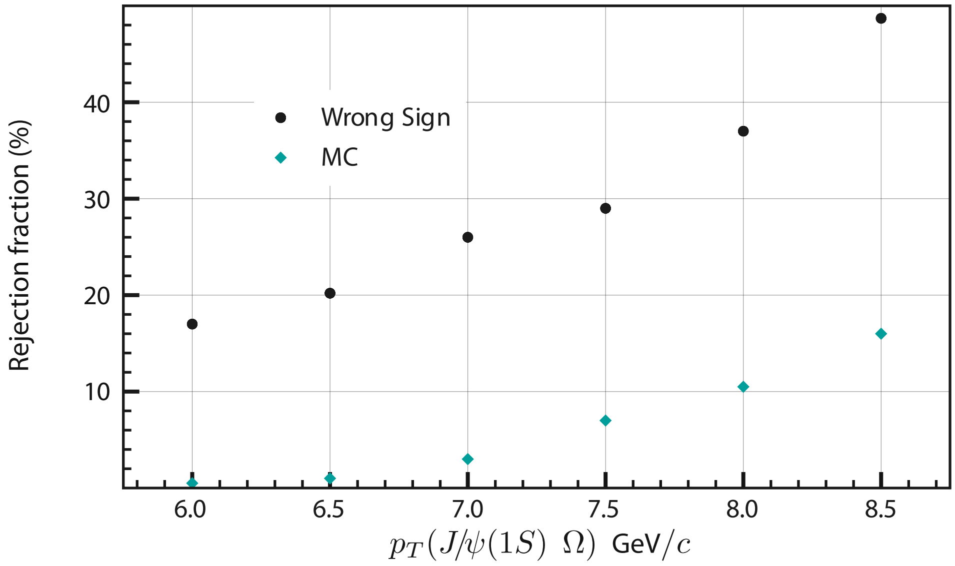


Figure 39: Rejection fraction comparison between MC and Wrong Sign events for different cuts on $p_T(\lambda_{J/\psi(1S)\ \Omega^-})$.

Control Samples

The selection method is first applied over different control samples in which is already known that a resonance of Ω_b^- is not present. This is done to be sure that this procedure is not going to create a resonance in an artificial way in both, real data and MC.

Three samples coming from real data were explored:

Λ Sidebands. Events from the Right Sign combination but having masses out of the 6·σ of the Gaussian peak on the Λ mass window, Fig. 40.

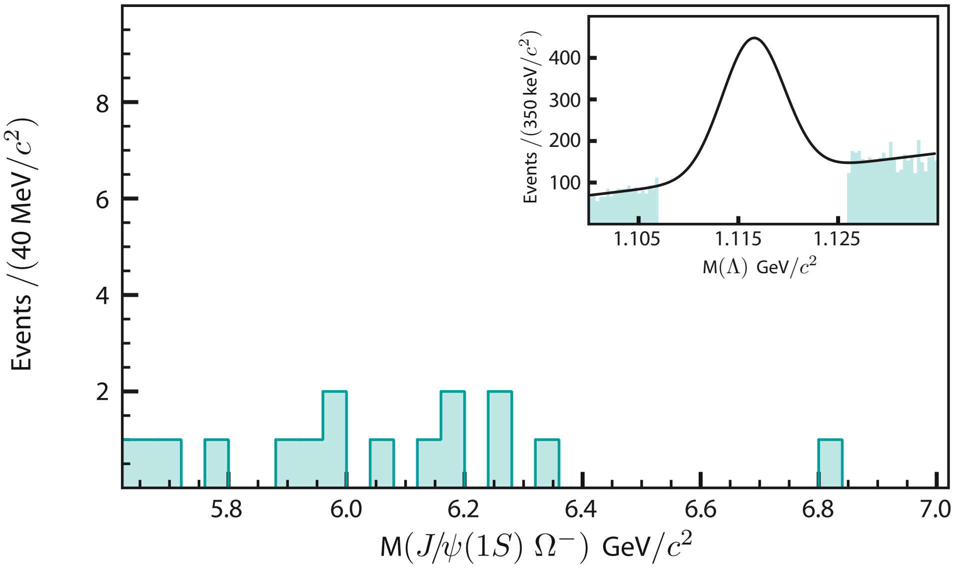


Figure 40: Method of selection applied to events coming from Λ sidebands. No resonance is seen on the events surviving the method of selection.

Ω^- Sidebands. Events from the Right Sign combination but having masses out of the 6·σ of the Gaussian peak on the Ω^- mass window, Fig. 41.

Wrong Sign. Events from the Wrong Sign combination that along the whole analyses have been used as background, Fig. 42.

In addition, the following high statistics MC samples with similar topologies are tested:

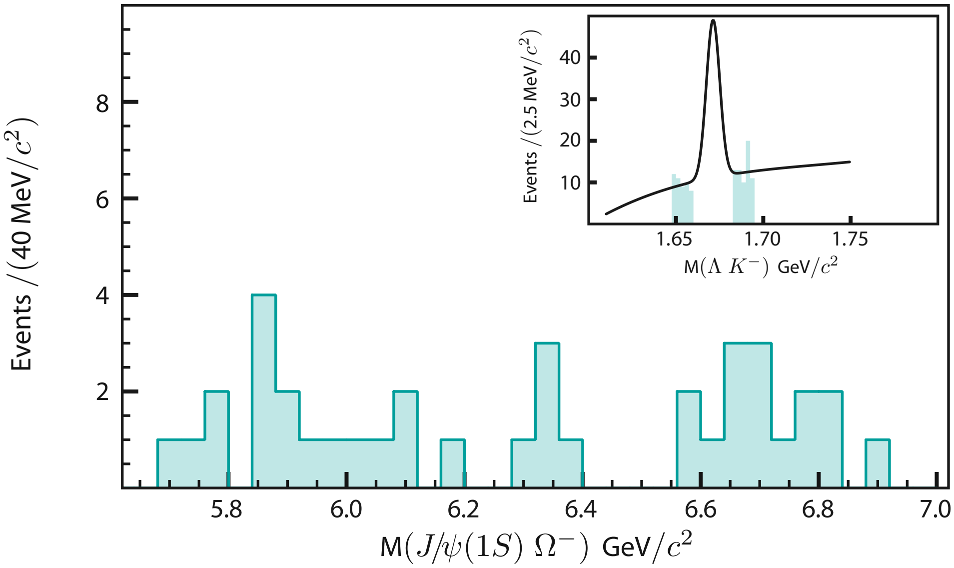


Figure 41: Method of selection applied to events coming from Ω^- sidebands. No resonance is seen on the events surviving the method of selection.

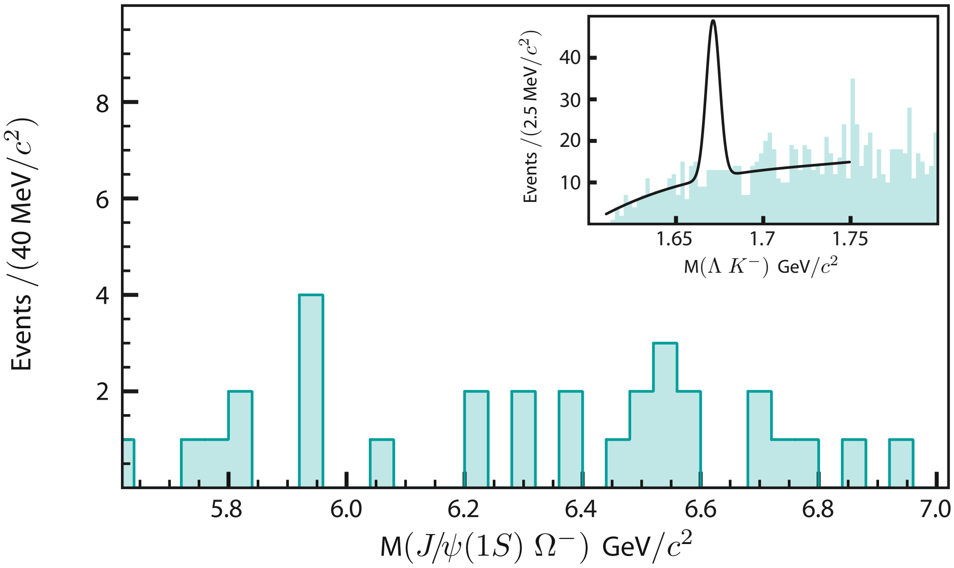


Figure 42: Method of selection applied to events coming from Wrong Sign. No resonance is seen on the events surviving the method of selection.

- $\Lambda_b \rightarrow J/\psi(1S)\ \Lambda \rightarrow (\mu^+\mu^-)\ (p\ \pi^-)$ (74k events),
- $\Xi_b^- \rightarrow J/\psi(1S)\ \Xi^- \rightarrow (\mu^+\mu^-)\ ((p\ \pi^-)\ \pi^-)$ (30k events), and
- $B^- \rightarrow J/\psi(1S)\ K^*(892)^- \rightarrow (\mu^+\mu^-)\ ((\pi^+\pi^-)\ \pi^-)$ (70k events).

As expected, none of these MC samples produce any resonance for the Ω_b^-, in fact, no events from these MC samples pass the Ω_b^- selection criteria.

All these test give confidence about the method of selection and is now safe go back to the Right Sign sample in order to look for a possible resonance.

Look at the Right Sign Sample

The optimized cuts are applied over the Right Sign sample in order to look for a possible resonance. Here, 79 events remain in the selected mass window with an excess located just below 6.2 GeV$/c^2$, Fig. 43. These events[5] are summarized in Table 8.

To measure the mass and the corresponding signal–to–background ratio of the observed excess of events, an unbinned extended log–likelihood fit in the mass window 5.6 to 7.0 GeV$/c^2$ is done. Unbinned log–likelihood fits are statistically more powerful than binned fits, *i. e.* will produce smaller errors on averages, and avoid any arbitrariness that is introduced by a choice of binning definition. These advantages are most visible when fitting small datasets, as in this case, and fitting multidimensional datasets. A Gaussian function is used to describe the peak and a constant value for the background.

It was not possible to fit the excess of events by allowing the width of the Gaussian to float, then it was fixed to the mean value of the width observed in MC events, 0.034 GeV$/c^2$, Fig. 37. The results of the unbinned fit are represented by the superimposed curve on Fig. 43, and summarized in Table 9.

To estimate the significance of the peak, a second unbinned fit is performed. In this case assuming only background in the same mass region and looking for its log–likelihood value. The result is:

---

[5] Data at the DØ Collaboration is taken in such way that to every recorded event is given a unique pair of integers for proper identification; the first one corresponding to the Run number from what it comes and the second is a sequential integer known as Event number.

Run	Event	Mass (GeV/c^2)	Run	Event	Mass (GeV/c^2)
152419	773856	6.335624	203613	50042678	6.55444
160585	554395	6.154512	203929	22881065	6.158335
163943	4409921	6.286062	203978	41379640	6.587322
167288	36793546	6.199677	204009	15108647	6.162176
167300	1360420	5.990438	204010	25983617	6.565955
168715	4944530	6.740241	204158	53612596	6.751408
170016	14492331	5.860953	204159	73163351	6.718659
173016	14553526	6.198785	206917	80642851	5.894412
174999	41823047	6.917226	207020	115025112	5.951568
176298	6684569	5.808267	207231	2188630	6.533908
176534	6644838	5.772659	207255	65033880	6.17786
176566	10253398	6.53304	207291	20553215	6.183106
177889	29415557	6.010407	207291	22632859	6.79282
177935	1100858	6.104838	207354	28412162	6.004997
178242	34631417	6.267342	207571	20570372	5.783729
178446	2143000	5.945344	207867	41954385	6.825897
179066	5036775	5.685159	208000	110255687	6.079188
179070	10936802	5.950622	208121	39312631	5.727999
179194	8663968	5.867919	208691	10703838	5.918028
179236	61796454	5.674977	209165	54484753	5.680678
179910	41520853	6.082665	209613	31938752	6.162704
180208	1713982	6.166889	210159	57181987	6.06177
180736	18427541	6.163343	210172	24829066	6.112304
187818	74249875	5.916054	210215	11632454	6.276017
188028	16573826	6.172063	210421	87644530	6.163563
189317	25412866	6.678204	210609	36508887	6.851988
189400	38941529	6.709374	210609	36508887	6.967876
191330	45751824	6.332727	211252	39151457	6.746307
191617	26759040	6.2051	211481	81822578	6.983675
191622	75966101	6.168653	211814	79361986	5.957745
192580	77820161	6.212005	211917	41206905	6.156171
192667	51883621	6.013875	212088	21960775	6.639544
192872	91767441	5.764635	213245	3711233	6.468697
193800	40131420	6.19771	213256	78789672	5.765941
194729	52839533	6.117869	213309	34358251	6.105085
195139	75494624	6.720045	213390	65577220	6.490448
195239	13018117	6.162319	213407	33689402	6.5522
195381	67225740	6.79625	213608	33401485	6.172512
195565	72670485	5.765104	215596	12128479	5.76468
203266	4351525	6.389002			

Table 8: Events found consistent with the reconstruction of an Ω_b^- baryon in RunIIa and their corresponding masses.

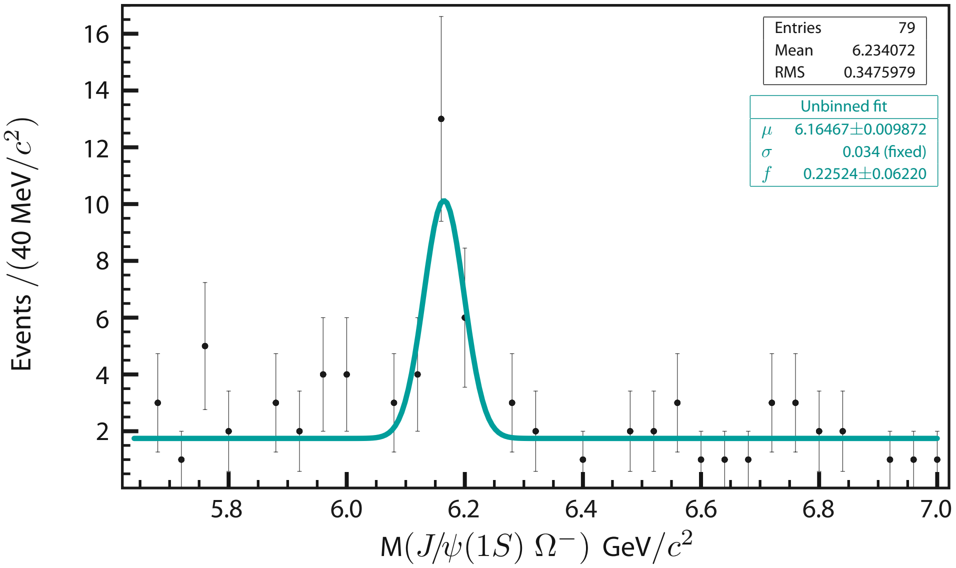


Figure 43: The 79 events from the Right Sign sample, Table 8.

Parameter	Mean Value	Error	Meaning
μ	6.16467	0.009872	Gaussian mean value (GeV/c^2)
σ	0.034	–	Gaussian fixed–width (MeV/c^2)
f	0.22524	0.06220	Signal fraction
$-\log\left(L_{S+B}\right)$	12.2855	–	negative of the log–likelihood value
S	17.79396	4.91443	Signal events

Table 9: $J/\psi(1S)$ Ω^- signal unbinned fit results. Fig. 43.

$$-\log\left(L_B\right) = 26.5813 \tag{3}$$

With the result shown in Table 9 ,and the following definition for the significance of the signal:

$$\sqrt{-2\Delta\log\left(L\right)} \equiv \sqrt{2\left\{\left[-\log\left(L_B\right)\right] - \left[-\log\left(L_{S+B}\right)\right]\right\}} = 5.347 \tag{4}$$

Two more values of signal width, 0.025 GeV/c^2 and 0.030 GeV/c^2, were also analyzed. Their results are shown in Fig. 44, and summarized in Table 10.

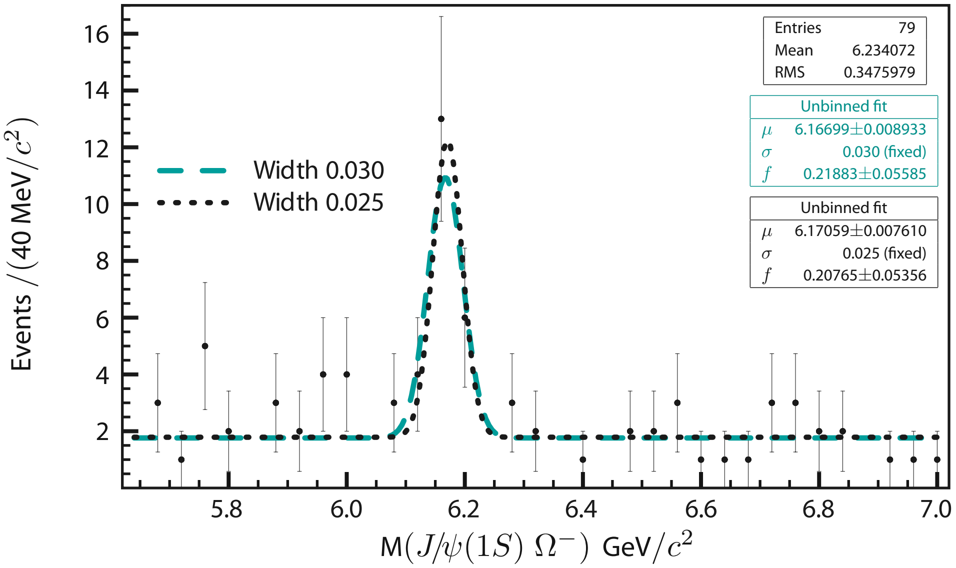


Figure 44: The 79 events from the Right Sign sample, Table 8, for two different signal widths.

Cross–checks

Ω_b^- signal with similar Ξ_b^- cuts

Table 11 shows the selection cuts used to select the Ω_b^- final events compared to those used for the Ξ_b^- observation [15]. The set of cuts labeled as CBA–1 belongs to a selection following very close the one used for the Ξ_b^- analysis, while in the one labeled as CBA–2, some cuts were modified to follow those in the BDT analysis. The last rows show the results of the unbinned fit for every distribution of events.

Ω_b^- signal with loose cuts

To investigate the effects of the selected cuts in the stability of the observed excess of events, we loose some of these cuts and repeat the unbinned likelihood fit to extract the mass, number of events, and significance of the observed peak. These test are:

Cuts 1. Remove the Ξ^- veto, *i. e.*, events with M$(\Lambda\ \pi^-) < 1.34$ GeV$/c^2$, Fig. 34, are **not** removed. The resulting distribution is shown in Fig. 45a.

Parameter	Mean Value	Error	Meaning
μ	6.17059	0.00761	Gaussian mean value (GeV$/c^2$)
σ	0.025	–	Gaussian fixed–width (MeV$/c^2$)
f	0.20765	0.05356	Signal fraction
$-\log\left(L_{S+B}\right)$	11.2864	–	negative of the log–likelihood value
S	16.40459	4.23135	Signal events
$\sqrt{-2\Delta\log\left(L\right)}$	5.53081	–	Signal significance
μ	6.16699	0.00893	Gaussian mean value (GeV$/c^2$)
σ	0.030	–	Gaussian fixed–width (MeV$/c^2$)
f	0.21883	0.05585	Signal fraction
$-\log\left(L_{S+B}\right)$	11.869	–	negative of the log–likelihood value
S	17.28789	4.41241	Signal events
$\sqrt{-2\Delta\log\left(L\right)}$	5.42445	–	Signal significance

Table 10: $J/\psi(1S)\ \Omega^-$ signal unbinned fit results for two different signal widths. Fig. 44.

Variable	CBA (Ξ_b^-)	BDT (Ω_b^-)	CBA–1 (Ω_b^-)	CBA–2 (Ω_b^-)
$p_T(\pi)$, π^- from Λ (GeV)	> 0.3	BDT input ($>$ 0.2)	> 0.3	> 0.2
$p_T(p)$, p from Λ (GeV)	> 0.7	BDT input ($>$ 0.2)	> 0.7	> 0.7
$p_T(\pi^-/K^-)$, π^-/K^- from Ξ_b^-/Ω_b^- (GeV)	> 0.2	BDT input	> 0.2	> 0.3
$p_T(\Lambda)$ (GeV)	> 0.7	Not used	> 0.7	Not used
$p_T(J/\psi(1S))$ (GeV)	> 5	Not used	> 5	Not used
Collinearity of Ξ_b^-/Ω_b^-	> 0.99	BDT input	> 0.99	> 0.99
Transverse decay length of Ξ_b^-/Ω_b^- (cm)	> 0.5	BDT input	> 0.5	> 0.5
Lifetime significance of Ξ_b^-/Ω_b^-	> 2	Not used	Not used	Not used
$\sigma(\lambda_{\Xi_b^-/\Omega_b^-})$[6] (cm)	Not used	< 0.03	Not used	< 0.03
M($\Lambda\ \pi^-$) (GeV$/c^2$) [Ξ^- veto]	Not used	< 1.34	< 1.34	< 1.34
Mass (GeV$/c^2$)	–	6.165±0.010	6.177±0.021	6.177±0.015
S	–	17.8±4.9	5.7±3.1	15.7±5.3
$\sqrt{-2\Delta\log\left(L\right)}$	–	5.35	2.41	3.94

Table 11: Comparison of Ξ_b^- selection cuts *vs.* different selections used in Ω_b^- . Cuts on CBA–1 follow the Ξ_b^- analysis and some in CBA–2 are modified to follow the BDT analysis.

Selection	Mass (GeV/c^2)	S	$\sqrt{-2\Delta\log(L)}$	Distribution
Cuts 1	6.154±0.015	16.1±5.7	3.64	Fig. 45a
Cuts 2	6.164±0.010	18.1±5.1	5.04	Fig. 45b
Cuts 3	6.168±0.010	20.1±5.5	5.63	Fig. 45c
Cuts 4	6.166±0.011	16.8±5.1	4.58	Fig. 45d
Cuts 5	6.177±0.015	15.7±5.3	3.94	Fig. 45e
Cuts 6	6.164±0.010	19.3±4.96	6.06	Fig. 45f

Table 12: Summary of unbinned fit results after loosing the cuts on the Ω_b^- event selection. See the text for explanation of each specific cut.

Cuts 2. Open the Λ and Ω^- mass windows to $\pm 4\cdot\sigma$ originally set to $\pm 3\cdot\sigma$. The resulting distribution is shown in Fig. 45b.

Cuts 3. Cut on $\lambda/\sigma_\lambda(\Lambda) > 5$, Fig. 26, originally set to > 10. The resulting distribution is shown in Fig. 45c.

Cuts 4. Accept events with BDT output –0.25 and above, Fig. 30, originally only 0 and above. The resulting distribution is shown in Fig. 45d.

Cuts 5. CBA–2 from Table 11. The resulting distribution is shown in Fig. 45e.

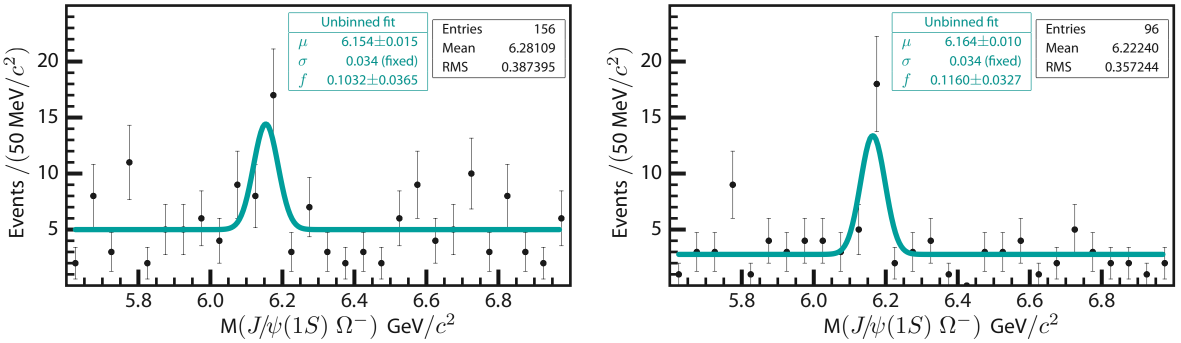
Cuts 6. Harder cut on $p_T(J/\psi(1S)\ \Omega^-)$, Fig. 39, from 6 to 7 GeV/c. The resulting distribution is shown in Fig. 45f.

The results of the unbinned fits to each one of this modifications in the selection criteria is summarized in Table 12.

Systematic Uncertainties

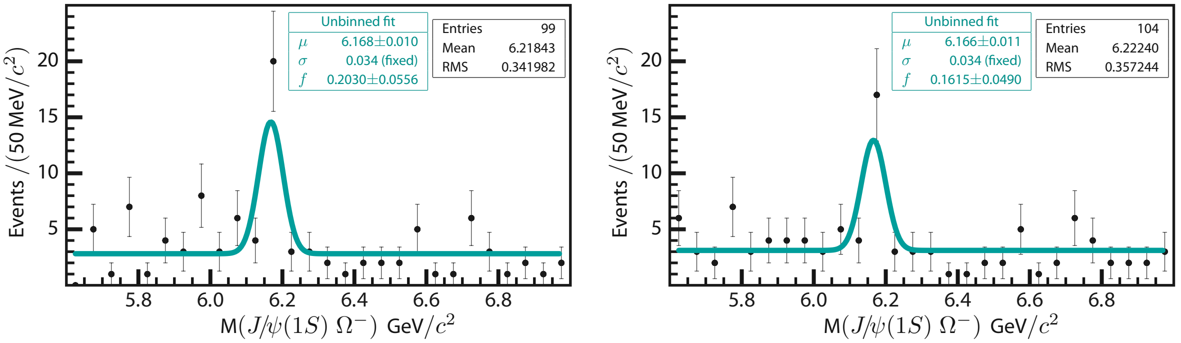
Mass

The identified sources of systematic uncertainties on the measured Ω_b^- mass and their contributions are the following:



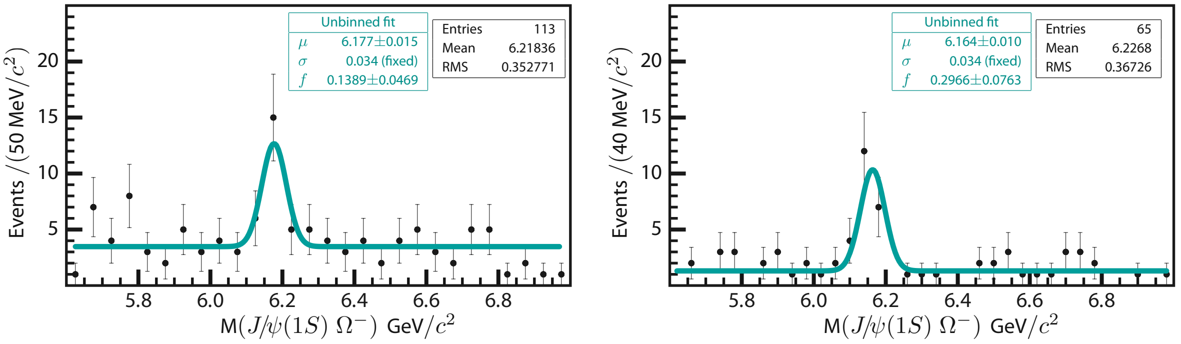
(a) No Ξ^- veto.

(b) Λ and Ω^- : $\pm 4\cdot\sigma$ signal windows.



(c) $\lambda/\sigma_\lambda(\Lambda) > 5$. Fig. 26.

(d) BDT output above −0.25.



(e) CBA–2 from Table 11.

(f) p_T cut above 7 GeV

Figure 45: Results of loosing the cuts applied to select the Ω_b^- signal. All yield a consistent signal.

Event Selection. Varying the selection criteria and applying a set of cuts on individual kinematic variables was previously done as a cross–check. The maximum change in the mass central value of 12 MeV/c^2 is achieved using the so called CBA–2, Fig. 45e.

Signal and Background Models. Using a linear function as the background model causes a negligible change in the mass. The unbinned fit yields 6.1653±0.01000 GeV/c^2 with 17.39±4.54 events, for a contribution of just 0.59 MeV/c^2. Fig. 46.

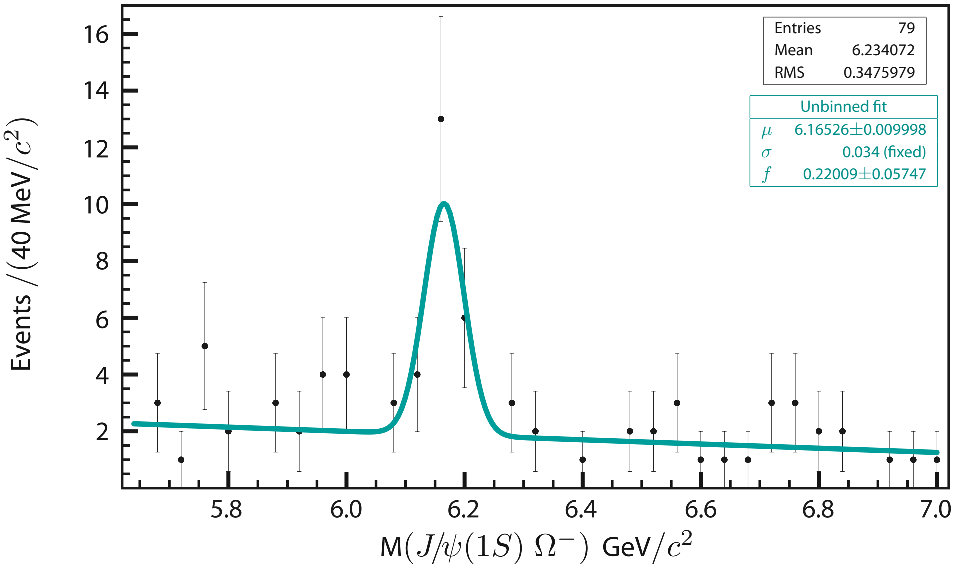


Figure 46: When the model to describe the background is replaced by a straight line, the global result remains practically unchanged.

Gaussian Width. Variation in the Gaussian width is allowed on the 0.034±3·0.002 GeV/c^2 range, according to the results of the unbinned fit over the distribution on Fig. 37.

Table 13 shows the results from the unbinned fit for different Gaussian widths. The maximum variation on the mass mean value results when the width is fixed to 0.028 GeV/c^2 and its contribution is 3.7 MeV/c^2.

Tighter Selection. When a tighter selection is applied to enhance signal over background, letting the width to float instead of fixing it to the MC value, the mass shift is 2 MeV/c^2 and the signal width is 0.033±0.010 GeV/c^2, which is consistent with the MC expectation.

Width (GeV/c^2)	Mass (GeV/c^2)	f
0.028	6.1684±0.00845	0.21456±0.05505
0.029	6.1677±0.00869	0.21674±0.05547
0.030	6.1670±0.00893	0.21883±0.05585
0.031	6.1664±0.00917	0.22083±0.05622
0.032	6.1658±0.00941	0.22273±0.05656
0.033	6.1652±0.00964	0.22453±0.05689
0.034	6.1647±0.00987	0.22524±0.06221
0.035	6.1642±0.01010	0.22786±0.05751
0.036	6.1638±0.01033	0.22939±0.05780
0.037	6.1634±0.01055	0.23084±0.05808
0.038	6.1631±0.01077	0.23220±0.05835
0.039	6.1628±0.01099	0.23349±0.05862
0.040	6.1626±0.01121	0.23472±0.05888

Table 13: Summary of unbinned fit results when the width of the Gaussian is allowed to float between the MC–signal uncertainties, Fig. 37.

Momentum Scale. To study the effect of the track momentum scale uncertainty on the measured Ω_b^- mass, the $\Lambda_b \rightarrow J/\psi(1S)\ \Lambda$ high statistics decay is reconstructed. The Λ_b mass is measured for different minimum p_T requirements on the Λ_b daughter particles. The comparison between these measurements to the world average value for the Λ_b mass [21], yields a maximum deviation of 0.004 GeV$/c^2$ which is quoted as a systematic uncertainty.

Table 14 contains the summary of the systematic contributions to the mass of the Ω_b^-. The total systematic computed is 13.34 MeV$/c^2$.

Number of Events

The systematic uncertainty on the number of signal events is estimated from the variation of both: the signal and background models. The signal model was modified to let the width of the signal float between the uncertainties of the MC–signal value on Fig. 37, and the corresponding results were summarized on Table 13. On the other hand, the new background model was a straight line instead of a flat distribution, Fig. 46.

Source	Contribution (GeV$/c^2$)
Event Selection	0.012
Linear Background	0.00059
Signal width variation	0.0037
Tighter selection	0.002
Momentum scale correction	0.004
Total	0.01334

Table 14: Summary of the systematic uncertainties considered on the mass of the Ω_b^-.

The maximum change on the number of events achieved from those changes comes from the Gaussian width fixed to 0.028 GeV$/c^2$, yielding to 16.95$\pm$4.35 signal events, which gives the maximum change of 0.84 events to the 17.79$\pm$4.91 signal events reported as result, Table 9. The systematic uncertainty on the number of events quoted is then 0.84.

Once these systematic uncertainties are included, the resonance is found having the following characteristics:

$$\begin{aligned} &\mathrm{M}(\Omega_b^- \rightarrow J/\psi(1S)\ \Omega^-) = 6.165 \pm 0.010\ \text{(stat.)} \pm 0.013\ \text{(syst.)}\ \text{GeV}/c^2\text{, and} \\ &\mathrm{S}(\Omega_b^- \rightarrow J/\psi(1S)\ \Omega^-) = 17.79 \pm 4.91\ \text{(stat.)} \pm 0.84\ \text{(syst.)}. \end{aligned} \tag{5}$$

Production Ratio

Finally, the production rate relative to that of the Ξ_b^- [15] is computed. First, the selection efficiency ratio

$$\frac{\epsilon\left(\Omega_b^- \rightarrow J/\psi(1S)\ \Omega^-\right)}{\epsilon\left(\Xi_b^- \rightarrow J/\psi(1S)\ \Xi^-\right)} = 1.5 \pm 0.2\ \text{(stat.)}, \tag{6}$$

assuming inclusive Ω^- and Ξ^- decays.

The higher efficiency for the Ω_b^- is due primarily to a harder p_T spectrum of the kaon from the Ω^- decay than that of the pion from the Ξ^- decay and a shorter lifetime of the Ω^- compared to the Ξ^-.

From the previously reported observation of the Ξ_b^-, [15], is known that the number of events on the same sample used for this analysis is $\mathsf{S}(J/\psi(1S)\ \Xi^-) = 15.2 \pm 4.4$ (stat.)$^{+1.9}_{-0.4}$ (syst.), while in the case of the Ω_b^-, the yield appears on Eq. 5.

From these yields and Eq. 6, the estimation of

$$\mathcal{R} = \frac{\mathsf{S}(\Omega_b^- \to J/\psi(1S)\ \Omega^-)}{\mathsf{S}(\Xi_b^- \to J/\psi(1S)\ \Xi^-)} \cdot \left\{ \frac{\epsilon\left(\Omega_b^- \to J/\psi(1S)\ \Omega^-\right)}{\epsilon\left(\Xi_b^- \to J/\psi(1S)\ \Xi^-\right)} \right\}^{-1} = 0.80 \pm 0.32\ \text{(stat.)}^{+0.14}_{-0.22}\ \text{(syst.)} \tag{7}$$

is obtained. The systematic uncertainty includes contributions from both: the signal yields as well as the efficiency ratio. On the other hand,

$$\mathcal{R} = \frac{f\left(b \to \Omega_b^-\right) \mathrm{Br}\left(\Omega_b^- \to J/\psi(1S)\ \Omega^-\right)}{f\left(b \to \Xi_b^-\right) \mathrm{Br}\left(\Xi_b^- \to J/\psi(1S)\ \Xi^-\right)} \tag{8}$$

being $f\left(b \to \Omega_b^-\right)$ and $f\left(b \to \Xi_b^-\right)$ the fraction of b quarks that hadronize into Ω_b^- and Ξ_b^- respectively.

Using the value

$$\frac{\Gamma\left(b \to \Omega_b^-\right)}{\Gamma\left(b \to \Xi_b^-\right)} = 9.8, \tag{9}$$

found on [11], the central values for $\tau\left(\Xi_b^-\right) = 1.42^{+0.28}_{-0.24}$ ps [21], the Eq. 7, and the prediction for $\tau\left(\Omega_b^-\right)$ of 0.83–1.67 ps [3, 4, 6, 7, 8, 9, 10], the following result is found

$$\frac{f\left(b \to \Omega_b^-\right)}{f\left(b \to \Xi_b^-\right)} \approx 0.07\text{–}0.14. \tag{10}$$

Conclusions and Final Comments

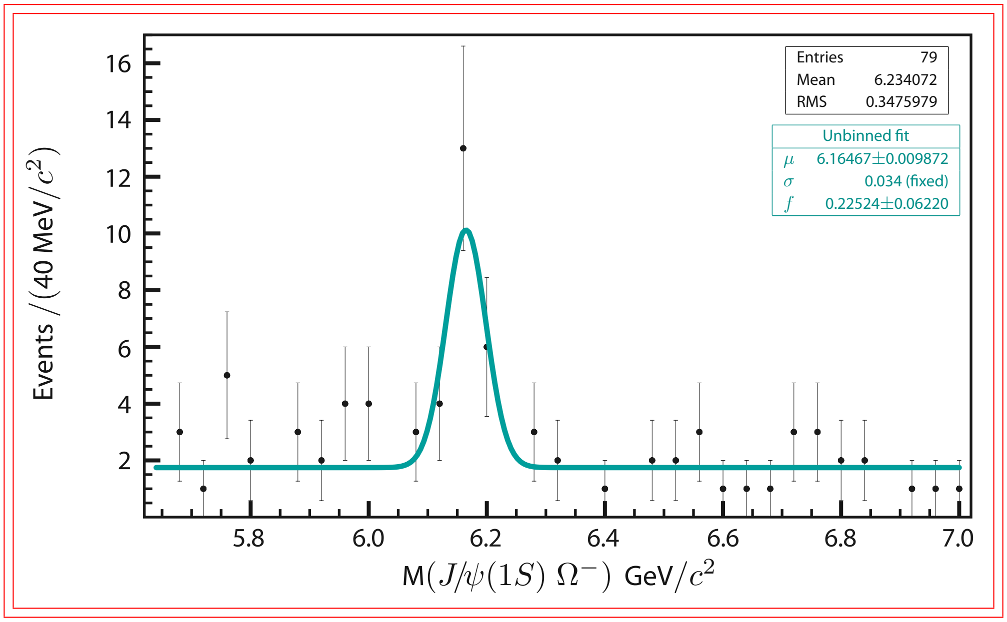


Figure 47: The 79 events consistent with the reconstruction of $\Omega_b^- \rightarrow J/\psi(1S)\ \Omega^-$ on 1.3 fb^{-1} of integrated luminosity collected with the DØ detector.

The search for Ω_b^- decays in data collected with the DØ detector between April 2002 and February 2006, equivalent to approximately 1.3 fb^{-1} integrated luminosity, yield to the direct observation of events matching the reconstruction in the following decay channel:

$$\Omega_b^- \to \overbrace{J/\psi(1S)}^{\mu^+\,\mu^-}\quad \underbrace{\Omega^-}_{\underbrace{\Lambda}_{p\,\pi^-}\;\;K^-}\quad .$$

Blind selection criteria were developed using MC as signal and Wrong Sign events as background, to look for the $b-$baryon. The method of selection was then tested over various samples where the absence of Ω_b^- events was already known, in order to avoid artificial creation of any signal. None of these samples produced any resonance.

In the sample of search, the blind selection was used to isolate only the events consistent with the above decay and 79 were found matching all selection criteria. Table 8 lists the events and shows the calculated mass for each one of them. Fig. 47 (above), shows the distribution of the events and the superimposed curve is the result of the unbinned extended log–likelihood fit to the masses reported on Table 8.

The results from this fit and the analysis of systematics yield to the statistically significant (5.4 $\cdot\,\sigma$) observation of 17.8 $\pm$ 4.9 (stat) $\pm$ 0.8 (syst) Ω_b^- signal events at 6.165 $\pm$ 0.010 (stat) $\pm$ 0.013 (syst) GeV/c^2. The significance of this signal means that the probability of this resonance may come from a fluctuation in the background is as small as 6.7×10^{-8}.

Then, an extensive number of cross–checks were made in order to further support the resonance found. In all cases, the results were consistent with the reported signal, and the computation of the production ratio of the Ω_b^- using the previously observed Ξ_b^- [15] yield to

$$\mathcal{R} = 0.80 \pm 0.32\ (\text{stat.})^{+0.14}_{-0.22}\ (\text{syst.}) \quad \text{and} \quad \frac{f\left(b \to \Omega_b^-\right)}{f\left(b \to \Xi_b^-\right)} \approx 0.07\text{–}0.14.$$

Final Comments

I would like to comment briefly on two subjects regarding the observation of the Ω_b^- baryon.

First. The analysis of the remaining data at the DØ Collaboration was the next natural step as the ongoing performance of Tevatron is accumulating large quantities of data with time. Fig. 48.

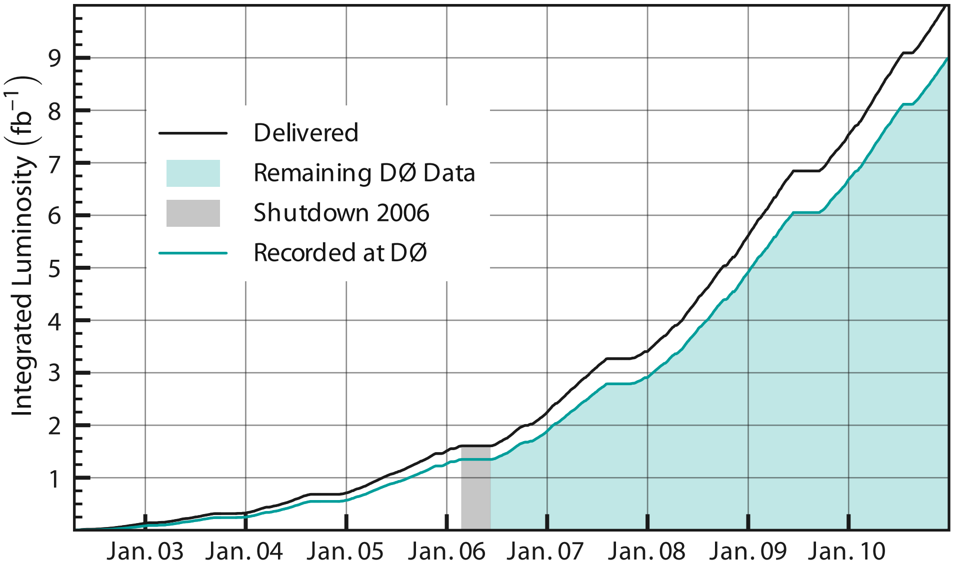


Figure 48: The search for the Ω_b^- on the remaining 7.64 fb^{-1} already recorded (shaded in green),

It is not a straight forward process to apply the method of selection that yield to the observation of the Ω_b^- in RunIIa to the rest of the data. The reason, due to the modifications done during the shutdown in 2006, many changes in the global data taking were introduced, from detector upgrades at hardware level, to algorithms, to significantly higher luminosities and their corresponding multiplicity of tracks per event, to triggering, among others.

The search is still in progress, but some difficulties have arised and the final reconstruction with a strong evidence of an observed resonance is yet to be seen.

The nature of the problems keeping the DØ Collaboration from seeing the Ω_b^- in RunIIb is out of the scope of this thesis, but in summary, the reconstructed yields of some of the intermediate particles on the decay chain result to be lower than the expected in the special reprocessing done to improve the efficiency in the reconstruction of long lived particles, taking into account the yields observed in RunIIa.

For this reason, until a fully validated reprocessing is produced for the rest of the data, the DØ Collaboration is not in a position to make a statement about the observation of the Ω_b^- in the RunIIb data set.

Second. The year after the announcement of the Ω_b^- observation by DØ, the CDF Collaboration at Fermilab made public the result about their own observation of this baryon [28]. The controversy started once the result was unveiled.

Analyzing data equivalent to an integrated luminosity of 4.2 fb^{-1}, with a cut–based analysis the CDF collaboration reports 16^{+6}_{-4} Ω_b^- events at 6.0544 $\pm$ 0.0068 (stat) $\pm$ 0.0009 (syst) GeV/c^2 on the same decay channel.

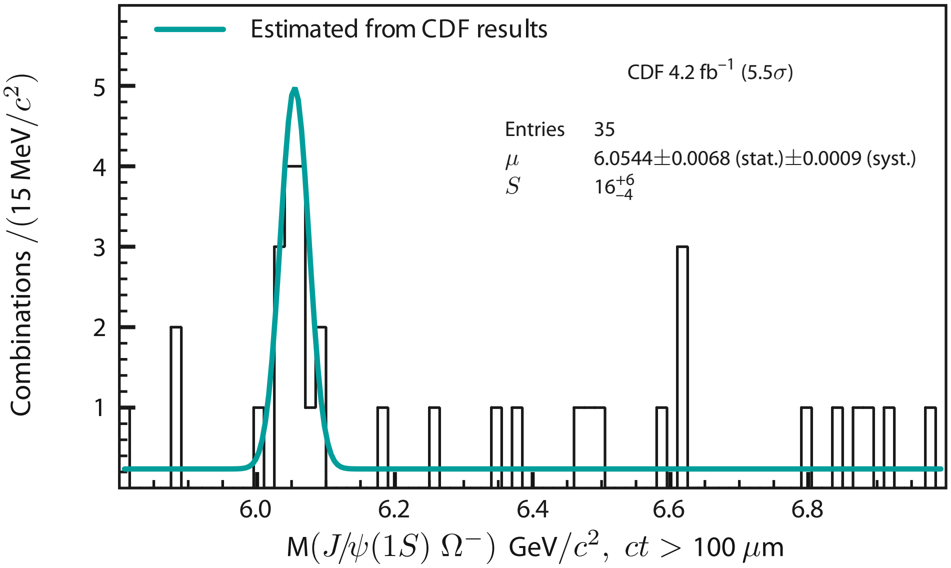


Figure 49: The observation of $\Omega_b^- \rightarrow J/\psi(1S)\ \Omega^-$ on 4.9 fb^{-1} of integrated luminosity collected with the CDF detector.

The mass difference between the CDF and DØ measurement is not consistent within quoted uncertainties. To this respect, DØ has studied heavy baryon mass reconstruction with known resonances, such as the Λ_b and the Ξ_b^-. In all cases, the results are consistent with the PDG values [21]. The reconstructed decay corresponds to one of the decay modes of the Ω_b^-. The Ω_b^- mass in MC events is consistent with the input value within systematic uncertainties. All these and the extensive number of test showed on this analysis gives confidence that the resonance found is actually the Ω_b^-.

The update of the analysis to the whole data set already recorded probably will bring new insights to the controversy on the difference between the two results.

Finally, the efforts of this analysis and their results, constitute the first experimental observation of the Ω_b^- baryon, possible by the collaboration between the High Energy Physics group from the CINVESTAV Physics Department and the DØ Collaboration at the Fermilab Tevatron Collider.

References

[1] V. E. Barnes *et al.*,
"Observation of a Hyperon with Strangeness −3"
Phys. Rev. Lett. 12, 204 (1964).

[2] M. Gell–Mann and Y. Ne'eman,
"The Eightfold Way"
W. A. Benjamin, New York (1964).

[3] X. Liu, H. X. Chen, Y. R. Liu, A. Hosaka and S. L. Zhu,
"Bottom baryons"
Phys. Rev. D 77, 014031 (2008) [arχiv:0710.0123 [hep–ph]].

[4] M. Karliner, B. Keren-Zur, H. J. Lipkin and J. L. Rosner,
"Predictions for masses of bottom baryons"
[arχiv:0708.4027 [hep–ph]].

[5] M. Karliner, B. Keren-Zur, H. J. Lipkin and J. L. Rosner,
"The Quark Model and b Baryons"
Annals Phys. 324, 2 (2009) [arχiv:0804.1575 [hep–ph]].

[6] E. E. Jenkins,
"Model–Independent Bottom Baryon Mass Predictions in the $1/N_c$ Expansion"
Phys. Rev. D 77, 034012 (2008) [arχiv:0712.0406 [hep–ph]].

[7] R. Roncaglia, D. B. Lichtenberg and E. Predazzi,
"Predicting the masses of baryons containing one or two heavy quarks"
Phys. Rev. D 52, 1722 (1995) [arχiv:hep–ph/9502251].

[8] N. Mathur, R. Lewis and R. M. Woloshyn,
"Charmed and bottom baryons from lattice nonrelativistic QCD"
Phys. Rev. D 66, 014502 (2002) [arχiv:hep–ph/0203253].

[9] D. Ebert, R. N. Faustov and V. O. Galkin,
"Masses of heavy baryons in the relativistic quark model"
Phys. Rev. D 72, 034026 (2005) [arχiv:hep–ph/0504112].

[10] T. Ito, M. Matsuda and Y. Matsui,
"New possibility of solving the problem of Lifetime Ratio $\tau\left(\Lambda_b\right)/\tau\left(B_d\right)$"
Prog. Theor. Phys. 99, 271 (1998) [arχiv:hep–ph/9705402].

[11] H. Y. Cheng,
"Nonleptonic weak decays of bottom baryons"
Phys. Rev. D 56, 2799 (1997) [arχiv:hep–ph/9612223].

[12] S. Abachi *et al.* [DØ Collaboration],
"The DØ detector"
Nucl. Instrum. Meth. A 338, 185 (1994) [Fermilab Pub. 93—179—E].

[13] V. M. Abazov *et al.* [DØ Collaboration],
"The upgraded DØ detector"
Nucl. Instrum. Meth. A 565, 463 (2006) [arχiv:physics/0507191 [physics.ins–det]].

[14] Status of Run II Data Taking at DØ
http://d0server1.fnal.gov/Projects/Operations/D0RunII_DataTaking.htm

[15] V. M. Abazov *et al.* [DØ Collaboration],
"Direct Observation of the Strange b Baryon Ξ_b^-"
Phys. Rev. Lett. 99, 052001 (2007) [arχiv:0706.1690 [hep–ex]].

[16] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin,
"High–energy–physics event generation with PYTHIA 6.1"
Comput. Phys. Commun. 135, 238 (2001) [arχiv:hep–ph/0010017].

[17] D. J. Lange,
"The EvtGen particle decay simulation package"
Nucl. Instrum. Meth. A 462, 152 (2001).

[18] R. Brun and F. Carminati [CERN],
"GEANT – Detector Description and Simulation Tool"
CERN Program Library Long Writeup W5013 (1993).

[19] The BANA package at DØ
http://www-d0.fnal.gov/Run2Physics/ckm/d0_private/bgv/aa/bana.htm

[20] $J/\psi(1S)$ & dimuon mass continuum at DØ
http://d0server1.fnal.gov/users/nomerot/Run2A/BANA/Dskim.html

[21] C. Amsler *et al.* [Particle Data Group],
"Review of particle physics"
Phys. Lett. B 667, 1 (2008) [PDG server, 37.68 MB].

[22] ROOT
User's Guide v5.20 (2008).

[23] L. Breiman *et al.*,
"Classification and Regression Trees"
Wadsworth International Group, Belmont (1984).

[24] A. Höcker, P. Speckmayer, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss,
"TMVA – Toolkit for Multivariate Data Analysis"
User's Guide v3.8 (2007) [arχiv:physics/0703039 [physics.data–an]].

[25] W. Verkerke and D. Kirkby,
"RooFit"
Users Manual v2.07—29 (2006).

[26] D. Kirkby *et al.* [BaBar],
"A Users Guide to the RooFitTools Package for Unbinned Maximum Likelihood Fitting"
Analysis Document #18, Version 12, (2001).

[27] F. James [CERN],
"MINUIT – Function Minimization and Error Analysis"
CERN Program Library Long Writeup D506, Reference manual v 94.1, (1994).

[28] T. Aaltonen *et al.* [CDF Collaboration],
"Observation of the Ω_b^- and Measurement of the Properties of the Ξ_b^- and Ω_b^-,"
Phys. Rev. D 80, 072003 (2009) [arXiv:0905.3123 [hep–ph]].

Multivariate Analysis: BDT

The following information was found on [24] and is reproduced here as reference.

Decision Trees and Boosting

A decision tree is a binary tree structured classifier like the one sketched in Fig. 50. Repeated left/right (yes/no) decisions are performed on a single variable at a time until some stop criterion is reached. Like this the phase space is split into regions that are eventually classified as signal or background, depending on the majority of training events that end up in the final leaf nodes. The boosting of a decision tree (BDT) represents an extension to a single decision tree. Several decision trees (a forest), derived from the same training sample by re–weighting events, are combined to form a classifier which is given by a (weighted) majority vote of the individual decision trees. Boosting stabilizes the response of the decision trees with respect to fluctuations in the training sample.

Description and Implementation

Decision trees are well known classifiers that allow straightforward interpretation as they can be visualized by a simple two dimensional tree structure. They are in this respect similar to rectangular cuts. However, whereas a cut–based analysis is able to select only one hypercube as region of phase space, the decision tree is able to split the phase space into a large number of hypercubes, each of which is identified as either “signal–like” or “background–like”.

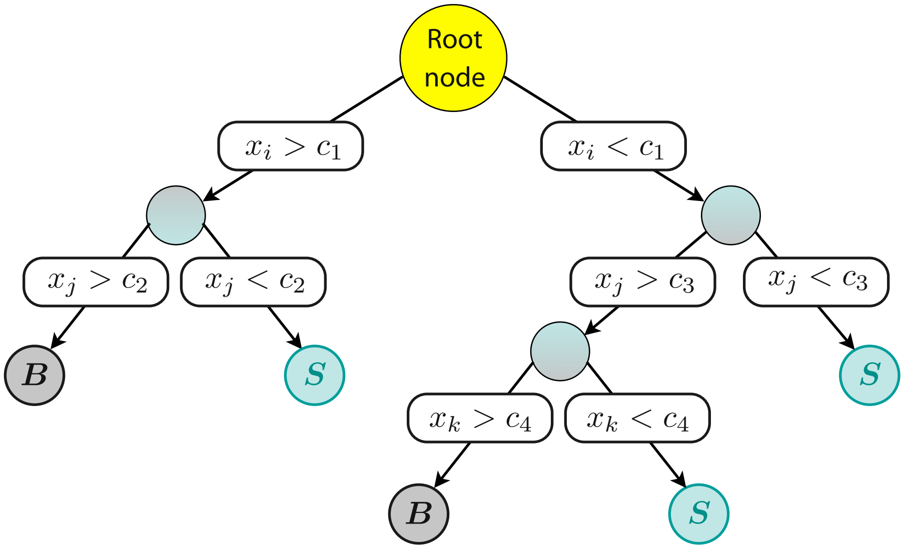


Figure 50: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_q is performed. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes.

The path down the tree to each leaf node represents an individual cut sequence that selects signal or background depending on the type of the leaf node. A shortcoming of decision trees is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. For example, if two input variables exhibit similar separation power, a fluctuation in the training sample may cause the tree growing algorithm to decide to split on one variable, while the other variable could have been selected without that fluctuation. In such a case the whole tree structure is altered below this node, possibly resulting also in a substantially different classifier response.

This problem is overcome by constructing a forest of decision trees and classifying an event on a majority vote of the classifications done by each tree in the forest. All trees in the forest are derived from the same training sample, with the events being subsequently subjected to so–called boosting, a procedure which modifies their weights in the sample. Boosting increases the statistical stability of the classifier and typically also improves the separation performance compared to a single decision tree.

However, the advantage of the straightforward interpretation of the decision tree is lost. While one can of course still look at a limited number of trees trying to interpret the training result, one will hardly be able to do so for hundreds of trees in a forest.

Nevertheless, the general structure of the selection can already be understood by looking at a limited number of individual trees.

Boosting

Boosting is a general procedure whose application is not limited to decision trees. The same classifier is trained several times using a successively boosted (reweighted) training event sample.

The final classifier is then derived from the combination of all the individual classifiers. The most popular boosting algorithm is the so–called AdaBoost (adaptive boost), where events that were misclassified during the training of a tree are given a higher event weight in the training of the next following tree.

Starting with the original event weights when training the first decision tree, the subsequent tree is trained using a modified event sample where the weights of previously misclassified events are multiplied by a common boost weight α. The boost weight is derived from the misclassification rate err of the previous tree,

$$\alpha = \frac{1 - err}{err}. \tag{11}$$

The entire event sample is then renormalized to keep the total number of events (sum of weights) in a tree constant.

With the result of an individual tree $h\left(\mathbf{x}\right)$ ($\mathbf{x}$ being the tuple of input variables) encoded for signal and background as $h\left(\mathbf{x}\right) = +1$ and –1, respectively, the resulting event classification $y_{\text{BDT}}\left(\mathbf{x}\right)$ for the boosted classifier is given by

$$y_{\text{BDT}}\left(\mathbf{x}\right) = \sum_{i\,\in\,\text{forest}} \ln\left(\alpha_i\right) \cdot h_i\left(\mathbf{x}\right), \tag{12}$$

where the sum is over all trees in the forest.

Small (large) values for $y_{\text{BDT}}(\mathbf{x})$ indicate a background–like (signal–like) event. Eq. 12 is the default BDT boosting but it can be altered to use only the average of the individual trees without the weighting factors ln (α_i).

Another possible modification of Eq. 12, is to use the training purity[1] in the leaf node as respectively signal or background weights rather than relying on the binary decision. Such an approach however should be adopted with care as the purity in the leaf nodes is sensitive to overtraining and therefore typically overestimated.

Other boosting technique is a resampling technique, sometimes referred to as bagging. The resampling is done with replacement, which means that the same event is allowed to be (randomly) picked several times from the parent sample. This is equivalent to regarding the training sample as being a representation of the probability density distribution of the parent event ensemble. If one draws an event out of this ensemble, it is more likely to draw an event from a region of phase–space that has a high cross section, as the original Monte Carlo sample will have more events in that region. If a selected event is kept in the original sample (that is when the same event can be selected several times), the parent sample remains unchanged so that the randomly extracted samples will have the same parent distribution, albeit statistically fluctuated.

Training several decision trees with different resampled training data and combining them into a forest results in an averaged classifier that, just as for boosting, is more stable with respect to statistical fluctuations in the training sample.

Technically the resampling is implemented by applying random weights to each event of the parent sample.

Training (Building) a Decision Tree

The training, building or growing of a decision tree is the process that defines the splitting criteria for each node. The training starts with the root node, where an initial splitting criterion for the full training sample is determined. The split results in two subsets of training events that each go through the same algorithm of determining the next splitting iteration. This procedure is repeated until the whole tree is built.

---

[1] The purity of a node is given by the ratio of signal events to all events in that node. Hence pure background nodes have zero purity.

At each node, the split is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. The node splitting is stopped once it has reached the minimum number of events previously specified. The end- or leaf nodes are classified as signal or background according to the class the majority of events belongs to.

A variety of separation criteria can be specified to assess the performance of a variable and a specific cut requirement. Because a cut that selects predominantly background is as valuable as one that selects signal, the criteria are symmetric with respect to the event classes.

All separation criteria have a maximum where the samples are fully mixed, *i. e.*, at purity $p = 0.5$, and fall off to zero when the sample consists of one event class only.

Tests have revealed no significant performance disparity between the following separation criteria:

Gini Index. Defined by $p \cdot (1 - p)$. Used on the analysis presented here.

Cross Entropy. Defined by $-p \cdot \ln(p) - (1 - p) \cdot \ln(1 - p)$.

Misclassification Error. Defined by $1 - \max(p,\ 1 - p)$.

Statistical Significance. Defined by $S/\sqrt{S+B}$ or $\sqrt{p}$ in terms of the purity and in units of $\sqrt{S}$.

The splitting criterion being always a cut on a single variable, the training procedure selects the variable and cut value that optimizes the increase in the separation index between the parent node and the sum of the indices of the two daughter nodes, weighted by their relative fraction of events. The cut values are optimized by scanning over the variable range with a granularity that is set by the size of the partition made over that variable. A compromise should be achieved between the computing time and the step size.

In principle, the splitting could continue until each leaf node contains only signal or only back- ground events, which could suggest that perfect discrimination is achievable. However, such a decision tree would be strongly overtrained. To avoid overtraining a decision tree must be pruned.

Pruning

Pruning is the process of cutting back a tree from the bottom up after it has been built to its maximum size. Its purpose is to remove statistically insignificant nodes and thus reduce the overtraining of the tree. It has

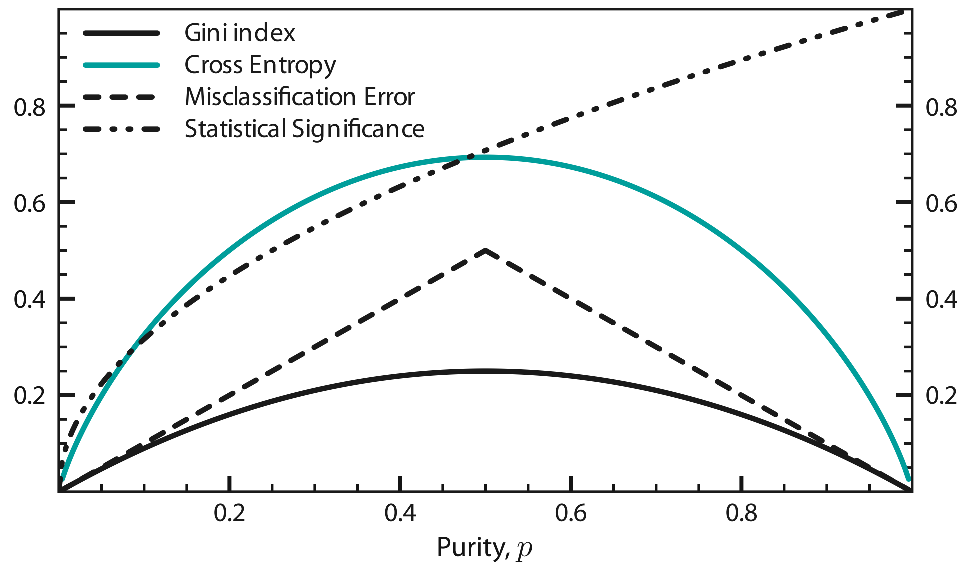


Figure 51: Main separation criteria used to evaluate the performance of a variable and a specific cut requirement.

been found to be beneficial to first grow the tree to its maximum size and then cut back, rather than interrupting the node splitting at an earlier stage. This is because apparently insignificant splits can nevertheless lead to good splits further down the tree. TMVA currently implements two tree pruning algorithms.

Expected Error Pruning. For the expected error pruning all leaf nodes for which the statistical error estimates of the parent nodes are smaller than the combined statistical error estimates of their daughter nodes are recursively deleted. The statistical error estimate of each node is calculated using the binomial error $\sqrt{p \cdot (1-p)/N}$, where N is the number of training events in the node and p its purity. The amount of pruning is controlled by multiplying the error estimate by some fudge factor.

Cost Complexity Pruning. It relates the number of nodes in a subtree below a node to the gain in terms of misclassified training events by the subtree compared the node itself with no further splitting. The cost estimate R chosen for the misclassification of training events is given by the misclassification rate $1 - \max(p,\ 1-p)$ in a node. The cost complexity for this node is then defined by

$$\rho = \frac{R_{\text{node}} - R_{\text{subtree below that node}}}{\#\text{nodes}_{\text{subtree below that node}} - 1}. \qquad (13)$$

> The node with the smallest ρ value in the tree is recursively pruned away as long as ρ is less than some fudge factor.

Note that the pruning is performed after the boosting so that the error fraction used by AdaBoost is derived from the unpruned tree.

Variable ranking

A ranking of the BDT input variables is derived by counting how often the variables are used to split decision tree nodes, and by weighting each split occurrence by the separation gain–squared it has achieved and by the number of events in the node. This measure of the variable importance can be used for a single decision tree as well as for a forest.

Performance

Only limited experience has been gained so far with boosted decision trees in HEP. In the literature decision trees are sometimes referred to as the best "out of the box" classifiers. This is because little tuning is required in order to obtain reasonably good results, mainly due to the simplicity of the method where each training step (node splitting) involves only a one–dimensional cut optimization. Decision trees are also insensitive to the inclusion of poorly discriminating input variables. While for artificial neural networks it is typically more difficult to deal with such additional variables, the decision tree training algorithm will basically ignore non discriminating variables as for each node splitting only the best discriminating variable is used.

Published paper

PRL **101**, 232002 (2008).

The Physical Review Letters reporting the results shown in this thesis is presented in the following pages.

El jurado designado por el Departamento de Física del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, aprueba la tesis:

Observación del Barión Ω_b^-
[Observation of the doubly strange $b-$Baryon Ω_b^-]

que presenta José de Jesús Hernández Orduña para su examen final de Doctor en Ciencias en la especialidad de Física, el día dieciocho del mes de Febrero del año dos mil once.

Dr. Salvador Carrillo Moreno

Dr. Omar Miranda Romagnoli

Dr. Luis Fernando Rojas Ochoa

Dr. Alberto Sánchez Hernández

Dr. Heriberto Castilla Valdez

Dr. Eduard de la Cruz Burelo