Analysis of Beauty Quark Hadronization in Vacuum and Quark-Gluon Plasma with CMS

by

Zhaozhong Shi

B.A., University of California, Berkeley (2016)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author ....
Department of Physics
September 5, 2021

Certified by ....
Yen-Jie Lee
Associate Professor
Thesis Supervisor

Accepted by ....
Nergis Mavalvala
Associate Department Head of Physics

Analysis of Beauty Quark Hadronization in Vacuum and Quark-Gluon Plasma with CMS

by

Zhaozhong Shi

Submitted to the Department of Physics
on September 5, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

An analysis of fully reconstructed B_s^0 and B^+ mesons decay into J/ψ and strange hadrons using Compact Muon Solenoid (CMS) Experiment 2017 pp dataset and 2018 PbPb data at the center of mass energy per nucleon $\sqrt{s_{NN}} = 5.02$ TeV at the Large Hadron Collider (LHC) is presented. We apply machine learning techniques along with multivariate analysis to obtain significant B-meson signals and extend the kinematic regime of B-meson measurements with higher precision. In our analysis, B_s^0 signal of greater than 5 σ significance is observed for the first time in heavy-ion collisions. The measured B_s^0/B^+ ratio in PbPb along with pp references are compared with theoretical model predictions. These results will help elucidate the beauty quark hadronizaton mechanisms in vacuum and quark-gluon plasma at the LHC energy. Significant B-meson signals have also been observed at low very p_T and high multiplicity in pp collisions, which will allow us study beauty hadrochemistry in small systems and energy loss mechanism in the future.

Thesis Supervisor: Yen-Jie Lee
Title: Associate Professor

Acknowledgments

This is the acknowledgements section. You should replace this with your own acknowledgements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Model of Particle Physics

There are four known fundamental forces in nature: gravitational force, electromagnetic force, strong force, and weak force. The gravitation force describes the interaction between two massive objects. The electromagnetic force describes the interaction between electrically charged objects. The strong force describes the interaction between nucleons. The weak force describe the radioactive decay of particles. The Standard Model (SM) of Particle Physics, first proposed and named by physicist Steven Weinberg in the 1960s [1], is based on theoretical frame of relativistic quantum field theory with a gauge symmetry of $SU(3) \times SU(2) \times U(1)$ [2]. It unifies the electromagnetic and weak interactions and include the strong interaction into a theory and describes all particles participating in these interactions. The ingredient of the standard model are leptons, quarks, gauge bosons, and the Higgs boson shown in Figure 1-1.

There are 19 parameters in the Standard Model: 6 quark masses, 3 lepton masses, 3 coupling strengths, 4 angles in the Cabibbo?Kobayashi?Maskawa Matrix, Higgs mass, vacuum expectation value, and QCD vacuum angle. These parameters are determined from the experiments. Physicists perform calculations based on the Standard Model and predict the cross section of different processes in high energy physics experiments. Since it is proposed in the 1970s, the Standard Model has been tested

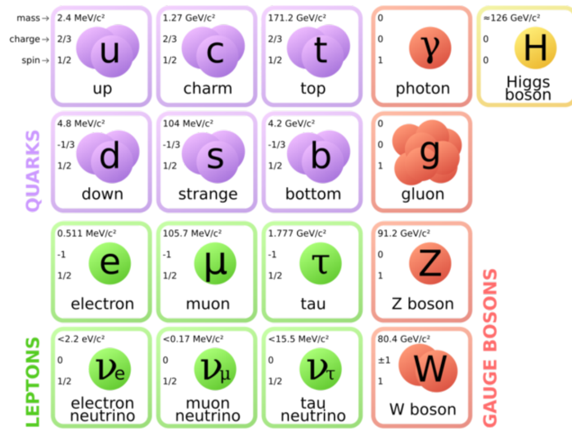


Figure 1-1: The 17 elementary particles, including 6 leptons, 6 quarks, 4 gauge bosons, and the Higgs boson, and their basic properties, such as mass, electric charge, spin, in the Standard Model of Particles Physics are shown above.

extensively in countless high-energy physics experiments. Its prediction holds for all of them with very few exceptions. The Standard Model consists of two sectors: the Electroweak theory (EW) and Quantum Chromodynamics (QCD). The Lagrangian of the Standard Model can be written as the sum of EW and QCD: $\mathcal{L}_{\mathcal{SM}} = \mathcal{L}_{\mathcal{EW}} + \mathcal{L}_{\mathcal{QCD}}$

1.2 Quantum Chromodynamics

1.2.1 QCD Lagrangian

QCD, a non-abelian gauge theory with $SU(3)$ symmetry, is the theory for the strong interaction between quarks and gluons. The QCD Lagrangian is shown as follows:

$$\mathcal{L}_{\mathcal{QCD}} = \bar{\Psi}^i i(\not{D})_{ij}\Psi^j - m\bar{\Psi}^i\Psi_i - \frac{1}{16\pi^2}G_a^{\mu\nu}G_{\mu\nu}^a \tag{1.1}$$

Where

$$\not{D} = \gamma^\mu\partial_\mu - ig_s\frac{\lambda}{2}\gamma^\mu A_\mu \tag{1.2}$$

$$G_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial_\nu A_a^\mu + g_s f_{abc}A_b^\mu A_c^\nu \tag{1.3}$$

Here, λ are the Gell-Mann Matrices. f_{abc} is the structure of constant of $SU(3)$. A^{μ} is the eight gluon field. g_s is the strong coupling constant. The color indices i and j run from 1 to 3, which stands for 3 colors: red, blue, and green. The gluon field indices a, b, and c run from 1 to 8, standing for the 8 gluon state (Gluon octet as the combination of 3 color and 3 anticolor: $3 \times \bar{3} = 1 \oplus 8$) living in the adjoint representation of $SU(3)$ of color.

1.2.2 Asymptotic Freedom

The running of the strong coupling constant $\alpha_s = \frac{g_s^2}{4\pi}$ according to the 1-loop calculations in the renomalization theory [3] is shown as follows

$$\alpha_s(Q^2) = \frac{12\pi}{(11N_c - 2N_f)\ln(\frac{Q^2}{\Lambda_{QCD}^2})} \tag{1.4}$$

We can see that as the energy scale increases, the coupling strength of the strong interaction decreases. This is in contrast to QED where the electromagnetic coupling strength increases as the energy scale increases. In the ultra-violate limit $Q^2 \to \infty$ and $\alpha_s \to 0$, quarks and gluons behave like free particles. This feature in QCD is called Asymptotic Freedom [5]. Meanwhile, in the infrared limit, the strong coupling constant increases. Near the $\Lambda_{QCD} \simeq$100 MeV, the strong coupling is greater than 1, where the perturbative expansion of QCD breaks down. Experimentally, physicists measure the strong coupling constant at different energy scales from different experiments at different colliders. Figure 1-2 [4] shows the running of strong coupling constant in experiment and comparison with the theoretical calculations

An excellent agreement between theoretical predictions and experimental results of the strong coupling constant is observed in Figure 1-2.

1.2.3 Perturbative QCD

It is mathematically proven that there is in general no closed form expression for the partition function $Z[J(x)]$ Standard Model Lagrangian under the Quantum Field

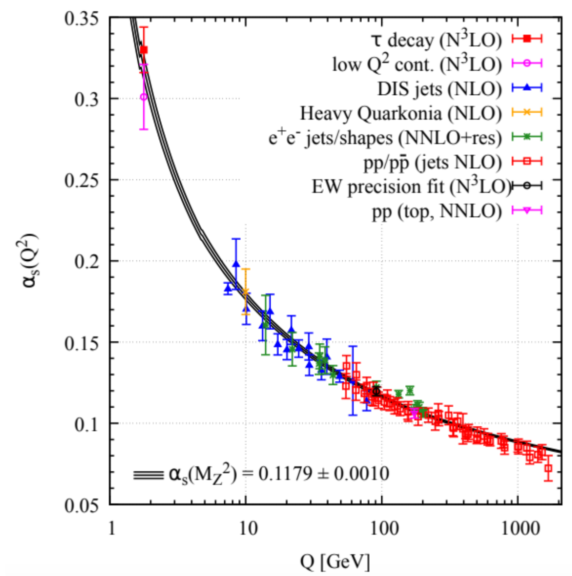


Figure 1-2: The running of the strong coupling constant α_s in different experiments at different energy scale Q and the comparison with QCD calculations are shown above. Image from: [4]

Theory framework.

$$Z[J(x)] = \int \mathcal{D}[\phi(x)] e^{iS+\int d^4x J(x)\phi(x)} \tag{1.5}$$

Therefore, physicists develop perturbation theory in Quantum Field Theory and apply it to the Standard Model. Physicist perform asymptotic expansions to obtain power series of the coupling constants and approximately calculate the expectation values of the observables to prediction experimental results.

For QCD, , perturbation theory is applicable to QCD in high energy and hard scattering processes since the coupling constant is much less than 1. Feynman rules and diagrams are applicable in the matrix element to evaluate the cross section of hard parton-parton scattering. Perturbative QCD (pQCD) calculations have been tested with various experiments such as electron-positron annihilations, deep inelastic electron-proton scatterings, and high energy proton-proton collisions.

1.2.4 Non-perturbative QCD

For soft scattering processes at low energy, the strong coupling constant is greater than 1. Perturbation theory of QCD breaks down. Many low-energy QCD processes such as hadronization and hadron-hadron interactions are non-perturbative. Historically, physicists developed Lattice gauge theory where the spacetime is discretized into lattice with finite size to evaluate the path integrals in the partition function $Z[J(x)]$. Lattice QCD can be applied to calculate the mass of the proton [6]. Aside from Lattice gauge theory, effective theory is also to study non-perturbative QCD. For example, Chiral Perturbation Theory, where a low-energy effective Lagrangian in degree of freedom of hadrons is constructed by exploiting the approximate chiral symmetry while preserving other symmetries of parity and charge conjugation, has achieved some success to study pion-nucleon scattering [7]. Non-perturbative QCD has achieved many successes in hadronic physics. Currently, some novel developments applying non-perturbative QCD to understand nuclear structure and nucleon spin structure are being carried by physicists. For instance, Chiral Perturbation Theory has been applied to study light nuclei structure such as ${}^{6}_{3}Li$ and ${}^{10}_{5}B$ [8] and Lattice is employed to investigate nucleon spin structure [9].

1.2.5 QCD Factorization Theorem

The QCD factorization theorem states that in events with hadrons as incoming particle involving both hard and soft QCD processes, hard and soft process are mathematically factorized in the cross section computation shown schematically below [10]:

$$\sigma = PDF \otimes Diagrams \otimes FF \tag{1.6}$$

The hard processes are encoded in the factor of partonic cross sections while the soft processes are measured in experiments. Physicists employ parton distribution function (PDF), defined as the probability of finding a particle with a certain longitudinal momentum fraction x at resolution scale of μ^2, to describe the initial kinematic of partons inside hadrons [11] and fragmentation function (FF), defined as the prob-

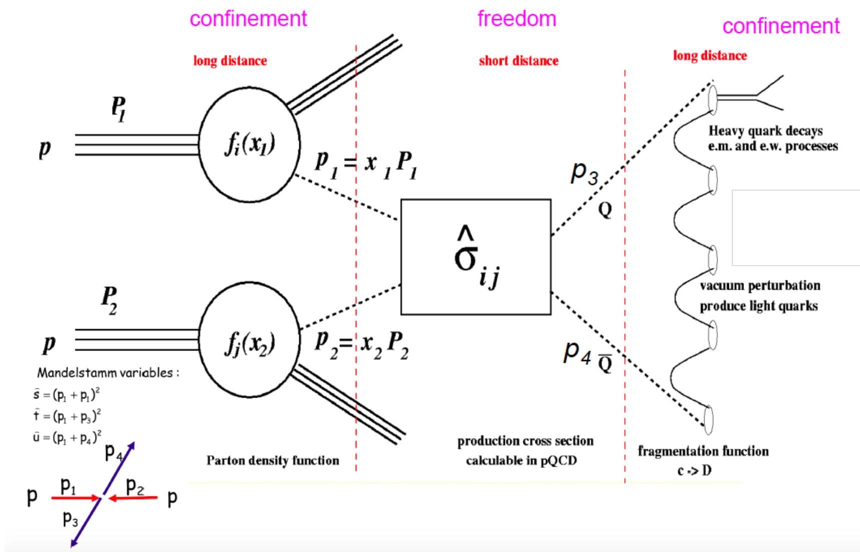


Figure 1-3: The QCD factorization theorem applied to a pp collision event involving in soft and hard processes are shown above.

ability of a parton turn into a hadron $D_i^h(z)$ for a given energy fraction of the parton z at resolution scale of μ^2, to describe the hadronization process of partons [12].

In addition to PDF, we could also define nuclear PDF (nPDF) for [?] to describe the parton kinematics inside nucleus. nPDF could be understood as the PDF of nucleons modified by the nuclear environment. Both parton distribution function and fragmentation function are extracted in experiments.

Physicists apply QCD factorization theorem to perform pQCD calculations and compare them with hadron spectra in electron-positron (e^+e^-), electron-proton (ep), and proton-proton (pp) collisions.

1.2.6 Color Confinement

Another feature of QCD as a non-abelian gauge theory is color confinement. The strong force carrier gluon itself is also color charged. Color charged partons, namely quarks and gluons, are never detected in isolation. In experiments, only color neutral hadrons are detected. Currently, the analytic explanation of color confinement is still not yet rigorously proven. The theoretical explanation of color confinement in QCD remains one of the unsolved problem in physics.

1.2.7 Hadronization

The formation process hadrons from partons is called haronization. Because in experiments we can only measure final state hadrons, in order to study the interactions and dynamics of quarks and gluons during partonic stage from hadron spectra, we also need to understand hadronization mechanisms. However, hadronization is in general non-perturbative and cannot yet be described by first principle QCD calculations. Therefore, physicists make phenomenological models such as the Statistical Hadronization Model [14], Lund String Model [15], Cluster Hadronization Model [16], Quark Coalescence Model [17] to study hadronization. Figure 1-4 schematically shows the hadronization of heavy quarks via fragmentation and recombination mechanism.

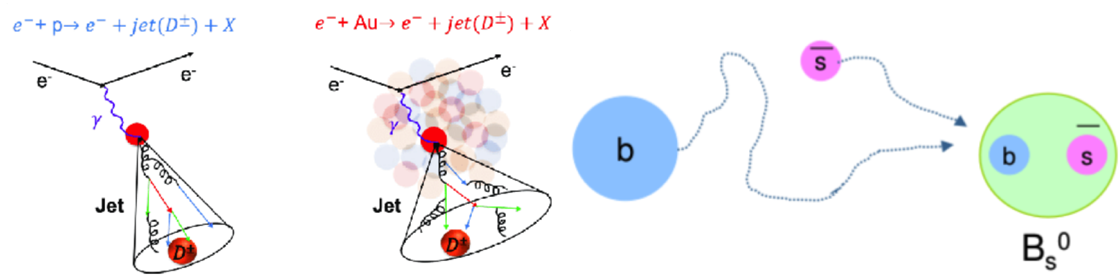


Figure 1-4: The fragmentation process of charms quarks hadronize into $D^{\pm}$ (left) and the coalescence process of beauty quark combining with a strange quark nearby to form a B_s^0 are shown above.

1.2.8 Initial State and Final Effect

In high energy proton-nucleus (pA) collisions, protons scatter off nucleons inside nuclei. Assuming QCD factorization still holds, nPDF could be applied to calculate the cross section of particle production. The initial state effects of nuclei including event-by-event geometry fluctuations due to nuclear dynamics [64], nuclear shadowing effect [?], EMC effect [48] will modify the PDFs of nucleons in nuclei compare to PDFs of nucleons in vacuum.

In the final state, the struck parton will lose energy from the interaction with the nuclear fragment and modify the final state hadron spectra. Because the parton-hadron interaction is generally non-perturbative, to retain the formula of QCD factorization theorem, the parton FF is modified [49]. These initial stage and final stage effects in pA collision are call cold nuclear matter effects.

1.3 Hot QCD

1.3.1 QCD in Finite Temperature

In a system of dense and energetic quarks and gluons confined in a given size of volume, they scatter with each other and exchange momenta via the strong interaction. Many-body dynamics between quarks and gluons become relevant. In the limit of large number of quarks and gluons, after a sufficiently long period of time, the system eventually converges to the thermal equilibrium state via strong interaction [20–22] regardless of its initial states. Therefore, a description based on thermodynamics can be formulated to study such systems [24]. We call this thermalized many-body system of quarks and gluons to be QCD matter. Therefore, a thermodynamic variable temperature (**T**) can be introduced to characterize these many-body QCD systems. The study of many-body QCD in finite temperature is called hot QCD.

1.3.2 Temperature Dependence of QCD Static Potential

.

If we consider two color charged quarks in the limit of infinite mass and are essentially at rest in the lab frame, we can define a QCD static potential between these two quarks due to the strong interaction. In vacuum, such a potential is called "Cornell Potential" [34]. The potential as a function of the distance between two quarks is shown as follows:

$$V(r) = -\frac{\alpha_{eff}}{r} + \sigma r \tag{1.7}$$

Here, α_{eff} is the effective strong coupling coupling between the two quarks and $\sigma \simeq 0.184$ GeV/c is the string coupling constant [35].

Now if we consider a thermalized system in finite temperature T, the potential becomes:

$$V(r) = -\frac{\alpha_{eff}}{r}e^{-m_D r} + \frac{\sigma}{m_D}(1 - e^{-m_D r}) \tag{1.8}$$

Here, $m_D \sim g_s T$ is the Debye mass due to Debye color screening effect [36], which essentially modifies the gluon propagator by inserting a finite mass term: $-i\frac{g^{\mu\nu}}{q^2} \rightarrow -i\frac{g^{\mu\nu}}{q^2 - m_D^2}$. We have observed that as $V(\infty) = \frac{\sigma}{g_s T}$, which is finite for T >0. In fact, Equation (2) reduces to the Cornell potential when $T = 0$. The QCD static potential is shown below in Figure 1-5 [37]

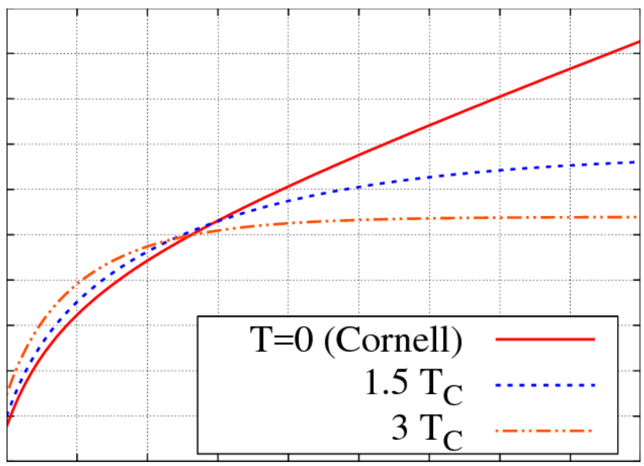


Figure 1-5: The QCD potential $V(r)$ from at zero and at finite temperatures as a function of distance r is shown above. Here, the critical temperature $T_c = 192$ MeV. We can see that the QCD saturates at a finite value at finite temperature.

1.3.3 Color Deconfinement

As mentioned in the sections above, at finite temperature, the QCD static potential is screened and color degree of freedom become relevant in the system. As the temperature of the system increases, the quarks and gluon inside color-neutral hadrons will have more available space to move around and start to deconfine [40]. At some critical temperature T_c, quarks and gluons confined in hadrons will melt and form a new state of color deconfined QCD matter, which is called Quark-Gluon Plasma

(QGP) [?]. The typical temperature of QGP is in the order of a few hundred MeV or about 10^{12} K, which is about hundreds of thousands times hotter than the core of the Sun. It is belived that QGP existed in the early universe several microsecond after the Big Bang []. Cosmologically, the study of QGP will help us understand the quark epoch and quark-hadron phase transition to under the history of the universe.

1.3.4 QCD Phase Diagram

Similar to form everyday matters such as metal, water, wood, glass, and plastic, which are formed by electromagnetic interaction and could all be described macroscopically by equations of states that are parameterized by thermodynamic variables. Similarly, thermodynamical variables, for instance temperature (**T**) and baryon chemical potential (μ_B) can be introduced to characterize the equation of state of QCD matter formed via the strong interaction between many quarks and gluons. Like our everyday matter which has gas, liquid, and solid phases at different pressure and temperature, QCD matter also has different phases at different temperature and baryon chemical potential. and can be describe by QCD phase diagrams. Figure 1-7 shows the QCD phase diagram at different temperature and baryon chemical potential:

We consider a system of free up and down quarks, antiquarks and gluons in temperature T and baryon chemical potential μ_B. According to MIT Bag Model [41], its equation of state is given by

$$\epsilon(T, \mu_B) = \frac{37\pi^2}{30}T^4 + \frac{\mu_B^2}{3}T^2 + \frac{\mu_B^4}{54\pi^2} - \mathcal{B} \tag{1.9}$$

Here p is the pressure and $\mathcal{B}$ is the bag constant, which can be understood as the pressure of the vacuum on the quarks and gluons to make them form hadrons with finite volume.

In a system of interacting quarks and gluons at $\mu_B = 0$, based on lattice QCD calculations [42], the reduced energy density ϵ/T^4 as a function of the temperature T is shown below

A steep increase of the ϵ/T^4 near the critical temperature at around $T_c = 173$

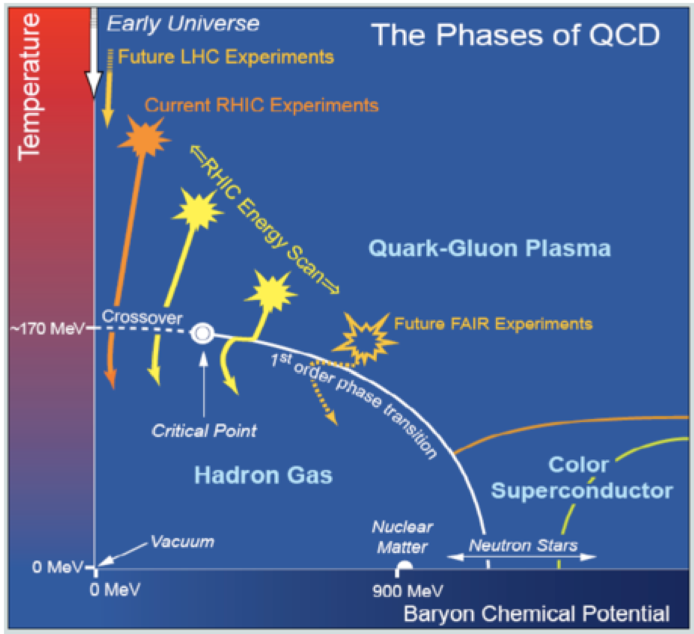


Figure 1-6: The theoretical QCD phase diagram of different QCD matter, including hadron resonance gas, quark-gluon plasma, neutron star, and color superconductor, as function of temperature and baryon chemical potential is shown above. The solid line indicates the conjecture of first order phase transition between quark-gluon plasma and hadron gas while the dash line is a smooth crossover. Thermodynamically, a critical points must exist in the boundary of smooth crossover and first order phase transition.

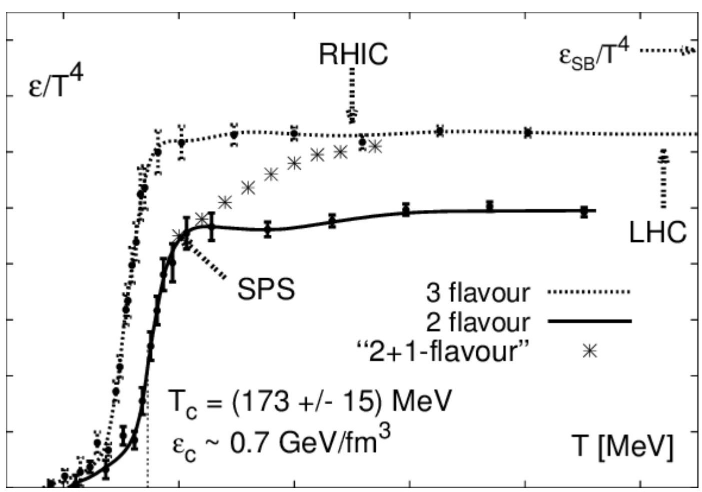


Figure 1-7: The reduced energy density ϵ/T^4 as a function of temperature T for different number of flavor scenarios from the lattice QCD calculations (data points) and the interpolation curves are shown above.

MeV is observed, which signals the transition from hadron gas to the QGP [43]. Experimentally, the critical point is estimated to be around $T_c = 175^{+1}_{-7}$ MeV and $\mu_B = (22 \pm 4.5)$ MeV [44] .

1.4 High Energy Nuclear Physics

Nuclear Physics is the study of atomic nuclei and their constituents and interactions. The typical energy scales of nuclear physics range from MeV to GeV. High Energy Nuclear Physics is a subfield of Nuclear Physics at an energy scale on the order of GeV. Its main goal is to under the physics of QCD matter from various approaches such as collider experiments, astrophysical observations, physics simulations, and theoretical modelings. In this thesis, I will focus on the research of QGP from the experimental approach using high energy heavy-ion colliders.

1.4.1 Laboratories

In laboratories, high energy nuclear physicists accelerate and collide heavy ions (A $>$ 56) at center of mass high energy per nucleon at grater than 1 GeV to create extremely hot and dense conditions and study QGP. Relativistic heavy-ion collision is also known as "The Little Bang" compared to "The Big Bang" in cosmology [45]. Historically, many colliders, such as the Alternating Gradient Synchrotron (AGS) at Brookhaven National Laboratory (BNL), in Upton, Long Island, New York and Super Proton Synchrotron (SPS) at European Center for Nuclear Research (CERN) in Meyrin, Switzerland, and GSI at Helmholtz Centre for Heavy Ion Research with both proton-proton and relativistic heavy-ion collision capabilities, have been built and established high-energy nuclear physics research programs. Today, two active colliders facilities, the Relativistic Heavy Ion Collider (RHIC) at BNL and the Large Hadron Collider (LHC) at CERN, are running at a wide range energies with various nuclei species and different impact parameters. In the future, another collider, called Facility for Antiproton and Ion Research (FAIR) running at relatively low energies, is being constructed at Darmstadt, Germany to map the location of the critical point in the QCD Phase Diagram.

In addition to collider facilities, QGP might also be studied from astrophysical observations. For instance, strange stars, a quark star made of strange quark matter, may come from stable strangelet according to Bodmer–Witten conjecture [50] or exist in the core of neutron stars under extreme pressure and temperature. It is believe there are several potential strange stars candidates according to telescope observations and gamma ray burst analysis [51–53].

1.4.2 Relativistic Heavy Ion Collider (RHIC)

Located at BNL in Upton, Long Island, New York, United States of America, RHIC is one of the major high energy accelerator facilities and currently the highest energy collider in America. It is a circular collider with a circumference of 3.843 kilometers and can provide proton energy up to 500 GeV and gold ion energy up to 200 GeV [54]. It was built in 2000 in order to search for a strongly interacting hot and dense state of nuclear matter created under ultra-relativistic heavy-ion collisions, currently known as QGP, with hints from the measurements at AGS and SPS. Moreover, RHIC provides physicists with a wide range of energies and a variety of ion species from proton to deuteron and cooper to uranium to create different sizes of systems at different temperature and baryon densities. In addition, taking the advantage of its highly polarization beam with high luminosity, RHIC has great machine capabilities for cold QCD physics. Figure 1-8 below shows a sky view of RHIC at BNL:

Here is how RHIC accelerates charged particles to the energy scale of GeV per nucleon. For instance, if we consider the acceleration of a typical ion source gold ($^{197}_{79}Au$) ion, we first use a cesium sputter ion source operated in the pulsed beam mode and point it to the gold metal to produce the Au^{-} ion [56]. Then, the Au^{-} will undergo a series of electron stripping processes to reach the Au^{79+} ion [57]. First, 13 electrons are stripped by the carbon foil in the Terminal Stripping (S1) after the acceleration of tandem Van der Graaf generator to turn Au^{-} into Au^{12+}. Then, the Au^{12+} ion will go through the Object Foil (S2) at the second stripping stage and becomes Au^{31+}. Next, the Au^{31+} will go through the third stripping station BTA

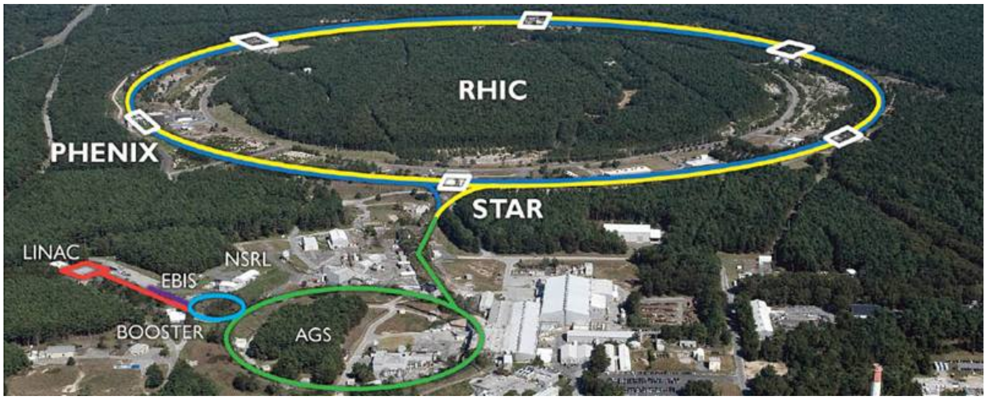


Figure 1-8: The view of RHIC at BNL from the sky is shown above. The actual locations of other facilities at BNL, including Linac, Booster, EBIS, NSRL, AGS, and the experiments at RHIC such as, STAR and PHENIX, are also labelled.

foil (S3) made of aluminum and vitreous carbon between the Booster Synchrotron and AGS and becomes Au^{77+}. Finally, two more electrons of the gold ion Au^{77+} are removed at the fourth stripping station ATF foil (S4) made of thin tungsten, located in between the AGS and RHIC. The fully stripped gold ions Au^{79+} will then be injected to the blue and yellow rings at RHIC. For polarized protons, H^- pass a single stripping stage called located in the Booster Synchrotron. The stripping station is called Linac-to-Booster (LTB) stripper made of carbon foils with special geometry and converts polarized H^- to H^+. Figure 1-9 schematically shows the accelerating process of gold ions at RHIC [58]

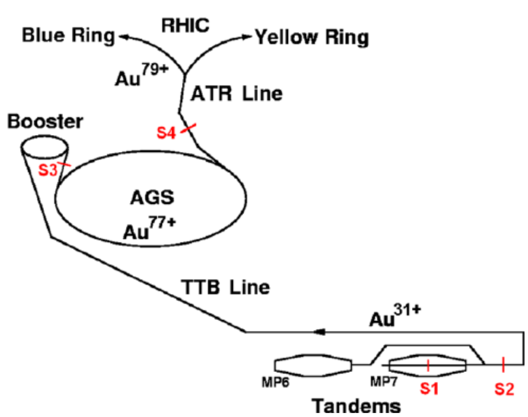


Figure 1-9: The acceleration of gold ions for RHIC is shown above.

At RHIC, we will accelerate the Au^{79+} ions in the superconducting Radio Frequency (RF) cavity under perpendicular electric and magnetic fields until they reach the energies up to about 100 GeV/c per nucleon. Subsequently, we collider them via bunch crossing at the interaction points of the experiments to perform relativistic heavy-ion collisions and study high energy nuclear physics. RHIC usually operates in the first six months of a calendar year. At RHIC, the energy can also be lower where the ion beam collides with ions at a lower energy in the laboratory frame. The STAR experiment at RHIC has already finished beam energy scan and is currently taking data in the fixed target program.

1.4.3 Large Hadron Collider (LHC)

Located at the border between Switzerland and France, LHC is one of the major high energy accelerator facilities in Europe and currently the highest energy collider in the world. It is a circular collider with a circumference of 26.7 kilometers and can provide proton energy up to 14.0 TeV and lead ion energy up to 5.02 TeV [59]. It was built in 2008 with the main purpose to discover the Higgs Boson, perform precision measurements on SM, and search for Physics beyond SM. Due to its high energy ion capabilities, high energy nuclear physicists also use the existing general purpose detectors designed for high energy particle experiments at the LHC to conduct research on relativistic heavy-ion physics. LHC ion physics runs usually start at the end of the year and lasts for about a month. The photo taken from the sky to picture LHC is shown in Figure 1-10:

CERN usually uses is lead $^{208}_{82}Pb$ ions, which are stable and approximately spherical. In the 2017 ion physics run, it also used the xenon $^{131}_{52}Xe$. Currently, there is also a discussion of potential future usage of lighter ions such as oxygen $^{32}_{16}O$ [60]. Similar to RHIC, the lead ions at the LHC also undergo a series of stripping processes using stripping foils in to order to become partially ionized Pb^{81+} [61]. Also, the lead ions pass a series of energy boosting before reaching to the desired energies at the LHC. Lead ions start from a source of vaporized lead and enter Linac 3 before being collected and accelerated in the Low Energy Ion Ring (LEIR) at the energy from 4.2

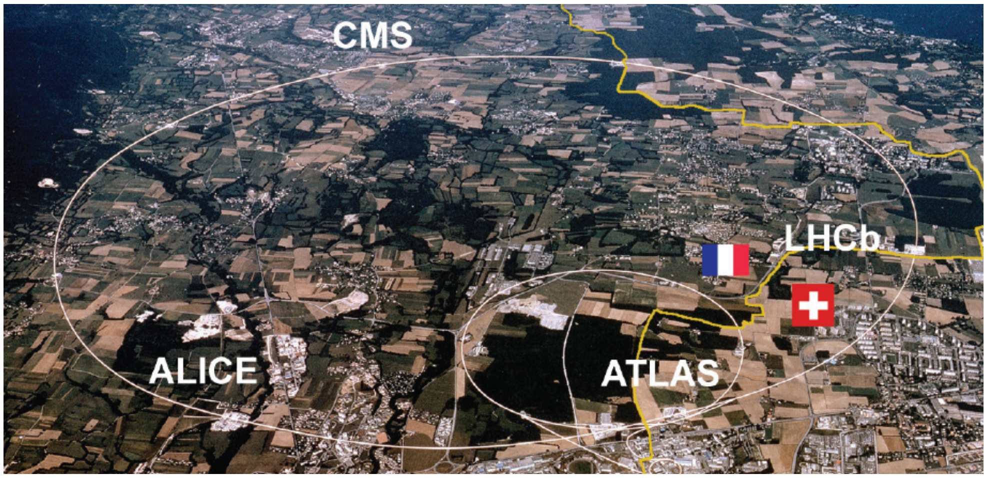


Figure 1-10: The sky view of LHC at CERN is shown above. The actual locations of the experiments at the LHC: ATLAS, CMS, ALICE and LHCb, as well as the French-Swiss border, are also displayed.

MeV to 72 MeV. Then, the lead ions will be injected to Proton Synchrotron (PS) to boost their energies. Next, they are sent to the Super Proton Synchrotron (SPS). Finally, the lead ions are injected to the LHC and increase their energies to TeV scale in two LHC rings with the RF cavity [59]. Finally, the energetic lead ion beams from two LHC rings will collide with a small crossing angle at the interaction points of the LHC experiments. The CERN accelerator complex is shown schematically in 1-11

After Run III, LHC will upgrade to high-luminosity (HL) LHC and allows physicists to collect huge datasets, which is crucial to for precision measurements in heavy-ion physics program. Because the beam energy at the LHC is higher than RHIC, the QGP created at the LHC has a higher temperature and a smaller baryon chemical potential than the one created at RHIC.

1.4.4 High Energy Physics Coordinates

As mentioned in the previous section, heavy-ion is in general highly relativistic. Therefore, Lorentz transformation will be relevant in our studies. In Cartesian coordinates $x^\mu = (t, x, y, z)$, under Lorentz transformation, if we boost the system by a speed β in the $+z$ direction. The Lorentz gamma factor will be given by $\gamma = \frac{1}{\sqrt{1-\beta^2}}$.

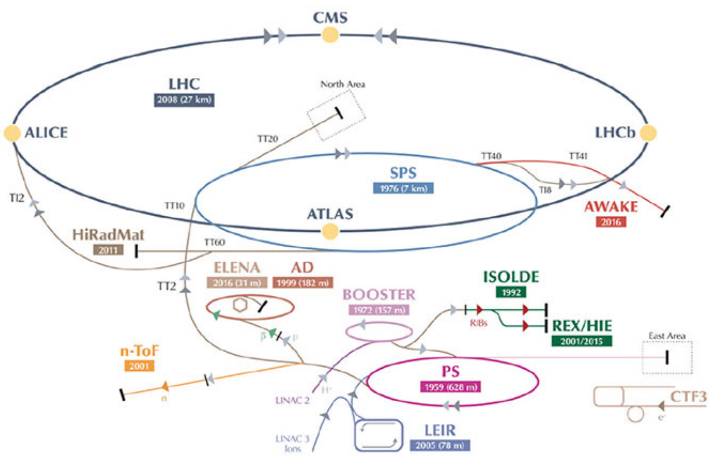


Figure 1-11: The schematic overview of CERN accelerator complex with the accelerators labelled is shown above. Proton and lead ions are accelerated using these facilities and boost their energies to the TeV scale.

The four vector $x^{\mu} \rightarrow x'^{\mu}$ transforms as follows

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} \tag{1.10}$$

The equation above is called the Lorentz Transformation. It is an orthogonal transformation preserving the Minkowski metric tensor $diag(1, -1, -1, -1)$ using particle physicists conventions.

Nowadays, heavy-ion detectors usually have 2π angular coverage in the transverse direction with some finite longitudinal acceptance along the beam line. They are essentially cylindrically symmetric. Hence, it is convenient and sensible to choose a cylindrical coordinate system and use Lorentz invariant kinematic variables. In general, we define the beam direction to be the z-direction of the coordinate system. Fort the standard cylindrical coordinates in the position space, the Lorentz four-vectors is $(t, x, y, z) \rightarrow (t, r, \phi, z)$.

The relativistic coordinate system for our analysis is shown below in Figure 1-12.

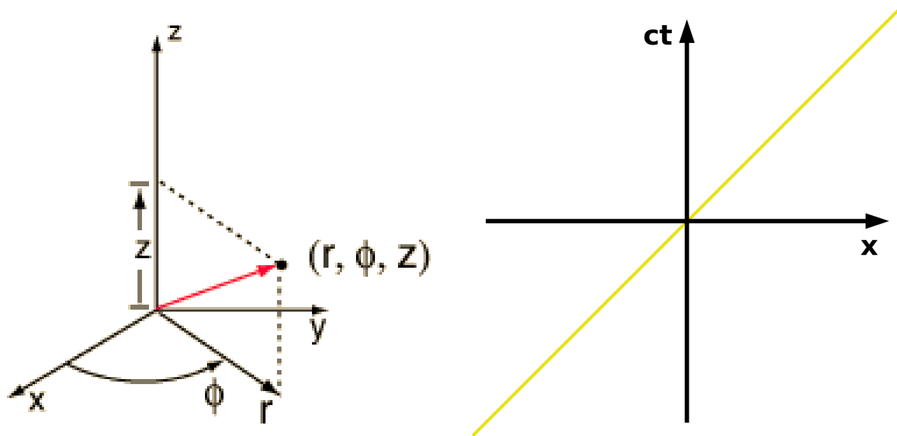


Figure 1-12: The cylindrical coordinate system in the position space (left) and the space time diagram (right) for relativistic heavy-ion physics analysis are shown above.

Thus, in the momentum space, we can use $p^{\mu} = (E, p_x, p_y, p_z) \rightarrow (E, p_T, \phi, p_z)$

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{1.11}$$

$$\phi = \arctan(\frac{p_y}{p_x}) \tag{1.12}$$

We also define rapidity y, a relativistic version of velocity that can be convenient add to the boost.

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{1.13}$$

Experimentally, we also use pseudo-rapidity η, which is more directly connected to the detector measurements assuming ultra-relativistic limit kinematics ($E \rightarrow p$). The definition of pseudo-rapidity η is shown as follows:

$$\eta = -\ln \tan(\frac{\theta}{2}) \tag{1.14}$$

Here θ is the angle labelled in the left of Figure 1-12. Particularly, $y = 0$ and $\eta = 0$ when $p_z = 0$. In addition, boosting by a speed β in the longitudinal z-direction, we found that the rapidity simply shift by a const number $y' = y + \tanh \beta$. We should note that the cylindrical coordinates (p_T,ϕ, p_z) are perfectly orthogonal while (p_T,ϕ, y) or (p_T,ϕ, η) are not.

For general collider experiments, two particles are moving toward each other with four-momenta p_1^μ and p_2^μ and interact with each other. It is also very convenient to use the Mandelstam variables s, t, u in our studies. They are defined as follows

$$s \equiv (p_1 + p_2)^2 \tag{1.15}$$

$$t \equiv (p_1 - p_2)^2 \tag{1.16}$$

$$u \equiv (p_1 - p_3)^2 \tag{1.17}$$

In the center of mass frame, the momentum three vector follows $\vec{p_1} = -\vec{p_2} = \vec{p}$. Therefore, $p_1^\mu = (E, \vec{p})$ and $p_2^\mu = (E, -\vec{p})$. Hence, $s \equiv (p_1 + p_2)^2 = 4E^2 = E_{CM}^2$. Hence, the center of mass energy of the collision system could be represented by the Mandelstam variable $\sqrt{s}$: $E_{CM} = \sqrt{s}$.

1.4.5 Stages of Heavy-Ion Collisions

In high energy heavy-ion collisions, both Electroweak and QCD processes occur in each event and contribute to the total cross section. We classify the events with elastic and inelastic reaction processes. For elastic processes, two nuclei scatter mainly electromagnetically with each via photon exchange without breaking themselves up or losing energy. For inelastic scattering, we classify diffractive and non-diffractive disassociation processes. In diffractive dissociation processes, the two nuclei may be slightly excited and lose a relatively small fraction of their energies, and produce relatively small number of particles. On the other hand, in non-diffractive dissociation processes, the nuclei lose a substantial fraction of their energies and produce a large number of particles [62].

Therefore, in events with significant contribution from non-diffractive dissociation, the interaction between two nuclei has multiple stages including both perturbative and non-pertubative QCD processes. We can define stages of heavy collisions and study the details of each stage. There are five stages: initial state of two highly Lorentz contracted nuclei before the collision, the very early pre-equilibrium stage when hard scatterings between partons inside nuclei begin, the rapid expansion of the fireball when the thermally and chemically equilibrated QGP is created, the hadronization stage after QGP expands and cools down, and the freeze-out stage when the inelastic scattering processes cease.

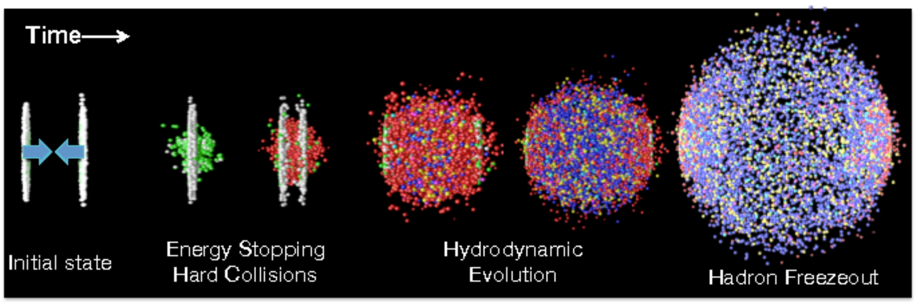


Figure 1-13: An event of a typical heavy-ion collisions event with different stages as time evolves is shown above.

Theoretically, many phenomenological models such as Ultra-Relativistic Quantum Molecular Dynamics (UrQMD) and A Multi-Phase Transport Model (AMPT) are developed to describe relativistic heavy-ion collisions.

1.4.6 Global Event Observables

Globally, we can define some physical quantities in heavy-ion collisions to generally characterize each event. Heavy-ion physicists define the impact parameter, centrality, number of participants, number of binary nucleon-nucleon collisions, and event multiplicity. We will discuss all of them below.

Impact Parameter: Prior to heavy-ion collisions, similar to other collider experiments, each event are prepared with the same unpolarized incoming particles with the same center of mass energy. Therefore, the incoming state $|i\rangle$ is used for each

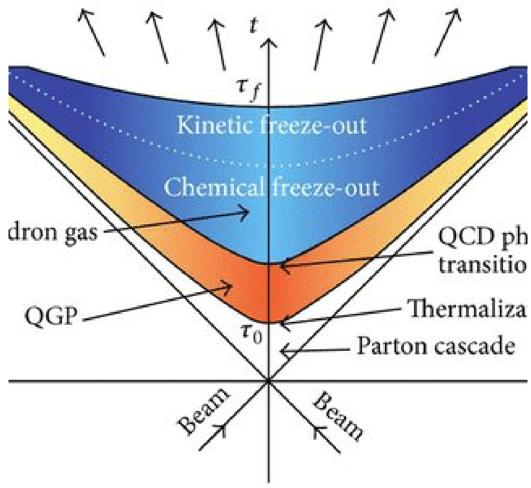


Figure 1-14: The space-time evolution of heavy-ion collisions is shown above. It consists of four stages: initial state before the collision, early stage of hard scattering processes, the hydrodynamic expansion of QGP, hadronization after QGP expands and cools down, and the freeze-out stage, first chemical freezeout when the particle species no longer change, and finally kinetic freezeout when the elastic scattering processes ceas.

event. However, different from e^+e^- and pp collision, in heavy-ion physics, we introduce another parameter called the impact parameter denoted b to the transverse distance between center of two nuclei to classify the events. Therefore, the incoming state can be rewritten as $|i(b)\rangle$. Figure 1-15 shows the definition of impact parameter in heavy-ion collision [63].

Number of Participants: Right at the end of heavy-ion collisions after two nuclei pass through each other, we can define the number of participants denoted N_{part}. The number of participants is essentially equivalent to the number of participating nucleons. The smaller the impact parameter, the more overlap volume between two nuclei, leading to a larger number of number of participating nucleons in the collision. The nuclear interaction system size is determined by the number of participating nucleons. However, due to event-by-event nuclei geometry fluctuations caused by the motion of nucleons inside nuclei [64], it is more proper to say that the average number of participant $\langle N_{part}\rangle$ is related to the impact parameter.

Number of Binary Nucleon-Nucleon Collisions: In addition to N_{part}, we can

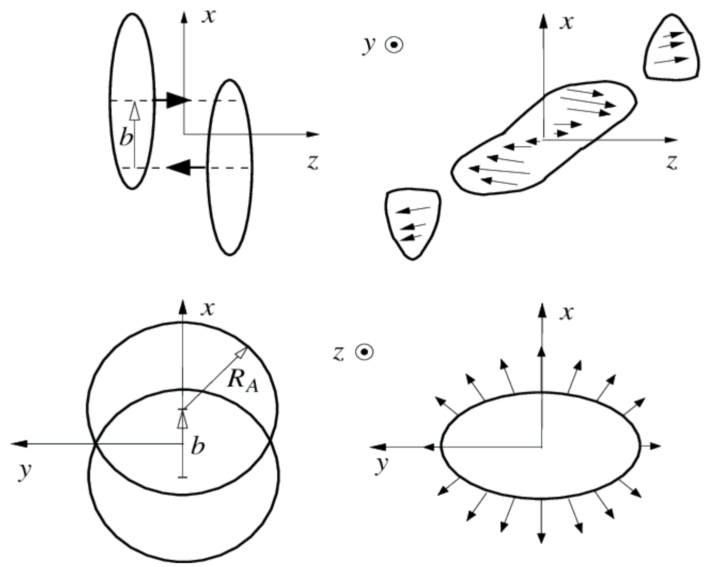


Figure 1-15: The definition of impact parameter b in heavy-ion collisions, the of overlapping interaction region, and the break up remnants of the two nuclei, which is called spectator, moving in the z-direction are shown above. An almond shape of the nuclear interaction region, which results in the azimuthal anisotropic emission of final state particles, is seen in heavy-ion collisions.

also define another quantity that characterizes the detailed interaction in the events at rather hard scales. The number of binary nucleon-nucleon collisions, denoted N_{coll}, is also related to the impact parameter. At higher energy, nucleons inside nuclei become the relevant degree of freedom to describe the cross section. We could treat the collisions of two nuclei as the superposition of the collisions between nucleons inside the nuclei. Since binary nucleon-nucleon collision has a rather small cross section, it dominates the total nucleon-nucleon cross section according to binomial principle. Higher order effects, such as ternary nucleon-nucleon collisions, are negligible. The Glauber model [65] is developed to study the relationship between b, N_{part}, and N_{coll} in nuclei collisions and will be discussed in the following subsection.

Centrality: Experimentally, it is difficult to directly measure the impact parameter of each collision. Therefore, we define another physical quantity called centrality to characterize the impact parameter. The centrality (C) is defined as the fraction of the total nuclear interaction cross section: $C = \int_0^b \frac{d\sigma}{dx} dx$. Centrality is expressed in terms of percentage [66]. It is related to the quantity: $\frac{\pi b^2}{4\pi R_A^2}$ where R_A is the radius of a nuclei defined above in Figure 1-15. When the impact parameter between two nuclei is 0, the centrality is at 0%. When the impact parameter between two nuclei is $2R_A$, the centrality is 100%. There is also a relationship between the centrality and the average number of participants. Heavy-ion experimental measurements are in general presented in terms of centrality or average number of participants. Experimentally, we look at the number of tracks and cluster energies of calorimeters at the forward direction, for instance, the forward hardonic calorimeters, to estimate the centrality [67–69].

Virtuality: Similar to deep inelastic scattering, we can also define the virtuality Q^2, which is the momentum transfer between the two nucleons in nucleon-nucleon collisions. To generate nucleon-nucleon collision event in Monte Carlo (MC) simulation, we used $\hat{p}_T$, defined as the transverse momentum of the hard subprocess, which is a quantity related to Q^2, developed by the high energy theory group of Lund University.

Event Multiplicity: We can also define the event multiplicity by counting the number of final state charged particles to quantify the activity of the event. Event multiplicity can be denoted as N_{trk}, number of tracks in the event, which is approximately proportional to the number of charged particle denoted as N_{ch}. Figure 1-16 shows the correlation between the number of participants, their cross section, and the impact parameter in heavy-ion collisions, defining the centrality classes [65].

The initial global parameters such as the collisions energy, impact parameter, collision nuclei species, and polarization can be treated as knobs for high energy nuclear physicists to play with in order to study relativistic heavy-ion collisions and create strongly interacting systems with different sizes, chemical potentials, and temperatures in the QCD phase diagram. Figure 1-17 shows an event display of thousands of tracks from a central Au + Au collision event at 200 GeV recorded by the Time Projection Chamber (TPC) of the STAR experiment at RHIC.

1.4.7 Glauber Model

The Glauber Model, named after physicist Roy Glauber [71], was originally developed to address high energy scattering problems with composite particles in the op-

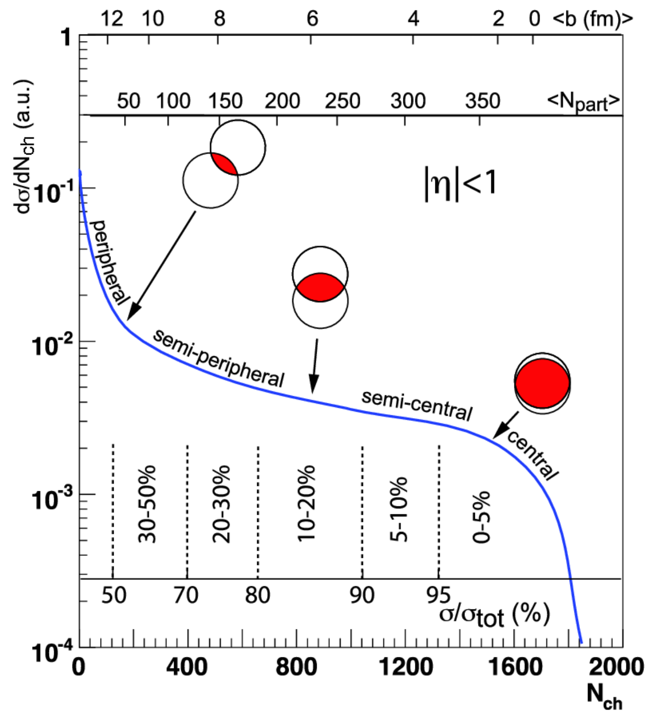


Figure 1-16: The plot showing relationship among number of charged particle, N_{ch}, related to the number of participants N_{part}, the differential cross section $\frac{d\sigma}{dN_{ch}}$, and the centrality, according to the Glauber Model calculations, is shown above.

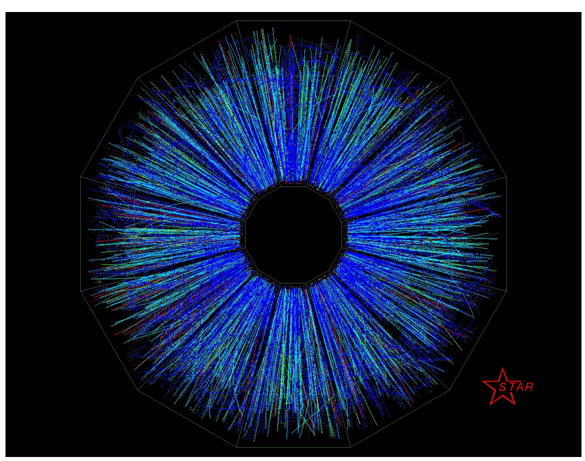


Figure 1-17: Two gold ions collide head-on in the STAR detector. The event with reconstructed tracks of final state particles are display by STAR TPC shown above. Image from [70]

tical limit where optical theorem is applicable [72, 73]. It is a model describing two composite objects collider inelastically with each other and decompose the total cross section to the cross section of collision between two point objects. The Glauber Model can be applied to study nucleon-nucleus (N-A) and nucleus-nucleus (A-B) collisions with nucleon-nucleon (N-N) collisions and determine relationship between the global observables mentioned in the previous subsection.

If we consider a spherically symmetric nucleus, the nuclear charge density can be described $\rho(r)$ by the Fermi distribution with three parameters below

$$\rho(r) = \rho_0 \frac{1 + w(r/R)^2}{1 + \exp(\frac{r-R}{a})} \tag{1.18}$$

The equation above is called the Wood-Saxon density formula. According to the Glauber Model [71], the N-N inelastic cross section is denoted as σ_{in}^{NN} and the effective thickness function of a nucleon is defined as a function of impact parameter in the transverse direction: $T(\vec{b})$. It is defined as follow

$$T(\vec{b}) = \int \rho(\vec{b}, z) dz \tag{1.19}$$

It is normalized to unity: $\int_0^{R_A} T(\vec{b}) d^2 b = 1$. $T(\vec{b})$ essentially depends on density of the nucleus $r(b)$. Therefore, the probability that a nucleon collides with a nucleon inside the nucleus is given by $\sigma_{in} T(\vec{b})$. Therefore, the probability of n nucleon collisions is given by

$$P_n = \binom{A}{n} \sigma_{in}^{NN} T(\vec{b})^n [1 - \sigma_{in} T(\vec{b})]^{A-n} \tag{1.20}$$

Hence, if we consider a constant fraction of μ ($0 \leq \mu \leq 1$) of particle produced after each collisions, we can calculation the average event multiplicity $\langle N(\mu) \rangle$:

$$\langle N(\mu) \rangle = \Sigma_n P_n \Sigma_0^{n-1} \mu^m = \Sigma_{n-1} P_n \frac{1-\mu^n}{1-\mu} = \frac{1}{1-\mu} \{1 - [1 - (1-\mu)\sigma_{in} T(\vec{b})]^A\} \tag{1.21}$$

It turns out that we have the following relationship between N_{part} and N_{coll} with

$\langle N(\mu) \rangle$ [71]

$$N_{part} = \langle N(\mu = 0) \rangle \tag{1.22}$$

$$N_{coll} = \frac{1}{2} \langle N(\mu = 1) \rangle = AT(\vec{b})\sigma_{in}^{NN} \tag{1.23}$$

In a more generalized case: A-B collisions, Figure 1-18 shows side view and beam-line view of heavy-ion collision of projectile B on target A

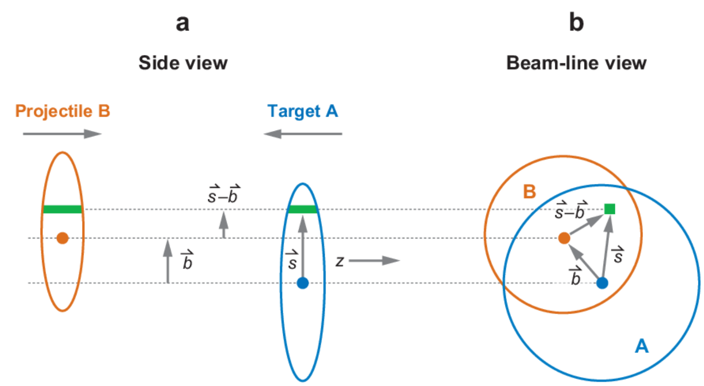


Figure 1-18: The A-B collision with the definition of the impact parameter vector $\vec{b}$ and the distance of nucleon to the center of projectile B $\vec{s}$ are shown above. The distance of the nucleon in B to center of the target A is $\vec{s} - \vec{b}$ according to vector subtraction rule. Here we assume both nuclei A and B are perfect spheres.

Using similar ideas [65], we could first calculate the effective thickness function T_{AB} as follows:

$$T_{AB}(\vec{b}) = \int T_A(\vec{s}) T_B(\vec{b} - \vec{s}) d^2 s \tag{1.24}$$

Now replacing T($\vec{b}$) in N-A by $T_{AB}(\vec{b})$ in A-B, we can obtain

$$\langle N(\mu) \rangle = \frac{A}{1-\mu} \int_0^b T_A(\vec{s}) \{1 - [1 - (1-\mu) T_B(\vec{b} - \vec{s}) \sigma_{in}^{NN}]\}^A d^2 s + \frac{B}{1-\mu} \int_0^b T_B(\vec{s}) \{1 - [1 - (1-\mu) T_A(\vec{b} - \vec{s}) \sigma_{in}^{NN} \tag{1.25}$$

To obtain N_{part}, evaluate at $\mu = 0$, we get

$$N_{part} = A\int_0^b T_A(\vec{s})\{1-[1-T_B(\vec{b}-\vec{s})\sigma_{in}^{NN}]^A\}d^2s + B\int_0^b T_B(\vec{s})\{1-[1-T_A(\vec{b}-\vec{s})\sigma_{in}^{NN}]^B\}d^2s \tag{1.26}$$

To obtain N_{coll}, evaluate at $\mu = 1$, we get

$$N_{coll} = ABT_{AB}(\vec{b})\sigma_{in}^{NN} \tag{1.27}$$

In a very special case, assume the nuclei are simply perfect rigid sphere with the same radius and collide with zero impact parameter is $b = 0$. That is $T_A\sigma_{in}^{NN} = T_B\sigma_{in}^{NN} = T_{AB}\sigma_{in}^{NN} = 1$, we get

$$N_{part} = A + B \tag{1.28}$$

$$N_{coll} = AB \tag{1.29}$$

The results above of N_{part} and N_{coll} agree to our expectation.

The comparison of the Glauber Model with simulations of the N_{part} and N_{coll} as a function of impact parameter b is shown on Figure 1-19 from the [65]

Therefore, we can apply the Glauber model to determine N_{part} and N_{coll} for a given centrality range of AA collision ($T_{AB} \rightarrow T_{AA}$), which will be used in our analysis to obtain the corrected yield. It has been reported that the production of light hadrons, such as pions and kaons, are scaled as N_{part} [74] while electroweak bosons, such as W and Z boson, are scaled as N_{coll} [75].

1.5 Characterization of Quark-Gluon Plasma

Equipped with the knowledge and collider technologies of heavy-ion collisions, we are ready to apply them to conduct scientific research on QGP in laboratories. The following subsections will describe the characterization of QGP from its predicted signatures to open questions today, which leads to my thesis research.

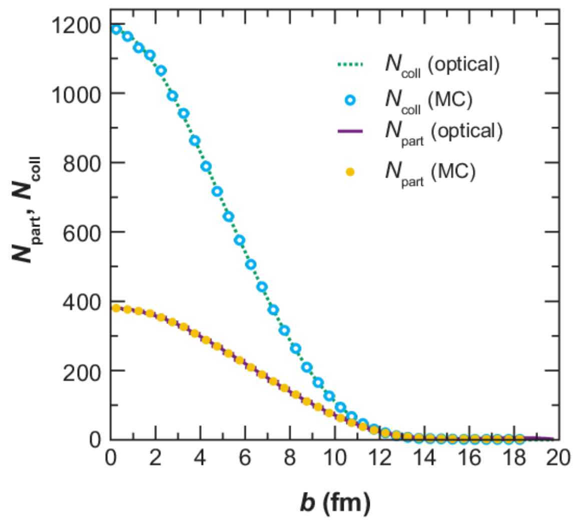


Figure 1-19: The N_{part} and N_{coll} as a function impact parameter calculated from the Glauber Model with optical approximation (lines) and from MC simulations (circles) are shown above. We can see they have almost perfect agreement with each other.

1.5.1 Signatures

QGP has been hypothesize long before its discovery as a color deconfined phase of quark matter named "quark gluon plasma" [76] and will demonstrate some specific benchmarks in experiments to prove the creation of QGP [77]. Here, four classic signatures of QGP will be discussed: J/ψ and Υ suppression, jet quenching, elliptic flow, strangeness enhancement.

1.5.2 J/ψ and Υ suppression

J/ψ meson, as a type of heavy quarkonia, is bound state of $c\bar{c}$, made of charm quark and an anti-charm quark, with mass heavier than the Λ_{QCD}. Therefore, we could approximately treat the interaction between charm and the anti-charm quark with a static the Cornell potential $V(r)$ in the non-relativistic quantum mechanical hamiltonian system [78]:

$$\hat{H} = \hat{T} + \hat{V} \tag{1.30}$$

$$\hat{H}\left|\psi\right\rangle = i\frac{\partial}{\partial t}\left|\psi\right\rangle \tag{1.31}$$

and solve Schrodinger equation the to describe J/ψ mesons in vacuum. As we have seen in Section 1.3.2, with the QGP medium, at a finite temperature T, the potential is modified due to color screening effect. The distance between two charm quarks $V(r) \to \frac{\sigma}{m_D}$, which does not diverge, as $r \to \infty$. Therefore, the $c\bar{c}$ system could be unbounded if they have sufficiently high energy. In the field theory picture, this could be understood as the color string breaking between charm and anti-charm quark [80], also known as quarkonia melting [81]. Hence, with the influence of QGP at $T > 0$, the production cross section of J/ψ will decrease compared to the vacuum at $T = 0$. Experimentally, we define an observable to quantify the modification of particle production cross section in AA collision compared to the reference pp collisions normalized by the number of binary nucleon-nucleon collisions N_{coll}, which is defined in the previous subsection. We called this observable as nuclear modification factor denoted R_{AA}. Mathematically, R_{AA} is defined as follows:

$$R_{AA} = \frac{1}{N_{coll}}\frac{\frac{d^2N_{AA}}{dp_T dy}}{\frac{d^2N_{pp}}{dp_T dy}} = \frac{1}{T_{AA}}\frac{\frac{d^2N_{AA}}{dp_T dy}}{\frac{d^2\sigma_{pp}}{dp_T dy}} \tag{1.32}$$

Therefore, $R_{AA} < 1$ means suppression. $R_{AA} = 1$ means no modification. $R_{AA} > 1$ means enhancement. Hence, in experiments, we expect to observe the $R_{AA} < 1$ a suppression of J/ψ production. Figure 1-20 shows the measurement of fully reconstructed J/ψ at RHIC and LHC [82]

In fact, we could see that $R_{AA} < 1$ for every data point, which indicates a clear suppression of J/ψ production from experiments at both RHIC and the LHC. However, we should note that the larger J/ψ R_{AA} observed at the LHC compared to RHIC could be explained by regeneration mechanism [83].

Similarly, we expect to see this in Υ, which is made of $b\bar{b}$. Indeed, they expect to have sequential suppression since 3 Υ states: $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$, could be observed in experiments. Because the total energy of the $b\bar{b}$ system or equivalently the rest mass: $m_{\Upsilon(3S)} > m_{\Upsilon(2S)} > m_{\Upsilon(3S)}$, a sequential suppression: $R_{AA}^{\Upsilon(1S)} > R_{AA}^{\Upsilon(2S)} >$

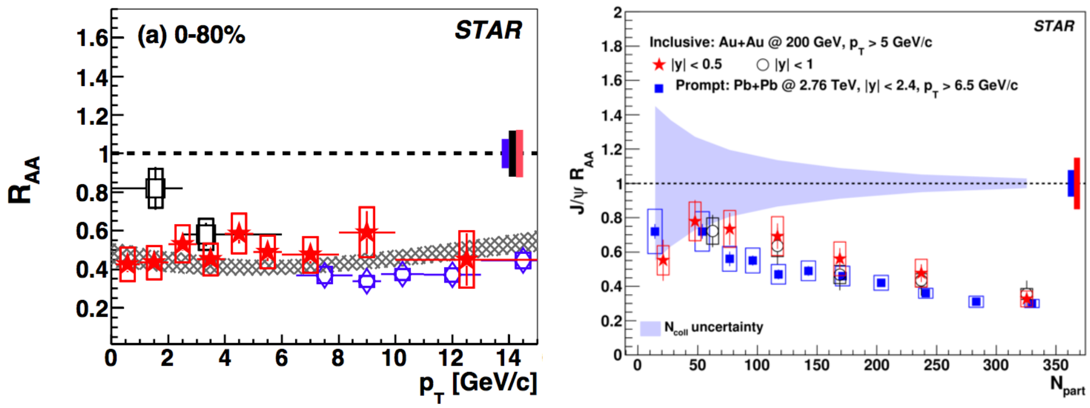


Figure 1-20: The nuclear modifications factor R_{AA} of fully reconstructed J/ψ as a function of p_T (left) and N_{part} (right) measured by the STAR experiment (red data points) at RHIC and CMS (blue diamond data points) and the ALICE (blue circle data points) experiment at the LHC are shown above. We can see that the J/ψ R_{AA} is below 1 for both p_T and N_{part}. There is no significant p_T dependence of J/ψ R_{AA}. The J/ψ R_{AA} decreases as N_{part} increases, consistent to the increasing creation probability of QGP with larger N_{part}.

$R_{AA}^{\Upsilon(3S)}$ should be observed if QGP is created. Figure 1-21 shows the measurements of fully reconstructed Υ states at RHIC and LHC [84, 85]

1.5.3 Jet Quenching

Experimentally, due to color confinement, it is impossible to directly detect and track the energetic partons. Therefore, physicists define jet as a spray of collimated hadrons within a narrow cone initiated from color charged partons. In nuclear and particle physics, jets are used to study the dynamics of partons before hadronization [86] and understand the properties of QGP. A schematic view of a di-jet production from di-qurark event in electron-positron collider $e^+e^- \to q\bar{q}$ is shown below in Figure 1-22

Since we know QGP a color deconfined state of matter, an energetic parton carrying color charge traveling through the QGP medium is expected to lose a substantial a mount its energy to the medium. This is similar the effect that an electron beam losing energy in the electron-ion plasma via electromagnetic interaction [87]. We call this effect as jet quenching. Figure 1-23 shows jet quenching in QGP in AA collisions compared to pp collisions

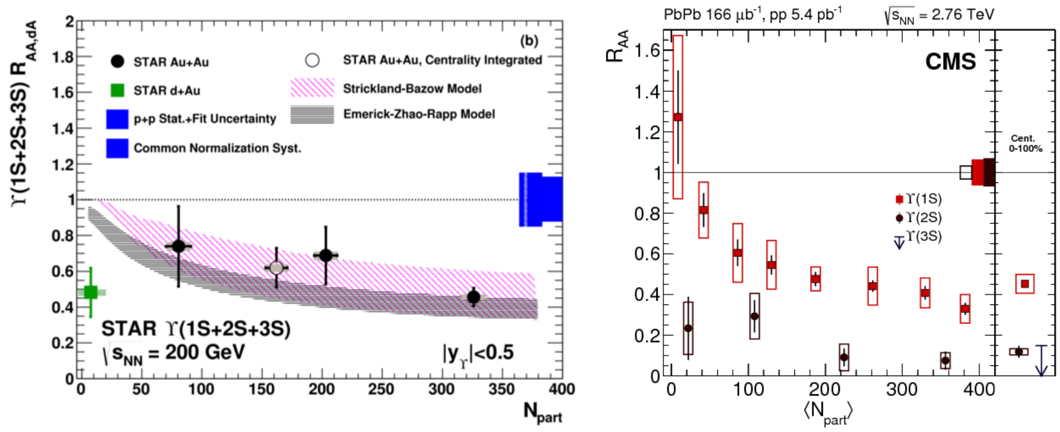


Figure 1-21: The nuclear modifications factor R_{AA} of fully reconstructed Υ as a function of N_{part} measured by the STAR experiment (left) at RHIC and CMS experiment (right) at the LHC are shown above. We can see that the R_{AA} of the three Υ states are below 1 when $N_{part} > 3$. The Υ R_{AA} decreases as N_{part} increases, consistent to the increasing creation probability of QGP with larger N_{part}. In addition, a sequential suppression of Υ R_{AA} is observed by the CMS experiment: $R_{AA}^{\Upsilon(1S)} > R_{AA}^{\Upsilon(2S)} > R_{AA}^{\Upsilon(3S)}$, which agrees with the expectation QGP color screening effect.

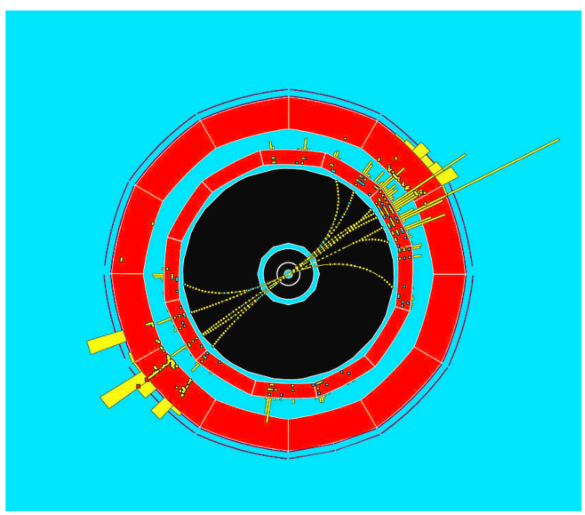


Figure 1-22: The schematic display of a di-jet event from the ALEPH (a particle detector at the Large Electron-Positron collider) Experiment at the Large Electron-Positron Collider (LEP) is shown above. We can see two sprays of back to back particles within narrow cone, representing a di-jet event.

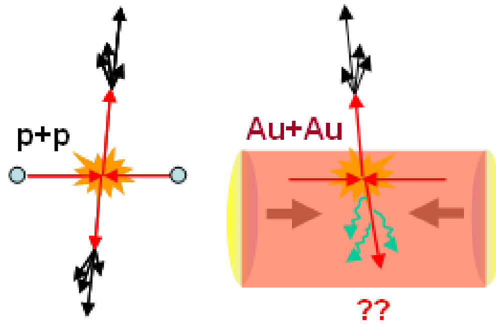


Figure 1-23: The schematic picture explaining jet quenching is shown above. Hard scatterings in pp collisions produce back-to-back "jets" of particles, but in Au + Au collisions, the presence QGP modifies the jets' properties.

Experimentally, compared to pp collision where the QGP is not expected to be created, the jet spectra is modified by the QGP medium. The angular distributions would be broaden due to interaction with the medium. The p_T spectra will be shifted to the left due to energy loss. This can be quantified by jet nuclear modification factor R_{AA} similar to the R_AA for quarkonium suppression mentioned previously. Figure 1-24 shows the hadron angular correlation with the STAR experiment at RHIC and jet R_{AA} as a function of p_T with the ALICE experiments at LHC [88, 89]:

The jet R_{AA} are all below 1 at RHIC and LHC [89,90], which suggest jet quenching in AA collisions, supporting existence of QGP.

1.5.4 Elliptic Flow

The reaction region in heavy-ion collisions, where the two nuclei overlap with each other, has an almond shape, which is azimuthally asymmetric. If the color deconfined matter QGP is created, particles emitted from the almond shape fire ball are expected to be azimuthally anisotropic due to differences of the pressure gradient of the QGP in the and their path length through QGP in the x and y directions. Experimentally, physicists Dr. Arthur Poskanzer (who sadly just passed away on June 30 2021) and Dr. Sergey Voloshin developed the event plane method to analyze the azimuthal anisotropy of particle emission in heavy-ion collisions [93]. The reaction plane is defined as the plane of the impact parameter and the x-axis. Figure 1-25 schematically

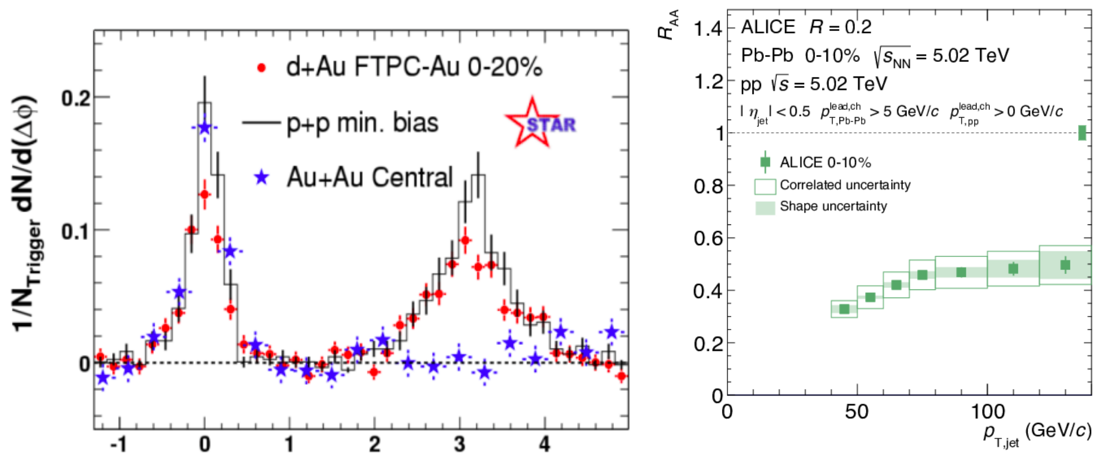


Figure 1-24: The comparison of two-particle azimuthal distributions for central d + Au collisions to those seen in pp and central Au + Au collisions measured with the STAR experiment and the jet R_{AA} as a function p_T measured by the ALICE experiment at LHC (right). From the STAR result, in central Au + Au collisions, the back-to-back peak has disappeared due to the redistribution of jet energy to the slow expanding medium constituents. The jet R_{AA} from ALICE measurement is clearly below 1, suggesting that jets lose significant fractions of energy in AA collision compared to pp.

shows the definition of reaction plane in heavy-ion collisions.

The particle spectra in heavy-ion collisions can be factorized as

$$E\frac{d^3N}{d^3p} = E\frac{1}{2\pi p_T}\frac{d^3N}{dp_T dy d\phi} = E\frac{1}{2\pi p_T}\frac{d^2N_1}{dp_T dy}\frac{dN_2}{d\phi} \tag{1.33}$$

Since the particle emission is azimuthally anisotropic, we can expand the $F(p_T, \phi, y) = \frac{dN_2}{d\phi}$ into a Fourier series [93]:

$$F(p_T, \phi, y) = \frac{x_0(p_T, y)}{2\pi} + \sum_{n=1}^{\infty}[x_n(p_T, y)\cos(n\phi) + y_n(p_T, y)\sin(n\phi)] \tag{1.34}$$

According to trigonometry, we get

$$F(p_T, \phi, y) = \frac{x_0(p_T, y)}{2\pi} + \sum_{n=1}^{\infty} 2v_n(p_T, y)\cos[n(\phi - \Psi_n)] \tag{1.35}$$

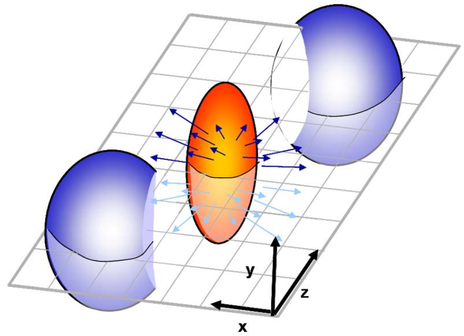


Figure 1-25: The figure above shows the ellipsoid of the overlapping nuclear reaction region of two nuclei in heavy-ion collisions. The reaction plane, which is the x-z plane shown above, is constructed by the beam direction and the impact parameter vector. The emissions of particles are azimuthally anisotropic in the x-y plane due to the geometry.

Here, $v_n = \frac{1}{2}\sqrt{x_n^2 + y_n^2}$ and $\Psi_n = \frac{1}{n}\arctan(\frac{y_n}{x_n})$.

To find the Fourier coefficients v_n, we can apply the Fourier tricks to find x_n and y_n.

Theoretically, because the function $\frac{dN_2(\phi)}{d\phi}$ is continuously analytical, we can use integral to find the Fourier coefficients [18]

$$x_n = 2\int_0^{2\pi} \frac{dN_2(\phi)}{d\phi}\cos(n\phi)d\phi \tag{1.36}$$

$$y_n = 2\int_0^{2\pi} \frac{dN_2(\phi)}{d\phi}\sin(n\phi)d\phi \tag{1.37}$$

Experimentally, because our data take on discrete values, we can convert the integral into a sum

$$x_n = \frac{2}{N}\sum_{i=1}^{N}\cos(n\phi^i) = 2\langle\cos n\phi\rangle \tag{1.38}$$

$$y_n = \frac{2}{N}\sum_{i=1}^{N}\sin(n\phi^i) = 2\langle\sin n\phi\rangle \tag{1.39}$$

Here, we sum up all tracks in the experiment to get the x_n and y_n. Then, we will be able to find

$$v_n = \frac{1}{2}\sqrt{x_n^2 + y_n^2} = \sqrt{(\langle\cos n\phi\rangle)^2 + (\langle\sin n\phi\rangle)^2}. \tag{1.40}$$

In heavy-ion physics, the first order Fourier coefficient v_1 is called the directed flow.

$$v_1 = \sqrt{(\langle\cos\phi\rangle)^2 + (\langle\sin\phi\rangle)^2}. \tag{1.41}$$

It can be connected to the initial tilting source of the colliding nuclei [91] and can be used to study Chiral Magnetic Effect [92].

The second order Fourier coefficient v_2 is called elliptic flow.

$$v_2 = \sqrt{(\langle\cos 2\phi\rangle)^2 + (\langle\sin 2\phi\rangle)^2} = \sqrt{(\langle\cos^2\phi\rangle - \langle\sin^2\phi\rangle)^2 + (2\langle\sin\phi\rangle\langle\cos\phi\rangle)^2}. \tag{1.42}$$

Assuming in initial stage before the collision, the sum of the momentum of two colliding nuclei $\vec{p_1}$ and $\vec{p_2}$ is exactly 0 without any fluctuation. That is

$$\vec{p_1} + \vec{p_2} = 0 \tag{1.43}$$

According to momentum conservation, for the final state particles, we have

$$\sum_i^N p_x^i = 0 \tag{1.44}$$

$$\sum_i^N p_y^i = 0 \tag{1.45}$$

Therefore, we have

$$\langle p_T\cos\phi\rangle = \langle p_x\rangle = \frac{1}{N}\sum_i^N p_x^i = 0 \tag{1.46}$$

$$\langle p_T\sin\phi\rangle = \langle p_y\rangle = \frac{1}{N}\sum_i^N p_y^i = 0 \tag{1.47}$$

But since the p_T and ϕ are completely orthogonal, the random variable p_T is uncorrected to ϕ. Therefore, we have

$$\langle p_T \cos\phi \rangle = \langle p_T \rangle \langle \cos\phi \rangle = 0 \tag{1.48}$$

$$\langle p_T \sin\phi \rangle = \langle p_T \rangle \langle \sin\phi \rangle = 0 \tag{1.49}$$

Finally, we know that $p_T > 0$, thus

$$\langle p_T \rangle > 0 \tag{1.50}$$

Hence,

$$\langle \cos\phi \rangle = 0 \tag{1.51}$$

$$\langle \sin\phi \rangle = 0 \tag{1.52}$$

Therefore, we have

$$v_2 = \sqrt{(\langle \cos^2\phi \rangle - \langle \sin^2\phi \rangle)^2 + (2\langle \sin\phi \rangle \langle \cos\phi \rangle)^2} = \langle \cos^2\phi \rangle - \langle \sin^2\phi \rangle. \tag{1.53}$$

In terms of momentum p_x and p_y, we can rewrite v_2 as

$$v_2 = \langle \cos^2\phi \rangle - \langle \sin^2\phi \rangle = \langle \frac{p_x^2}{p_T^2} \rangle - \langle \frac{p_y^2}{p_T^2} \rangle = \langle \frac{p_x^2 - p_y^2}{p_T^2} \rangle = \langle \frac{p_x^2 - p_y^2}{p_x^2 + p_y^2} \rangle. \tag{1.54}$$

Classically, we know that the momentum is proportional to the pressure gradient. Schematically, we could write

$$p_x \simeq \frac{m\tau}{\rho} \frac{\partial P}{\partial x} \simeq \frac{m\tau}{\rho} \frac{P}{L_x} \tag{1.55}$$

Where m is the mass of the particle, τ is the life time of the QGP, ρ is the density of the QGP, and L_x is the minor axis of the ellipse in the x direction according to the geometry of Figure 1-25.

Likewise, we have the same relation for p_y

$$p_y \simeq \frac{m\tau}{\rho}\frac{\partial P}{\partial y} \simeq \frac{m\tau}{\rho}\frac{P}{L_y} \tag{1.56}$$

Here, L_y is the major axis of the ellipse in the y direction according to the geometry of Figure 1-25. Apparently, $L_y > L_x$.

Hence, we can write v_2 as

$$v_2 = \langle\frac{p_x^2 - p_y^2}{p_x^2 + p_y^2}\rangle = \frac{\frac{1}{L_x^2} - \frac{1}{L_y^2}}{\frac{1}{L_x^2} + \frac{1}{L_y^2}} = \frac{L_y^2 - L_x^2}{L_x^2 + L_y^2} > 0 \tag{1.57}$$

In heavy ion collision, we define the eccentricity ϵ_s of an ellipse is defined as [94]

$$\epsilon_s \equiv \frac{L_y^2 - L_x^2}{L_x^2 + L_y^2} \tag{1.58}$$

Hence, we have

$$v_2 \simeq \epsilon_s \tag{1.59}$$

Therefore, we can see that v_2 is essentially proportional to the eccentricity simply based on ellipse geometry of reaction region. Historically, v_2 has extensively studied experimentally and theoretically. It turns out light hadrons demonstrate collectivity. Their elliptic flow v_2 could be calculated using relativistic viscous hydrodynamics [95]. If QGP is created, we expect v_2 of the light flavor hadrons to be positive as we derive above. Figure 1-26 show the v_2 as a function of p_T of charged light flavor hadrons in heavy-ion collisions at mid-rapidity measured by RHIC and LHC experiment [96,97]

We can clearly see positive v_2 of charged particles at both RHIC and LHC, which also supports the creation of QGP in high energy heavy-ion collisions.

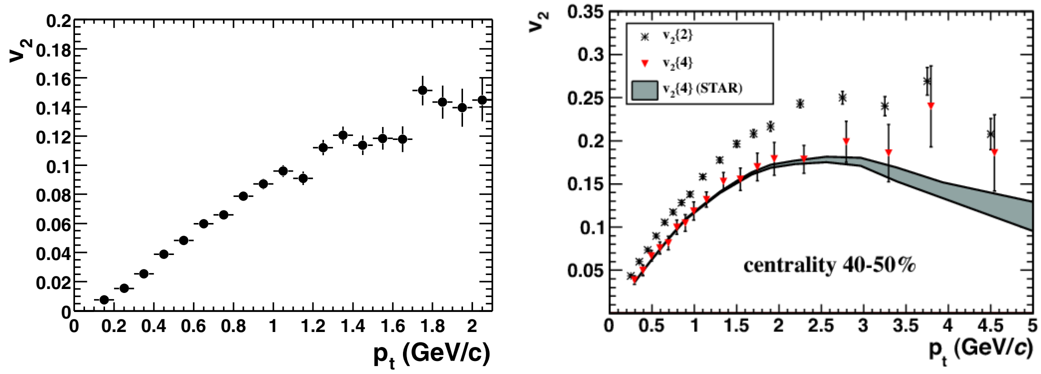


Figure 1-26: The elliptic flow v_2 of charged particles as a function of p_T in Au + Au collision measured by the STAR experiments at RHIC (left) and in PbPb collisions by the ALICE experiments at LHC (right) are shown above. Clearly, $v_2 > 0$ is observed in both experiments.

1.5.5 Strangeness Enhancement

As described in Section 1.4.6, the temperature of QGP is well above 100 MeV, which is much larger than the strange quark mass (about 95 MeV). Therefore, since $T_{QGP} > m_s$, in the thermally and chemically equilibrated QGP, strange quarks could be produced thermally via the pair production processes: $u\bar{u} \to s\bar{s}, d\bar{d} \to s\bar{s}$, and $gg \to s\bar{s}$, establishing the chemical abundance equilibrium [99]. Therefore, the strangeness content in the QGP is enhanced, which could be experimentally observed from enhancement of strange particle yields in AA collisions compared to pp collisions. A direct experimental observable is the ratio of strange hadron yield to pions in AA and pp collisions scaled by N_{part}. Figure 1-27 and 1-28 show the measurements on strange mesons and baryons to pion ratios in AA and pp collisions at RHIC and LHC

We can see that ϕ/π ratio increases as N_{part} and $\sqrt{s_{NN}}$ increases, which indicates strangeness enhancement in AA collisions compared to pp collisions. This again could be served as an evidence for the formation of QGP in heavy-ion collisions at RHIC and LHC.

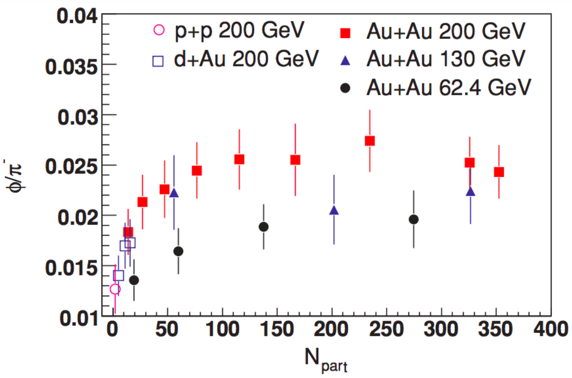


Figure 1-27: The yield ratio of ϕ/π as a function N_{part} in p + p, p + Au, and Au + Au from the STAR experiment at RHIC are shown above.

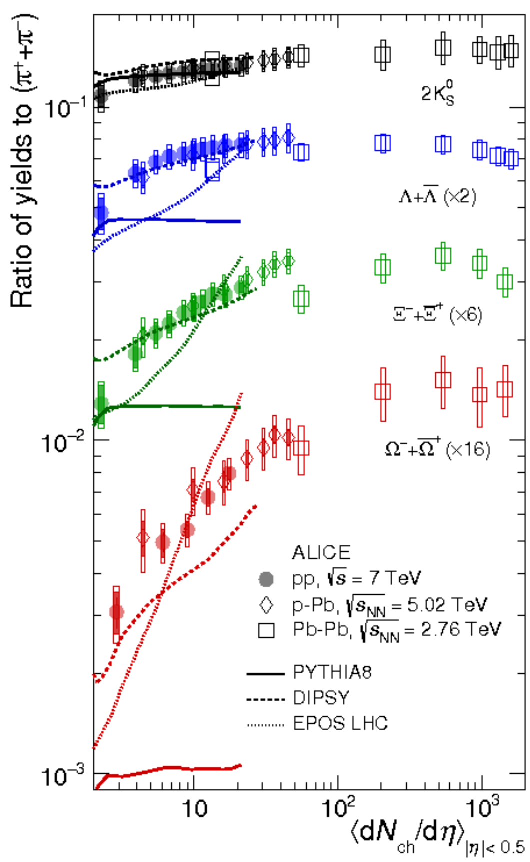


Figure 1-28: The yield ratio of strange hadrons $K_s^0, \Lambda^+, \Xi^0, \Omega^-$ as a function of $\langle dN_{ch}/d\eta \rangle$ from the ALICE experiment at LHC are shown above.

1.5.6 Macroscopic Properties

Physicists have conducted extensive studies to understand macroscopic properties of QGP. Below are some of the interesting properties of QGP observed in experiments:

Transient Lifetime: According to the experimental results at RHIC and LHC, QGP has a very short lifetime. It is on the order of 10 fm/c [108]. It is generally assumed that QGP reaches thermal [109] and near chemical equilibrium [110] via the strong interaction. So far, there is not sufficient experimental evidences to directly support this assumption.

Strong Interacting System: Moreover, QGP, as a deconfined state of matter, demonstrates a strongly interacting behavior, which contradicts to the prediction weak coupling according to the asymptotic freedom of quarks and gluons in QCD [5]. At $T \sim 1-3\ T_c$, the coupling strength of QGP is still strong: $g_s \sim O(1)$ [111]. Therefore, strong interaction between the QGP constituents is in general non-perturbative. The equation of state of strong interacting QGP, as an input for hydrodynamic calculations, can be calculated by models such as MIT Bag Model [41] or Lattice QCD [112].

Perfect Liquid Behavior: Finally, QGP demonstrates nearly perfect liquid properties. The expansion of QGP in the fireball stage is approximately isentropic and could be well described by hydrodynamics [115]. More specially, due to the relativistic nature of the strongly coupled near-perfect liquid system, assuming QGP reaches thermal [109] and near chemical equilibrium [110], relativistic viscous hydrodynamics [95] is the correct theoretical formalism to describe the dynamics of QGP. As an nearly perfect liquid, the shear viscosity to entropy density ratio of the QGP is very small: $\frac{\eta}{s} \sim (1-2.5)\frac{1}{4\pi}$ [113], approaching the quantum limit $\frac{\eta}{s} = \frac{1}{4\pi}$ predicted by the strongly coupled N =4 supersymmetric Yang-Mills plasma in Anti-de-Sitter Space/Conform Field Theory (AdS/CFT) correspondence [114].

Color Opaque Plasma: It is also interesting that QGP is a color opaque plasma [116]. This means that gluons propagating through the QGP will be absorbed by the plasma medium. Experimentally, the suppression of hadrons is a measure of the color opacity of the QGP [116]. Physicists have found that QGP is indeed highly color opaque [117].

1.5.7 Open Questions

However, although extensive studies have been carried out for many years, there are still many open questions, most of which are derived from the mysterious macroscopic behavior of QGP. Below is the list of selected open questions that are currently under active investigation by the heavy-ion physics community [118]:

1) **Thermalization of QGP:** How can QGP reach thermal equilibrium within such a short time, which is on the order 1 fm/c, from the non-equilibrium stage?

2) **Inner Workings of QGP:** What is the correct degree of freedom to describe the QGP? The inner workings of QGP, as a deconfined phase of matter, must lay between asymptotically free quarks and gluons and color neutral hadrons. That is also why the sPHENIX experiment at RHIC, as the next generation DOE flagship Heavy Ion Physics program in the U.S., is going to built at BNL and collect date to probe the inner workings of QGP by resolving its properties at shorter and shorter length scales.

3) **Smallest Droplet of QGP:** What is the smallest droplet of QGP that can be created? Can QGP be created in pPb, pp, or even e^+e^- collision systems? What are the limits of the applicability of hydrodynamics?

1.6 Heavy Flavor Physics

1.6.1 Open Heavy Flavor Physics

My graduate research focuses on answering the second question through the data analysis of fully reconstructed heavy flavor hadrons with the CMS experiment to understand transport properties and probe the microscopic structure of QGP. In this section, we will focus on discussing open heavy flavor physics where the only one heavy quark Q is in hadron. Open heavy flavor hadrons have ± 1 heavy flavor number. Quarkonia states $Q\bar{Q}$ are considered as hidden heavy flavor with a zero net heavy flavor quantum number. Their properties are very different from open heavy flavor hadrons. We will not discuss them in this thesis and focus only on open heavy flavor physics.

1.6.2 Heavy Quarks

Heavy quarks, such as charm and beauty quarks, whose masses are on the order of GeV, lie in a scale above the Λ_{QCD} and T_{QGP}. Therefore, they are predominantly produced in early stage of heavy-ion collisions where hard scattering processes occur. Their production could be calculated by perturbation QCD. Figure 1-29 shows the lowest order Feynman diagrams of heavy quark pair production in QCD.

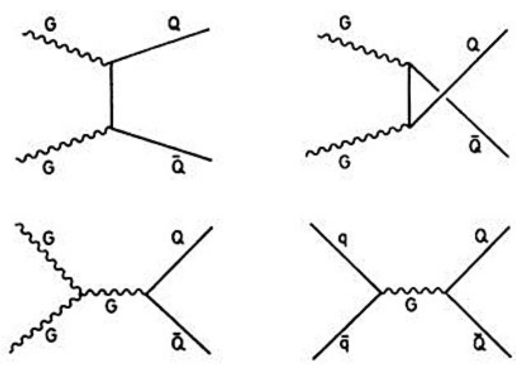


Figure 1-29: The four lowest order tree level Feynman diagrams of heavy quark pair production are shown above.

In general, due to their relatively small momentum transfer to the QGP medium constituents compared to their large masses [136], they should not reach complete thermalization via multiple scattering as they traverse through the QGP. In addition, since their lifetime is much longer than the QGP lifetime, they retain their identities and record the evolution of the QGP, which makes them excellent probes. Then, they travel through the medium, hadronize into heavy flavor hadrons, and decay weakly. Their decay product are detected and identified by particles detectors.

Experimentally, from the final stage decay products, we can fully reconstruct heavy flavor hadrons where the dynamics of heavy quarks is encoded with different transverse momenta to study their diffusion coefficients, hadronizaton mechanism, and energy loss to probe the microscopic structure of QGP via their scattering patterns with the QGP constituents at different wavelengths. Figure 1-30 below shows respectfully an event of beauty quark production and hadronization in vacuum and QGP.

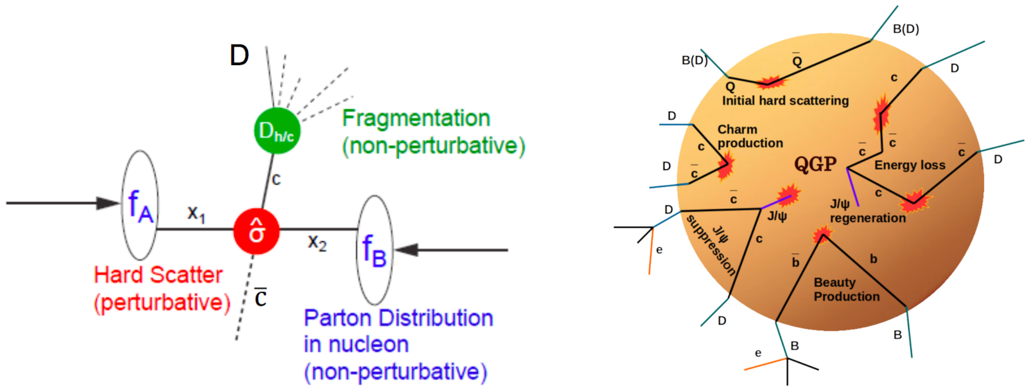


Figure 1-30: The schematic demonstrations of heavy quark production and hadronization in vacuum (left) and QGP (right) are shown above.

1.6.3 Heavy Flavor Physics in Vacuum

To use heavy quark to probe the QGP created in heavy-ion collisions, we first need to understand heavy quark physics in vacuum from pp collisions. In the process $pp \to Q\bar{Q}$, QCD factorization theorem could be applied. High precision pQCD calculations, including next-to-leading order (NLO), next-to-next-to-leading-order (NNLO), and Fixed-to-Next-to-the-Leading (FONLL), have been developed to describe heavy quark production. in pp collisions Here, we will show FONLL calculations of charm and beauty quarks spectra, schematically denoted as: $\frac{d^2\sigma^Q}{p_T dp_T dy}$, in $pp \to Q\bar{Q}$ at different energies [132]. Figure 1-31 shows the FONLL calculations of charm and beauty quarks spectra produced for pp collisions at the LHC energy $\sqrt{s} = 5.02$ TeV.

In vacuum, heavy quarks fragment into heavy flavor hadrons $Q \to H_Q$. We can defined the parton fragmentation function $D_i^{H_Q}(z, \mu^2)$ where is the probability for a quark q with energy E fragment into a hadron with energy zE $(0 < z < 1)$ at the factorization scale of μ^2 [12]. According to pQCD, $D_i^{H_Q}(z, \mu^2)$ is universal in vacuum

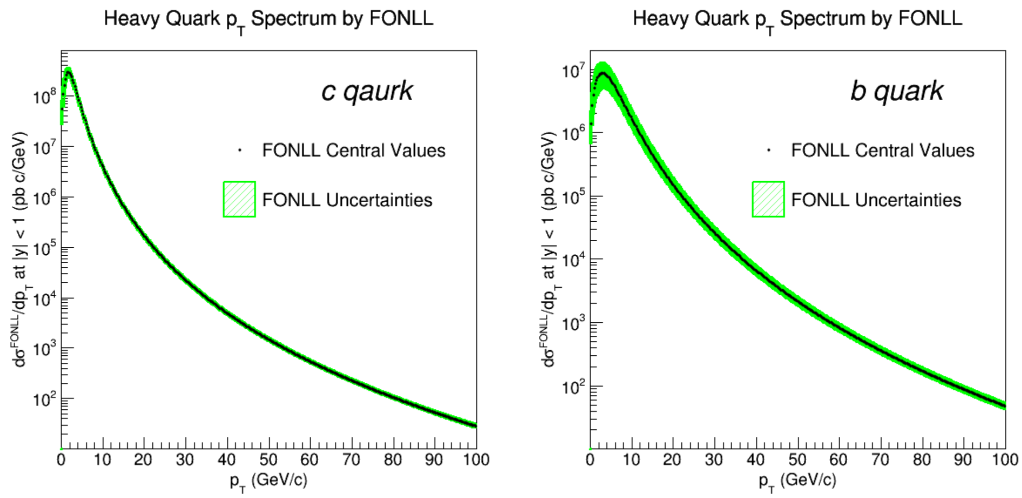


Figure 1-31: The charm quark (left) and beauty quark (right) differential cross section $\frac{d\sigma}{dp_T}$ as a function transverse momentum p_T at $|y| < 1$ from FONLL calculations, are shown above.

from e^+e^-, ep, and pp collisions. Figure 1-32 shows the scattering processes which fragmentation fraction is involved:

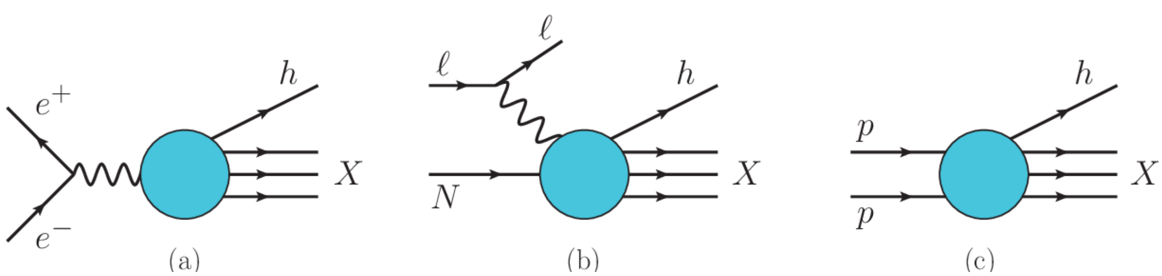


Figure 1-32: Single-inclusive hadron production process, where fragmentation function are involved, in (a) electron-positron annihilation, (b) deep-inelastic lepton-nucleon scattering, (c) proton-proton scattering are shown above.

Next, we are ready to define heavy quark fragmentation fraction $f(Q \to H_Q)$. First we know, the energy

$$E = \sqrt{m^2 + p_T^2 \cosh^2 y} \tag{1.60}$$

Ignoring the mass, we have

$$E \simeq p_T \cosh y \tag{1.61}$$

So energy of the hadron E^h where the quark with E^Q fragments into will be

$$E^h = zE^Q \tag{1.62}$$

So we have the transverse momentum of the hadron p_T^h

$$p_T^{H_Q} = zp_T^Q \tag{1.63}$$

With heavy quark spectra $\frac{d^2\sigma^Q}{p_T dp_T dy}$ and parton fragmentation function $D_i^{H_Q}(z, \mu^2)$, we let

$$\frac{d^2\sigma^Q}{p_T dp_T dy} = F^Q(p_T, y) \tag{1.64}$$

Hence, for a hadron with p_T, the heavy quark will have p_T/z with probability $D_i^{H_Q}(z)$ to fragment into this hadron. Therefore, the heavy flavor hadron spectra is given by:

$$\frac{d^2\sigma^{H_Q}}{p_T dp_T dy} = \int_{x_T}^{1} F^Q(p_T/z, y) D_i^{H_Q}(z, \mu^2) dz \tag{1.65}$$

Here $x_T = \frac{2p_T}{\sqrt{s}}$ [133].

Now if we consider a factorization scale near the heavy quark mass $\mu^2 \to m_Q^2$, according to PDG reference [4], solving the leading evolution equation, the heavy quark fragmentation function $D_Q^{H_Q}(z)$ is in a form of delta function and light quark q and gluons g $(i = g, q)$ will not contribute to produce heavy flavor hadrons. Hence, we could write

$$D_{q,g}^{H_Q}(z, \mu^2)|_{\mu^2 = m_Q^2} = 0 \tag{1.66}$$

$$D_Q^{H_Q}(z, \mu^2)|_{\mu^2 = m_Q^2} = f(Q \to H_Q)\delta(1 - z) \tag{1.67}$$

Here $f(Q \to H_Q)$ is the heavy quark fragmentation fraction and stands for the probability of a heavy quark Q hadronize into an open heavy flavor hadron H_Q. Indeed, according to the momentum sum rule that constrains the parton fragmentation function [12]

$$\sum_{H_Q} \int_0^1 z D_Q^{H_Q}(z, \mu^2) dz = 1 \tag{1.68}$$

$$\sum_{H_Q} \int_0^1 z f(Q \to H_Q) \delta(1-z) dz = 1 \tag{1.69}$$

$$\sum_{H_Q} f(Q \to H_Q) = 1 \tag{1.70}$$

This verifies that the sum of heavy quark fragmentation fraction over all heavy flavor hadrons is equal to unity. Next, we have

$$\frac{d^2\sigma^{H_Q}}{p_T dp_T dy} = \int_{x_T}^1 F^Q(p_T/z, y) D_i^{H_Q}(z, \mu^2) dz = \int_{x_T}^1 F^Q(p_T/z, y) D_Q^{H_Q}(z, m_Q^2) dz \tag{1.71}$$

Thus,

$$\frac{d^2\sigma^{H_Q}}{p_T dp_T dy} = \int_{x_T}^1 F^Q(p_T/z, y) f(Q \to H_Q) \delta(1-z) dz = f(Q \to H_Q) F^Q(p_T, y) \tag{1.72}$$

Hence, we have

$$\frac{d^2\sigma^{H_Q}}{p_T dp_T dy} = f(Q \to H_Q) \frac{d^2\sigma^Q}{p_T dp_T dy} \tag{1.73}$$

This means that the open heavy flavor hadron spectra $\frac{d^2\sigma^{H_Q}}{p_T dp_T dy}$ is essentially proportional to the heavy quark spectra $\frac{d^2\sigma^Q}{p_T dp_T dy}$ with heavy quark fragmentation fraction $f(Q \to H_Q)$ the as the coefficient of proportionality. Experimentally, charm and beauty fragmentation fractions have been measured at LEP, HERA, and the LHC and documented in PDG [4]. The fragmentation fraction is often treated roughly a constant independent to p_T, y, and $\sqrt{s}$ and is assumed to be universal in e^+e^-, ep, and pp collisions systems [4].

In terms of being a constant, according LHCb pp results [134], it appears that the fragmentation fraction has significant $\sqrt{s}$ and p_T dependence while no significant y_B (or η_B) dependence is observed. Figure 1-33 shows the beauty quark fragmentation fraction: $f_u = f(b \to B^+)$, $f_d = f(b \to B^0)$, and $f_s = f(b \to B_s^0)$

In terms of universality, according to Strangeness Quark Matter Conference (SQM) in 2021, a hadronization universality breaking is observed from the ALICE experiment at the LHC [135]. Figure 1-34 shows the hadronization universality breaking reported by the ALICE experiment in SQM 2021

Further investigations of these results are currently ongoing. However, we will not expand the discussions here. Now, equipped with the understanding of heavy flavor physics in vacuum from pp collisions as a reference, we are ready to use heavy quarks to study QGP created in heavy-ion collisions.

1.6.4 Heavy Quark Diffusion

In the limit of low p_T or equivalently long wavelength, for heavy quarks inside the QGP medium, the elastic collision cross section dominates. In elastic $Qq \to Qq$ process in the thermally equilibrated QGP medium, heavy quarks has the relatively small momentum transfers on the order of the temperature compared to the heavy quark mass: $m_Q > |k| \simeq T$. Considering the mean free time of HQ in the QGP medium is about $\tau \sim 0.44 fm/c$ [137]. Therefore, the number of scattering of heavy quarks in the QGP medium is about $n \sim \frac{\tau_{QGP}}{\tau_{HQ}} \simeq 23 \sim O(10)$.

Now, we can consider a simple binomial process to model the diffusion of heavy quark in the QGP medium. Assuming the momentum of the heavy quark at $t = 0$ is p, after the time τ_{HQ}, one scattering happens. The momentum of the heavy quark at $t = \tau_{HQ}$ either $p + k$ or $p - k$. Each has 1/2 probability. Next, after another τ_{HQ}, another scattering happens. The momentum of the heavy quark at $t = 2\tau_{HQ}$ either $p + 2k$, p or $p - 2k$ with 1/4, 1/2, and 1/2 probability respectfully. Therefore, the

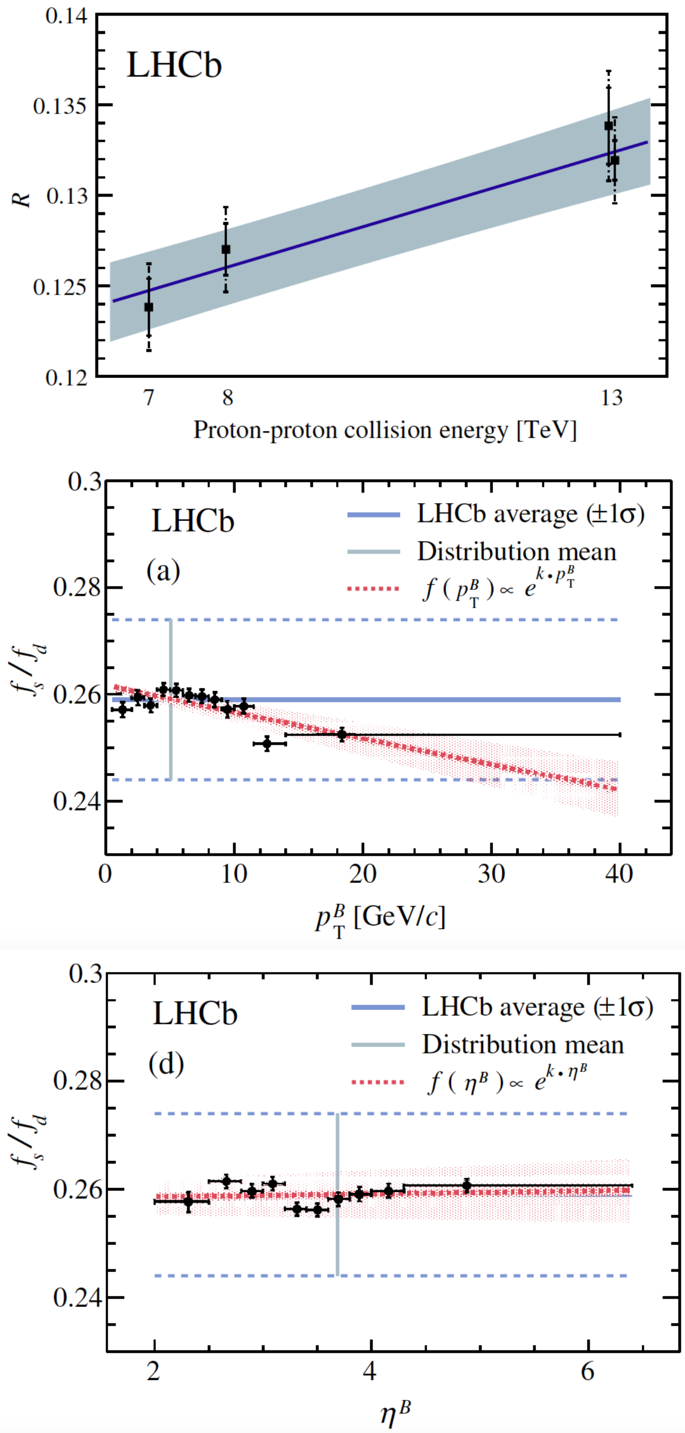


Figure 1-33: R, the corrected yield ratio of B_s^0/B^+, as a function the pp collision energy $\sqrt{s}$ (top), the f_s/f_d ratio as a function p_T (middle), and the f_s/f_d ratio as a function η_B (bottom), from the LHCb experiment are shown above.

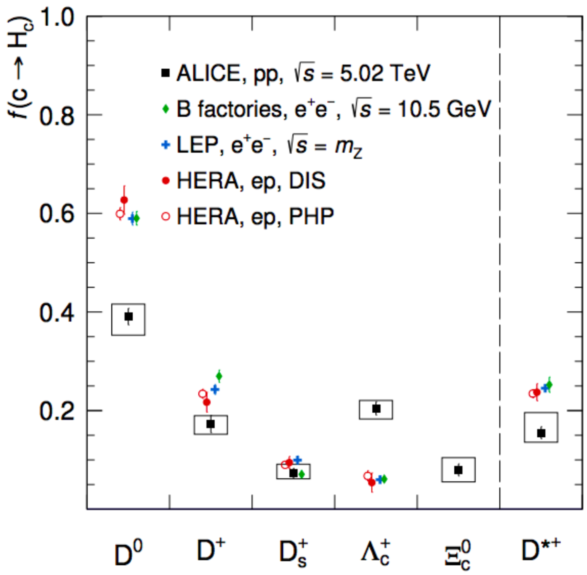


Figure 1-34: The charm quark fragmentation fraction to different charm hadrons species in e^+e^-, ep, and pp collisions are presented above. From the ALICE experiment, we can clearly see that the fragmentation fraction of D^0 has drop by about 40% while the Λ_c^+ has enhanced by about a factor of 4. Therefore, the hadronization universality is clear broken at the LHC energy in the charm sector.

standard deviation of binomial process $\sigma_p = \frac{\sqrt{n}}{2}k$. If we take $n = 25$, $\sigma_p = 2.5k \simeq 2.5T_{QGP} = 0.4$ GeV. Experimentally, we consider a heavy quark with momentum about $p > 1.5$ GeV/c $>> \sigma_p$. According to Figure 1-31, we could see that the heavy quark transverse momentum is well above 1 GeV/c. Hence, such heavy quarks still retain a lot of memory about its initial conditions even after multiple small scattering with QGP medium. Hence, in these conditions, heavy quark undergoes Brownian-like motion in the QGP medium [136]. Their motion in the QGP medium could be characterized by the Planck-Fokker Equation, which could be schematically written as follows [138]:

$$\frac{\partial}{\partial t} f_q(t, \vec{p}) = \frac{\partial}{\partial p_i}\{A_i(\vec{p}) f_q(t, \vec{p}) + \frac{\partial}{\partial p_j}[B_{ij}(\vec{p}) f_q(t, \vec{p})]\} \tag{1.74}$$

Here, $f_q(t, \vec{p})$ is the heavy quark phase space distribution function. If we ignore the modification of the cold nuclear matter effect on the heavy quark initial production spectra, then in heavy-ion collision:

$$F^Q(t = 0, p_T) \propto \frac{d\sigma_{FONLL}}{p_T dp_T} \tag{1.75}$$

The transport parameters $A_i(\vec{p})$ is related to the thermal relaxation rate and $B_{ij}(\vec{p})$ is related to the momentum diffusion of heavy quarks [136]. The heavy quark special diffusion coefficient D_s is related to the transport parameter as follows:

$$D_s = \frac{T}{m_Q A(p = 0)} \tag{1.76}$$

D_s characters the fundamental property of the QGP $\frac{\eta}{s}$ via the relationship

$$2\pi T D_s \simeq \frac{\eta}{s} \tag{1.77}$$

More detailed studies have been carried out to examine heavy quark coupling strength and quantify the information heavy quarks carry as they traverse through the QGP medium [139].

1.6.5 Heavy Quark Energy Loss

In the limit of high p_T or equivalently short wavelength, inelastic cross section starts to dominate [136]. Heavy quarks lose a substantial amount of energy as they travel fast through the QGP medium [140]. In a simplified schematization, there are two different pictures that describe the energy loss mechanism of heavy quark in the QGP medium. In the pQCD picture, the coupling of the constituents of the QGP is assumed to be weak. Therefore, the QGP is made of weakly coupled quasiparticles. Heavy quarks scatter off the constituents incoherently when propagating through the QGP medium. There are two energy loss mechanisms: collisional energy loss and radiative energy loss [138]. The collisional energy loss is given by $-\frac{dE}{dx} = \kappa_{coll} T^2$ and the radiative energy loss is given by $-\frac{dE}{dx} = \kappa_{rad} T^3 x$ [141, 142]. Figure 1-35 shows schematically heavy quark energy loss mechanism in the QGP medium

The other picture, AdS/CFT, takes the strong coupling limit. In this picture, QGP behave like liquid and heavy quarks scatter off the constituents coherently in

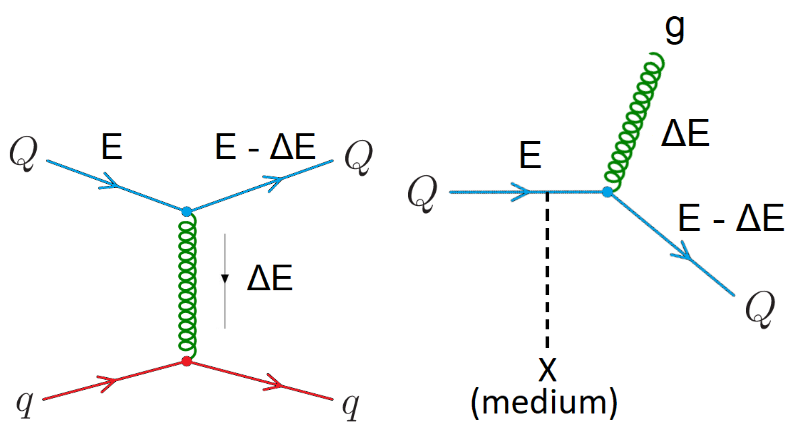


Figure 1-35: The schematic demonstration of the pQCD picture: collisional energy loss (left) and radiative energy loss (right) of heavy quarks in the QGP medium are shown above.

the QGP medium. The AdS/CFT model applies holographic drag force [143] to calculate the energy loss of heavy quark [144] in the QGP medium

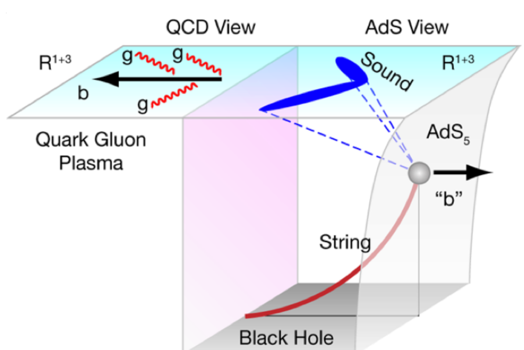


Figure 1-36: The schematic demonstration of ADS/CFT picture: a quark lose energy in the QGP medium holographically due ADS/CFT drag force.

In the pQCD picture, as $p_T \to \infty$, similar to electron Bremsstrahlung via QED radiation in the matter [145], for a heavy quark traveling through the QGP medium, its radiative energy loss via soft gluon radiation will dominate. The soft gluon radiation spectrum by a parton in the QGP medium is given by [146]

$$dP = \frac{\alpha_S C_F}{\pi} \frac{d\omega}{\omega} \frac{k_\perp^2 dk_\perp^2}{(k_\perp^2 + \omega^2\theta_0^2)^2} \tag{1.78}$$

Where

$$\theta_0 \equiv \frac{m}{E} \tag{1.79}$$

Here, ω is the energy of the gluon and $k_\perp$ is the transverse momentum of the gluon, C_F is color factor (Casimir) which is 3 for gluons with one color and one anti-color charges and 4/3 for quarks with one color charge. From Equation 1.78 above, a suppression of radiation at a small angle $0 < \theta < \theta_0$ is expected. This is effect is known as the dead cone phenomenon [146]. In Equation 1.79, that as m increases, the dead cone angle $\theta_0 = \frac{m}{E}$ will decrease as the parton mass increases. Figure 1-37 schematically shows a charm quark radiate gluon in the medium with a dead cone in the small angle:

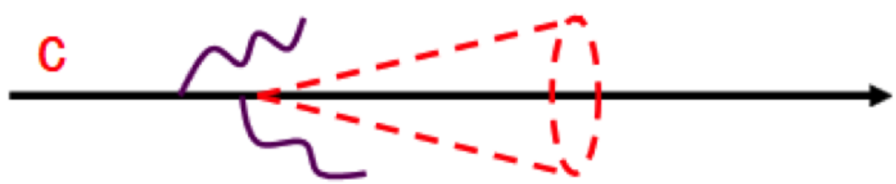


Figure 1-37: The schematic demonstration of a charm quark radiation is shown above. A suppression in small angle due to the dead cone effect in the QGP medium is highlighted.

Since we have the follow mass hierarchy for quarks and gluons:

$$m_g < m_q < m_c < m_b \tag{1.80}$$

We should expect the energy loss to follow

$$\Delta E_g > \Delta E_q > \Delta E_c > \Delta E_b \tag{1.81}$$

We call the inequality above to be the flavor dependence of energy loss, which is an important feature of heavy quark energy loss mechanism in the QGP medium. The studies of heavy quark energy loss mechanism in QGP will help us determine the fundamental jet transport coefficient $\hat{q}$ that characterizes the scattering power of the medium [136]. which relates to the mean free path and the momentum diffusion coefficient of heavy quarks [147]. The determination of $\hat{q}$ will be crucial for us decipher the inner workings of the QGP [148].

1.6.6 Heavy Quark Hadronization

After heavy quarks traverse through the medium, it will hadronize into heavy flavor hadrons, which could be fully reconstructed from their final state decay products in experiments. As described in Section 1.2.7, in general, hadronization is non-perturbative. Considering heavy quark dynamics and apply hadronization models, physicists develop theoretical models to describe heavy quark hadrochemistry. Below, I will present three model candidates, the Texas A&M University (TAMU) Model [149], the model developed from Cao et. al. [152], and the Equal Velocity Recombination (EVR) model [?] to describe beauty quark production and hadronization in vacuum:

TAMU Model

The TAMU Model uses a thermodynamic T-matrix formulism in terms of "ladder diagrams" to compute the heavy quark in-medium scattering amplitude and determine the non-perturbative transport parameters A_i and B_{ij} in the Planck-Fokker equation shown in Eq 1.81 [149]. Figure 1-38 shows schematically the "ladder diagram" describing the dynamic evolution of a heavy quark in the QGP medium

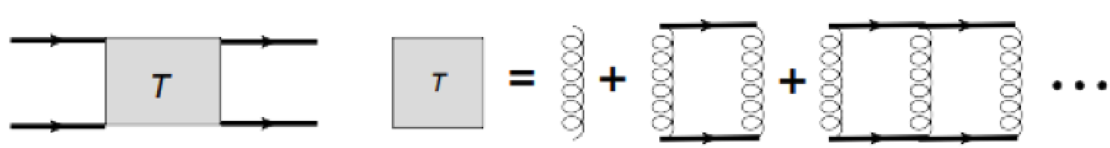


Figure 1-38: The ladder diagram used by the TAMU model to describe heavy quark diffusion in the QGP medium is shown schematically above.

The input of T-matrix uses a lattice QCD potential [150] corrected with relativistic effects to model the non-perturbative interaction between heavy quarks and partons in the medium and make it consistent with heavy flavor spectroscopy in vacuum to determine the thermal relaxation rate coefficient $A_i(p,T)$. Only elastic collisional energy loss is included in the calculation. Resonance recombination model of heavy quark with a light quark nearby is applied to describe heavy quark hadronization [151]. Finally, effective hadronic scattering amplitudes are used to model heavy flavor hadronic rescattering with other hadrons before kinetic freezout stage. The background parton composition and kinematics are modeled by the standard hydrodynamic simulations of the bulk medium in nuclear collisions

Cao, Sun, Ko Model

The Cao, Sun, Ko Model employs an advanced Langevin-hydrodynamics approach [154, 155] incorporating both elastic and inelastic energy loss of heavy quarks inside the dynamical QGP medium. The equation below schematically shows relativistic Langevin equation to simulate the dynamics of heavy quarks in the QGP medium

$$\Delta\vec{p} = -\gamma\frac{T^2}{M}\vec{p}\Delta t + \vec{\xi}(t) \tag{1.82}$$

And

$$\Delta\vec{x} = -\frac{\vec{p}}{E}\Delta t \tag{1.83}$$

The noise is modeled by the Gaussian diffusion function

$$P(\vec{\xi}) \propto \exp[\frac{\vec{\xi}^2}{2D_p\Delta t}] \tag{1.84}$$

The dimensionless γ factor is defined as

$$\gamma = \frac{M}{\tau_{HQ}T^2} \tag{1.85}$$

The Cao, Sun, Ko Model uses a comprehensive coalescence model with strict energy-momentum conservation and PYTHIA fragmentation simulation [157] with the default Peter fragmentation function. The coalescence probability is determined from resonant scattering rate of heavy quarks in the QGP according to the resonant recombination model [151, 156]. In this model, if heavy quarks do not coalesce, they will hadronize via fragmentation mechanism.

EVR Model

In this model, the transverse momentum distribution of initially produced heavy quark is calculated by FONLL [153]. The jet quenching effect in heavy-ion collisions is considered according to R_{AA} measurement of B^+. The transverse momentum distributions of light-flavor quarks are obtained from data of light hadrons in the model. This model is particular designed to study low p_T and mid-rapidity charm quarks produced at the LHC energy. It considers the equal-velocity combination of bottom quark with light-flavor anti-quarks to form B mesons, a framework based on of co-moving quark recombination model (QCM).

In addition to TAMU, Cao, Sun, Ko, and EVR Models, there are many other theoretical models attempting to describe heavy quark hadrochemistry in heavy-ion collisions. Nevertheless, due to the large discrepancies among different hadronization models, our ability to interpret the heavy flavor data is significant limited. Therefore, heavy-ion experimentalists precisely measure heavy flavor physics observables with different species and over broad kinematic ranges and provide constrains for models to reduce the theoretical uncertainties in hadronization.

Chapter 2

Review of Heavy Flavor Results

In the previous section, we have introduced the relativistic heavy-ion physics and heavy flavor physics in vacuum and QGP. Physicists propose many experimental observables to study open heavy flavor physics and test theoretical models in heavy ion collisions. Traditionally, heavy flavor hadron v_2, R_{AA}, and production yield ratio are extensively studied. In this section, we will review selected experimental and theoretical results results in this section and discuss the physics messages from the measurements.

2.1 Elliptic Flow

In the QGP medium, heavy quarks are diffused by the color force and multiple scatter with medium constituents, which could generate sizable azimuthal anisotropy v_2 [136]. In addition, due to the azimuthal anisotropic expansion of the medium, the different path lengths of heavy quarks are different in different direction, which will also contribute to . Experimentally, we scale the v_2 and the hadron kinetic energy $K_T = \sqrt{m^2 + p_T^2} - m^2$ of heavy quarks with n_q according to the Number of Constituent Quark (NCQ) Scaling in quark coalescence model [?]. Figure 2-1 shows the comparison of the v_2/n_q as a function of K_T/n_q of D^0 ($c\bar{u}$) meson with light flavor hadrons with STAR experiments at RHIC [158] and the CMS experiment at LHC [159].

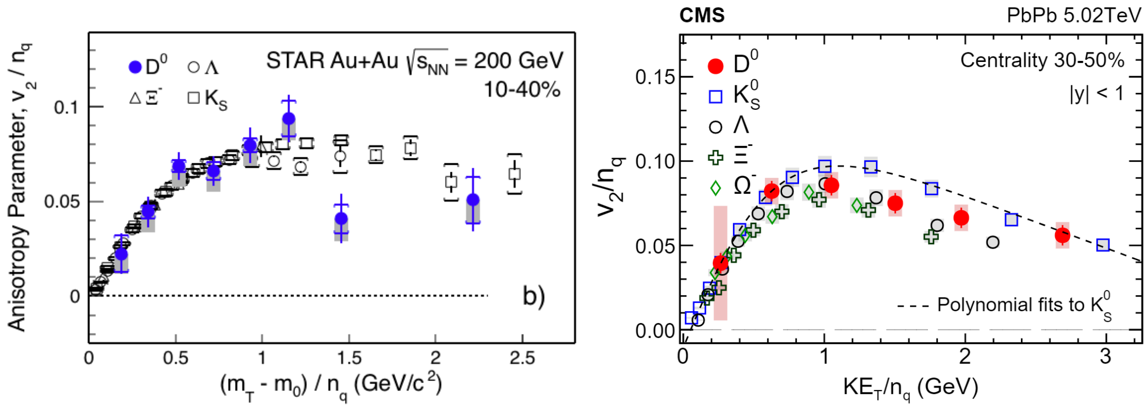


Figure 2-1: The NCQ scaled D^0 $v2/n_q$ vs K_T/n_q and the comparison light hardons measured by the STAR experiment at RHIC (left) and the CMS experiment at LHC (right) are shown above.

We could see a reasonably good NCQ scaling behavior of D^0 meson with other light flavor hadrons, which suggests sizable collectivity of charm quarks in the QGP medium. To study azimuthal anisotropy of beauty quarks, an indirect approach is employed. Figure 2-2 shows the elliptic flow of electron from beauty hadrons $b(\rightarrow c) \rightarrow e$ measured by the ALICE experiment [160] and muon from beauty hadrons $b \rightarrow \mu$ measured by the ATLAS experiment [161]

Comparing Figure 2-2 with Figure 2-1, we can see that beauty quarks does not demonstrate as much anisotropy as charm quarks in heavy-ion collision. However, so far, fully reconstructed B-meson v_2 has not been measured by any experiment.

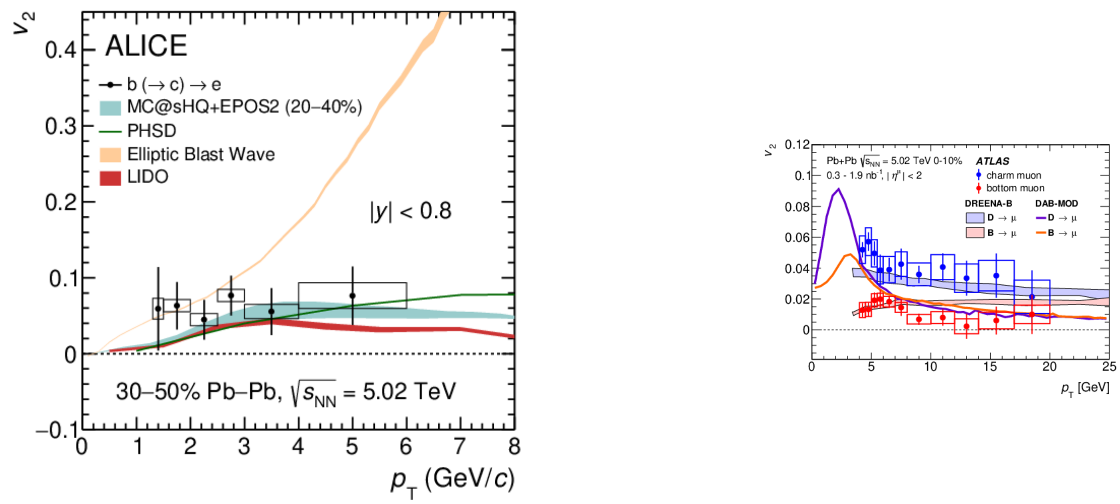


Figure 2-2: The v_2 of the electrons from b hadron decays as a function of electron p_T measured by the ALICE experiment (left) and the v_2 of the muon from b hadron decays as a function of muon v_2 measured by the ATLAS experiment (right) are shown above.

2.2 Nuclear Modification Factor

As mentioned previously, the hadron nuclear modification factor R_{AA} can describe the modification of collisions system compared to the pp collision in the hadron spectra. To study medium modification of heavy quarks, we first would like to investigate the cold nuclear matter effect in pA collisions. Figure 2-3 and Figure 2-4 show the prompt D mesons and B mesons nuclear modification factor R_{pA} in pPb collisions measured by the ALICE experiment [162] and the CMS experiment [163] respectfully

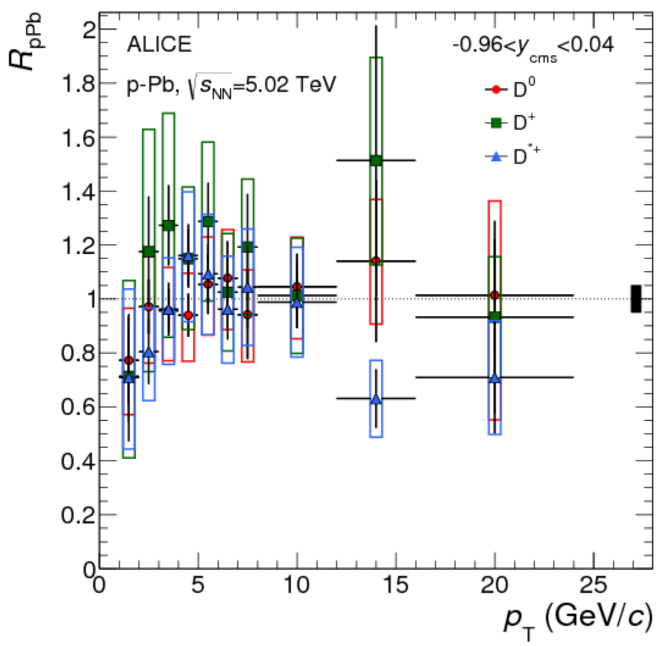


Figure 2-3: The R_{pA} as a function of p_T of prompt D mesons measured by the ALICE experiment is shown above.

Therefore, we can see that there is no significant modification of the charm quarks in cold nuclear matter since the R_{pA} of D^0 is overall unity within experimental uncertainties. Hence, any modification of D-mesons observed in the AA collisions should come from final state QGP effect instead of initial state effect of nPDF of Pb ions.

Next, we investigate B and D mesons R_{AA} the AA collisions. Figure 2-5 R_{AA} heavy flavor hadrons measured with experiments at RHIC and LHC.

We could see that R_{AA} of D^0 and B^+ are both below 1, which suggest charm

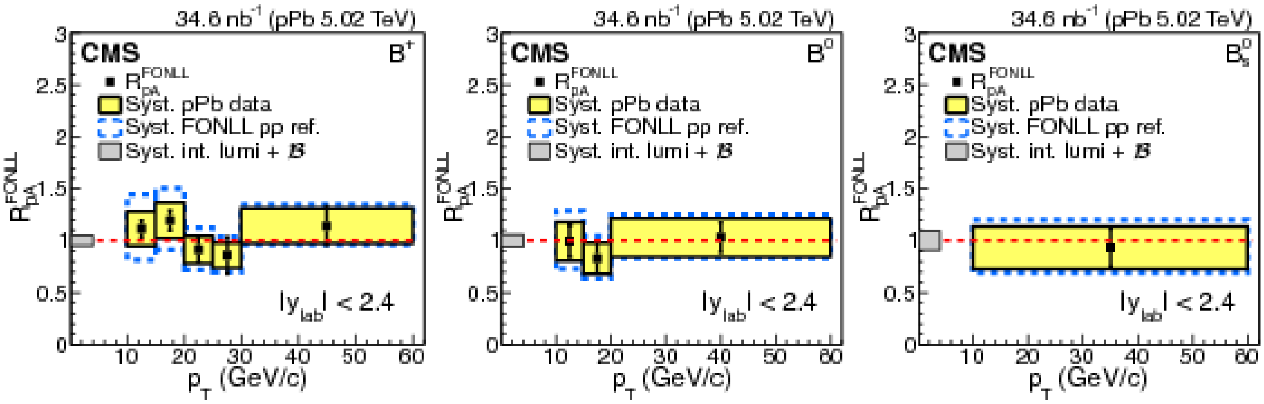


Figure 2-4: The R_{pA} as a function of p_T of B^+, B^0, and B^0_s mesons measured by the CMS experiment is shown above.

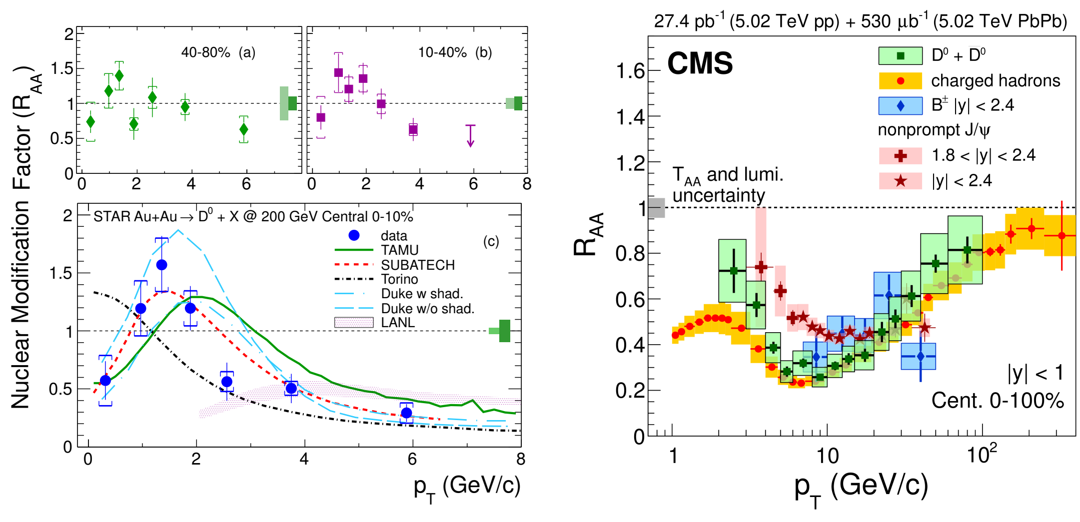


Figure 2-5: The D^0 R_{AA} vs p_T with the STAR experiment in 0 - 10%, 10 - 40%, and 40 - 80% centrality at RHIC and the D^0, B^+, non-prompt J/ψ and charged hadrons R_{AA} vs p_T at at 0 - 100% centrality with the CMS experiment at LHC are shown above.

and beauty quarks lose a significant fraction of energy to the QGP medium. As p_T increases, the R_{AA} of light and heavy flavor hadrons converge to the same value and approach 1, which Lorentz γ factor come into play where the mass of the hadron become irrelevant. In addition, the CMS results above indirectly agree with the expectation of the flavor dependence of energy loss: $R^h_{AA} < R^D_{AA} < R^B_{AA} < 1$. The R_{AA} results are in reasonable agreement with most theoretical model calculations.

To better constrain theoretical model calculations and understand the interaction mechanism of heavy quarks with the QGP medium, we need to perform more precise measurements of B and D mesons R_{AA} and v_2 down to lower p_T where the mass heavy quarks become important and theory diverge. The ALICE experiment has performed measurement of prompt and non-prompt D mesons R_{AA} down to very low p_T shown in Figure 2-6 below

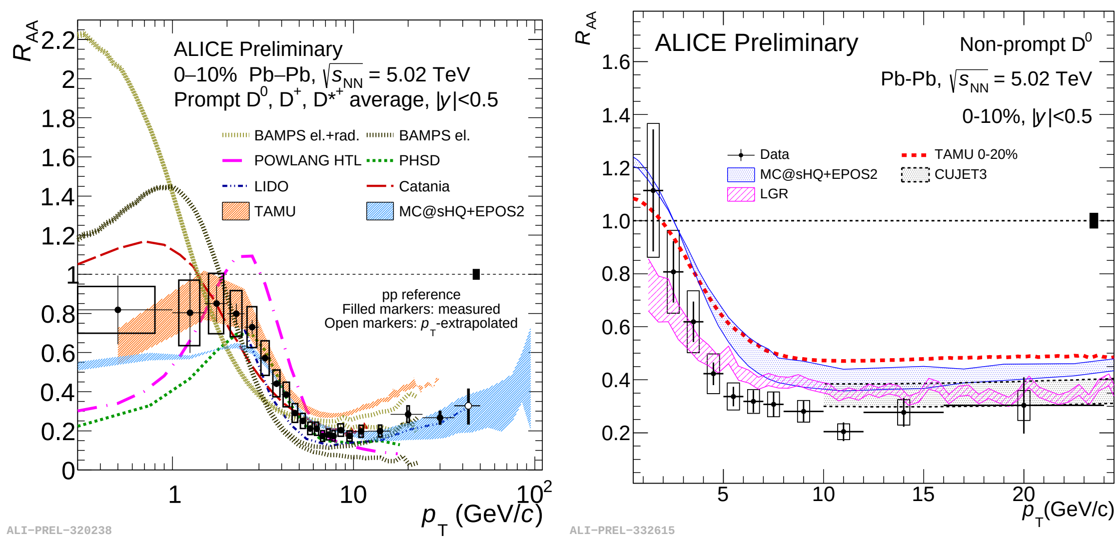


Figure 2-6: The prompt D mesons R_{AA} vs p_T down to $p_T = 0$ (left) and non-prompt D mesons down to $p_T = 1$ GeV/c (right) are shown above.

From the ALICE measurement of prompt and non-prompt D mesons R_{AA} down to very low p_T, we could see that very few model can simultaneously describe D^0 R_{AA} at both low and high p_T. Nonetheless, the fully reconstructed B meson R_{AA} from exclusive b decay down to very low p_T is still missing. We should try to perform B meson R_{AA} measurement down to low p_T to provide a complete picture to constrain the jet transport coefficient $\hat{q}$ and probe the QGP at longer wavelengths. Also, good to measure inclusive beauty production cross section in pp collisions to test pQCD calculations.

2.3 Production Yield Ratio

According to the theoretical reviews of heavy quarks hadrochemistry in heavy-ion collisions [167,168], the strange-to-non-strange meson (H_s/H^0) and baryon-to-meson (Λ_Q/H^0) ratios are excellent observables to test hadronization models. Both RHIC and LHC have carried out extensive measurements fully reconstructed charm hadron yield ratios. Figure 2-8 shows the fully reconstructed prompt D_s^+/D^0 ratio measured by the STAR [] and ALICE [] experiments

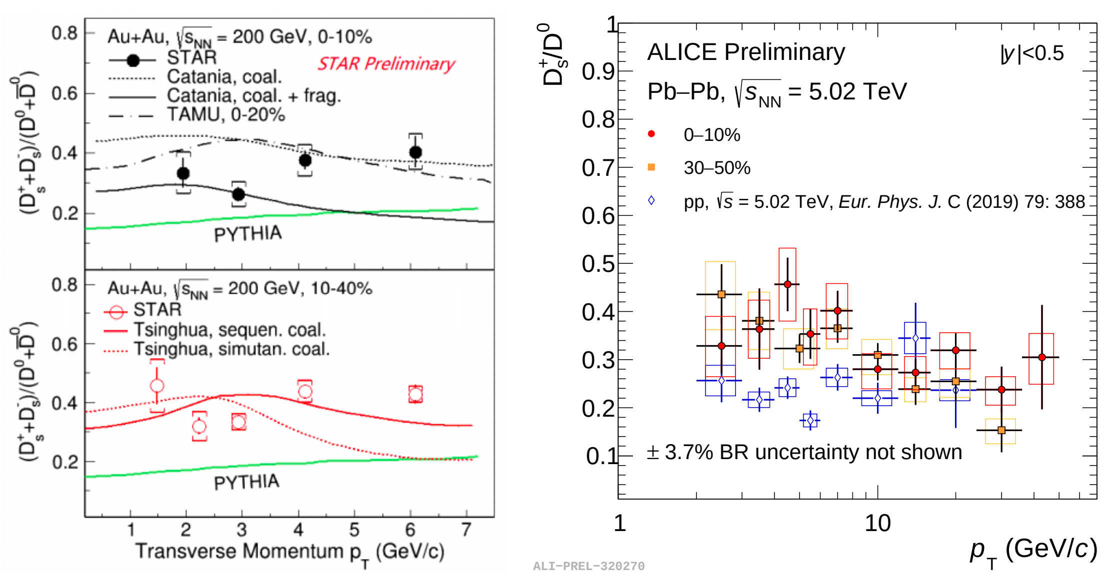


Figure 2-7: The fully reconstructed D_s^+/D^0 ratio in Au + Au measured by the STAR experiment at RHIC (left) and in PbPb the ALICE experiment at LHC (right) as functions of p_T are shown above.

Figure 2-8 shows the fully reconstructed Λ_C^+/D^0 ratio measured by the STAR and ALICE experiments

The ALICE experiment also performs a comprehensive study on charm quark hadronization in pp, pPb, and PbPb. Figure 2-9 shows the D_s^+/D^0 and Λ_c/D^0 ratios

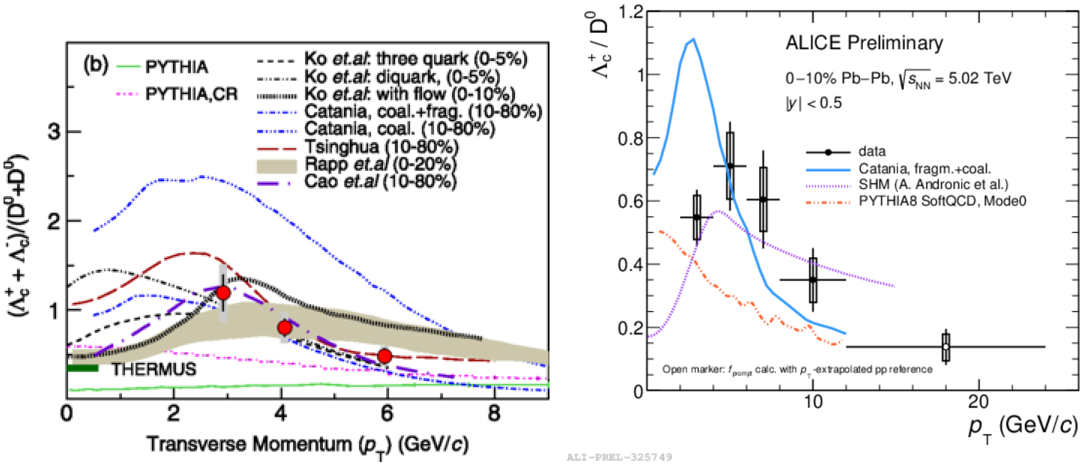


Figure 2-8: The fully reconstructed Λ_C^+/D^0 ratio in pp and heavy-ion collisions measured by the STAR experiment at RHIC (left) and the CMS experiment at LHC (right) are shown above.

as functions of event multiplicity from small to large collision systems.

We can see that in general, D_s^+/D^0 while Λ_C^+/D^0 ratio in heavy-ion collisions lies above its ratio in *pp* collisions. In the multiplicity studies, an overall increasing trend of both D_s^0/D^0 and Λ_C^+/D^0 ratios in higher multiplicity is observed. Many different theoretical predictions agree reasonably well with the experiments due to their large uncertainties. However, such large discrepancies among hadronization models significantly limit our ability to interpret heavy flavor experimental data. Moreover, this is only in the charm sector, such fully reconstructed b-hadron measurements to study beauty hadrochemistry are still missing.

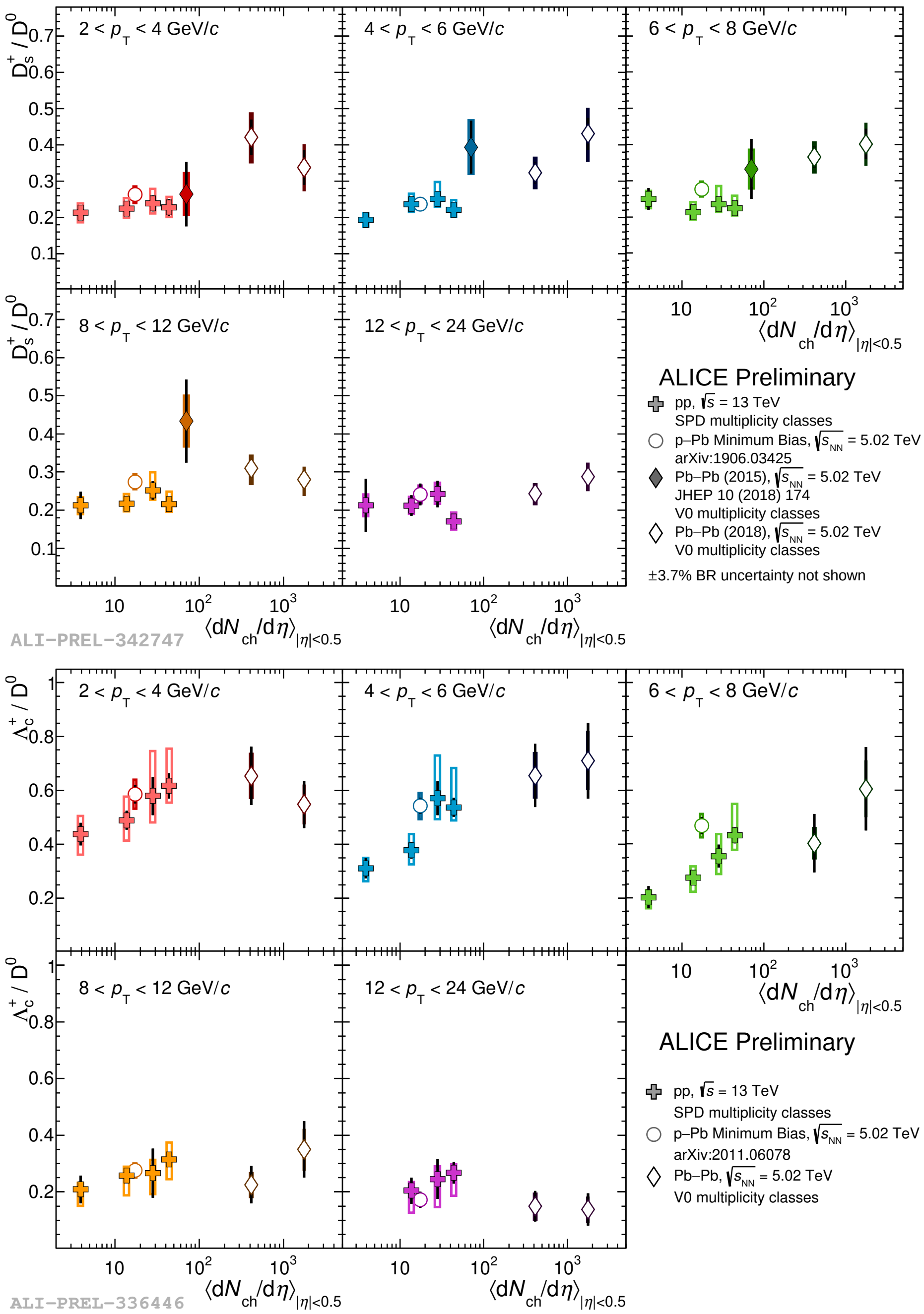


Figure 2-9: The fully reconstructed D_s^+/D^0 (top) and Λ_C^+/D^0 ratio (botton) as a function of event multiplicity $\langle dN_{ch}/d\eta\rangle$ within $|\eta| < 0.5$ in p_T from 2 - 4, 4 - 6, 6 - 8, 8 - 12, and 12 - 24 GeV/c in pp, pPb, and PbPb collisions measured by the ALICE experiment are shown above.

Therefore, it is crucial to have more precise measurement over a wide range of p_T and multiplicity in both beauty and charm sectors to constrain theoretical models. Currently, the only published fully reconstructed b-hadron measurement in heavy-ion collision is the B_s^0/B^+ based on CMS 2015 PbPb dataset [166]. Figure 2-10 shows the B_s^0/B^+ R_{AA} ratio in pp and PbPb collisions

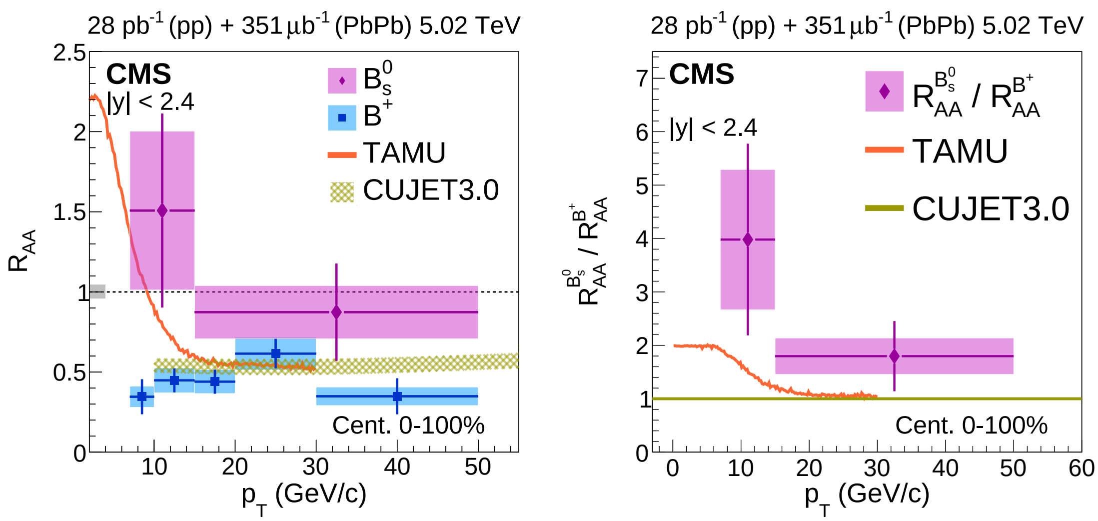


Figure 2-10: The fully reconstructed B_s^0 and B^+ R_{AA} (left) and B_s^0/B^+ R_{AA} ratio (right) as a function of p_T using the 2015 CMS pp and PbPb datasets are shown above.

This first fully reconstructed B-meson measurement in heavy-ion collision is good. Nevertheless, the measure has relatively large uncertainties. The B_s^0 significance is still below 5 σ. In order to better constrain model calculations and better, we should perform more differentiated measurement with improved precision.

In the baryon-to-meson ratio studies, LHCb has conducted fully reconstructed Λ_b/B^+ ratio in pp and pPb collision [172] shown below in Figure 2-11

We can see over near unity of Λ_b/B^0 double ratio in pPb to pp in the forward region from LHCb measurement. No significant p_T and y dependence is observed. It would be interesting to conduct similar measurement in the mid-rapidity region in pp and pPb collision. However, so far no fully reconstructed Λ_b measurement have been carried out in heavy-ion collisions due to the limited statistics and large combinatorial background of Λ_b.

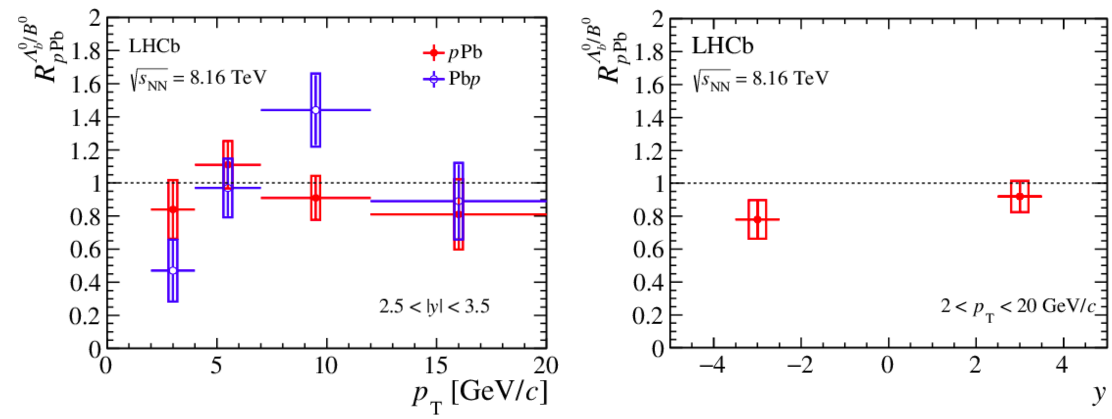


Figure 2-11: The fully reconstructed Λ_b/B^+ R_{pA} ratio as a function p_T (left) and y (right) in pp, pPb, and Pbp collisions measured by the LHCb experiment are shown above.

2.4 Heavy Flavor Hadron-Jet Angular Correlations

Aside traditional heavy flavor observables: R_{AA}, v_2, and cross section ratio, modern observables, such as heavy flavor hadron-hadron and heavy flavor hadron-jet angular correlations, have higher differentiation to provide more insight to understand the dynamics and interaction mechanism of heavy quarks in the QGP medium. The measurements of angular correlations between heavy flavor hadrons and jets can be used to constrain parton energy loss mechanisms and to better understand the heavy-quark diffusion inside the medium. From the D-jet angular correlation, we can quantify the medium modification to the radial profile of charm quarks and shed light on the interaction mechanism of charm quark with the medium. Figure 2-12 shows the measurement of D-jet angular correlation in PbPb and pp collisions with the CMS experiment [174]

We can see that the D^0 meson production is pushed radially outward in PbPb collisions compared to pp, which shows effects of charm quark diffusion with the present of the QGP medium. While the CCNU model is in reasonably good agreement with the PbPb/pp ratio, its prediction is lower when compare to measurement of $\frac{1}{N_{jD}}\frac{dN_{jD}}{dr}$ pp and PbPb collisions.

2.5 Heavy Flavor Hadron-Hadron Correlations

Another observable is heavy flavor hadron-hadron correlation, which is even better to tag the heavy quark $Q\bar{Q}$ pair the produced back to back in the early stage of hard scattering processes and understand the modification effect as they propagate thorough vacuum and medium. Experimentally, the observable is a fully reconstructed open heavy flavor hadron correlate with associated hadrons produced within the same event and subtract the background in mixed events. In the analysis, $\Delta\eta and \Delta\phi$ distributions of the heavy flavor hadron from associated hadron are used to quantify the correlation. Figure 2-14 shows the D meson-hadron angular correlation measured with the ALICE experiment [175]

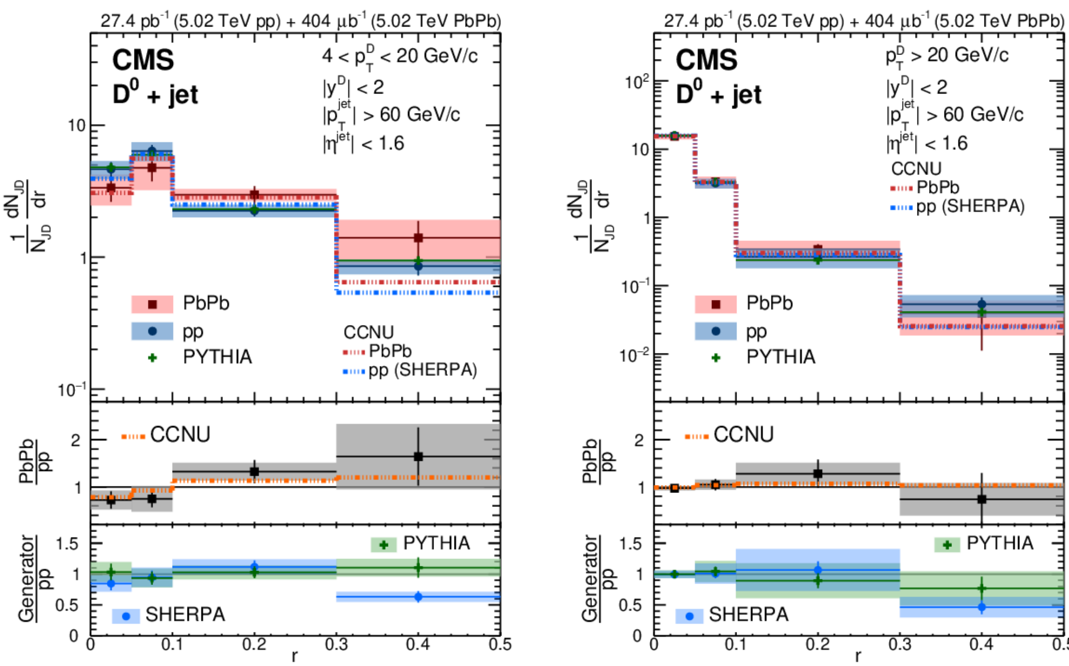


Figure 2-12: Distributions of fully reconstructed D^0 mesons in jets, as a function of the distance from the jet axis (r) for jets of $p_T^{jet} > 60$ GeV=c and $|\eta^{jet}| < 1.6$ measured in pp and Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, for $4 < p_T^D < 20$ GeV/c and $p_T^D > 20$ GeV/c are shown above. The jet radius is defined as $r = \sqrt{(\Delta\phi_{jD})^2 + (\Delta\eta_{jD})^2}$ where ϕ_{jD} and η_{jD} are the η and ϕ of the D^0 meson with respect to the jet axis.

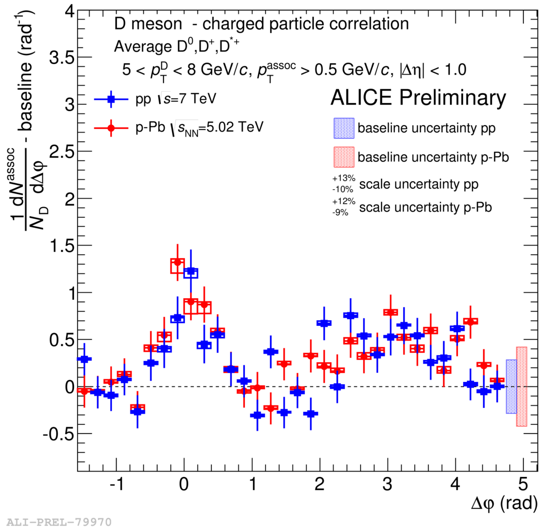


Figure 2-13: The ALICE D-hadron angular correlated in both pp (blue) and pPb (red) collision (left) and the comparison of pp data with PYTHIA calculations (right) are shown above.

In the D-hadron correlation, there are two peaks at $\Delta\phi = 0$ and π. At $\Delta\phi$ =0, hadron are produced along with the charm quark via fragmentation mechanism. At $\Delta\phi = \pi$, the hadron are produce from back-to-back jets the to maintain momentum conservation. The pp measurements are overall consistent to PYTHIA calculations. From the comparison of the results in the pp and pPb, the D-hadron angular correlation distribution are compatible with each other with uncertainties. Consequently, no evident effects on the charm fragmentation and hadronization due to cold nuclear matter can be claimed.

STAR has perform the 2D $\Delta\eta\phi$ measurement of fully reconstructed D^0-hadron correlation in Au + Au collision shown in Figure **??**

As expected, the D^0-hadron angular correlation get broaden and the peak near π disappear in more central Au + Au collisions because QGP are more likely to create and redistribute the energy among particles. In the beauty sector, so far there is no such measurement carried out in heavy-ion collision. The B-hadron correlation measurement, along with the D-hadron correlation measurement, will be crucial to provide deeper insights on heavy quark diffusion and energy loss in the QGP medium.

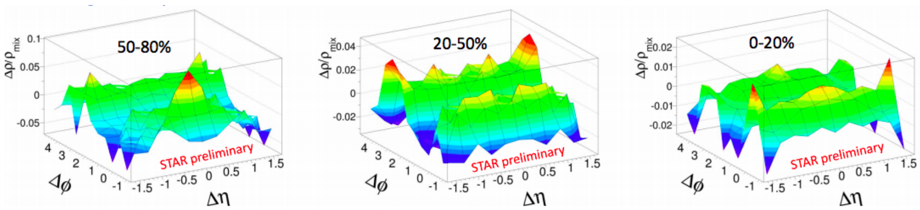


Figure 2-14: The 2D $\Delta\eta\phi$ distributions of D^0 meson and associated hadrons in Au + Au collision centrality 0 - 20%, 20 - 50%, and 50 - 80% at $\sqrt{s_{NN}} = 200$ GeV measured by STAR experiment are shown above.

2.6 Some Questions in Heavy Flavor Physics

As seen in section 2.3, extensive studies on fully reconstructed charm hadrons have been carried out at RHIC and the LHC. In addition, many measurements of fully reconstructed b-hadrons produced in pp and pPb collisions have been carried out by LHCb experiment. In heavy-ion collision, only one measurement of fully reconstructed b-hadron has also been published. Hence, to have a more comprehensive understanding of heavy flavor physics, we should perform more precise and differential measurement on fully reconstructed b-hadrons.

Hence, as seen above, Figure 2-1 and Figure 2-2 show that charm quarks not only have non-zero v_2 but also have less than unity R_{AA} in heavy-ion collisions. Such results could be interpreted as a hint of partial thermalization of charm quarks in the QGP medium [?], which could make charm quark not an ideal probe to the QGP medium.

However, results from Figure ?? show that beauty quarks have much smaller probability being thermalized in the QGP medium compared to charm because they are heavier. Hence, beauty quarks are more desired probes for QGP. However, so far, due to relatively small production cross section, lower reconstruction efficiency, and larger combinatorial background level in its decay chain, the fully reconstructed b-hadron measurements via exclusive production in heavy-ion collision are challenging. and still very limited.

These all leave us with some questions related to beauty hadrochemistry. They are listed as follows:

- Can we confirm the observation of fully reconstructed B_s^0 production in nucleus-nucleus collisions?
- Can perform more differential and precise measurements to study beauty energy loss mechanism in QGP?
- Do our measurements have enough precision to constrain theoretical model predictions?
- Can our measurements provide enough information to understand beauty quark hadronization mechanism from vacuum to QGP through our measurements?
- Does strangeness enhancement also occur in b-hadron production in PbPb collisions?
- How much information about heavy quark diffusion coefficients can we provide?
- Does hadronization universality breaking also occur when a beauty quark move slowly in high color density environment in small systems?

2.7 Motivation of This Thesis

To answer these questions, we propose to perform fully reconstructed B_s^0 and B^+ measurements of R_{AA} and B_s^0/B^+ using the 2018 CMS PbPb at $\sqrt{s_{NN}}$ =5.02 TeV dataset, which has about 3 times as much statistics as the 2015 PbPb dataset, and 2017 pp at $\sqrt{s_{NN}}$ =5.02 TeV dataset, which has more than 10 times statistics than the 2015 pp dataset. Our goal is to perform a better measurement than the published results using the 2015 datasets [166]. In order to improve our results, machine learning techniques along with multivariate analysis approach will be applied in the B-meson analysis. Our measurements may help elucidate the questions above and shed light on beauty quark hadronization mechanism in vacuum and QGP.

Chapter 3

The CMS Detector

3.1 Overview

The Compact Muon Solenoid (CMS) Detector is a general purpose high-energy physics detector located 100 meters underground on the French side of the LHC [179]. Overall, the complete detector is 21 m long, 15 m wide and 15 m high with a weight of 14 kiloton, heavier than the Eiffel Tower in Paris. It functions as a giant, high-speed camera, taking 3D “photograph” of particle collisions from all directions up to 40 million times each second. Figure 3-1 shows the photo taken for the CMS detector at the underground collision hall.

The CMS detector is made of sub-detectors including silicon strip and pixel trackers, the preshower made of silicon strips, the crystal electromagnetic calorimeter (ECAL), the superconducting solenoid with 3.8 T of magnetic field strength, the inner hadronic calorimeter (HCAL), the steel returning yoke to enhance the magnetic field strength, the outer hadronic calorimeter, the muon chambers, and the forward hadronic calorimeter [179]. Figure 3-2 shows schematic view of the CMS detector

The CMS detector is built, operated, and maintained by the CMS Collaboration. The CMS Collaboration consists of over 4000 members including scientists, engineers, technicians, students, and administrative assistants from 200 institutes and universities in 40 countries around the world. Physicists take data from the CMS detector and share data with each other with online system. The data are store in tapes and

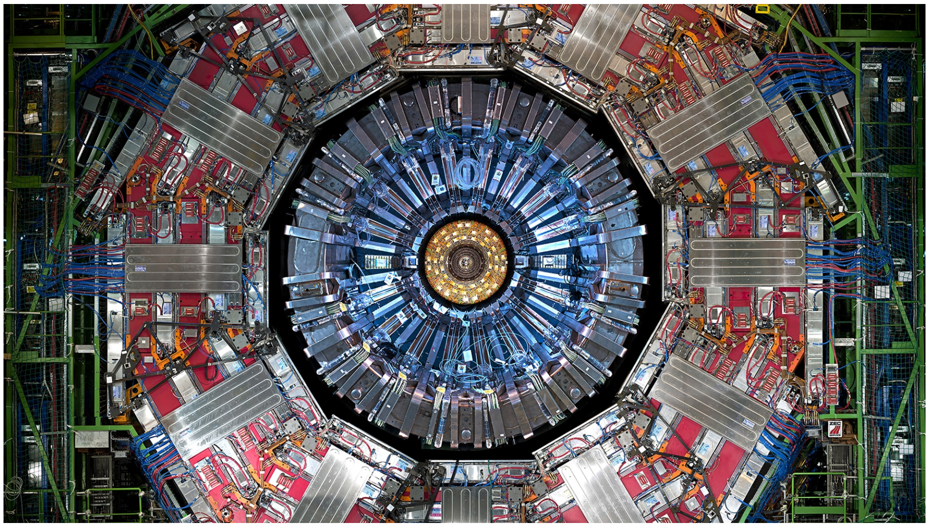


Figure 3-1: The front view of the CMS detector at the underground collision hall is shown above.

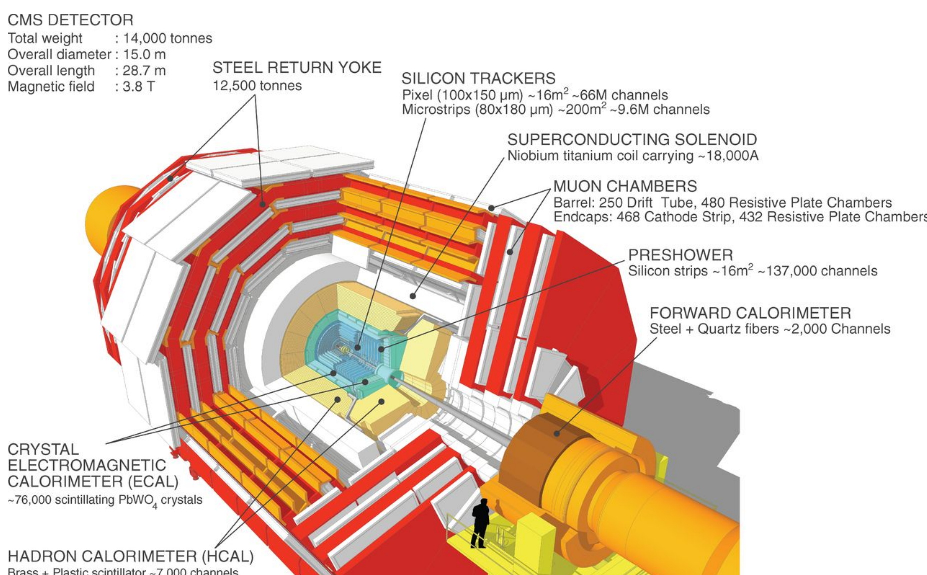


Figure 3-2: The schematic view of the CMS detector with brief descriptions of all its components is shown above. Image from [180]

kept at different institutions. Members of the CMS experiment collaborate with each other on detector studies and data analysis to produce important scientific results and have published in more than 1000 papers in internationally recognized journals.

In the following sections, I will describe in more details the CMS experiment including the trigger system for data acquisition, the tracking system to track charged particles, the muon system for muon detection, identification, and reconstruction, and the calorimeter system to measure the energy of the particles.

3.2 Triggers

The CMS experiment develops triggers to acquire experimental data [181]. Its main purpose is to select events of potential physics interests from approximately one billion events per second the particles collisions at the LHC. The CMS trigger system consists of two levels of triggers: hardware level 1 (L1) trigger and the software high level trigger (HLT). Different triggers encoded in the L1 and HLT are designed and fire to collect datasets for specific physics studies.

3.2.1 L1 Trigger

In the CMS experiment, an event is defined as a snapshot of one collision at the LHC. In the L1 trigger, physicists develop algorithms according to detector electronics response to decide if an event is accepted or rejected within the L1 trigger latency time. Figure 3-3 shows the schematic overview of L1 trigger making its decision online to select events based on the information from the calorimeter and muon systems.

In the interest of heavy-ion studies, physicists develop a set of dedicated triggers algorithms in the L1 trigger to build datasets. The minimum biased (MB) trigger is designed to collect minimum bias data for elliptic flow, D^0 meson, and charged particle multiplicity analyses while the single muon trigger is designed to select events muons for heavy flavor and electroweak physics analyses. We will describe the MB trigger since we will need to use it to determine the number of MB events in our analysis.

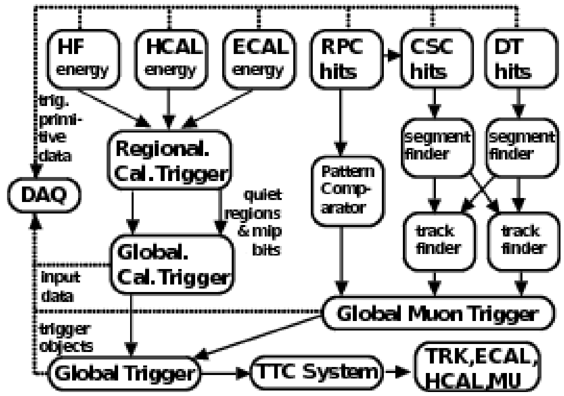


Figure 3-3: The figure above demonstrates how the CMS L1 hardware trigger function schematically.

3.2.2 MB Trigger

By definition, an MB event corresponds to a non-single diffractive inelastic interaction [182]. A totally inclusive trigger, or called zero bias (ZB) trigger, corresponds to a randomly reading out from the detector whenever a collision is possible. MB trigger is algorithm to determine interesting MB events based on the response from forward HCAL located at $3 < |\eta| < 5$. It is put a fixed analog to digital converter (ADC) threshold in the HCAL response to reject background noise and collect MB events from ZB trigger. There is also an essentially linear relation between the maximum ADC with the actual energy response of the forward HCAL. Figure 3-4 shows the ADC distribution and HF energy as a function of ADC in 2018 PbPb run.

The MB trigger consist “MB OR”, which requires the ADC threshold on either one of the forward HCAL (HF) out of both forward ECAL in both positive and negative sides, and “MB AND”, which requires the ADC threshold on both of HFs out of both forward ECAL in both positive and negative sides. Figure 3-5 shows the L1 MB trigger analysis of Run 326791 in the 2018 CMS PbPb data taking

In the 2018 CMS PbPb data taking, to reject the noisy background, the max ADC of each event is required to be greater than 15 with MB AND along with the HLT trigger of at least one pixel track are applied to select MB events, as seen above from Figure 3-5 in the max ADC distribution of MB evens in green. A total number

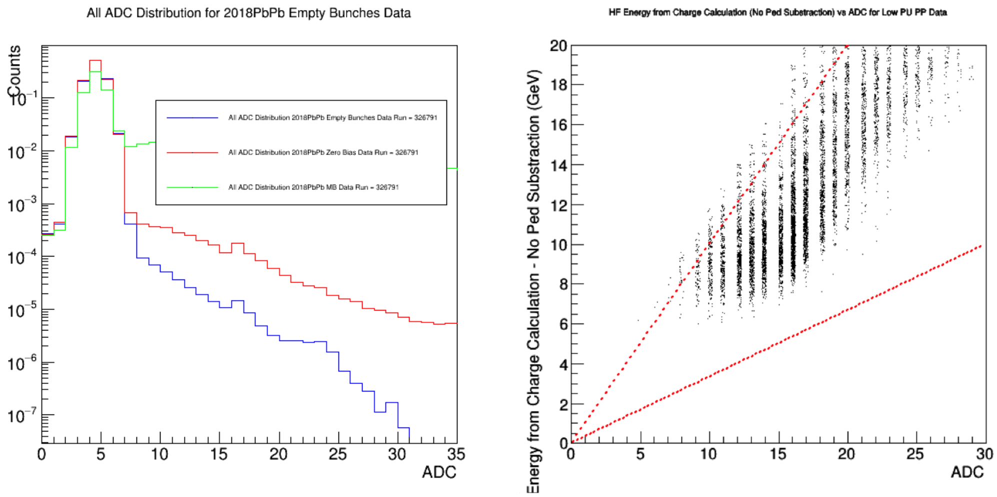


Figure 3-4: In the CMS 2018 PbPb Run 326791, the ZB data (red), Empty Bunches (blue), and MB data (green) ADC distributions (left), and the HF energy according to the charge collected as a function of ADC (right) are shown above. We can see that the HF energy is about (0.5 - 1) conversion factor to the ADC.

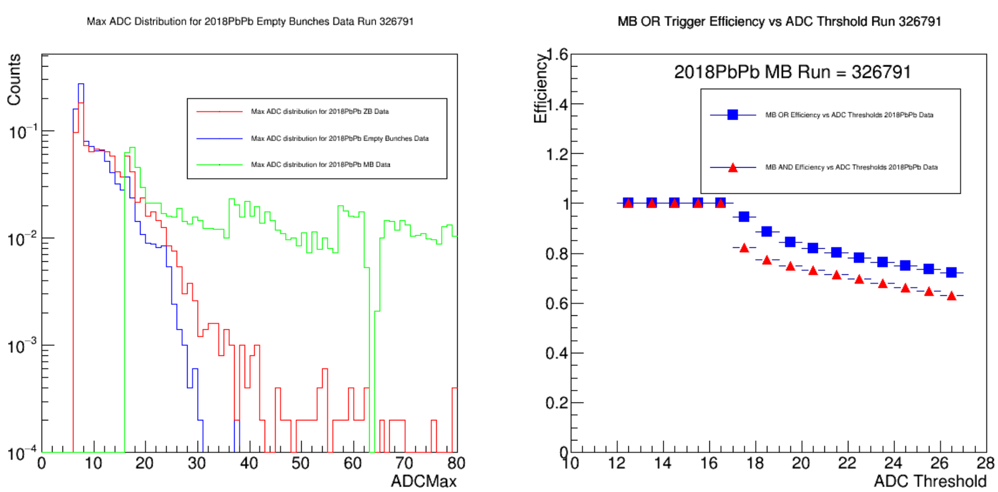


Figure 3-5: In the CMS 2018 PbPb Run 326791, the ZB data (red), Empty Bunches (blue), and MB data (green) maximum ADC distributions (left) and the efficiencies of MB OR (blue) and MB AND (red) as a function ADC threshold (right) are shown above.

of about 2.4 billion MB events corresponding to a luminosity about 1.7 nb^{-1} have been collected by CMS during the 2018 LHC PbPb run from November to December 2018. Figure 3-6 shows the MB events and corresponding luminosity as a function day throughout the 2018 CMS PbPb data taking period

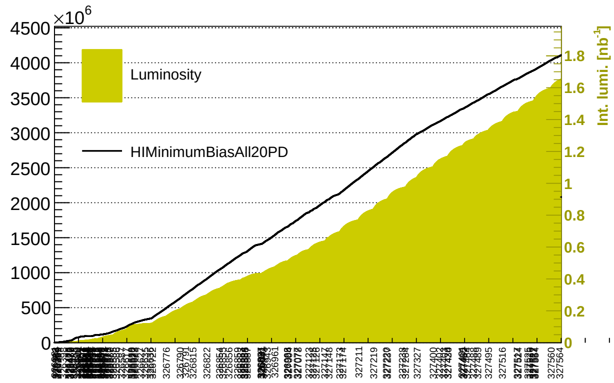


Figure 3-6: The figure above shows the total number of 20 PbPb MB events from and corresponding luminosity how the as a function Run ID from November 15 to December 2 2018.

3.2.3 Centrality Efficiency with MB Trigger

In addition to overall efficiency vs the ADC with the MB trigger, we also study the centrality efficiency with different ADC thresholds. Figure 3-7 shows the centrality as a function of efficiency using MB OR and MB AND with different thresholds

Because other physics trigger are mainly based on the MB datasets, in the physics analyses using 2018 CMS PbPb datasets, it is recommended to remove the very peripheral centrality range from 90 - 100%, which is not fully efficient (efficiency < 100%). Therefore, the most of the CMS heavy-ion physics results using the 2018 PbPb dataset will be presented in the centrality range of 0 - 90%.

3.2.4 HLT Trigger

The HLT software trigger is an array of commercially available computers running high-level physics algorithms [181]. Unlike the online L1 hardware trigger which runs

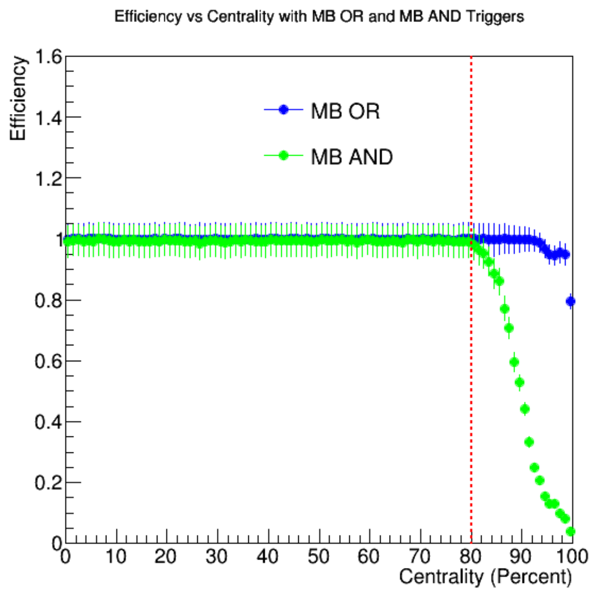


Figure 3-7: The efficiency vs centrality with ADC $>$ 16 for MB OR (blue) and MB AND (green) are shown above.

on-the-go during the data taking process, HLT is an offline software trigger that runs after the data are acquired. In the HLT trigger, more sophisticated analyses are performed to determine if the event is accepted or rejected for a specific dataset. The event data are stored locally on disk and eventually transferred to downstream systems, the CMS Tier-0 computing center, for offline HLT processing and permanent storage [181]. There are many trigger paths in the HLT such as the high multiplicity trigger to specifically collect events with many tracks, the D meson trigger to select high p_T D mesons, and the dimoun trigger to enrich Drell-Yen events, are designed and encoded in the HLT trigger. In the following, we will describe the dimuon trigger in details because the dimuon dataset will be used to fully reconstruct B mesons in this thesis.

3.2.5 DiMuon Trigger

The dimuon trigger, as it is named, is a trigger based on the information of two muons tracks. HLT is able to quickly reconstruct the invariant mass of two oppositely charged muons $m_{\mu\mu}$. Figure 3-8 shows the $m_{\mu\mu}$ reconstructed by the CMS HLT with 2018 pp dataset.

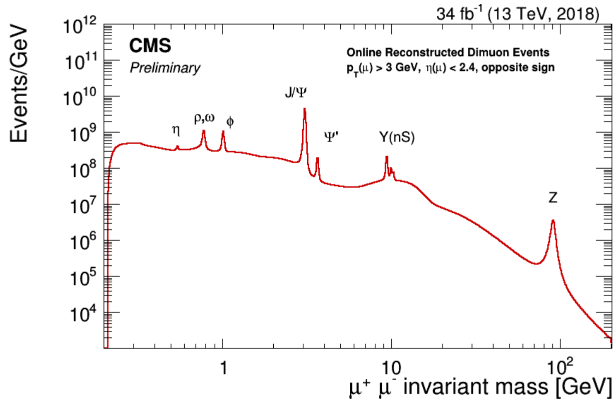


Figure 3-8: The dimuon invariant spectrum $m_{\mu\mu}$ reconstructed by CMS HLT trigger in the 2018 pp dataset is shown above. We can identify the neutral vector boson resonances shown above.

In the 2018 PbPb run, the dimuon trigger requires the presence of two muon candidates, with no explicit momentum threshold and with the HLT reconstructed dimuon invariant mass of 1.0 GeV/c^2 $< m_{\mu\mu} <$ 5.0 GeV/c^2, near the J/ψ PDG mass $m_{J/\psi} = 3.0969$ GeV/c^2 [4], in coincidence with lead bunches crossing at the interaction point. Moreover, One of the trigger-level muons is reconstructed using information both from the muon detectors and the inner tracker with requirement of more than or equal to 10 hits (named as L3 muon), while for the other only information from the muon detectors is required (named as L2 muon) [183].

3.3 Tracking System

3.3.1 Silicon Detectors

The CMS tracking system applies solid state semiconductor technologies. It consists of the 3 layers of silicon pixel tracker and 10 layers of silicon strip detector including 4 inner barrel layers and 6 outer barrel layers [184]. It have a $\phi = 2\pi$ and $|\eta| < 2.4$ acceptance coverage. Figure 3-9 shows the CMS tracking system schematically

In nuclear and particle physics, a tracker is a detector that measures the trajecto-

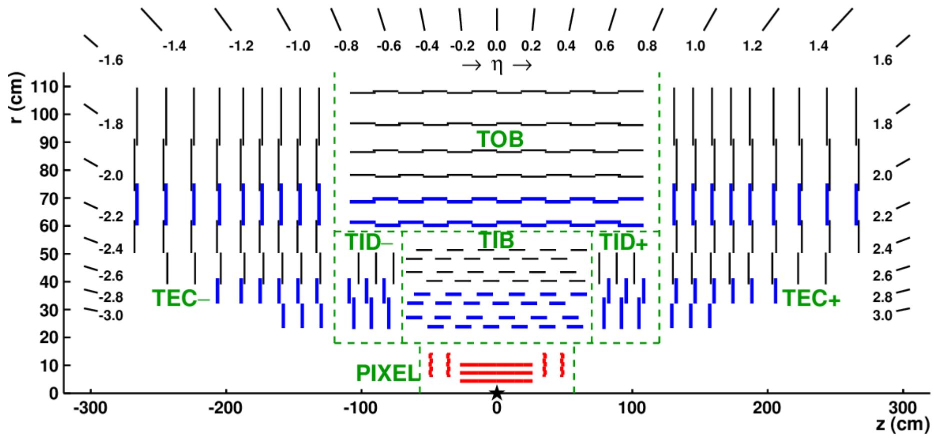


Figure 3-9: The schematic view of the CMS tracking system is shown above.

ries of charged particles via ionization. In general, it does not destroy or significantly change the energy of the particle. With the external magnetic field, the tracker can measure the momentum, the charge, and the mass of the particle by the studying the electric charges collected from electron avalanche or electron-hole pair. The CMS tracking systems provides physicists with excellent tracking capabilities. The CMS silicon tracker is solid state detector employing semiconductor technologies. The silicon tracker is operated at a reserve bias mode with a depletion voltage of about 600V. High energy charged particles passing through the silicon tracker has an energy loss of $dE/dx \simeq 0.5keV/\mu m$ [4]. Therefore, for a 320 μm thick silicon sensor, the charged particle will lose about 160 keV. The electron-hole pair in silicon is about 3 eV per pair. Therefore, the charged particle will produce roughly on the order of 10^4 electrons. The hit resolution in $r\phi$ direction of the silicon strip is about $10 - 40$ μm [185]. Figure 3-10 shows schematically how a high energy charged particle ionized a electron-hole pair in the depletion region of a silicon P-N junction diode operated at a reverse biased mode

However, in the CMS silicon tracker, due to the small number of electrons produced in the silicon sensor, the energy loss dE/dx vs momentum p of charged particle is does not good enough resolution to separate and identify electron, pion, kaons and protons. Therefore, we generally do not perform particle identification (PID) for hadrons with CMS detector in physics analyses.

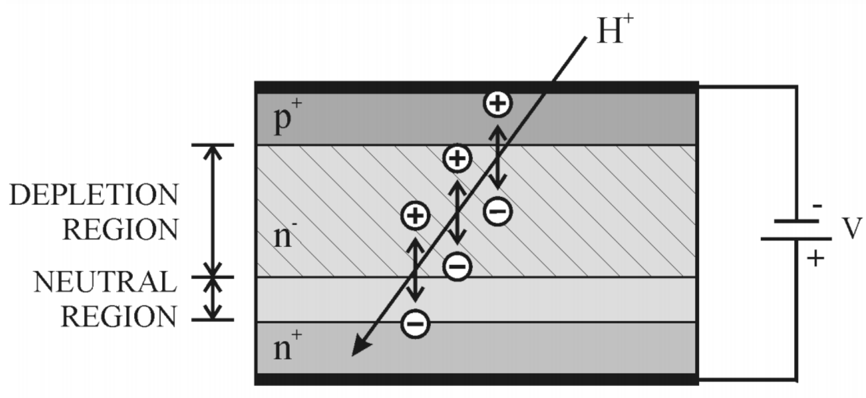


Figure 3-10: The schematic plot explaining how a silicon tracker detector charged particles is shown above.

3.4 Muon System

Named as “Compact **Muon** Solenoid”, the study on muon is one of the most important physics tasks of the CMS experiment. The CMS muon system has 1400 muon chambers including 250 drift tubes and 540 cathode strip chambers to track the positions of the muons and provide a trigger and 610 resistive plate chambers form a redundant trigger system with an acceptance coverage of $|\eta| < 2.4$. Due to the small energy loss of muon in ECAL and HCAL [4], the muon produced from the collisions usually penetrates through the trackers and calorimeters. Therefore, the muon system is located at the outer of the CMS detector. Figure 3-11 shows the particles produced at the interaction points and pass through the CMS detector

The muon system employs gaseous detector technology. Physical modules of drift tubes, cathode strip proportional planes, and resistive plates are called “chambers”. When a muon pass through the chambers, it will ionize electrons of the gas atom. Under a strong electric field, the avalanche electrons will be drifted to the anode and the gas ion will be drifted to the cathode. Electronic signal will be generated as this occurs. Figure 3-12 shows schematically how electron avalanches works in a gaseous detector to detect charged particles as well as the design of CMS drift tube to detect muons.

Therefore, with both the tracking system and the muon chambers, the CMS de-

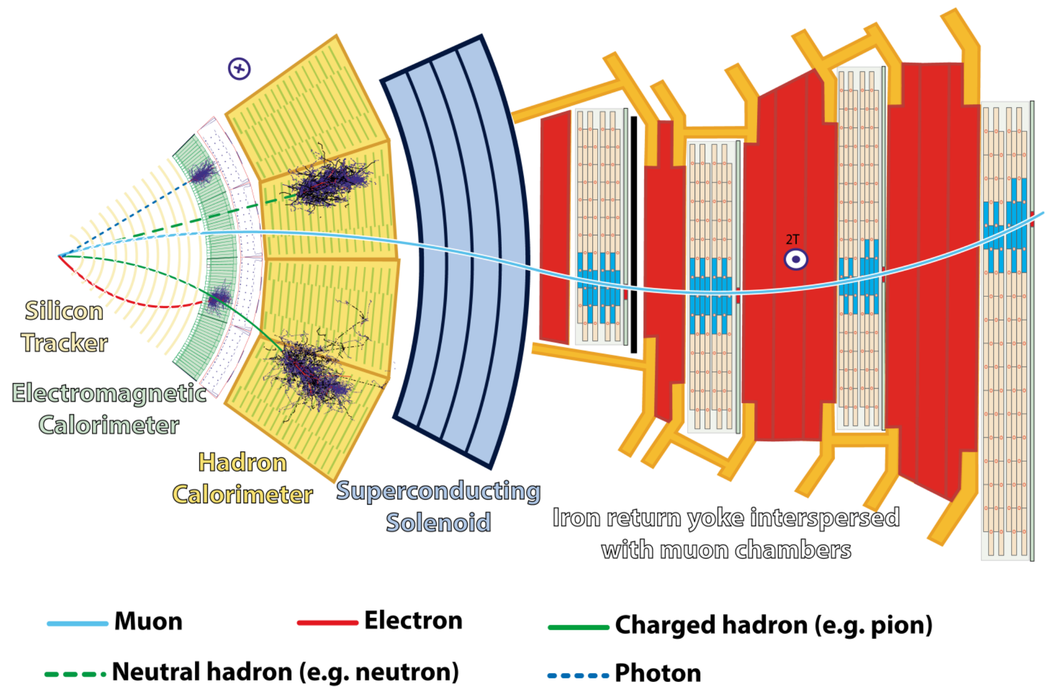


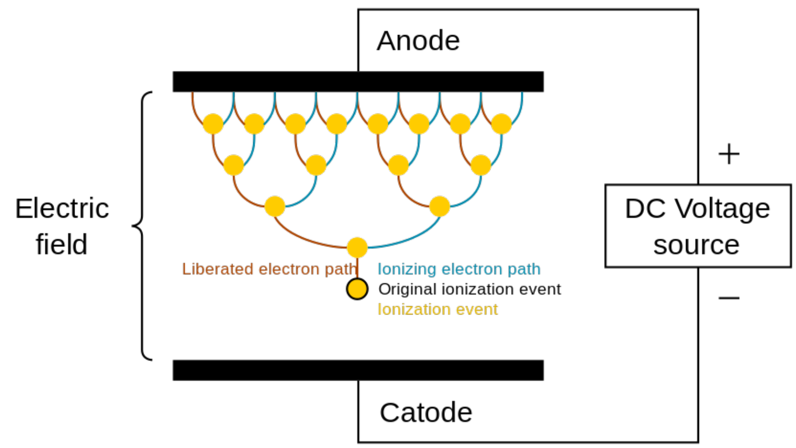
Figure 3-11: The particle flow of long life particles, such as electrons, muons, photons, charged hadrons: π, K, p, and neutral hadrons: neutrons, in the CMS detector are shown above.

tectors has excellent capabilities of detecting, identifying, and reconstructing muons, which is crucial for heavy flavor physics studies.

3.5 Calorimeter System

In nuclear and particle physics, a calorimeter is a detector that completely stops particles and measure the total energy deposited. According to the particles, calorimeter can be divided into electromagnetic calorimeter (ECAL or EMCAL) to measure the energy of electron and photons and hadronic calorimeter to measure the energy of charge and neutron hadrons. The CMS calorimeters system includes both ECAL and HCAL. It is located in between the tracker and the muon chambers as shown in Figure 3-2.

According to measurement of charged particle shower energy, calorimeter can typically be classified as sampling calorimeter and homogenous calorimeter. The



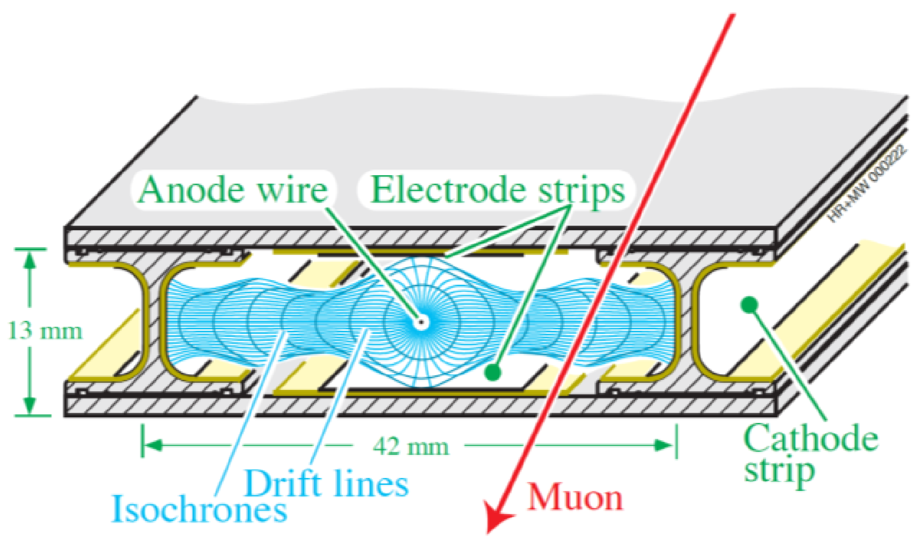


Figure 3-12: A visualization of Townsend Avalanche (top) and schematic plot of the CMS drift tube detecting a muon (bottom) are shown above.

sampling calorimeter has two components: absorber and scintillator. Absorber is generally made of metals and produces the shower. The scintillator collects a fraction of the total energy from the shower (visible energy) and then corrects the visible energy back to the total energy based on the light collection efficiency. On the other hand, the homogenous calorimeter collects all the energy deposited. Its material producing the particle shower also measures the energy deposition.

3.5.1 ECAL

The CMS ECAL is made of lead tungstate ($PbWO_4$) crystal and is a homogeneous type calorimeter. High energy electrons and photons interact with the CMS ECAL and undergo bremsstrahlung to produce electron, positron and photons and deposit energy to the ECAL. It has an acceptance coverage of $|\eta| < 1.48$ with a high granularity of $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ in the barrel region and $1.5 < |\eta| < 3.0$ in the endcap region. In addition, the ECAL has an excellent energy resolution of $\frac{\Delta E}{E} = \frac{2.83\%}{\sqrt{E}} \oplus \frac{12.0\%}{E} \oplus 0.26\%$ where E is in the unit of GeV [191] to precisely measure the energy of electrons and photons. It is capable of identifying electrons and detecting photons, which is crucial for heavy flavor physics studies and photon-jet analysis.

3.5.2 HCAL

The CMS HCAL is a sampling type calorimeter made of 926 tons of steel or brass. Over a million World War II brass shell casements are from the Russian Navy. Hadrons interact with the HCAL brass and steel nuclei and produce hadronic showers. A fraction of the shower energy is sampled by the tiles of plastic wavelength shifting scintillators and transferred readout boxes. Generally, all particles except muons and neutrinos will not be able to penetrate the HCAL. The CMS HCAL system consists of the inner HCAL with barrel (HB) and Endcap (HE), the outter HECAL (HO), and the forward HCAL (HF). The acceptance coverages of HB are $|\eta| <1.39$, $|\eta| <1.26$, $1.31 < |\eta| <3.0$, and $2.85 < |\eta| <5.19$ respectfully. The HO and HB have a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The overall energy resolution of HCAL is

$\frac{\Delta E}{E} \approx \frac{100\%}{\sqrt{E}}$ [192], which is excellent for jet physics studies.

3.5.3 HF

The forward HCAL is a special component of the CMS HCAL system. It is segmented into 36 × 13 towers in the $\eta - \phi$ plane. Figure 3-13 s schematic plot of HF shows schematic and physical views of the CMS HF detector [193]

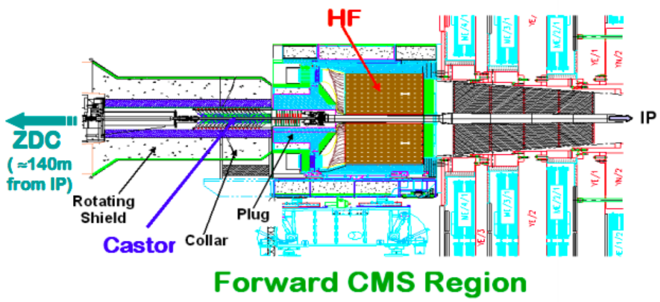


Figure 3-13: The schematic view of the CMS forward region including HF, CASTOR, and ZDC (left) and the physical view of the HF (right) are shown above.

As mentioned above, we have developed the L1 MB trigger based on HF response to select MB events. In addition, in CMS, centrality is defined based on the activities in the HF [194]. The more activity in the HF, the more remnants of colliding nuclei, the more central the collision event. Figure 3-14 shows the determination of centrality range from the HF response

In addition to HF, CASTOR ($-6.6 < \eta < -5.2$) and ZDC ($|\eta| > 8.1$) are also calorimeters which are located at the very forward region [195] as shown above on Figure 3-13. They can help select MB events and trigger ultra-peripheral collision (UPC) events. Figure 3-15 shows the pictures of CASTOR and the ZDC in the very forward direction of the CMS detector

3.6 Relevant Detector Components

In the data analysis of this thesis, the most relevant CMS sub-detectors are the silicon pixel and strip trackers and the muon chamber. We also use HF information to

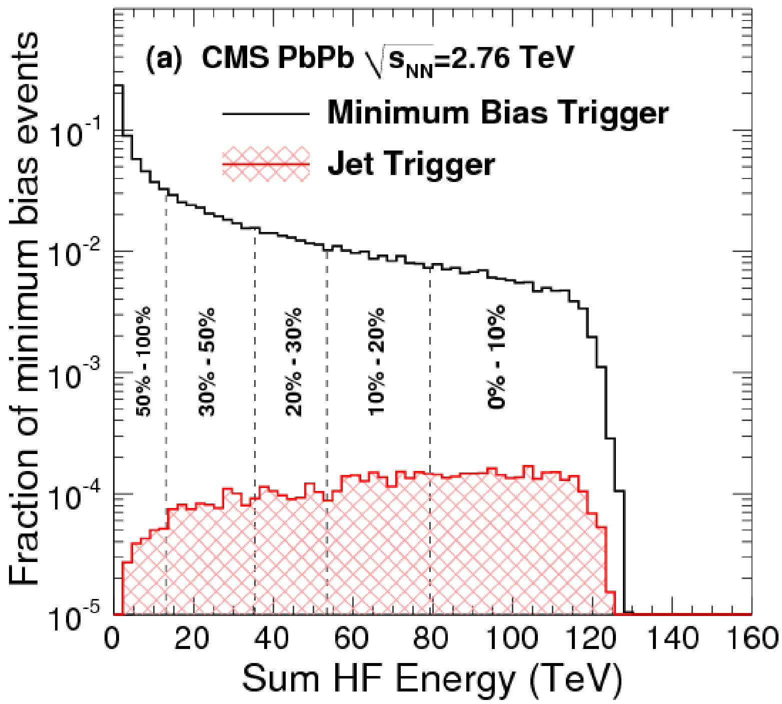


Figure 3-14: The distribution of sum of HF energy using Minimum Biased Trigger and Jet Trigger with the classification of centrality binning is shown above. As we can see, the energy of the HF increase as the collision events become more central, which is within our expectation.

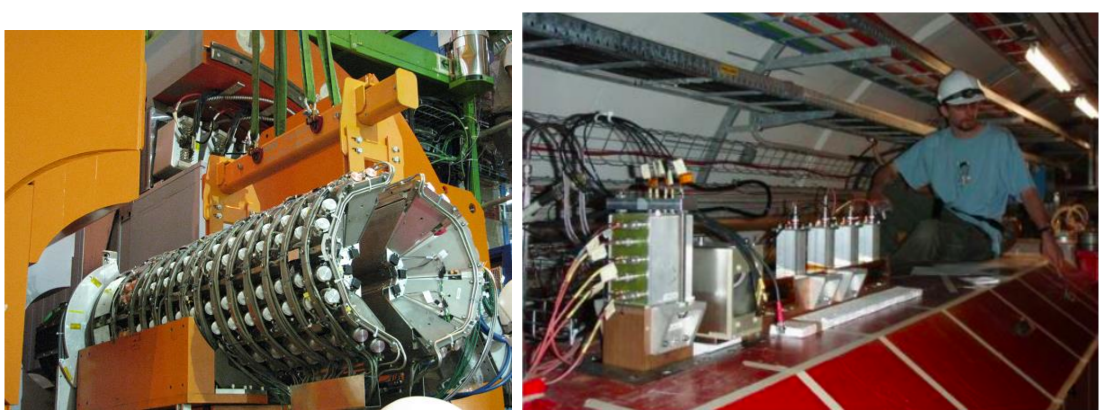


Figure 3-15: The picture of the CASTOR (left) at the CMS underground collision hall and ZDC (right) at 140 m away from the CMS beam interacting point are shown above.

select high quality events. The datasets we used in the analysis are dimuon triggered datasets. We also use the MB trigger samples to estimate the total number of MB events in order to determine the cross section in our analysis. In the next chapter, we will describe in details the physics objects obtained from the detectors and used in our analysis to fully reconstruct B mesons and measure theirs cross sections.

Chapter 4

Reconstructed Objects

The state-of-the-art CMS detector take a snapshot of each event and saves the detailed information of the collisions into datasets. In the datasets, we can access to event information with fully reconstructed objects including hits, tracks, muons, and vertex, which will be crucial for our data analysis to study B meson physics in heavy-ion collisions. Below, we will describe, in principle, how these objects with physical meaning are reconstructed from electronic signal in the CMS detector.

4.1 Event

As mentioned previously, an event is defined as a snapshot of one collision at the LHC. Many particles are produced in the collisions and then decay before they are detected in an event. Theoretically, to obtain the complete information of an event, we only need to know the position and momentum of each particle. Experimentally, we detect final state particles and record their kinematics. In high energy physics experiments, the particles reaching the detectors are $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, p, n, K_L^0, γ. All other particles already decayed into these particles before they can be detected. In order to study them, they need to be reconstructed. Historically, this is used to be done by fast camera with high resolution. The Figure 4-1 shows a famous Ω^{-} baryon (Strangeness -3: sss) event reconstructed from one of picture taken in the bubble chamber [?].

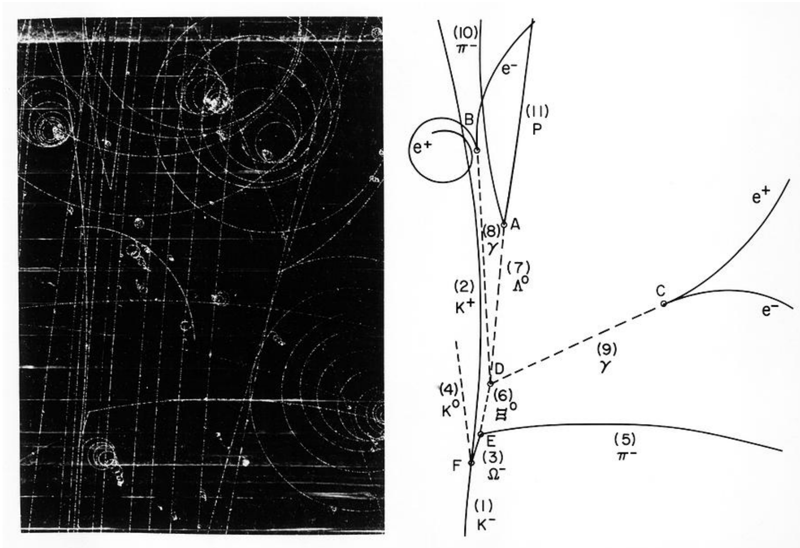


Figure 4-1: The bubble chamber picture of the an Ω^- baryon reconstructed from an event: $K^-p \to K^0K^+\Omega^- \to \Xi^0\pi^- \to \Lambda^0\gamma\gamma \to \pi^-p$ taken from the group led by Nicholas Samios at BNL is shown above.

Nowadays, the high speed electronics and semiconductor technologies have advanced. With the development of computing, detector hardware and readout electronics, high energy physics experiments are able to collect many events with higher precision of measurements. For instance, the CMS experiment has an event trigger rate of 100kHz, which corresponds to a rate of 100000 events per second with 100 GB/s information [196]. Experimental data have become more digital and abstract instead of pictorial and intuitive. All events information is stored at a file format instead of a photograph. Physicists use computer to read the experimental data and develop software to perform analysis of each event, extract the physics information from the analysis, and interpret the physics results.

In the following subsections, for simplicity, I will explain the reconstructed objects of events with only one charged particle.

4.2 Hit

All reconstructed objects start from hits as the energy deposition of particle passing through the detectors. Here I will explain the concept of hits based on CMS silicon pixel tracker. Figure 4-2, the schematic view of a chip with silicon pixels in the CMS tracker

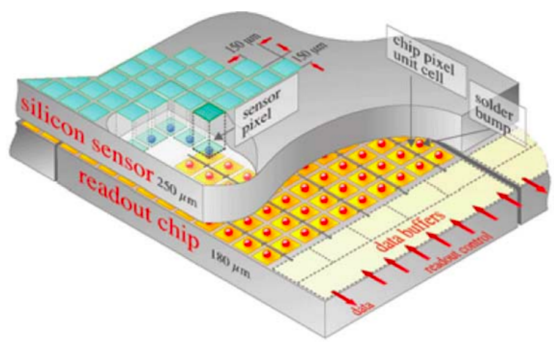


Figure 4-2: The schematic plot of a CMS silicon chip with pixel sensor is shown above.

When a charged particle pass through a layer of the CMS silicon pixel detector, we can look at the charges collected by each pixel on that layer due to the ionization of electron-hole pairs by the high energy charged particle. Ideally, if a particle enter the tracker at a normal angle, only one pixel is fired. However, in reality, its neighboring pixels may also have some response. When the particle enter the tracker with a small angle particularly when the part. Figure 4-3 schematically demonstrates the firing pixel when a particle passing the layer

Here we call each firing pixel as a hit, which is demonstrated above in Figure 4-3 in red. In CMS pixel tracker, the probability of a pixel firing when a charged particle passing through is greater than 99% [185], which means that it is very unlikely a hit is missing when a particle pass through the pixel.

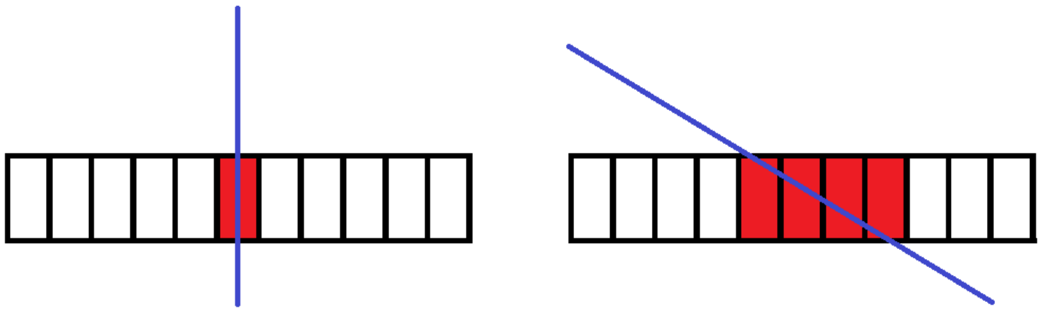


Figure 4-3: The schematic views of a charged particle (blue line) entering the silicon pixel layer (black) at a normal angle (left) and a small tilting angle (right) with the pixels fired (red) are shown above. The left cluster has 1 hit and the right cluster has 4 hits.

4.3 Cluster

Therefore, there should be at least one hit for each layer when a single particle pass through. We call the collection of the adjacent pixel hits in a layer due to one particle as a cluster [185]. Local hits reconstruction algorithm is implemented to obtain clusters. The number of electric charges Q is associated to each hit. We can design an algorithm determine the center of a cluster according to the charges of each hit. A simple algorithm is to calculate the center of gravity of the cluster taking the weighted averaging of the charge and the position of each hit. In this case, for a cluster with a single hit, its position is simply the center of the pixel. For clusters with many hits, we develop a dedicated algorithm to estimate its position [185]. The position of a cluster is a measurement of the particle trajectory.

However, in an event with many particles, the occupancy of each layer will be busy and the clusters will become complicated. The CMS collaboration develop In CMS terminology, the conversion of electronic signal of pixels to clusters is called DIGI.

4.4 Track

4.4.1 Overview of Basic Principles

In a uniform external magnetic field, the trajectory of a charged particle will be a helix in 3 dimensions. Geometrically, five parameters are needed to parametrize a helix. A parametric curve of a helix moving in the Cartesian coordinates moving in the z direction is written as follows

$$x(t) = R\cos(\omega t) + a \tag{4.1}$$

$$y(t) = R\sin(\omega t) + b \tag{4.2}$$

$$z(t) = vt + c \tag{4.3}$$

Therefore, we need at least 3 clusters to determine the all 5 parameters. 3 clusters can determine the radius R and the center of the circle (a,b) and also can determine the straight line in the z-direction. Figure 4-5 shows the helix path of a charged particle in a uniform magnetic field and the fit to determine the center and the radius of the helix.

Moreover, we can determine transverse momentum of the charged particle according to the R fitted from fit to the center of 3 clusters.

$$p_T = qRB \tag{4.4}$$

In general, the charges of the particles produced in the collision and pass through the tracker are $q = e$. Hence, $p_T = eRB$. For p_T in the unit of GeV, R in the unit of meter (m), and B in the unit of tesla (T), we have

$$p_T \simeq 0.3RB \tag{4.5}$$

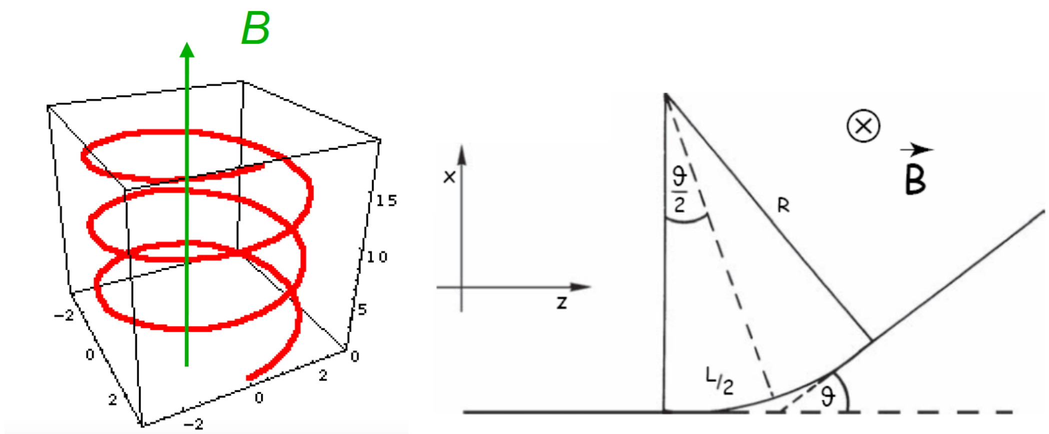


Figure 4-4: The helix motion of a charged particle under a constant and uniform magnetic field $\vec{B}$ pointing in the +z direction (left) and the fit to 3 points to determine the center and the radius of a circle (right) are shown above.

Therefore, as seen above from Figure 4-5, the transverse momentum resolution is driven by the determination of R assuming we have a perfect measurement on the magnetic field B.

According to Figure 4-5 on the right, at high p_T, essentially in parallel, for a 3 cluster fit. In addition, we know that the layers in the pixel track has equal spacing Δr between layers. For CMS pixel tracker, its inner most 3 layers has equal distance $\Delta r_{12} = \Delta_{23} = 2.9$ cm [197].

Hence, we can see that $L/2 = \Delta r$, which assume fixed with no uncertainties. Hence, we have

$$\frac{L}{2} = R \sin\frac{\theta}{2} \tag{4.6}$$

Again, at high p_T, the angle θ will be very small since the radius of the circle $R >> \Delta r$, $\sin\theta \simeq \theta$ and $\cos\theta \simeq 1 - \frac{\theta^2}{2}$. Hence, we can use small angle approximation

$$L = 2R \sin\frac{\theta}{2} \simeq R\theta \tag{4.7}$$

Therefore,

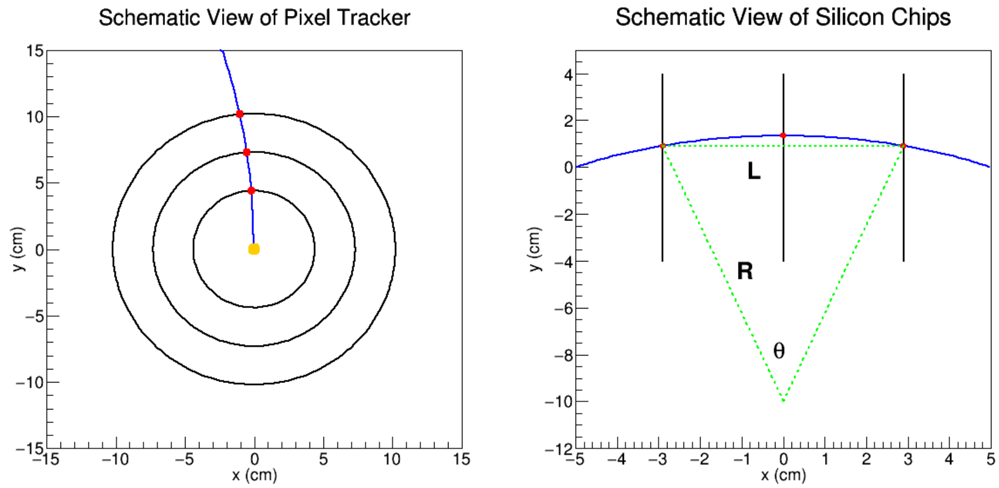


Figure 4-5: A track (blue) initiated from the beam spot (orange) passing through 3 layers of pixel detectors (black) with 3 clusters (red) is shown on the left and the circular fit to the 3 clusters with the definition of R, L, and θ is shown on the right.

$$p_T \simeq 0.3RB = 0.3\frac{BL}{\theta} \tag{4.8}$$

Hence, geometrically, we have

$$s = R - R\cos\frac{\theta}{2} = R(1-\cos\frac{\theta}{2}) = R(1-\cos\frac{\theta}{2}) \simeq \frac{L}{\theta}\{1-[1-\frac{1}{2}(\frac{\theta}{2})^2]\} = \frac{L\theta}{8} = \frac{0.3BL^2}{8p_T} \tag{4.9}$$

Thus, the uncertainties on both sides go as

$$\sigma_s = \frac{0.3BL^2}{8p_T^2}\sigma_{p_T} \tag{4.10}$$

Hence, the transverse momentum resolution $\frac{\sigma_{p_T}}{p_T}$ is given by

$$\frac{\sigma_{p_T}}{p_T} = \frac{8\sigma_s}{0.3BL^2}p_T \tag{4.11}$$

Here, σ_s is effective the position resolution of the silicon pixel detector. We can see that the transverse momentum resolution gets worse as p_T increases in the high p_T region. Figure 4-6 shows the $\frac{\sigma_{p_T}}{p_T}$ as a function p_T

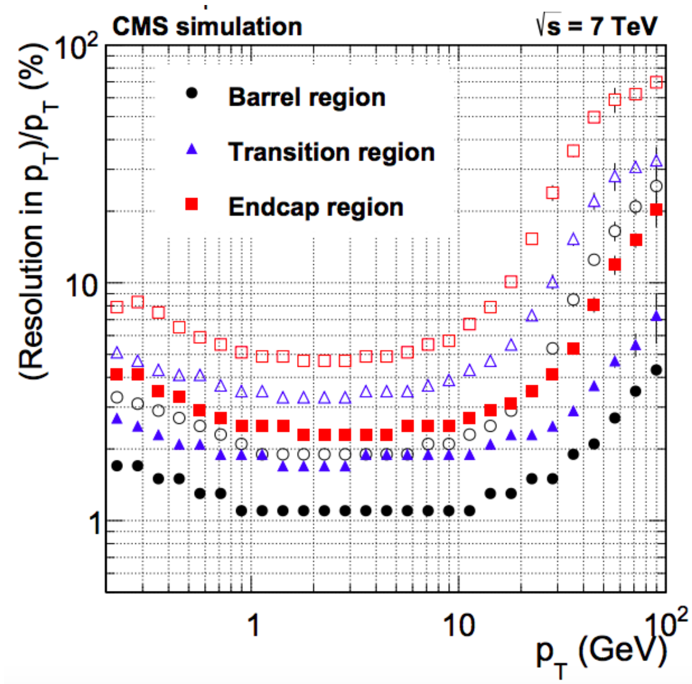


Figure 4-6: The transverse momentum resolution $\frac{\sigma_{p_T}}{p_T}$ of a track as a function of transverse momentum p_T is shown above.

We can see that a good agreement with linear growth of $\frac{\sigma_{p_T}}{p_T}$ for $p_T > 20$ GeV/c in the high p_T region.

Longitudinally, p_z can be determined by the p_T and the angle $\Delta\theta$ in the transverse direction

$$p_z = \frac{\Delta z}{\frac{R\Delta\theta}{p_T}} = 0.3B\frac{\Delta z}{\Delta\theta} \tag{4.12}$$

At this point, we have obtain the trajectory with the complete kinematic information about a particle except its mass which will require particle identification in order to determine.

4.4.2 CMS Tracking Algorithm

Because the CMS silicon tracker has 3 pixel and 10 strip layers, a charge particle passing through all 13 layers should leave 13 clusters, which is much more than required to determine the helix. Moreover, in reality, collision events at the LHC, many tracks are produced at multiple vertices. In collider physics, we call use the concept of pileup events (PU), which is defined as events with more than one vertices. Figure 4-7 shows the number of vertices and the number of tracks in pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV

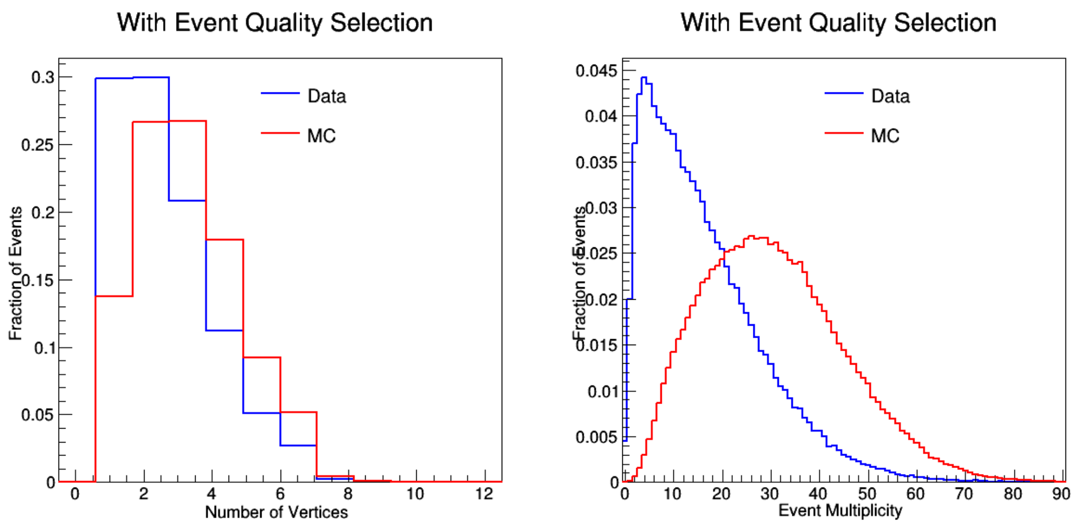


Figure 4-7: The Data (blue) and MC (red) of the number of primary vertex distribution (left) and event multiplicity (right) are shown above. We can see that an event could be more than 1 vertices with more than 100 tracks, which make it very challenging to perform tracking.

Hence, the CMS collaboration has developed the state-of-the-art tracking algorithm to reconstruct the paths and primary vertices of the collisions from the electronic readout signals. CMS tracking algorithm employs the Combinatorial Track Finder (CTF), an adaptation of the combinatorial Kalman filter [186–188], which in turn is an extension of the Kalman filter [189] to allow pattern recognition and track fitting to occur in the same framework. The collection of reconstructed tracks is produced by multiple passes (iterations) of the CTF track reconstruction sequence, in a process called iterative tracking [185]. The CMS tracking workflow and its performance are shown in Figure 4-8 and Figure 4-12

Track Seeding: After obtain the clusters and reconstruct the hits, the tracking is in the track seeding stage. A dedicated seeding algorithm is designed to select the clusters, either a triplet or a pair, from the pixel layers and other combinations of

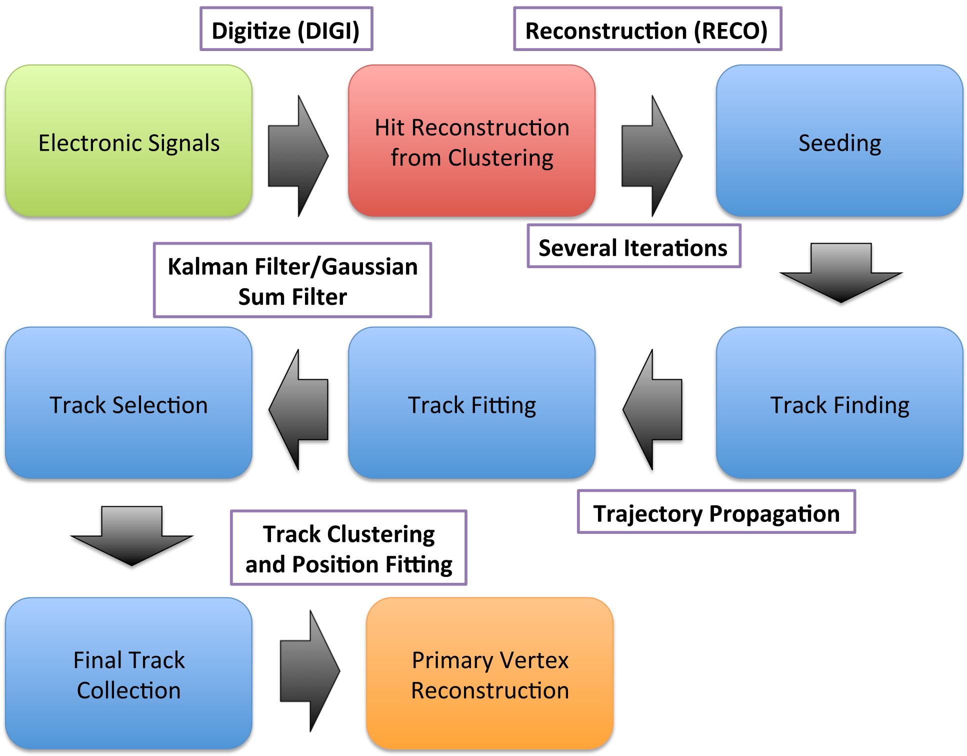


Figure 4-8: The schematic block diagram of CMS tracking workflow is shown above.

pixel and strip layers before fitting [185]. After these steps, preliminary fits to the seeds named trajectory seeds are created.

Track Finding: Then, it moves on to the track finding stage. A six-step iteration process, which includes navigation, hit search, hit grouping, and trajectory update, is implemented with the application of CFT algorithms based on Kalman Filter to build track candidates. A schematic overview of the track finding process is shown below Figure 4-10

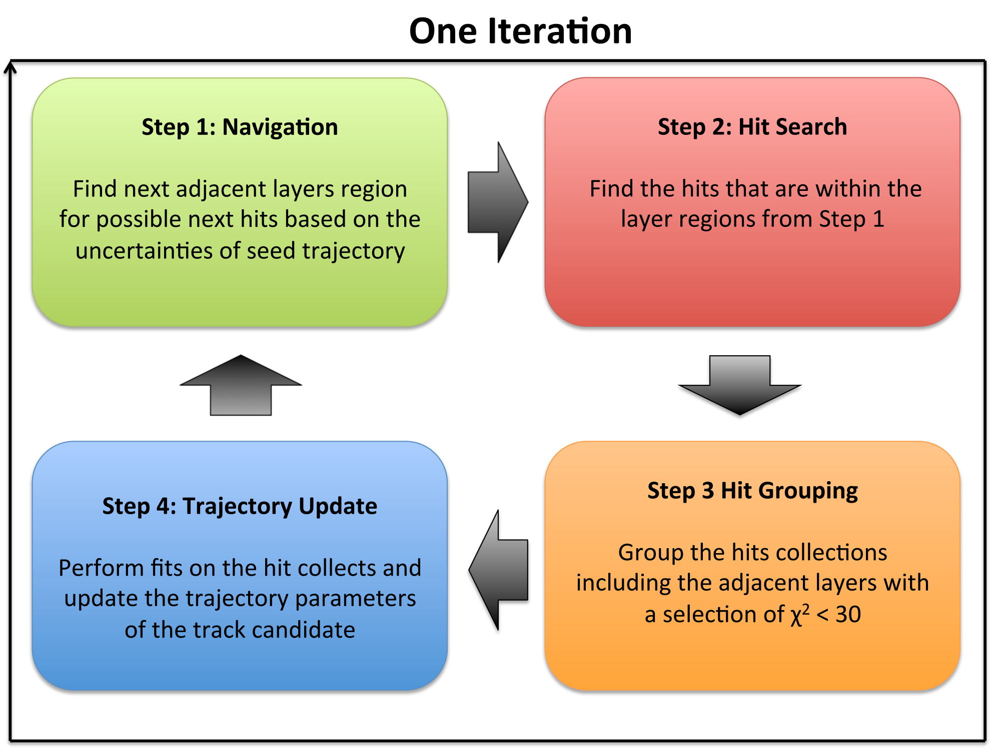


Figure 4-9: The four steps of CMS track finding workflow (left) and the schematic demonstration of each step (right) are shown above.

Track Fitting: Next, the tracking is in the stage of track fitting. Kalman filter [189] is applied to improve fitting performance. It starts from the innermost location with typically four hits [186–188]. When extrapolating the trajectory from one hit to the next, the filtering and smoothing procedure is carried out with a Rugga-Katta propagator to obtain the best precision. $\chi^2 < 20$ is required of each fit in order to improve its precision and reject fake tracks. Figure 4-11 schematically shows how the

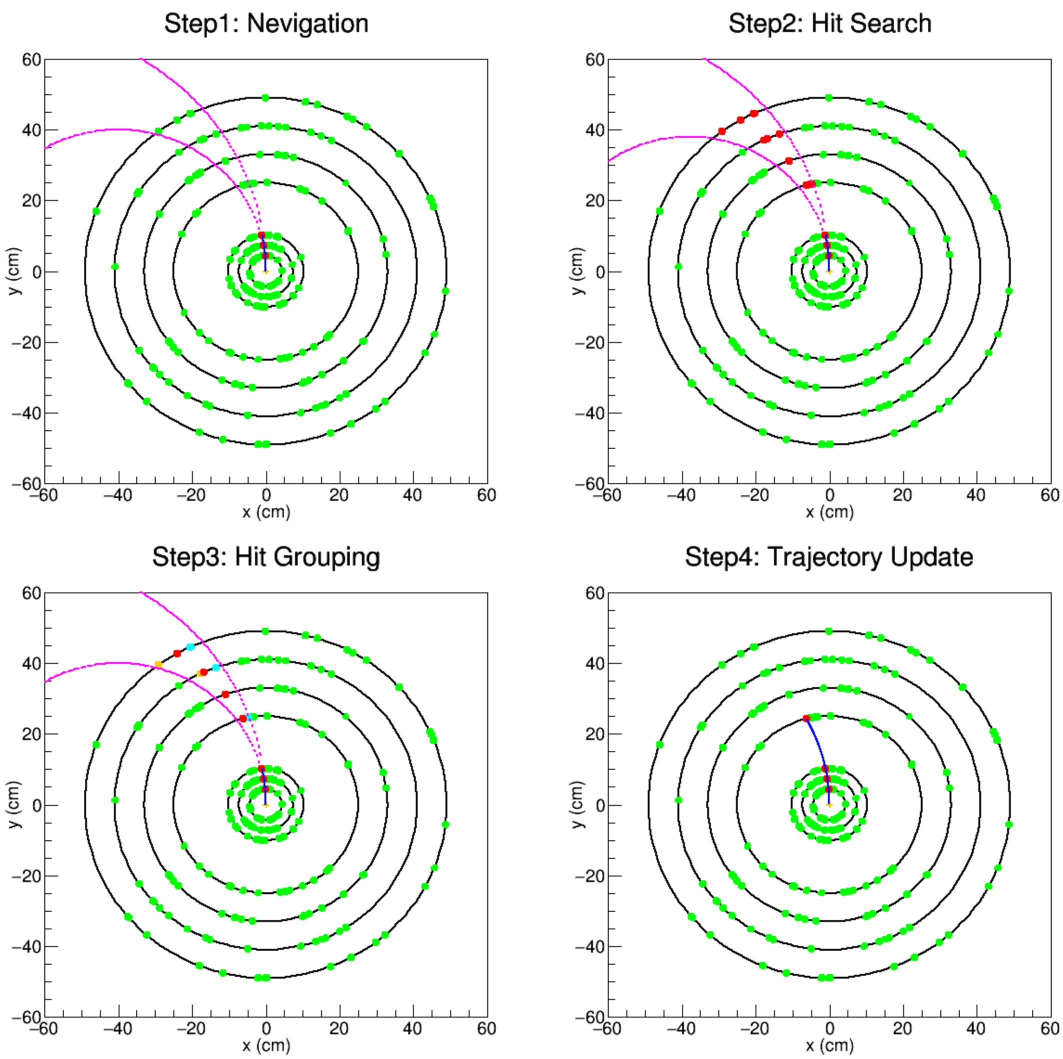


Figure 4-10: The four schematic plots demonstrating each of the four steps for track finding are shown respectfully above.

Kalman filter fit along with Rugga-Katta propagator is applied in the iterative fitting algorithm for CMS tracking

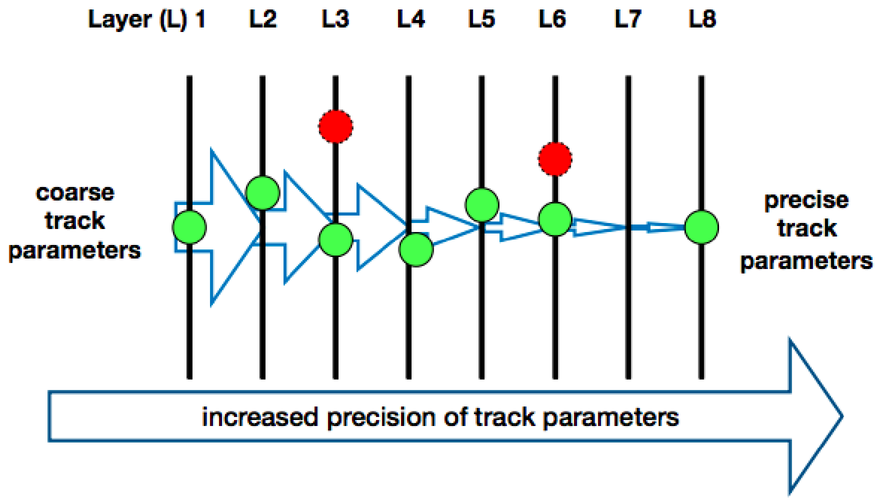


Figure 4-11: The schematic demonstration Kalman filter along with Rugga-Katta propagator to improve the tracks fitting is shown above.

Track Selection: Subsequently, the tracking is in the stage of track selection. At this point, we have already obtained a preliminary track collection of one event. To improve the track quality and reject fake tracks, further selection based on the track properties will be applied. The following selection criteria are applied to select high quality tracks [185]

- Minimum number of layers in which the track has at least one associated hit
- Minimum number of layers in which the track has an associated 3-D hit
- Maximum number of layers that has no associate hits
- $\chi^2/dof < \alpha_0 N_{layers}$
- $|d_0^{BS}|/\sigma_{d_0(p_T)} < (\alpha_1 N_{layers})^\beta$
- $|z_0^{PV}|/\sigma_{z_0(p_T,\eta)} < (\alpha_2 N_{layers})^\beta$
- $|d_0^{BS}|/\delta d_0 < (\alpha_3 N_{layers})^\beta$

- $|z_0^{PV}|/\delta z_0 < (\alpha_4 N_{layers})^{\beta}$

Here, α_n and β are configurable constant depend on the selection efficiency and purity requirements. d_0^{BS} is the closest transverse distance of the track to the beam spot and δd_0 is its associated error. z_0^{PV} is the distance along the beam-line from the closest pixel vertex and δz_0^{PV} is its associated error. Hence, $|d_0^{BS}|/\sigma_{d_0(p_T)}$ and $|z_0^{PV}|/\sigma_{z_0(p_T,\eta)}$ are expressed in terms of significance. $\sigma_{d_0(p_T)}$ and $\sigma_{z_0(p_T,\eta)}$ are essentially the associated errors of d_0^{BS} and z_0^{PV} parametrized by track p_T and η.

Final Track Collection: Finally, after we apply the selections, we have obtained a final track collection for one event. Figure 4-12 shows the general performance of CMS tracking algorithm

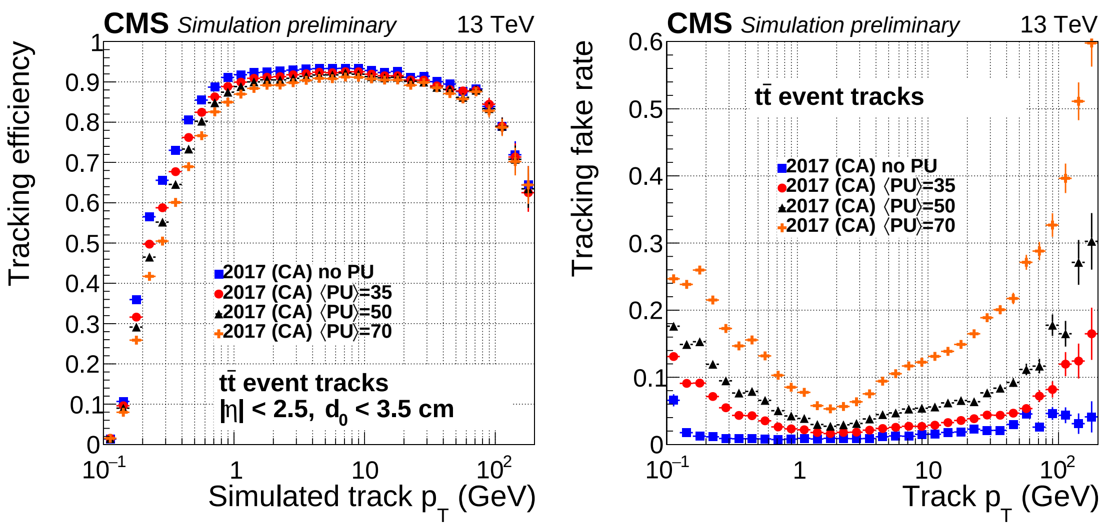


Figure 4-12: The CMS tracking efficiency (left) and fake rate (right) as a function of p_T from simulations of $t\bar{t}$ events at 13 TeV with different pileup conditions are shown above.

We should note that a modified version of Kalman filter named Gaussian Sum Filter [198] is applied to improve the tracking performance of electrons [185].

4.5 Muon

The muon in the tracker uses an essentially the same tracking algorithms as other charged particles [185]. Tracking performance of muon is excellent. For isolated muons with $1 < p_T < 100$ GeV/c, the tracking efficiency is $> 99\%$ over the full η-range of tracker acceptance and does not significantly depend on p_T while the fake rate is negligible [185]. We can require hits on the outer most muon chambers to identify muons because other charge particles will be stopped by the calorimeter and should not be able to enter the muon system as shown on Figure 3-11. The muon reconstruction workflow with the muon chambers are shown below in Figure 4-13

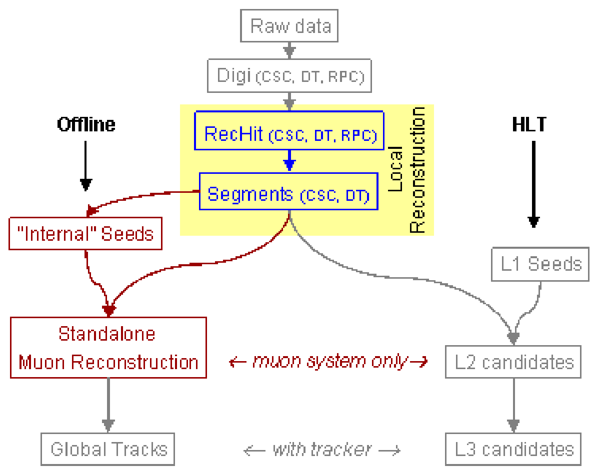


Figure 4-13: The schematic block diagram of muon reconstruction in the CMS muon system is shown above.

In addition, since mouns also deposit some energy to the ECAL and HCAL, we can also access calorimeter information for the muons. Therefore, the CMS muon system has excellent capabilities of detecting, identifying, and reconstructing muons, which is crucial for heavy flavor physics studies.

In CMS terminology, there are many types of muons based on their selection requirement in reconstruction. They are classified as follows:

- **Standalone Muons:** the muon segments reconstructed from muon chambers only.

- **Tracker Muons:** the muon segments reconstructed from tracker only but also valid with the information from calorimeter and muon systems

- **Global Muons:** the muons reconstructed from the fits on the hits of both trackering and muon systems

- **RPC Muons:** the muons reconstructed with both inner tracker and the resistive plate detector only

- **Calorimeter-based Muons (Calo Muons):** the muons reconstructed with both inner tracker and calorimeters

The relationship between stand alone muons, tracker muons, global muons, and calo muons are shown below in Figure 4-14

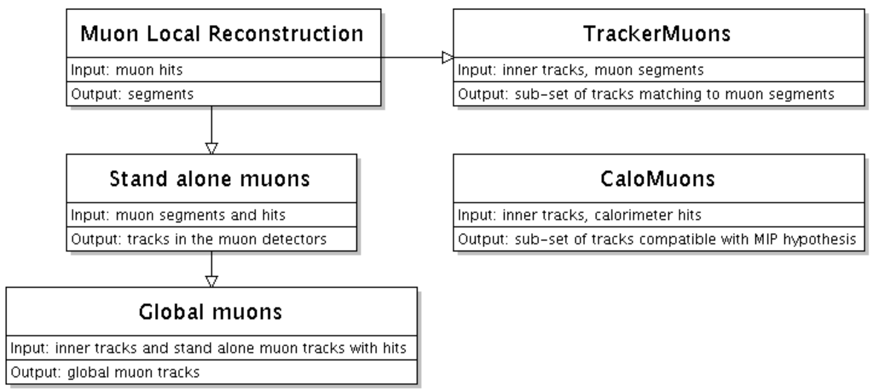


Figure 4-14: The relationship between different reconstructed muon in CMS is shown above

For physics analysis, addition selections on muons including trigger, identification, and acceptance will be applied. We will further discussed them in the analysis chapter.

4.6 Vertex

In particle physics, the term vertex is similar to a vertex in Feynman diagram where old incoming particles are destroyed and new outgoing particles are created via an interaction at a space-time point. This has already shown in Figure 4-1 of the Ω^- baryon event where we can clearly see the vertices of Ω^- baryon creation and decay in the reconstructed plot on the right.

4.6.1 Primary Vertex

By definition, in collider experiments, the primary vertex is assume where the hadron interact. Therefore, all particles produced in the collisions in that event originate from the primary vertices. With the final track collection for each event, assuming all the tracks are promptly produced at a given interaction point, we can determine the primary vertices by selecting the tracks, performing track clustering, and fitting for the position of each vertex using its associated tracks [185]. The selection criteria for the track to perform primary vertex are as follows:

- $|d_0^{BS}|/\Delta d_0 < 5$
- $N_{pixel} > 2$
- $N_{strip} > 5$
- $\chi^2/dof < 20$

The selection above make sure the tracks used have high quality and are indeed come from primary vertices. The deterministic annealing algorithm [190] is track clustering algorithm that CMS is currently using. An iterative process of minimize the annealing function of longitudinal distance between the tracks and the vertices gradual temperature reduction until dropping to the critical temperature is carried out to determine the number of vertices and their z coordinates [185]. A track can be used for more than one vertices during tracking clustering. Figure 4-15 is a *pp* collision event display of the CMS detector with reconstructed tracks and primary vertices

After that, we have determined the vertex candidates with z coordinates. Then, for the vertices candidates with at least two tracks, adaptive vertex fitter [199] is applied to compute the best estimate of vertex parameters, including its x, y and z position, covariance matrix, and parameters characterizing the fitting performance such as the χ^2, n_{dof}, vertex fitting probability, and the weights of the tracks used in the vertex. At this point we have obtain the complete information of an event with

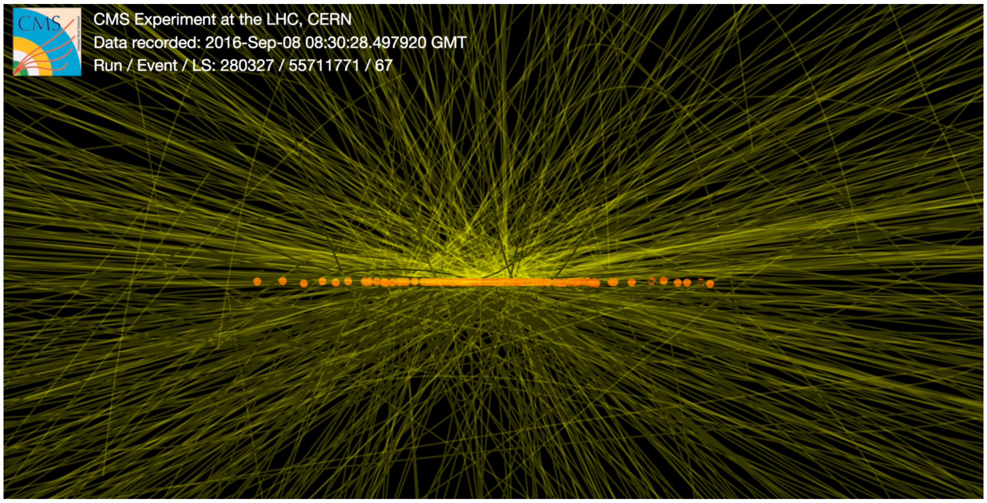


Figure 4-15: A pp collision event display of the CMS detector with reconstructed tracks (green curves) and primary vertices (yellow dots) is shown above.

a final collection of tracks with best reconstructed vertices. The track and primary vertex information of each event will be saved in datasets for further physics analyses.

4.6.2 Secondary Vertex

Another object we should mention, particularly in the context of this thesis, is secondary vertex. Again, as seen from Figure 4-1, there are many vertices in the Ω^- baryon event in its decay chain. We can the displaced vertex from the primary vertex due to the decay of a short-life particle produced at high energy as the secondary vertex. Figure 4-16 below is schematic plot of the decay topology of a prompt D^0 decay via the channel: $D^0 \to K^-\pi^+$ showing the primary vertex and secondary vertex

The decay length is defined as the distance between primary and secondary vertex. In LHC collisions, since the energy is very high, the particle produced from the primary vertex generally has a large momentum, which correspond to a large Lorentz gamma factor $\gamma\beta \simeq 5$. For B mesons produced at the LHC, we estimate their length a $\gamma\beta c\tau \simeq 3$ mm due to its long lifetime $c\tau = 500$ μm [4]. Therefore, the B meson secondary vertex is well seperated from the primary vertex and could even even viewed by eye. The relatively long B meson decay lengths make them ideal candidates to be fully reconstructed and study thanks to the excellent vertexing and tracking

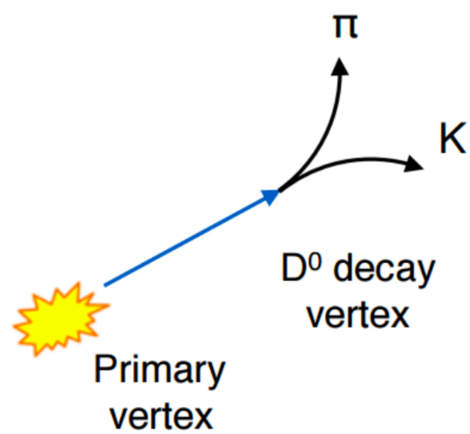


Figure 4-16: The decay of $D^0 \to K^-\pi^+$ with the definition of secondary vertex is shown above

capabilities of the CMS detector.

Finally, we can also call the secondary vertex for a particle as the reconstructed mother particle's vertex. For instance, $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$. We can identify 3 secondary vertices. We call them B_s^0, J/ψ, and ϕ vertices.

Chapter 5

Data Analysis

5.1 Analysis Tools

In this thesis, the analysis tools we use is based on Macbook computers, and CMS high computing, linux shell script, C++ programing language, and ROOT analysis package for high energy physics experiments. The machine performing the analysis includes CERN lxplus, MIT Tier-2 submit, Data file are saved as ROOT files format. Data samples are processed with crab CERN lxplus and MIT submit condor job submission with parallel computing framework. The core software for the data process and analysis is the CMS software (CMSSW_10_3_4). The codes for the analysis software is documented on github and gitlab. Throughout the analysis, we use Poison statistics to interpret the statistical uncertainties of the data. In the limit of high statistics, it is approximately equivalent to Gaussian Statistics. In the limit of low statistics, it is different from Gaussian Statistics. These tools are crucial for me to finish the analysis and present the results.

5.2 Analysis Strategies

5.2.1 Physics Goals

The exclusive productions of b hadrons in different collision systems are necessary to study beauty energy loss and hadronization mechanisms. In this thesis, we propose to fully reconstruct B_s^0 ($b\bar{s}$) and B^+ ($\bar{b}u$) mesons in pp and PbPb collisions at $\sqrt{s_{NN}} =$ 5.02 TeV with the CMS experiment. We aim at measuring precisely their cross section, yield ratios, and nuclear modification factor of fully reconstructed B_s^0 and B^+ mesons via the decay channels of to investigate beauty quark production and hadronization in vacuum and QGP. We hope to have a conclusive measurement of B_s^0/B^+ ratio to pinpoint the effect of strangeness enhancement in QGP on beauty hadronization and test theoretical model calculations [167] with both fragmentation [] and quark coalescence [] mechanisms. Therefore, we would like to present our experimental measure over a width range of p_T and centrality. We are particularly interested in the very low p_T region where the slow moving beauty quark will pick up the nearby light quarks in the QGP while such mechanism is not expected to occur in the vacuum []. The R_{AA} down to low p_T will constrain understand beauty quark energy mechanism in the QGP medium. These will studies will be crucial for us to understand beauty quark diffusion coefficient and probe the inner workings of QGP in order to provide insights with the one of the open question in high energy nuclear physics.

5.2.2 General Workflow

Figure 5-1 shows workflow we designed to fully reconstructed B_s^0 and B^+ mesons from final state particles via the exclusive decay modes. The B_s^0 is fully reconstructed from the decay channel of $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$, which has a fragmentation fraction of $f(b \to B_s^0) = 0.103$ and a decay branching fraction $BR(B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-) = 3.17 \times 10^{-5}$ $B^+ \to J/\psi K^+ \to \mu^+\mu^-K^+$ with the CMS detector. The B^+ is fully reconstructed from the decay channel of $B^+ \to J/\psi\phi \to \mu^+\mu^-K^+$, which has a fragmentation fraction of $f(b \to B^+) = 0.401$ and a decay branching fraction

$BR(B^+ \to J/\psi\phi \to \mu^+\mu^- K^+ K^-) = 6.02 \times 10^{-5}$ $B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+$ with the CMS detector.

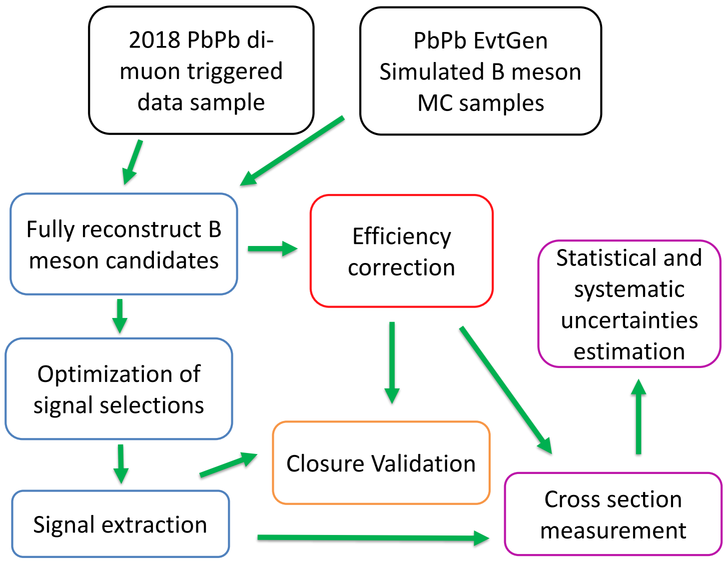


Figure 5-1: The block diagram of the workflow with major steps for both B-meson cross section measurement is shown above.

Figure 5-2 shows pictorially the decay topology of fully reconstructed B meson and our reconstruction strategies

In this thesis, we propose to measure B_s^0 and B^+ cross section in the p_T bin of [7, 10, 15, 20, 50] GeV/c within centrality [0, 90] and centrality bin of [0, 30, 90] with p_T [10, 50]. The rapidity range of B-meson measurements are confined in $B|y| < 2.4$.

5.2.3 Technical Challenges

Despite the excellent muon, tracking, vertexing capabilities of the CMS detector, there are still many challenges for the analysis. Below is a list of challenges in the B-meson analysis

- The small B meson decay branching ratio, which on in the order of 10^{-5}, and limited luminosity of the sample: $S \downarrow$

- Huge combinatorial background without hadron particle identification, particularly in PbPb collisions and at low p_T: $B \uparrow$

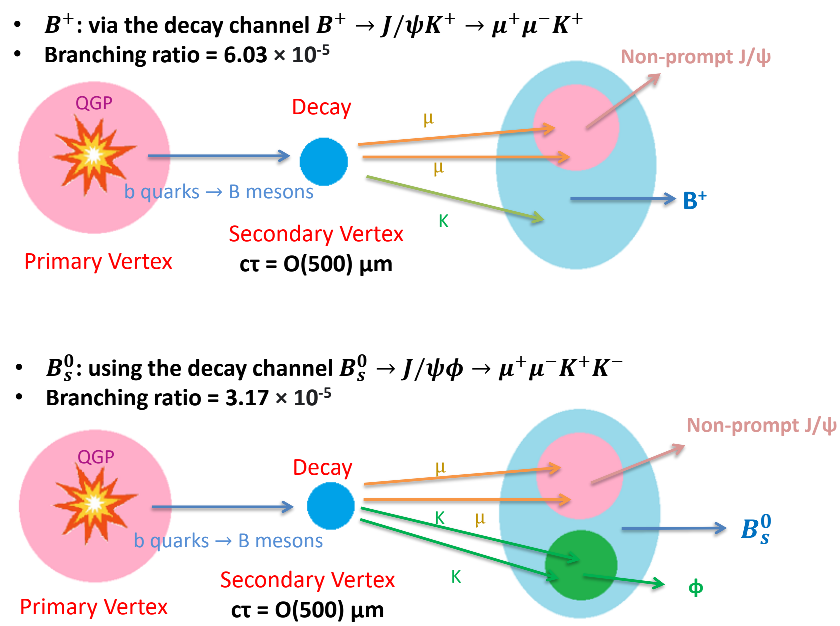


Figure 5-2: The strategies to fully reconstruct B_s^0 and B^+ in the selected exclusive decay modes are shown above.

- Low muon acceptance of at very low p_T: $S \downarrow$

- High fake track rate with low tracking efficiency at very low p_T: $B \uparrow$

Here, S stands for signal and B stands for background. These factors will all lower the signal-to-background ratio, which makes challenging to fully reconstruct B mesons, particularly at very low p_T. In this thesis, to reduce the signal to background ratio and the systematic uncertainties, we will employ a novel approaching machine learning with multivariate analysis and the elaborated single particle efficiency correction method to perform the measurements.

5.3 Analysis Samples

5.3.1 Dimuon Triggered Datasets

In this part of the thesis, I focus on the studying beauty production and hadronization in QGP. Therefore, this analysis is performed using the 2018 PbPb data at $\sqrt{s_{NN}}$=5.02 TeV, which has an integrated luminosity of 1.7 nb^{-1}. The analysis uses the dimuon primary datasets (*DoubleMu* PD). The full name of the used datasets and their corresponding luminosity can be found in Table 5.1.

Table 5.1: List of PbPb HLT datasets and triggers with the corresponding integrated luminosities used in the analysis.

System	Primary dataset	Trigger	Luminosity
PbPb	/HIDoubleMuonPsiPeri/HIRun2018A-04Apr2019-v1/AOD	HLT_HIL3MuONHitQ10_L2MuO_MAXdR3p5_M1to5_v1	522 nb^{-1}
PbPb	/HIDoubleMuon/HIRun2018A-04Apr2019-v1/AOD	HLT_HIL3MuONHitQ10_L2MuO_MAXdR3p5_M1to5_v1	1124 nb^{-1}
PbPb	Combined All		1.657 ($\sim$ 1.7) nb^{-1}

The details of the dimuon trigger selection to collect the data sample is explained in 2.2.5. In addition, an Muon JSON to select good luminosity sections in the PbPb dataset is applied. Both B_s^0 and B^+ data come from this sample. However, in the later stage, the B meson candidates are saved in different channels based on the reconstruction.

5.3.2 Monte Carlo Simulations Samples

Dedicated PbPb B_s^0 and B^+ samples are generated in order to estimate the acceptance and selection efficiencies, to study the background components, and to evaluate systematic uncertainties. PYTHIA8 Tune CUETPM8 [157, 200], set to generate inclusive (all quark/antiquark, as well as gluon initiated) QCD processes, was used to generate at 5.02 TeV the signal. Several preselections at the generation steps are applied in order to optimize the generation process and conserve resources.

For B_s^0, only signal events were kept with at least one B_s^0 (forced to decay through the channel $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$ by means of the EVTGEN package [201]), with $p_T > 5.0$ GeV/c, and $|\eta| < 2.4$. In addition, the J/ψ and ϕ meson, are forced to decay in the two muons and two kaons respectively. Final state radiations are generated using PHOTOS [202]. The selected signal B mesons PYTHIA8 events were embedded into a PbPb background simulated with the HYDJET (version 1.8, tune "Drum" for the prompt and non-prompt J/ψ MC and tune "Cymbal5Ev8" for the B_s^0 signal MC) [203] event generator.

For B^+, similar requirements for MC generation are applied except a different decay channel $B_s^0 \to J/\psi K^+ \to \mu^+\mu^-K^+$ is used.

For B_s^0 and B^+, around fifty thousand events were generated in 5 $\hat{p}_T$ bins, with boundaries of $\hat{p}_T > 0$, 5, 15, 30, 50, in both signal only, and embedded samples. The high $\hat{p}_T$ selections are used to enrich the high p_T B meson statistics in order to perform efficiency correction.

We should note that in there are two components in the MC sample. The truth information about the particles generated in the simulation, which is call generated (GEN), and the reconstructed one smeared according to the CMS detector effects, which is called reconstructed (RECO). Due to the nature of MC generation, we will need to reweigh on MC in order to model the data.

In addition to the B_s^0 and B^+, a sample of inclusive b hadron to J/ψ (non-prompt) MC is also simulated to study the possible background contribution to the B-meson analysis due to

5.3.3 $\hat{p}_T$ Reweighing

As we mention above, different $\hat{p}_T$ cut is applied to generate the MC samples. When merging the samplings, a $\hat{p}_T$ weight based on beauty production cross section is required to apply to the MC in order to obtain a smooth distribution that can model the real data. Figure 5-3 shows the generated p_T (Gp_T) distribution of J/ψ, B^+ , and B_s^0 before and after applying the $\hat{p}_T$ weight:

5.3.4 RECO B-meson p_T Reweighing

Then, we also check if this smooth Gp_T shape in fact correspond to a good agreement between the data and MC in the RECO side. Therefore, we take the ratio of the normalized data raw yield to the normalized MC raw yield and perform a variety of functions to fit the distribution. In our studies, we use Linear ($y = p_0 + p_1 x$), Quadratic ($y = p_0 + p_1 x + p_2 x^2$), Linear + Inverse ($y = p_1 x + \frac{p_2}{x}$), Linear + Square Root ($y = p_0 + p_1 x + p_2 \sqrt{x}$), Linear + Log ($y = p_0 + p_1 x + p_2 \log x$). The data vs MC raw yield shape and our fitting results on spectra ratio are show as follows 5-4 and 5-5 for B_s^0 and B^+ respectfully

5.3.5 Centrality Reweighing

Because the MC simulations employ PYTHIA embedded into a PbPb background simulated, they do not model the centrality of nucleus-nucleus collision well. Therefore, the MC simulations are also reweighed in order to match the centrality distribution in data. In the middle panel of Figure 5-6, the centrality distribution of the MC simulation (red) is compared to the one in data (blue), before the re-weighting. Each unit (hiBin) on the x-axis represents 0.5% centrality. The number of binary collisions N_{coll} was used as the weight to scale the MC centrality, and the distribution presented in the right panel of Figure 4 was obtained.

A better centrality agreement between the data and the MC is seen after the reweighing process.

5.3.6 PV_z Reweighing

In addition to above pt shape and centrality reweighing, there must be a primary vertex z position (PV_z) reweighing due to incorrect modeling of the primary vertex location and resolution in the MC simulation. In fact, it is known that the MB samples use for embedding for PbPb signal MC samples (with Cymbal5Ev8 tune) has a PV_z offset. Also, the offsets between data and MC in the X and Y directions are observed in the 2018 PbPb collisions. To remedy this, a Gaussian fit is applied to both the data and MC PV_z distributions, as showed in Fig. 5-7. The black markers represent the distribution points for MC (left), and data (right), while the red line represents the fit result. Then, the ratio between the two fit results is taken as the weighting function. The result after this weighting can be found in Fig. 5-7 But we should note that this analysis is not sensitive to the absolute value of the PV position because the reconstruction of the B-meson rely only on the relative distance between PV and B-meson reconstructed vertex which will be presented in the later sections.

A almost perfect MC-data agreement after PV_z reweighed is observed above. After these standard reweighing procedures, the residue disagreement between MC and data will be considered as a source of systematic uncertainties.

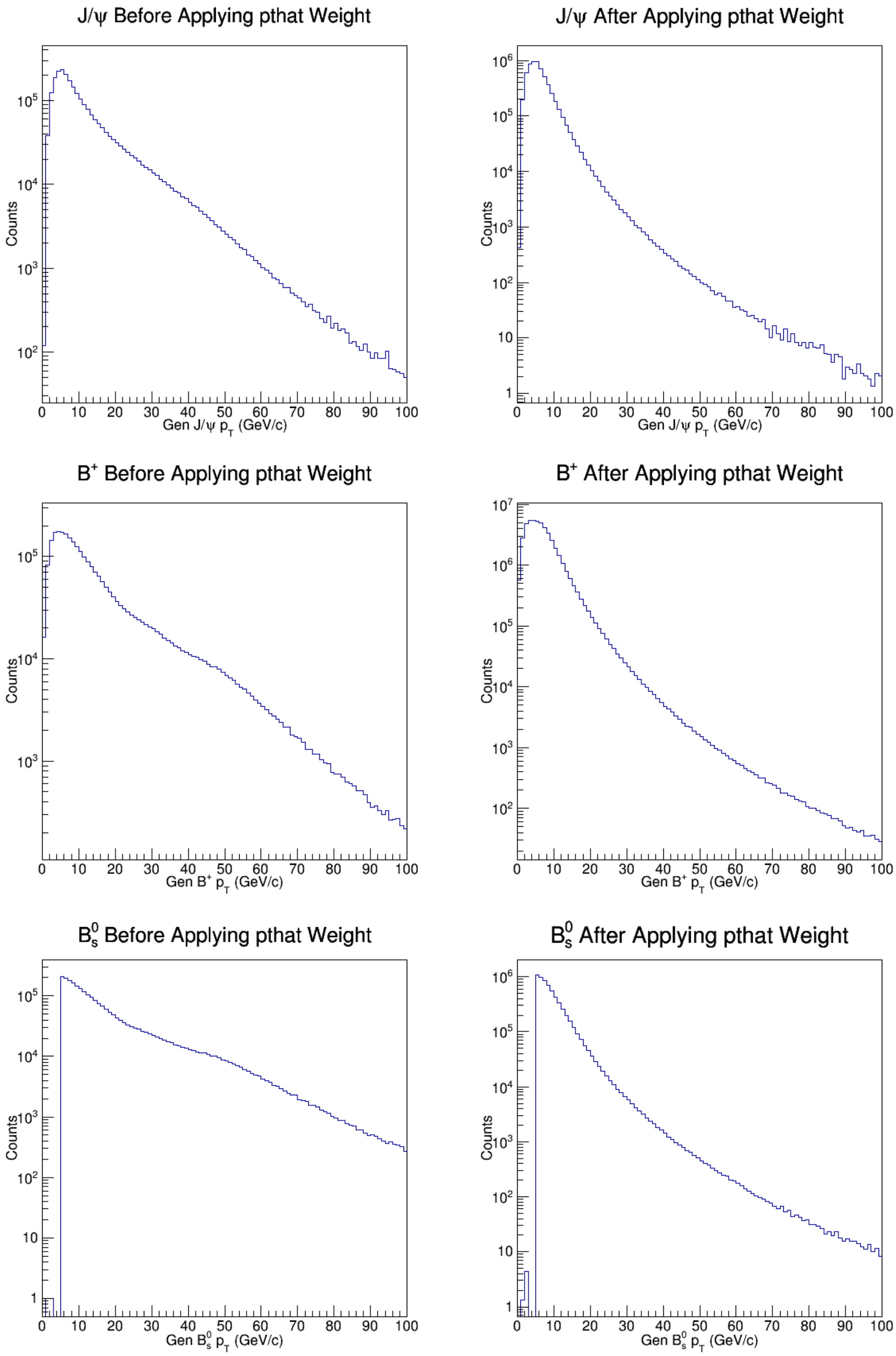


Figure 5-3: J/ψ generated p_T distribution before (upper left) and after (upper right) $\hat{p}_T$ reweighing, B^+ generated p_T distribution before (middle left) and after (middle right) $\hat{p}_T$ reweighing, and B_s^0 generated Gp_T distribution before (lower left) and after (lower right) $\hat{p}_T$ reweighing are shown above.

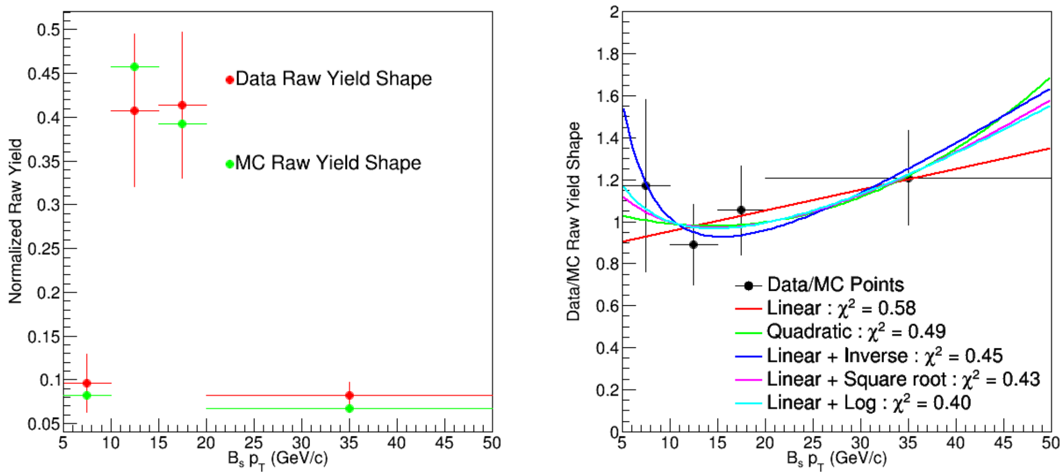


Figure 5-4: B_s^0 p_T normalized raw yields obtained in PbPb MC and Data are shown above on the top left panel. The data/MC ratio and different fitting functions: Linear (Red), Quadratic(Green), Linear + Inverse (Blue), Linear + Square Root (Purple), and Linear + Log (Cyan) and their χ^2 are shown above on the top right panel. The bottom plots are the data/MC reweighed yields with different functions from the fit on the top right panel.

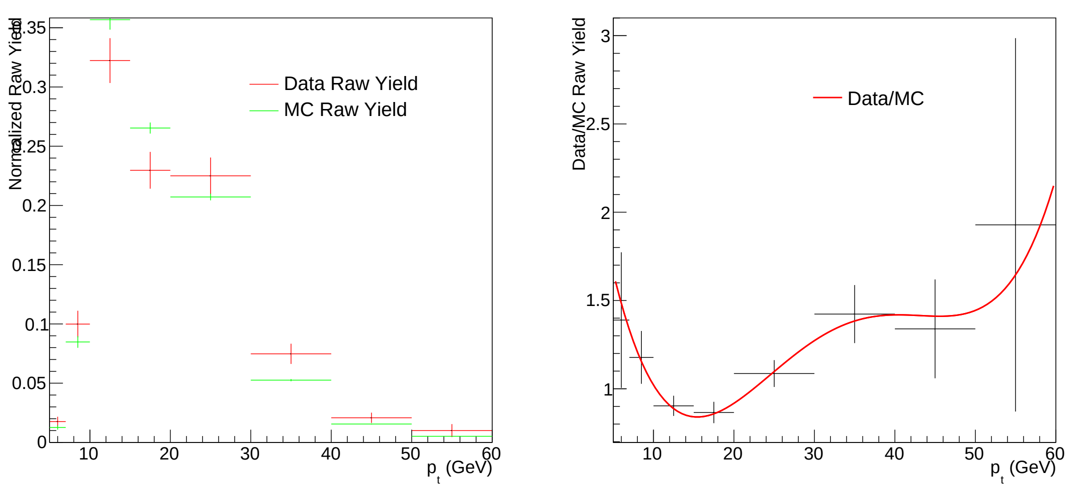


Figure 5-5: The normalized B^+ raw yield in MC (green) and Data (red) as a function RECO B^+ p_T (left) and the fourth order polynomial fits to their ratio (right) are shown above.

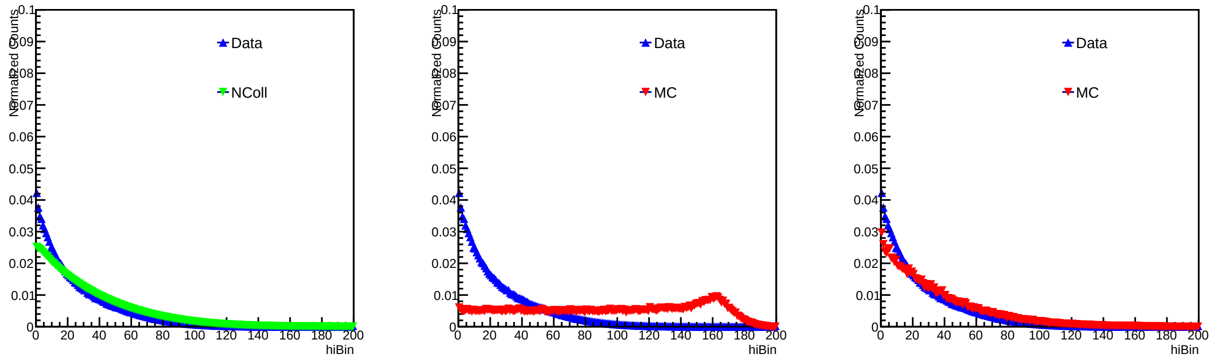


Figure 5-6: The comparison between N_{coll} and Data vs hiBin (left), centrality distribution of MC (red) and data (blue) in PbPb collisions in the centrality interval 0-100% without N_{coll} weight (middle), and with N_{coll} weight (right) are shown above.

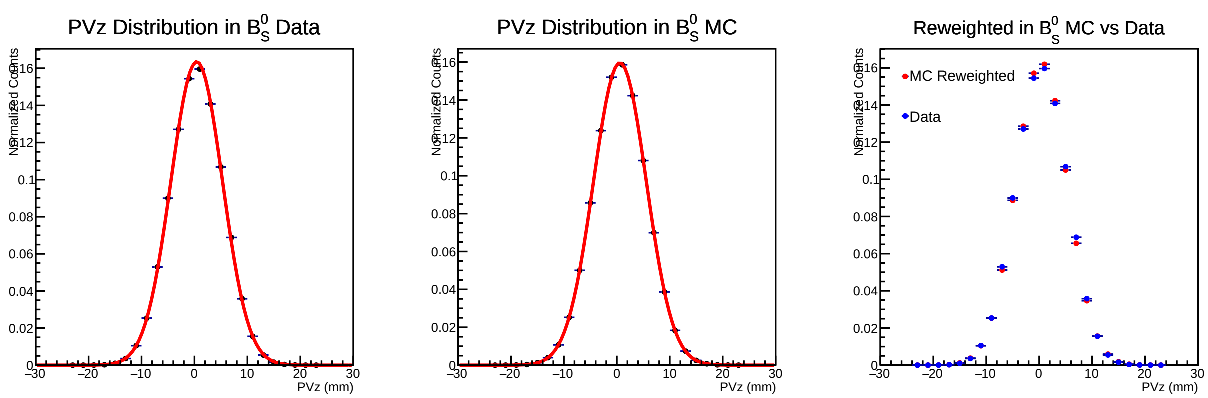


Figure 5-7: Primary vertex z position (PV_z) distribution, fitted with a gaussian function in PbPb MC simulations (left), in PbPb data (middle), PV_z reweighed MC to data with ratio of data-to-MC Gaussian Fits (right) are shown above. The PV_z distributions are well described by the Gaussian function and reweighing reduces the MC-data discrepancy.

5.4 Global Event Observables

The global event observables characterize general conditions of heavy-ion collision. In this analysis, we decide to use another set of quantities including total number of MB events to represent the luminosity and average number of participants N_{parts} to represent centrality. In addition, to compare with *pp* collisions, we also need to scale the cross section in PbPb according to . Therefore, we will determine the all the global event observable including total number of minimum bias events (N_{MB}), centrality, number of participant nucleons N_{part}, number of binary collisions N_{coll}, and nuclear overlapping function T_{AA} in dimuon PbPb dataset in the follow subsections.

5.4.1 Total Number of Events

As seen in Table 5.1, the nominal luminosity of the dimuon PbPb dataset 1.7 nb^{-1}. However, this nominal luminosity has large uncertainties and should be used in the analysis to measure the cross section. As mentioned previous in Chapter 2.2, the dimuon trigger, based on the MB trigger, will not save events that does not pass trigger selections. Hence, we can use the events of 1 PD MB datasets (PD0) via the follow formula to determine the actual number of MB events corresponding to the dimuon PbPb datasets:

$$N_{MB} = \frac{N_{MB}^{\mu\mu json}}{\mathcal{L}_{MBtrigger}^{\mu json}} \mathcal{L}_{\mu\mu trigger}^{\mu\mu json} \tag{5.1}$$

The definition of the variables in the formula are as follows:

N_{MB}**:** The number of minimum bias events in dimuon PD with muon json.

$N_{MB}^{\mu json}$**:** The number of event of all MB PDs with muon json.

$\mathcal{L}_{MB}^{\mu json}$**:** The luminosity of all MB PDs with muon json.

$\mathcal{L}_{\mu\mu trigger}^{\mu\mu json}$**:** The luminosity of dimuon PD with muon json.

For 0 - 90%, $N_{MB}^{\mu json}$ is 161507974. The number of events can then be computed as follows:

$$N_{MB} = \frac{N_{MB}^{\mu json}}{\mathcal{L}_{MBtrigger}^{\mu json}} \mathcal{L}_{\mu\mu trigger}^{\mu json} = \frac{1657.1320\mu b^{-1}}{24.0748\mu b^{-1}} \times 161507974 = 1.1823737719 \times 10^{10} \simeq \textbf{11.82 billion} \tag{5.2}$$

Hence, the number of MB event for the dimuon PbPb data is N_{MB} =11.82 billion. Below, in Table 5.2, we compile the number of minimum biased events N_{MB} in 0 - 30%, 30% - 90%,and 0 -90%.

Table 5.2: Summary table of the total number of MB events and their uncertainties vs centrality

Centrality	N_{MB} (billion)	Uncertainties
0-30%	3.941	1.26%
30-90%	7.882	1.26%
0-90%	11.82	1.26%

5.4.2 Centrality Definition

For the 2018 PbPb dataset, the centrality is given in hiBin with 0.5% increment. The hiBin is defined based on the HF response (hiHF). According to the Global observable, 5-8 is the hiHF as a function of centrality with uncertainties.

We can compute the percent deviation nom of final results to estimate systematics due to uncertainties of centrality.

5.4.3 $\langle N_{part} \rangle$, $\langle N_{coll} \rangle$, $\langle T_{AA} \rangle$ vs Centrality

As we discussed in the Glauber Model [65,71] section 1.5.7, the number of participant nucleons N_{part}, the number of binary nucleon-nucleon collisions, and nuclear over lapping function T_{AA} are all functions of the event centrality. The CMS Global Observable group has computed the average N_{part}, N_{coll}, T_{AA} and their uncertainties for different centrality bins based on the Glauber Model. The selected results are shown Table 5.3 below

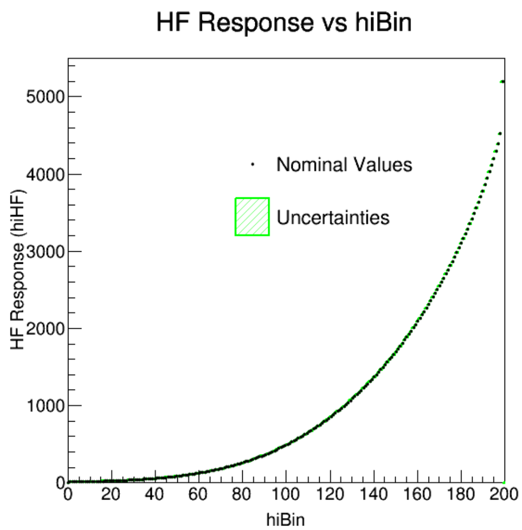


Figure 5-8: The nominal (black) and uncertainties band (green) hiHF vs hiBin for CMS 2018 PbPb dataset are shown above.

Table 5.3: Summary table of the total number of MB events vs centrality is shown below. The uncertainties are represented in terms of percentage in the parenthesis.

Centrality	$\langle N_{part} \rangle$	$\langle N_{coll} \rangle$	$\langle T_{AA} \rangle$
0-30%	269.1 (0.39%)	1042 (2.0%)	15.42 (2.0%)
30-90%	54.45 (1.5%)	115.2 (3.6%)	1.704 (3.6%)
0-90%	126.0 (0.67%)	424.1(2.2%)	6.274 (2.2%)

The global observables N_{MB}, N_{part}, N_{coll}, and T_{AA} will be used inputs for our B-meson analysis.

5.4.4 Event Multiplicity

Aside from N_{MB}, N_{part}, N_{coll}, and T_{AA}, event multiplicity is also an event observable. We count the number of tracks in each event with some track quality selections and use it to interpret the event multiplicity, which characterizes the event activity. The following is the selection criteria for

Nevertheless, the event multiplicity is not used in PbPb analysis. It will be used in *pp* analysis to study the B^0_s/B^+ ratio as a function of event multiplicity in *pp* collisions.

5.5 B meson Reconstruction

Now we look into each events of the PbPb dimuon dataset. It turns out that there is no PU in any event. Therefore, only one primary vertex for each event. We can then reconstruct the B meson candidates according to the final state muons and kaons tracks. In CMS, a dedicated software named "*Bfinder*" is developed to perform B-meson reconstruction. Figure 5-9 and 5-10 show the workflow to fully reconstruct *Bfinder* for B_s^0 and B^+ respectfully.

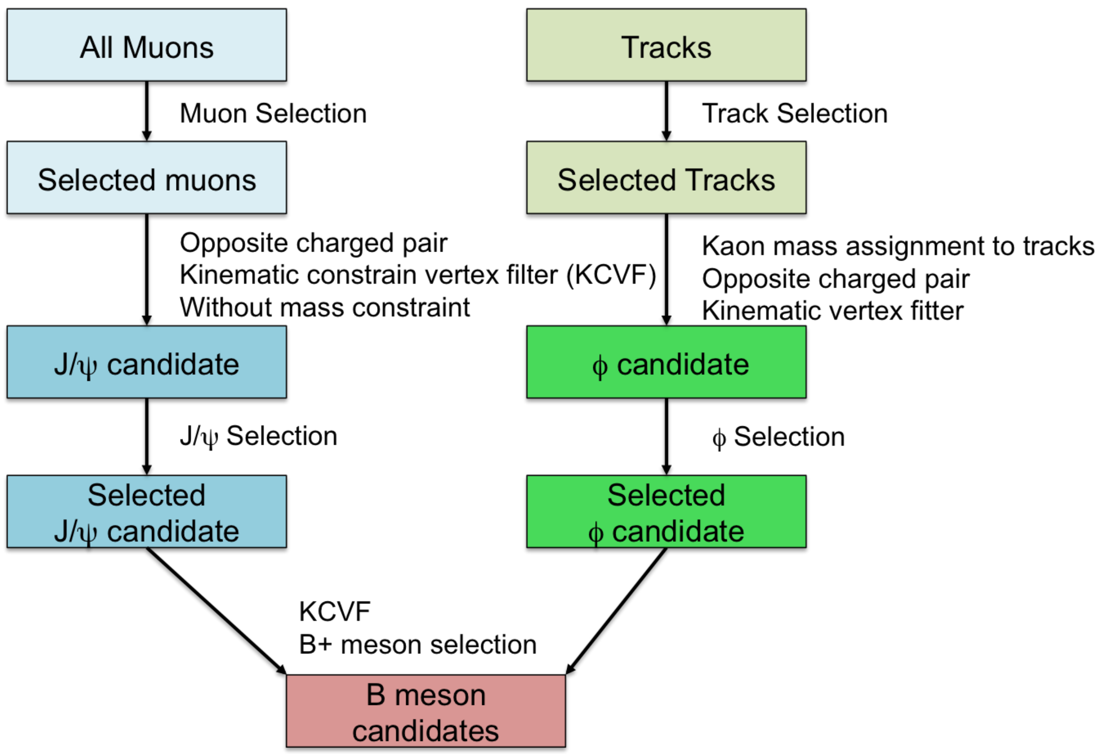


Figure 5-9: The schematic block diagram of the full reconstruction workflows for B_s^0 via the decay channel of $B_s^0 \rightarrow J/\psi\phi \rightarrow \mu^+\mu^-K^+K^-$ in the *Bfinder* is shown above.

Here, we should note that since we do not have hadronic PID for the kaons, we assume the track to be kaons and assume the charged kaon PDG mass to the tracks [4] to the tracks. Also, the invariant mass of muon pair is constrained to the nominal J/ψ meson PDG mass ($m_{J/\psi} = 3.096916$ GeV/c^2) [4] instead of a distribution of the dimuon mass $m_{\mu\mu}$. The output file format of *Bfinder* is an Ntuple. Finally, we do not distinguish particle and anti-particles during the B-meson reconstruction. Therefore,

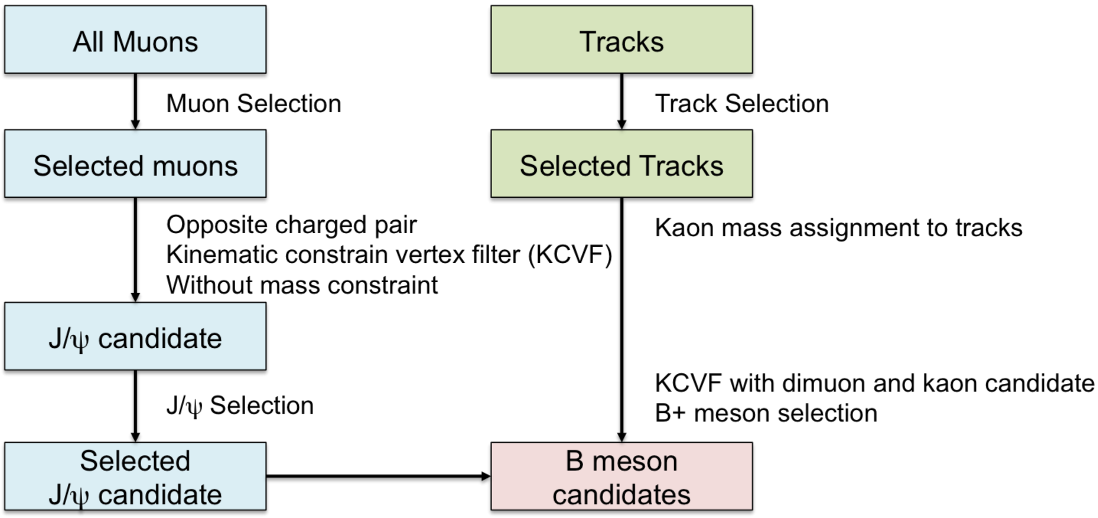


Figure 5-10: The schematic block diagram of the full reconstruction workflows for B^+ via the decay channel of $B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+$ in the *Bfinder* is shown above.

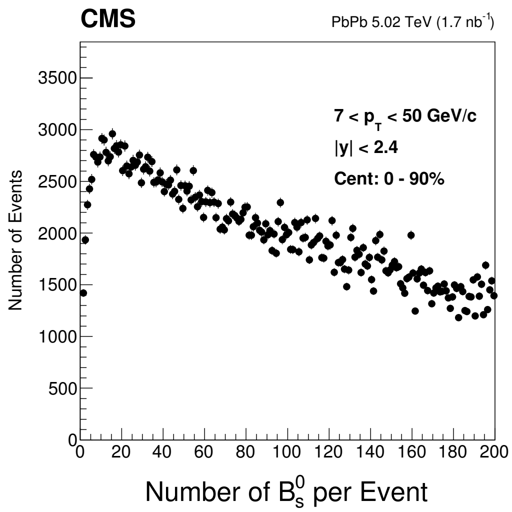
both B_s^0 and $\bar{B}_s^0$ as well as B^+ and B^- are reconstructed. Here, for simplicity, we only mention the B_s^0 and B^+ throughout the thesis. Each event will have multiple B-meson candidates 5-11.

Their information, including invariant mass, p_T, and y as well as their daughter particles kinematics such as p_T and η, is saved as a form of vector in each event. In this thesis, we use B-meson Ntuples to perform our analysis.

5.5.1 Event Selections

To ensure the quality of inelastic hadronic collisions events for B-meson reconstruct, we apply the follow selections

- At least one reconstructed primary interaction vertex, formed by two or more tracks
- The longitudinal distance from the center of the nominal interaction region of less than 15 cm along the beam axis: $|\mathrm{PV}_z| < 15$
- Compatible shapes of the clusters in the pixel detector with those expected from particles produced by a PbPb collision [204]



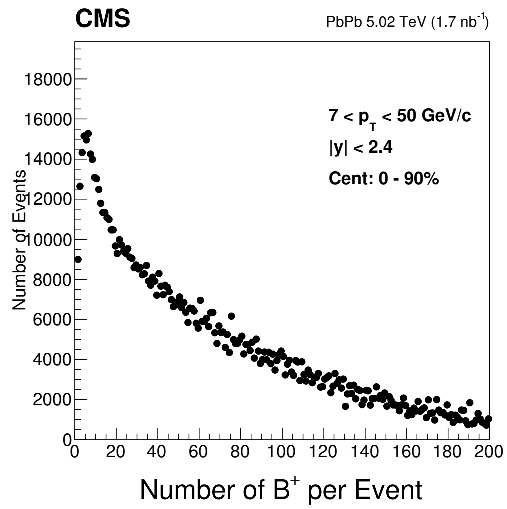


Figure 5-11: The number of reconstructed B-meson candidates per event distribution in the dimuon PbPb data for B_s^0 (left) and B^+ (right) are shown above. Multiple B-meson candidates are reconstructed in one event.

- At least two towers in each of the HF detectors with energy deposits of more than 4GeV per tower

5.5.2 Track Selections

In addition to event selection, we also apply track selections to improve the quality of the tracks and reject fake tracks. For B^+ we have the following selections

- General Tracks passing high purity selection (describe in section 3.4.2)
- $|\eta| < 2.4$ and $p_T > 1$ GeV/c
- p_T momentum resolution: $\frac{\sigma_{p_T}}{p_T} < 0.1$
- At least 10 hits in the pixel + strip tracker layers: $N_{hit} > 10$
- (Track χ^2/ndf)/(pixel + strip hits) > 0.18
- Vertex probability > 0.05

For B_s^0, since we expect it to have a ϕ resonance in the decay chain, we require the mass of the reconstructed dikaon candidate $|p_{K^+} + p_{K^-}| = m_{KK}$ to be 0.015 GeV/c^2 within the ϕ meson PDG mass ($m_\phi = 1.019455$ GeV/c^2): $|m_{KK} - m_\phi| < 0.015$ GeV/c^2

5.5.3 Muon Selections

The muon candidates are selected according to the *hybrid-soft muon* selection, developed for the muon analysis using CMS 2012 7 TeV *pp* data [205]. It is adapted from the soft-muon ID developed in the BPH group, with two modifications: a) the purity selection is removed, and b) the muon is required to be also *global*. This selection will be updated for the one developed in 2018. The *hybrid-soft muon* selection includes the following cuts:

- Requirement to be Global Muon and Tracker Muon (described in section 3.4.3)
- At least one good muons
- Transverse impact parameter $D_{xy} < 0.3$ cm
- Longitudinal impact parameter $D_z < 20$ cm
- At least 1 muon hits on pixel tracker layers and 5 hits on both the pixel + strip tracker layers

In addition, an muon acceptance selection to ensure the muon candidate to have a total efficiency: $\epsilon^\mu > 10\%$. Table 5.4 shows acceptance cuts, designed by the CMS muon analysis group, are also applied:

Table 5.4: Summary table of the muon acceptance selection for muon: $|\eta^\mu|$ as a function p_T^μ.

Centrality	$\langle N_{part} \rangle$	$\langle N_{coll} \rangle$	$\langle T_{AA} \rangle$
$\lvert\eta^\mu\rvert$	$0 - 1.2$	$1.2 - 2.1$	$2.1 - 2.4$
p_T^μ (GeV/c)	> 3.5	$>$ 5.47 - 1.89	$\eta > 1.5$

Table 5.4 comes from the muon analysis in the 2018 PbPb dataset. Figure 5-13 shows the muon reconstruction, identification, and trigger efficiency as a function of p_T^μ and η^μ [206]

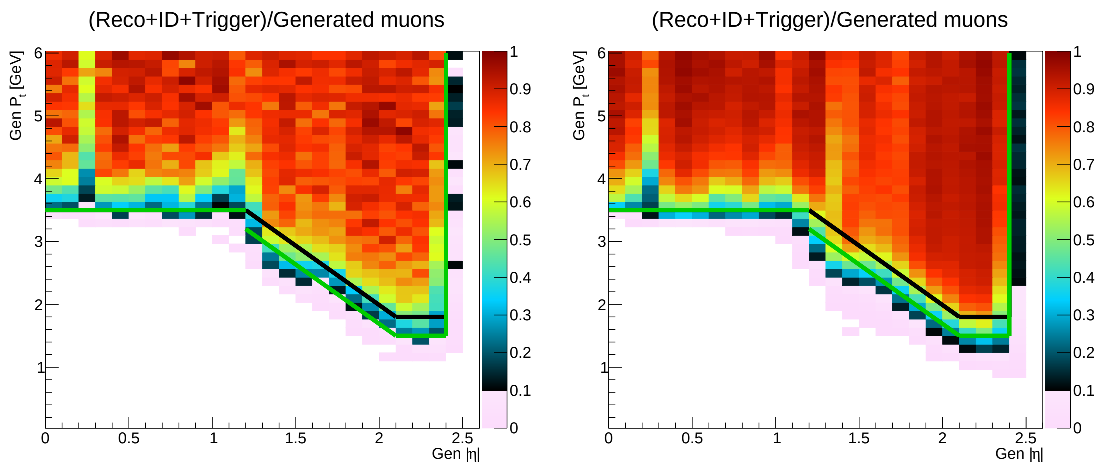


Figure 5-12: The total efficiency, including reconstruction, identification, and trigger, of a single muon in 2018 PbPb (left) and 2017 pp (right) are shown above. The black curve is the 2015 PbPb and pp 90% muon efficiency boundary while the green curve is the 2017 pp and 2018 PbPb 90% muon efficiency boundary. The green boundary is translated to numerical values in Table 5.4

We should note that there is a discontinuity of the muon acceptance selection at $|\eta| = 1.2$. Aside from the single muon selections, the following selections are applied to the reconstructed dimuon candidates

- Two muons have opposite charges
- Two mouns are tracker muons
- Dimuon invariant mass about 0.15 GeV/c^2 near the J/ψ PDG mass ($m_{J/\psi}$ = 3.096916 GeV/c^2): $|m_{\mu\mu} - m_{J/\psi}| < 0.15$ GeV/c^2
- One muon is L2 muon and the other one is L3 muon (described in section 2.25)
- Probability of the two muon tracks to originate from the same decay vertex > 1%

In addition, a B-meson invariant mass window of $4 < m_B < 6$ GeV/c^2 is applied since the B_s^0 mass is $m_{B_s^0} = 5.367$ GeV/c^2 and the B_s^0 mass is $m_{B^+} = 5{,}279$ GeV/c^2 [4]. Anything far away from the mass window should not be considered. After applying all these preliminary selections to improve the quality of our dataset for the analysis, we are ready to perform cut optimization to further reject background candidates based on the decay topology of the B_s^0 and B^+ decay chains.

5.6 Cut Optimization

Given the high combinatorial background, particularly in PbPb collision where we have thousands of tracks per event [207], it is not possible to observe B-meson resonance by simply applying the preselection presented in the previous section. Figure **??** shows the invariant mass distribution of fully reconstructed B_s^0 and B^+ at $7 < p_T < 50$ GeV/c after the event, track, and muon selections.

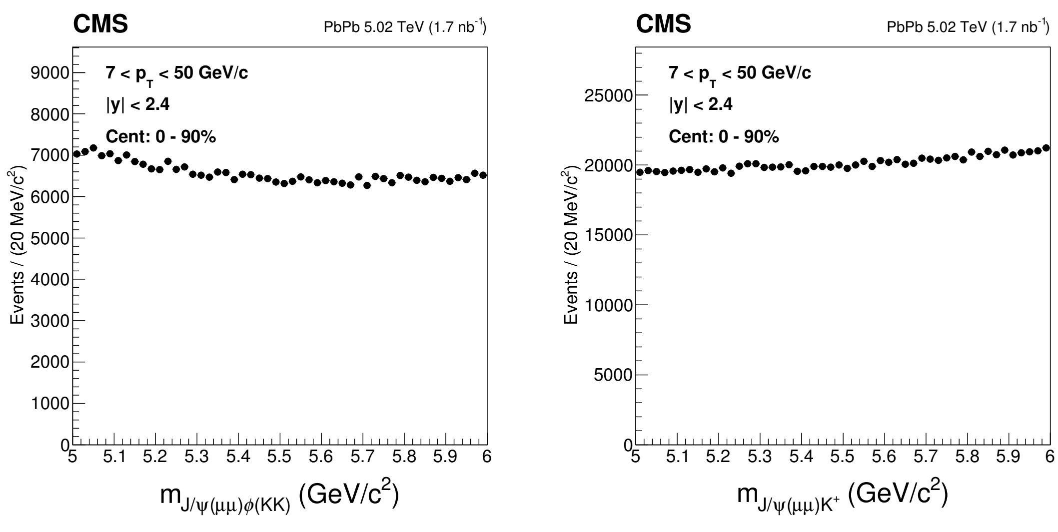


Figure 5-13: The invariant mass distributions of fully reconstructed B_s^0 (left) and B^+ (right) after preselection are shown above.

No B-meson signal is observed in the data. Therefore, aside from the preselections, a multivariate analysis (MVA) approach [208] is thus conducted in order to develop an optimal selection to separate signal B mesons from background and reconstruct a significant resonance in the invariant mass distribution in the data. The fitting performance is further related to both the amount of signal and background presented in the mass spectrum.

5.6.1 Topological Variables

By an MVA analysis, one can then find the proper selection criteria which is optimized for this purpose. Several variables related to kaon tracks and B mesons decay topology are applied in order to reduce the combinatorial background that arises from random combination of tracks and muons. The topological variables used in B-meson analyses to be optimized by are listed as follows:

Topological Variables for B_s^0:

- Kaon track p_T
- Kaon track transverse distance to closest approach (DCA) significance: $DCA_{xy}/\sigma_{DCA_{xy}}$
- Kaon track longitudinal distance to closest approach (DCA) significance: DCA_z/σ_{DCA_z}
- Dikaon invariant mass distance to the ϕ meson PDG mass: $|m_{KK} - m_\phi|$
- The B_s^0 meson decay length [or the distance between primary vertex (PV) and secondary vertex (SV)] significance: $|\vec{D}(SV, DV)|/|\sigma_{\vec{D}(SV,DV)}|$
- The open angle between the B-meson decay length vector and its three momentum: $\alpha : \cos(\alpha) = \frac{\vec{D}(SV,DV)\cdot\vec{p}}{|\vec{D}(SV,DV)||\vec{p}|}$
- The cosine angle of the opening angle in the transverse direction: $\theta_B : \cos(\theta_B) = \frac{D(SV,\vec{DV})_{xy}\cdot\vec{p_T}}{|D(SV,\vec{DV})_{xy}||\vec{p_T}|}$
- Vertex fitting probability: the χ^2 value of the vertex fitting

For B^+, we also apply some addition rectangular selections before cut optimization

Topological Variables for B^+:

- Kaon track p_T

- Kaon track $|\eta|$

- Kaon track transverse distance to closest approach (DCA) significance: $DCA_{xy}/\sigma_{DCA_{xy}}$

- The B^+ meson decay length [or the distance between primary vertex (PV) and secondary vertex (SV)] significance: $|\vec{D}(SV,DV)|/|\sigma_{\vec{D}(SV,DV)}|$

- The open angle between the B-meson decay length vector and its three momentum: $\alpha : \cos(\alpha) = \frac{\vec{D}(SV,DV)\cdot\vec{p}}{|\vec{D}(SV,DV)||\vec{p}|}$

- The cosine angle of the opening angle in the transverse direction: $\theta_B : \cos(\theta_B) = \frac{D(SV,\vec{D}V)_{xy}\cdot\vec{p_T}}{|D(SV,\vec{D}V)_{xy}||\vec{p_T}|}$

- Vertex fitting probability: the χ^2 value of the vertex fitting

Figure 5-14 and 5-15 show the definition of topological variables of B_s^0 and B^+ decay chains respectfully

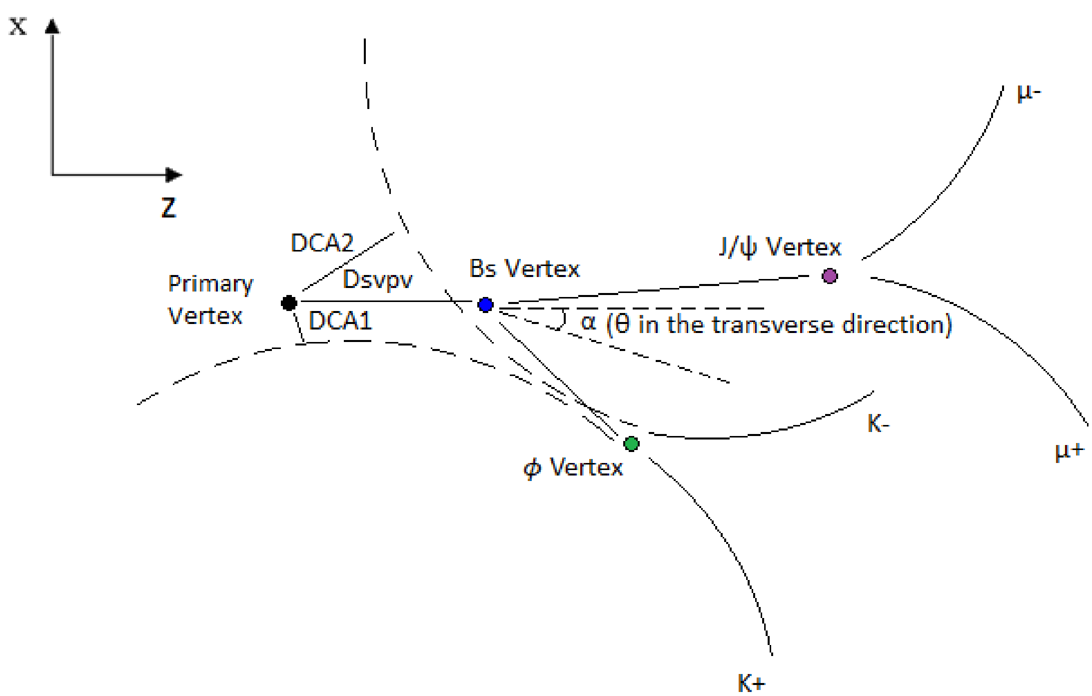


Figure 5-14: The definition of topological variables in the decay of $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$ (left) are schematically shown above.

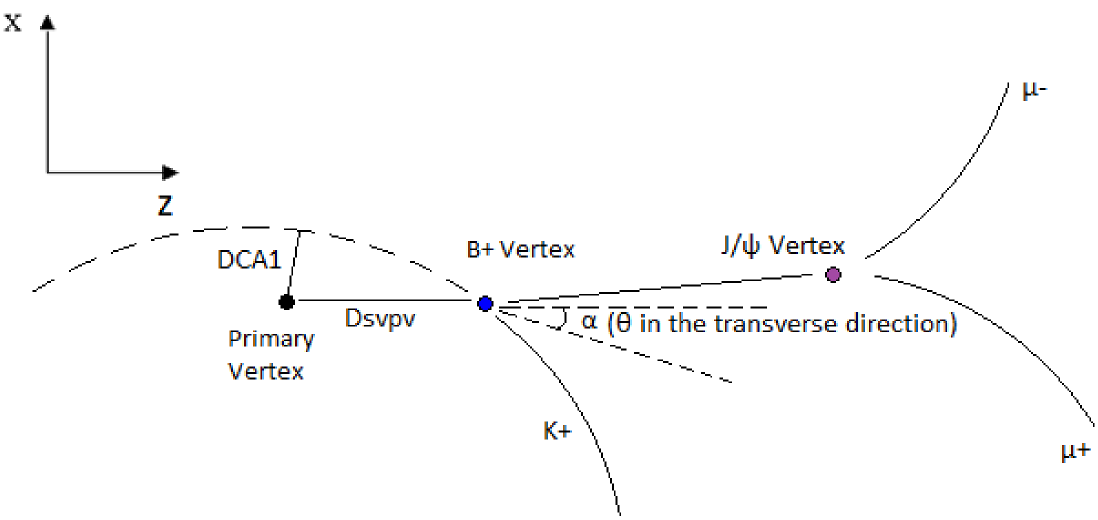


Figure 5-15: The definition of topological variables in the decay of $B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+$ are schematically shown above.

These topological variables will become the inputs to the multivariate analysis to optimize the signal significance.

5.6.2 Multivariate Analysis

In statistics, many data analysis techniques only focus on one or two variables individually. Multivariate analysis (MVA) analyzes more than two variables simultaneously to improve the data analysis. Figure 5-16 shows schematically the advantages of MVA to traditional statistical techniques in data analysis to separate signal from background with two variables.

We can see that in multivariate analysis, an MVA value as a function of two independent variables x and y: $MVA = f(x, y)$ is able to select signal out from background with higher purity (larger S/B ratio) than the rectangular selection function $x_1 < x < x_2$ and $y_1 < y < y_2$. Table 5.5 shows the performance of MVA and traditional rectangular selections

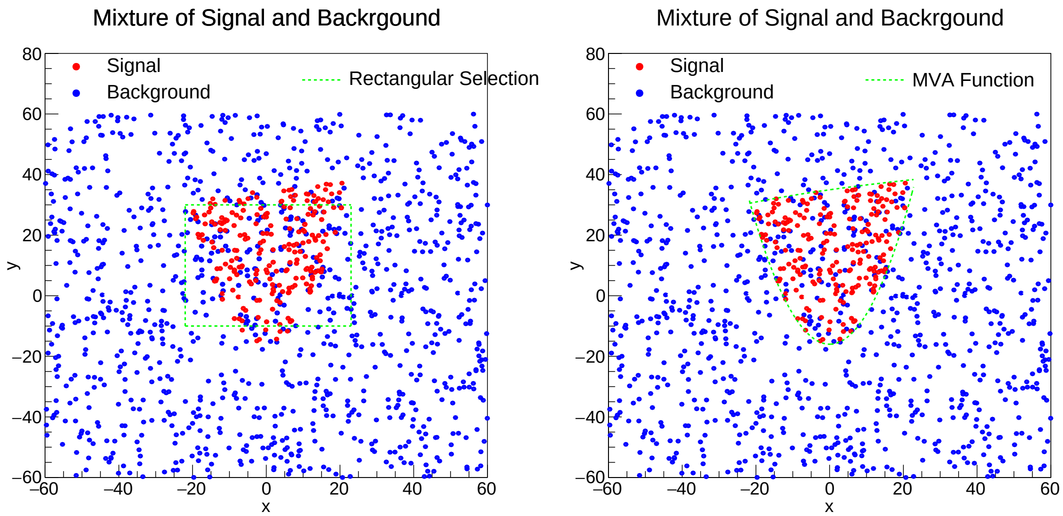


Figure 5-16: The performance of traditional rectangular selection with a range of x and y (left) compared to the MVA method of a curve as a function of X and Y (right) are shown above. Here, we have the total signal S = 215 and the background B = 1000.

Table 5.5: The comparison of the traditional rectangular selections to MVA for Figure 5-16.

Analysis Techniques	S	B	S/B
Rectangular	174	135	1.29
MVA	215	114	1.89

5.6.3 Machine Learning Techniques

Machine learning, as branch of artificial intelligence, is the science that gets the computers to learn what human beings do. It is an automating data analysis method for model building to solve practical problems. Figure 5-17 demonstrates the data analysis problems including classification, regression, and clustering, where machine learning could be applied.

In this analysis, our goal is to separate B meson signal out of the background. Therefore, it is a classification problem. Therefore, machine learning can be a power tool to solve our problems. We apply supervised machine learning to train the computer and let them find the optimal selections for B-meson analyses.

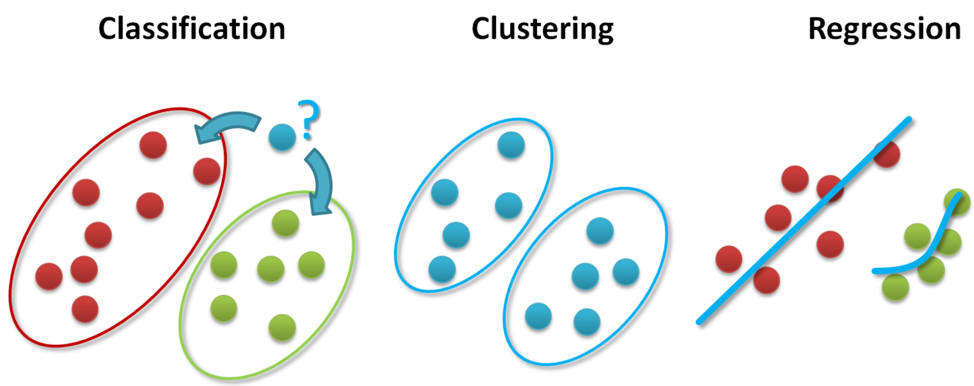


Figure 5-17: The solutions to classification, regression, and clustering problems with supervised and unsupervised machine learning approaches are shown schematically above.

5.6.4 Terminologies

The following list explains the technical jargons of machine learning techniques along with MVA that may be mentioned later in this thesis:

- **Training samples:** the input samples including both background and signal to train the computer. In this thesis, we use the B mesons candidates coming from our chosen B-meson decay channels (GEN Matched) in MC as the signal input and the B mesons candidates from the invariant mass sideband region with a distance of greater than 0.2 GeV/c^2 to the PDG mass of B mesons as background input to the *TMVA*

- **Testing samples:** the samples including both background and signal going to be tested with the output from the training. The testing sample should not be the same as the training sample.

- **Correlation matrix:** the linear correlation between the input variables for training

- **ROC curve:** the curve of signal efficiency as a function of the background rejection (= 1 − background efficiency) for a given MVA value. Here the efficiency is defined as: efficiency = the number of candidates with the given MVA cut/ number of candidates without the MVA cut.

- **Overtraining:** A Kolmogorov-Smirnov test [209] to compare the shape of the MVA distributions of training and test samples. It returns probabilities for both signal and background between 0 and 1. The closer to 0, the poor the matching, the more the overtraining will be.

Figure 5-18 shows the definition of signal and background regions in B_s^0 and B^+ invariant mass plot.

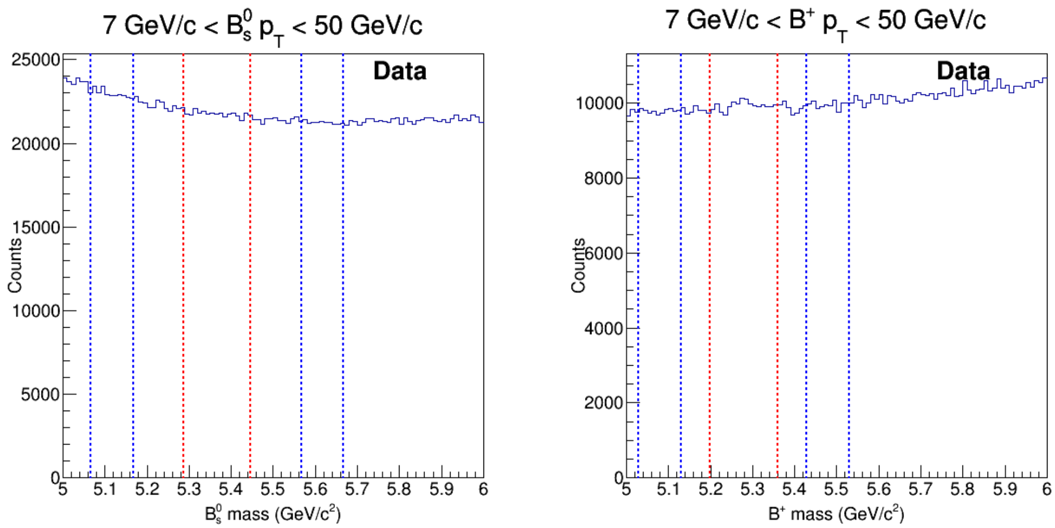


Figure 5-18: The definitions of signal (red band) and background region (blue band) in the fully reconstructed B-meson invariant distribution for B_s^0 (left) and B^+ (right) are shown above.

The signal region is $|m_B - m_B^{PDG}| < 0.08$ GeV/c and the background region is $0.20 < |m_B - m_B^{PDG}| < 0.30$ GeV/c^2 for B_s^0 and $0.15 < |m_B - m_B^{PDG}| < 0.25$ GeV/c^2 for B^+.

5.6.5 Boosted Decision Tree Algorithm

Nowadays, machining learning has become sophisticated. There are many well developed machining learning algorithms such as Rectangular Cut (CutsSA) Linear Discriminant (LD), Boosted Decision Tree (BDT), neutral network feed-forward Multilayer Perceptrons with recommended ANN with BFGS training method and bayesian regulator (MLPBNN), and Deep Learning in the market. Here, I will give a brief introduction about BDT algorithm in terms of *Boosted* and *Decision Tree*

Boosted: Boosted here refers to the "boosting" factor α in the hypothesis function to let the machine learn and correctly model the actual curve [?]. According to supervised machine learning, in order to use the boosting method, we first assign an ensemble of many weak learners to create of a strong learner. The idea of boosting is to keep reweighing the function to enhance the classification and regression and performance by applying an MVA algorithm subsequently to the reweighed version of the training data in a sequential matter. Eventually, the hypothesis will converge a function that can describe the actual data. Figure 5-19 below shows schematically the boosting scheme

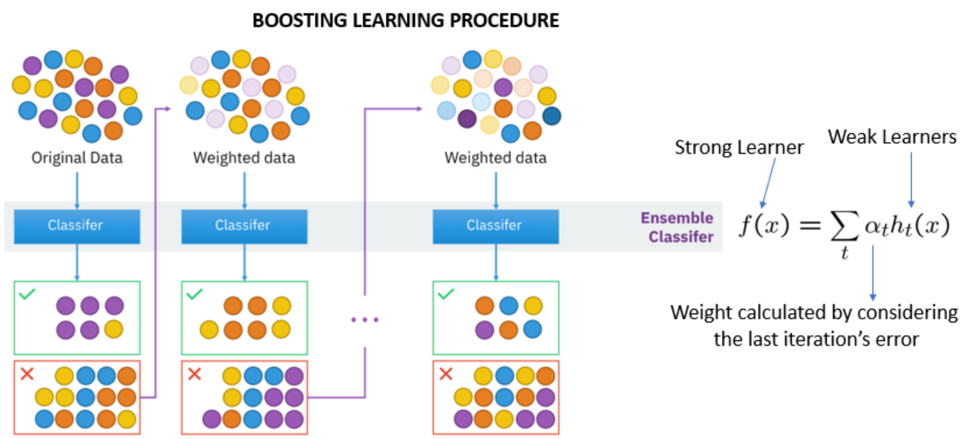


Figure 5-19: The schematic illustration of boosting procedures in machine learning is shown above.

Decision Tree: A decision tree, sometime called as a regression tree, is a binary decision support tool using a tree-like model of decisions and list their possible consequences. Starting from a sampling to analyze, it sets up criteria for selections to decide the likelihood of signal and background based on the pure signal and background input samples. Then, it makes binary decisions (Yes/No), in each branch to select subset of samples out of the current sample, can be in sequentially or in parallel. It will then iterate multiple times until the sub samples are considered as signal or background. In TMVA machinery, the number of iteration is called *NTree*. Figure 5-20 is a schematic demonstration of a decision tree with NTree = 4

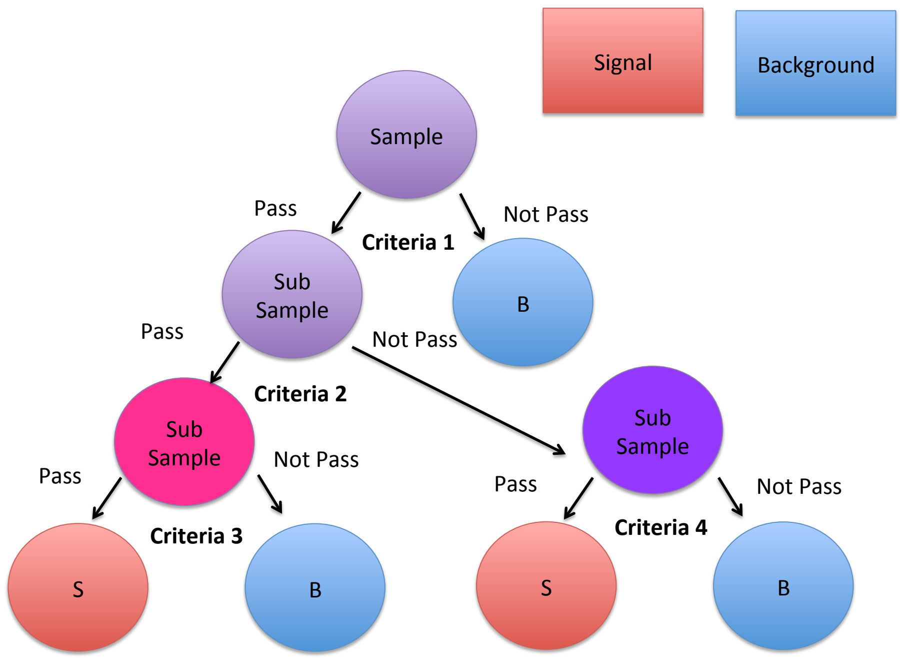


Figure 5-20: The schematic block diagram of a binary decision tree with NTree = 4 to separate signal and background in a training sample.

In general, the performance of BDT will improve as NTree become larger. However, we should note that normally NTree is required to be smaller than the sample statistics. Otherwise, overtraining may occur and could induces bias in the data analysis. We should note that the BDT score range from -1 to 1.

5.6.6 TMVA Training

To perform machine learning, I use the Toolkit for Multivariate Data Analysis with ROOT (*TMVA*) [211], a dedicated ROOT base machine learning software framework on C++ programing language, to train the computer to find the optimal MVA value as a function of the topological variables in our B-meson analyses. We propose to optimize B_s^0 in the p_T binning of [7, 10, 15, 20, 50] GeV/c and B^+ in the p_T binning of [5, 7, 10, 15, 20, 30, 40, 50, 60] individually. The following procedures are carried out:

Step 1: identify sufficient signal and background samples to train the computer within *TMVA* machinery.

Step 2: choose the training algorithms to use. Here, we choose CutsSA, CutsGA, BDT, MLP, and MLPBNN2 algorithms

Step 3: decide the training parameters for the each training algorithm.

Step 4: run the TMVA machinery and generate the performance plots

Step 5: choose the best algorithm according to the performance

5.6.7 Training Performance

After finishing the TMVA training procedure, we are ready to look at the training performance. First, we the correlation between the input topological variables. Figure 5-21 shows the correlation matrices of B_s^0 and B^+ topological variables for B-meson $10 < p_T < 15$ GeV/c.

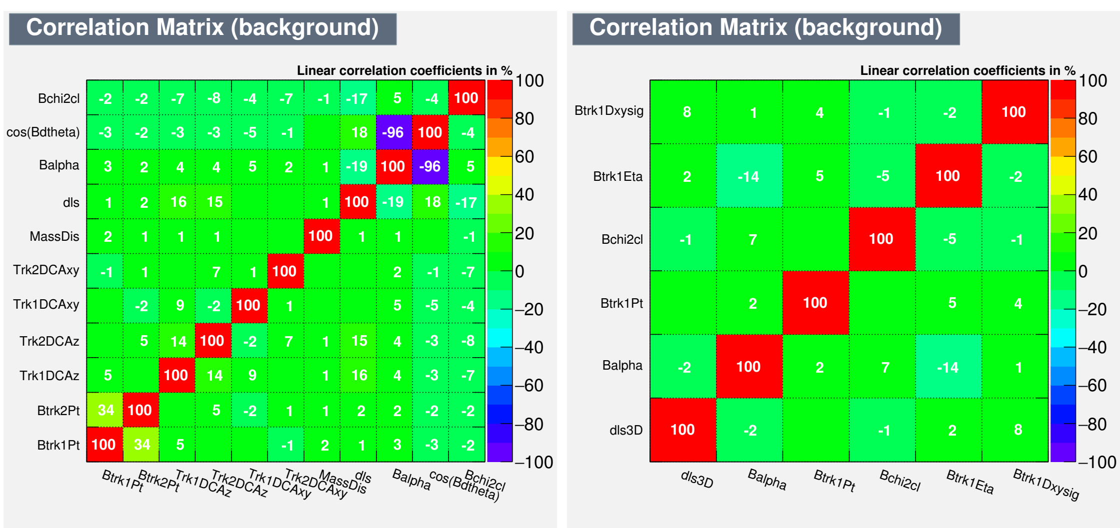


Figure 5-21: The correlation matrices of B_s^0 (left) and B^+ (right) in data at $10 < p_T < 15$ GeV/c are shown above.

We can see that there are no significant correlation among the topological variables except the open angle α and the cosine its transverse projection $\cos\theta$, which is know. Therefore, the variable sets we input to the TMVA training are good.

Next we compare the overall performance among the algorithms. Figure 5-22 below shows the ROC curves for B_s^0 *TMVA* training at $10 < p_T < 15$ GeV/c.

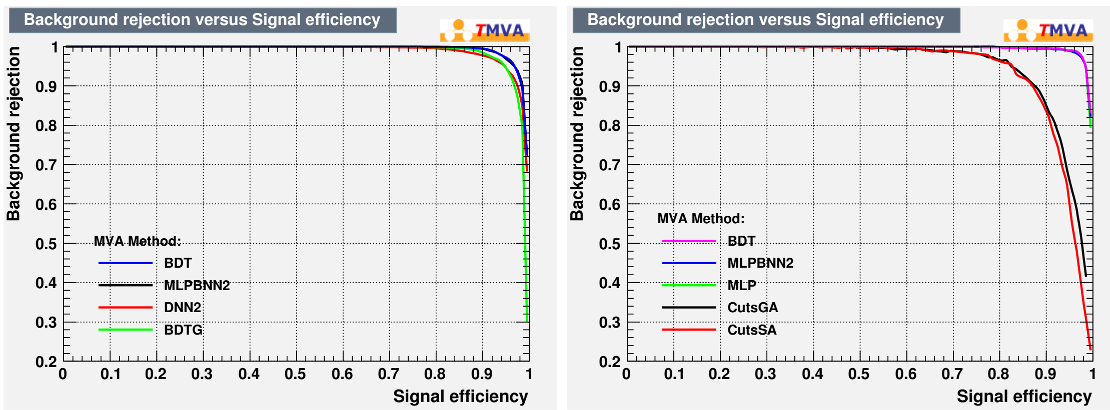


Figure 5-22: The B_s^0 ROC curves of CutsSA, CutsGA, BDT, MLP, and MLPBNN2 algorithms are shown above.

We should note that the number of tree for BDT here is NTree = 2000. From Figure 5-22, we can see that BDT has the best performance compared to other algorithms in the given parameter settings. Basically, for a given background rejection, the BDT curve has the highest signal efficiency. It is closer to the upper right corner, which is the perfect algorithm. Therefore, we decide to use BDT algorithm to look for the optimal selections in all p_T bins in both B_s^0 and B^+ analysis.

Finally, before, implementing the BDT algorithm to the analysis, we also check the overtraining and make sure that no significant overtraining is observed. Figure 5-23 show the overtraining test for both B_s^0 and B^+ BDT at $10 < p_T < 15$ GeV/c.

According to Figure 5-23, neither signal nor background of B_s^0 and B^+ BDT is vanishing. Hence, no significant overtraining is observed. The BDT trainings are valid to use in the B_s^0 and B^+ analysis.

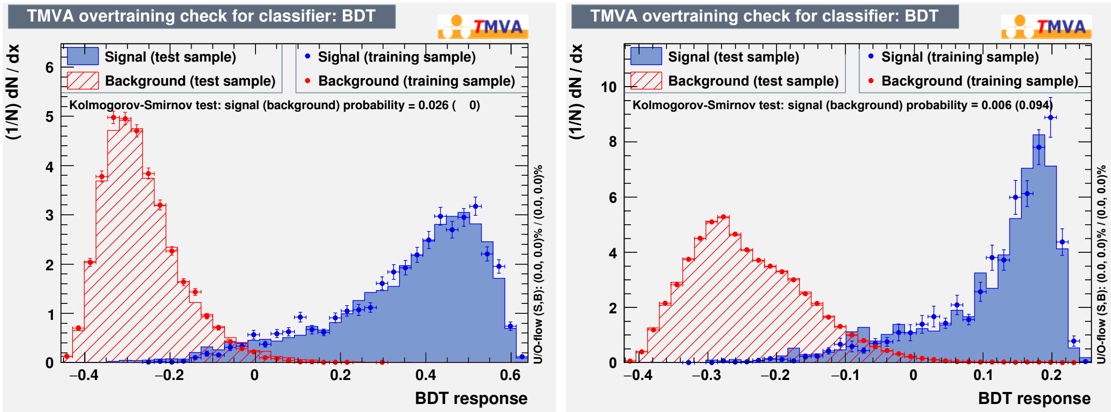


Figure 5-23: The Kolmogorov-Smirnov overtraining tests on the signal (blue) and background (red) in B_s^0 (left) and B^+ (right) at $10 < p_T < 15$GeV/c are shown on the right. It looks like they both pass the tests.

5.6.8 Working Point Determination

Now with the BDT training results, next step is to choose the BDT selection that can give us the best analysis results. We decide to use the statistical significance, which is as follows

$$Sig = \frac{S}{\sqrt{S+B}} \tag{5.3}$$

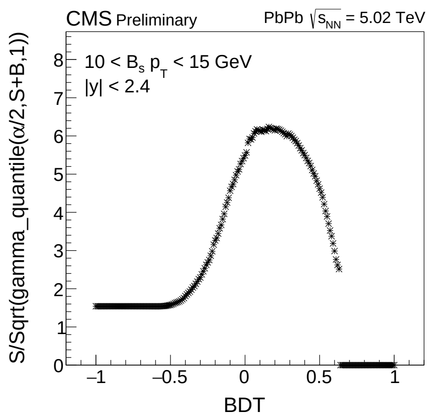
as the figure of merit. We estimate S and B for set of BDT cuts and choose BDT score that return to the maximum statistical significance in each p_T bin. We can directly estimate the background B according to the data sideband region width scaled to the signal region band width

$$B = \frac{w_S}{w_B} N_B \epsilon_B \tag{5.4}$$

To determine the signal S, we do not directly look at the data. Instead, we use FONLL cross section [132] shown in 1-31 and the MC to determine the signal S as follows

$$S = 2R_{AA}^{2015Ref} L \sigma_{FONLL}^{pp \to b\bar{b}} \epsilon_{pre} \epsilon_S f(b \to B) BR \tag{5.5}$$

According to our calculations, for B_s^0 at $10 < p_T < 15$ GeV/c, before the BDT selections, S = 47 and B = 20672. Now, we scan through a series of BDT values from -1 to 1 and compute their corresponding statistical significance. Figure 5-24 shows the statistical significance as a function of BDT for B_s^0 and B^+ at $10 < p_T < 15$ GeV/c



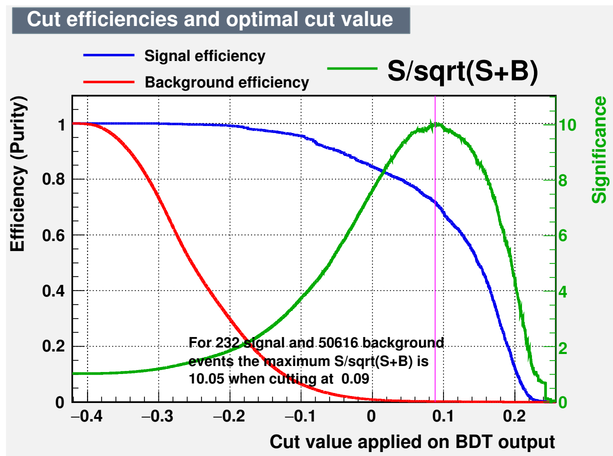


Figure 5-24: The significance: $Sig = \frac{S}{\sqrt{S+B}}$ as a function of BDT in B_s^0 at $10 < p_T < 15$ GeV/c are shown above. We can see that B_s^0 BDT peaks near 0.32 while B^+ peaks near 0.09.

For B_s^0 Our optimal selection returns us with an S = 38 and B = 5, which has a remarkable background-to-signal rejection of more than $10^3 : 1$.

Table 5.10 and 5.7 documents the optimal BDT selections maximizing the statistical significance in each p_T bin for B_s^0 and B^+ respectfully.

Table 5.6: The summary of optimal BDT selection maximizing the B_s^0 statistical significance.

B_s^0 p_T (GeV/c)	5 – 10	10 – 15	15 – 20	20 – 50
BDT Cut Values	> 0.32	0.29	0.35	0.33

Table 5.7: The comparison of the traditional rectangular selections to MVA for Figure 5-16.

B^+ p_T (GeV/c)	$5-7$	$7-10$	$10-15$	$15-20$	$20-30$	$30-40$	$40-50$	$50-60$
BDT Cut Values	> 0.02	0.03	0.09	0.07	0.10	0.16	0.20	0.27

5.6.9 Optimal Selection Performance

To check the performance BDT selections, we look at the dimuon and dikaon invariant mass distributions to see if J/ψ and ϕ resonance are observed. Figure 5-25 shows the after apply the selections

We can see clear J/ψ and ϕ peaks after apply the selections in both data and MC, which suggests that our selections are reasonable. Now we can also look at the invariant mass distributions of B_s^0 and B^+ in Figure 5-26

Again, we can see very clear signal after the optimal BDT selections in both B_s^0 and B^+. Now we are ready to study its background and signal before extracting the raw yield for B-meson cross section measurement.

5.7 Background Studies

5.7.1 Overview

The production of J/ψ mesons occurs in three ways. The prompt J/ψ produced directly in the proton-proton collision or indirectly via the decay of heavier charmonium states, and non-prompt J/ψ from the decay of a b hadron. Non-prompt J/ψ lead to a measurement of the b-hadron cross section. According to PDG, so far physicists have observed thousands of known decay modes of b hadrons [4]. Without hadronic PID, we can envision that there are potential background feed-down sources coming from other B meson decays in the B-meson invariant mass spectrum. For instance, the decay of $B^0 \to J/\psi K^{*0}(892) \to \mu^+\mu^- K\pi$ could contribute to the $B_s^0 \to J/\psi\phi \to \mu^+\mu^- K^+K^-$ due to misidentification of π to K. We call such back-

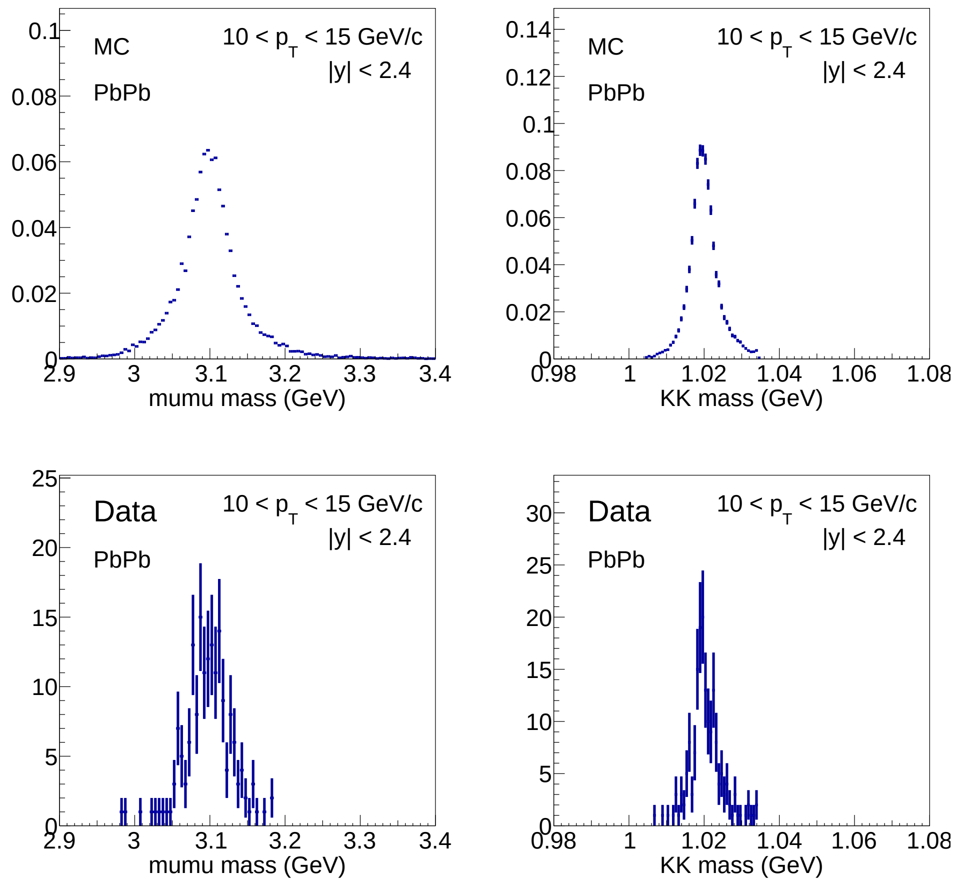
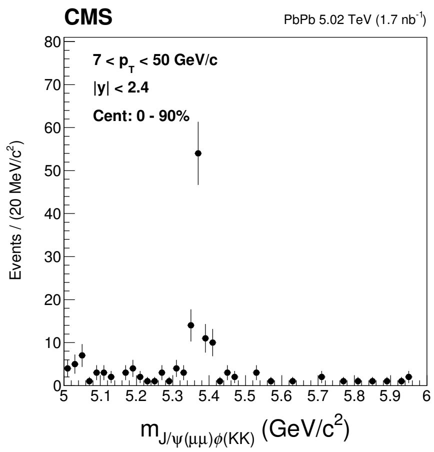


Figure 5-25: The J/ψ (left) and ϕ (right) meson mass distributions after applying BDT > 0 for MC (top) and data (low) in B_s^0 analysis are shown above.

ground as "non-prompt (NP) background", to distinguish them from the combinatorial background due to random combination of decay daughters when reconstructing B mesons. NP background generally form a peaking structure in the region of interests. A dedicated inclusive NP J/ψ from b hadron decay MC is simulated to determine NP background component near our B meson invariant mass region. We then classify each reconstructed B-meson candidate by their GEN-level particle, e.g. whether it is coming from a B^0, B^+ or other decays that falls into the B-meson reconstruction work flow, in order to measure their individual contribution to the peaking structure.



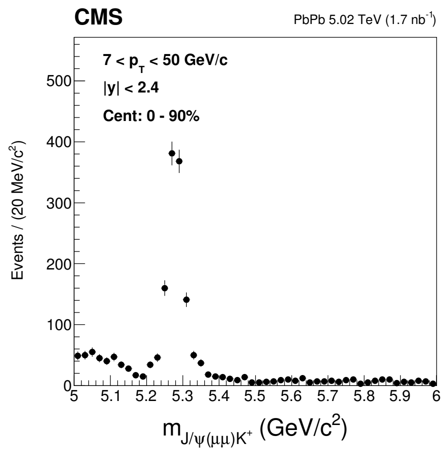


Figure 5-26: The B_s^0 (left) and B^+ (right) invariant mass distributions after applying optimal BDT selections from $10 < p_T < 50$ GeV/c in data are shown above.

5.7.2 Individual Channel NP Background Studies

There are many small contributions that gives the form to the peaking background structure. We can not identified each of them individually but most of the contributed ones are determined. Below is a list of example processes that compose of the majority of the peaking background:

- Case 1: $X \to J/\psi \pi^- K^+$, here pion is mis-identified as kaon
- Case 2: $B_s^0 \to J/\psi K^+ K^-$, in which both Kaons are not coming from the decay of an intermediate ϕ meson resonance.
- Case 3: $B^+ \to J/\psi K^+$, (added extra K^-)
- Case 4: $X \to J/\psi \pi^+ \pi^-$, pions mis-identified as Kaons.
- Case 5: $B_s^0 \to J/\psi K^+ K^- X$
- Case 6: $B_s^0 \to J/\psi \phi \pi \pi$
- Case 7: $B^0 \to J/\psi K^{*0}$

Figure 5-27 shows the contribution of the determined channels with respect to the total background after applying optimal cuts for PbPb for $10 < p_T < 15$ GeV/c, 10

$< p_T < 20$ GeV/c, $20 < p_T < 50$ GeV/c, as well as the inclusive $10 < p_T < 50$ GeV/c. The peak at the B_s^0 signal region has been investigated and found out that mostly contributed channels are $B^0 \rightarrow J/\psi K^{*0}(892)$ in grey color and $B_s^0 \rightarrow J/\psi K^+K^-$ in black color.

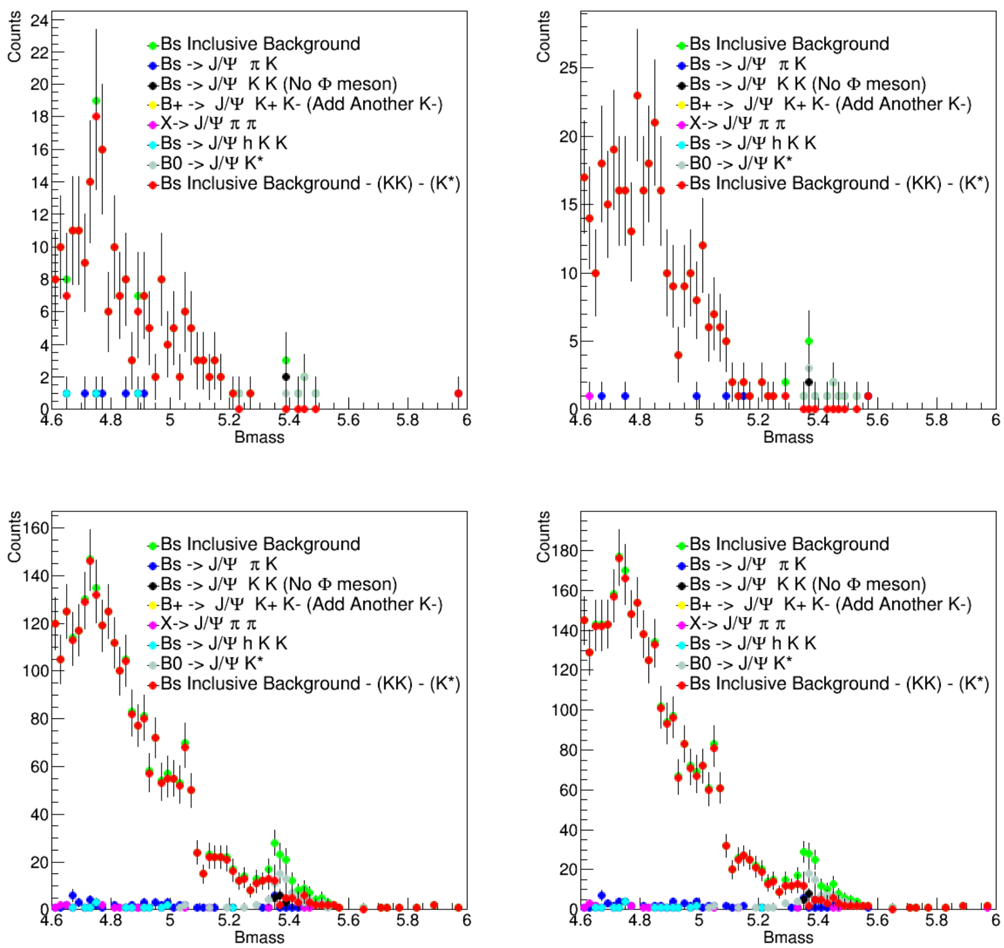


Figure 5-27: Individual NP background contributions with respect to the total background components for 10 to 15 GeV (top left), 15 to 20 GeV (top right), 20 to 50 GeV (bottom left), and 10 to 50 GeV (bottom right) for PbPb sample. We can see that the non-prompt background from all channels listed above are negligible compared to the inclusive background. Also, the no peak near the B_s resonance is observed when the inclusive background subtract the $B^0 \rightarrow J/\psi K^{*0}$ and $B_s \rightarrow J/\psi K^+K^-$ components.

Figure 5-28 also includes the signal component in the distribution. We can see that the non-prompt background is insignificant compared to the total inclusive background and the inclusive background is low comparing to the $B_s \rightarrow J/\psi\phi \rightarrow$

$J/\psi K^+K^-$ signal in our studies.

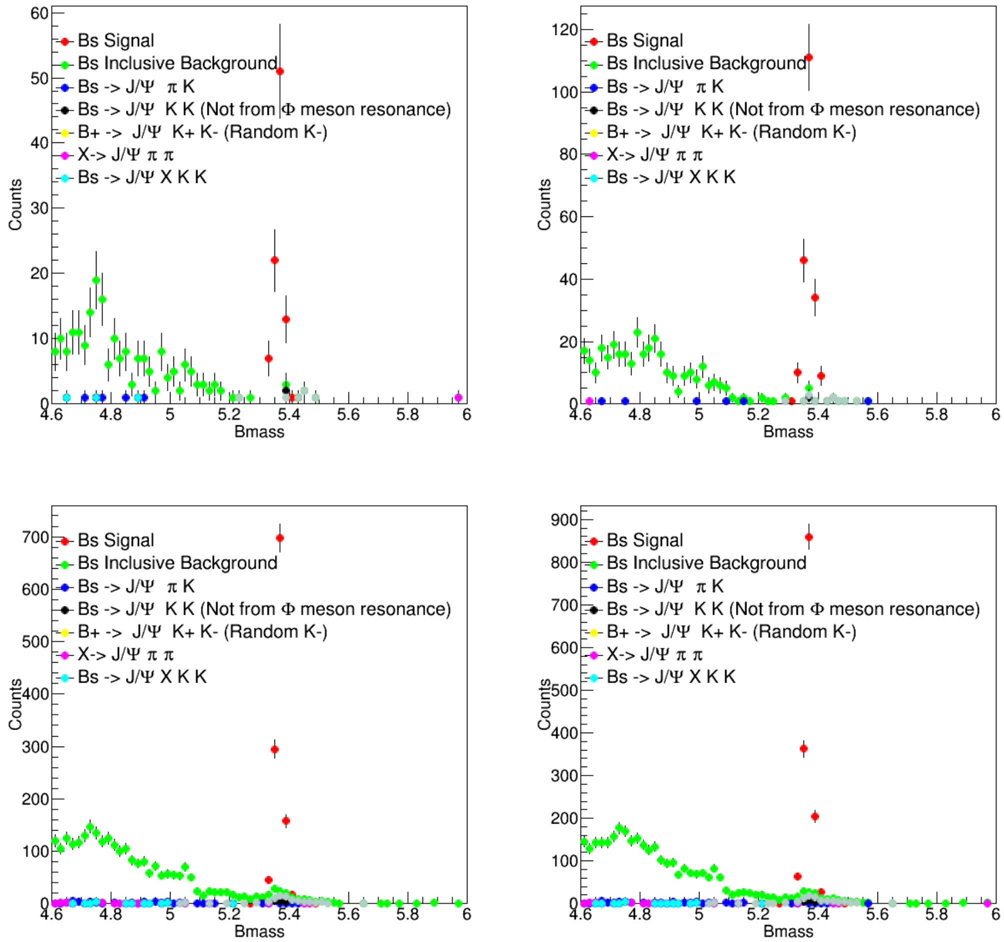


Figure 5-28: Individual Non Prompt background contributions with respect to the total background components and the signal channel for 10 to 15 GeV (top left), 15 to 20 GeV (top right), 20 to 50 GeV (bottom left), and 10 to 50 GeV (bottom right) for PbPb sample. We can see that the inclusive background is small compared to the signal we used in our studies.

5.7.3 B meson contribution of NP Background to B_s^0

From the non-prompt background studies above, we can see that there are many small contributions. Those small components together summed up constructing a large background and we can not identify all of them individually. Under this circumstances, another definition of individual component defined. The new definition of the NP background components are listed below;

- Case 1: $B^+ \to J/\psi X$, all contribution from B+
- Case 2: $B_s^0 \to J/\psi X$, all contribution from B_s^0
- Case 3: $B^0 \to J/\psi X$, all contribution from B^0
- Case 4: Other contributions

In Figure 5-29 the signal region is dominated by the B^0, B^+, and B_s^0 decays while the other channels are contributed on the left side of the mass spectrum. For both pp and PbPb data, the B_s^0 to J/ψ + X (orange region) makes a big contribution.

The conclusion is that the NP peaking background contribution to the B_s^0 is negligible due to the $K^{*0}(892)$ veto cut: $|m_{KK} - m_\phi| < 0.015$ GeV/c^2. We estimate that it may only contribute around 4% uncertainties to the signal yield, which is negligible compared to the statistical uncertainties of the B_s^0 signal, which is on the order of 10%. Hence, there is no need to develop a specific function to model the NP background component for B_s^0.

5.7.4 B meson contribution of NP Background to B^+

We veto the candidates from the inclusive NP J/ψ MC sample that are matched to a genuine B^+ signal. The resulting B candidate mass spectrum in the inclusive p_T range (5-100 GeV/c) is shown in Figure 5-30 for PbPb MC samples.

It is clear that these sources create a peaking structure in the region of M_{inv} <5.20 GeV/c^2 as seen in Figure 5-26 after applying the optimal selection. This structure can be nicely fit with an error function as done previously in B proton-proton analyses [212]. In addition, there is a minor peak on the right shoulder (≈5.34 GeV/c^2) of the nominal signal (≈5.28 GeV/c^2), and this can be fit with an Gaussian function. There is additional combinatorial background which is fitted with a linear function. This contribution is absorbed in the total combinatorial background of our nominal channel of the main analysis. This will be described in details in Section 4.6. The

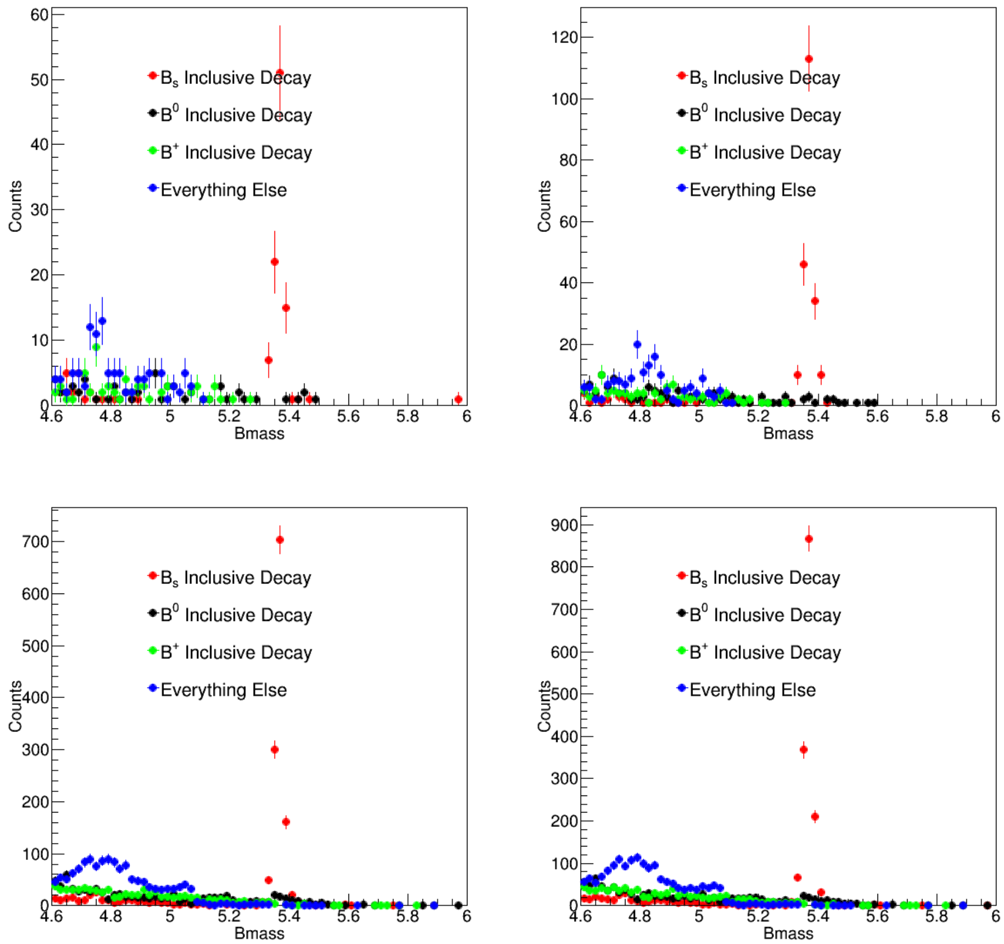


Figure 5-29: B^+, B_s^0, B^0 channels make nearly equal contribution in the signal region for all p_T bins 10 - 15 GeV/c, 15 - 20 GeV/c, 20 - 50 GeV/c, and 10 - 50 GeV/c.

shape of the NP background model function is used as template in the fit extraction procedure.

Further MC studies have been done in order to identify the different channels that give rise to the non-prompt peaking structure in the B^+ invariant mass spectrum. Several main processes have been identified as follows:

- Case 1: 4-body B^+ decays which occur via resonant decay channels e.g. $B^+ \to J/\psi K^{*+}(892)$. In these cases, we distinguish the kaons coming from the $K^{*+}(892)$ decays as coming from a signal $B^+ \to J/\psi\ K^+$ decay.

- Case 2: 4-body B^0 decays channels e.g. $B^0 \to J/\psi K^{*0}(892)$.

- Case 3: $B^+ \to J/\psi\pi^+$ decays in which we have misidentified the π^+ as a K^+.

The different contributions in PbPb are presented in Fig.5-31. The contribution from $B^+ \to J/\psi\pi$ clearly form a peaking structure on the right shoulder of the nominal decay channel B^+ decay. However, the overall magnitude of this component is tiny compared to the other two sources, and negligible compared to the nominal signal. As a consequence, we can barely see the contribution of this peaking structure in the invariant mass plot of B^+ nominal channel.

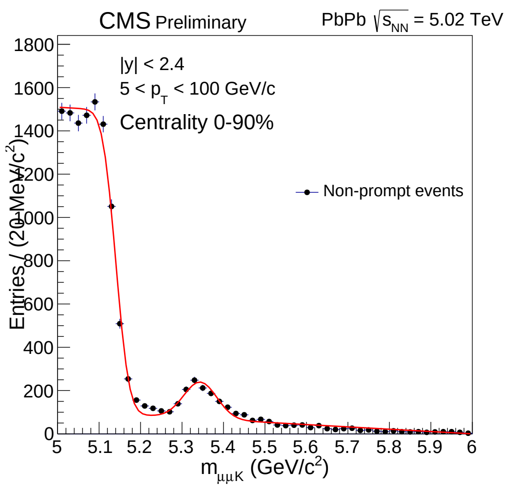


Figure 5-30: B^+ candidate mass spectrum obtained in inclusive B meson MC production after vetoing the contribution of genuine B^+ signal candidates in PbPb.

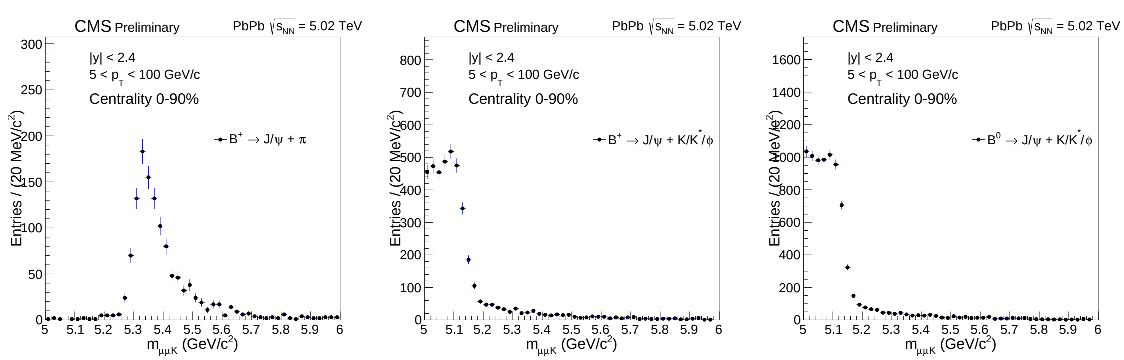


Figure 5-31: Peaking background contribution from $B^+ \to J/\psi\pi$ and from K resonant decay channels of B^0 and B^+ in PbPb MC.

The conclusion from these studies is that we need to use a function to model the

NP background component. Its template shape should be determined according to the B^+ NP MC sample and applied to the data with a scale parameter. According to our studies, such function is the error function in the left hand shoulder (low than the B^+ PDG mass) and a double Gaussian in the B^+ signal region.

5.8 Signal Extraction

Now, equipped with the optimal selection and the NP background studies, we are ready to extract the signal raw yield from the B-meson invariant distribution and measure the cross section. Also, since we see very clear B_s^0 signal, we can estimate the B_s^0 significance and check if there will be an observation.

5.8.1 Fitting Models

Raw yields are extracted through extended unbinned maximum likelihood fits to the invariant mass of reconstructed B_s^0 meson candidates, performed using the *Roofit* package [213]. The unbinned fit can reduce the potential bias due to the binning artifact. We develop the probability density function (PDF) to fit to the B-meson invariant mass distributions to extract their signal raw yields. In the PDF, the signal region for both B_s^0 and B^+ are described by Gaussian functions with the same means but different widths and the combinatorial background is modeled with an exponential decay function. For B^+, an additional error function in the left hand shoulder (low than the B^+ PDG mass) and a double Gaussian in the B^+ signal region to fit NP background according to the template fits to inclusive NP J/ψ MC sample in Section 4.7.

Hence, the generic event likelihood in data is described by the formula below

$$\mathcal{L}(m; N_S) = N_S \cdot (\alpha G(m; M, \sigma_1) + (1-\alpha)G(m; M, \sigma_2)) + N_B \cdot E(m; \lambda_m) \qquad (5.6)$$

$$G(m; M, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(m-M)^2}{2\sigma^2}} \tag{5.7}$$

$$E(m; \lambda_m) = \exp^{-\lambda_m m} \tag{5.8}$$

Where m is the candidate mass (input); M and σ_i are the signal mass mean and widths (resolution); G and E denote respectively Gaussian and Exponential functions, normalized in the fitting mass window; N_S denotes the signal raw yield (the parameter of interest), N_B is the background yield, while α and λ_m are nuisance parameters (describing the signal fractions and exponential decay slope).

It should note that the main reason for using double Gaussian functions instead of a single Gaussian is that the reconstructed B_s^0 signal width varies as a function of B-meson pT due to the p_T dependence of the track p_T resolution. Figure 5-32 shows the B_s^0 and B^+ invariant mass width as a function of p_T in the MC

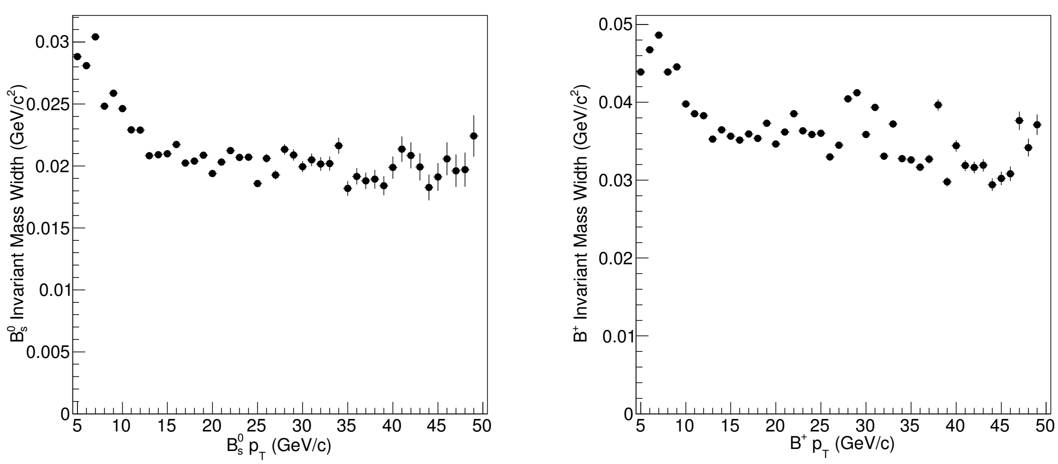


Figure 5-32: The signal B_s^0 (left) and B^+ (right) invariant mass width as a function of p_T are shown above.

5.8.2 Raw Yield Extraction

To obtain the signal raw yields and their statistical uncertainties from B_s^0 and B^+ invariant mass distribution, the following fitting procedures are carried out within

B-meson invariant mass window $5 < m < 6$ GeV/c^2.

- First, a fit is performed, with a double gaussian function to the MC invariant mass distribution of the genuine B-meson signals.
- For B^+, the shape of the non-prompt background component is obtained using the dedicated non-prompt J/ψ MC samples.
- The fit is performed to the data with the fixed the shape (widths and relative proportion of the two Gaussians) same with the Gaussians obtained from the MC fit.
- For systematic uncertainty check, add a free parameter (a), that is commonly multiplied to the widths of the signal Gaussians, serving as a scale factor of the resolution that parametrizes data and MC signal shape difference. However in the nominal fit, this is set to be unity ($a = 1$), which means the widths of the data signal are set to be identical to the ones of the MC signal.
- The parameters of the background PDF, the mean of the signal Gaussians are the free parameters of the fit.

Figure 5-33 shows the fitting results of B_s^0 in the p_T range of [7, 10, 15, 20, 50] GeV/c.

Figure 5-34 shows the fitting results of B_s^0 in the centrality range of [0, 30, 90] in PbPb collisions.

Figure 5-35 shows the fitting results of B^+ in the p_T range of [7, 10, 15, 20, 50] GeV/c.

Figure 5-36 shows the fitting results of B_s^0 in the centrality range of [0, 30, 90] in PbPb collisions.

Table 5.8 and 5.9 below summarize the selected fit parameters and their error extracted from the fits of B_s^0 and B^+ respectfully

At a glance, we can see that our fits looks good. In addition to the fits, the pull, defined as the ratio of the difference between the data and the fit to the statistical

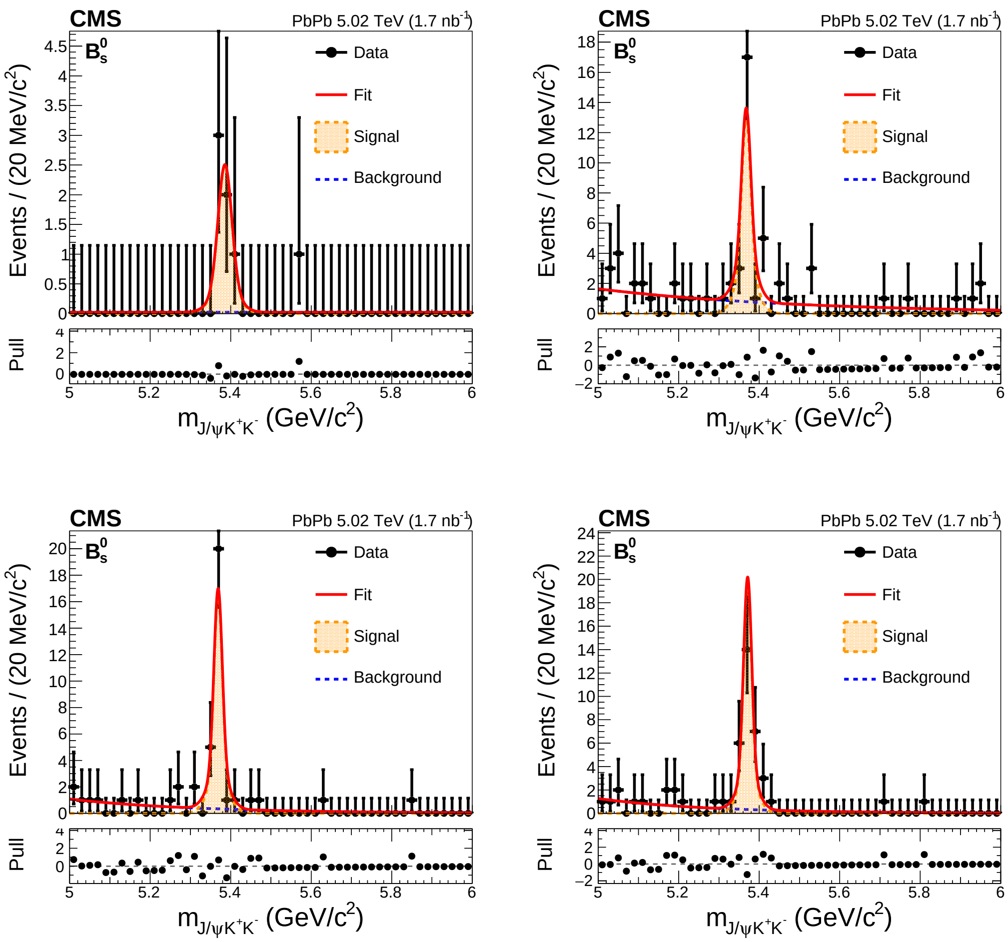


Figure 5-33: The B_s^0 invariant mass distributions as well as the fits to extract the signal raw yield N_S in different p_T bins are shown above.

uncertainties of the data, are also shown above in Figure **??**. We can see that the pull is basically consistent with 0 with a 2 σ fluctuation, which suggests that our fit also looks good. A dedicated closure test on the fits will be conduct later to further validate our fits.

5.8.3 Signal Significance Estimation

The significance (Z) is calculated through a likelihood method, which follows the formula below:

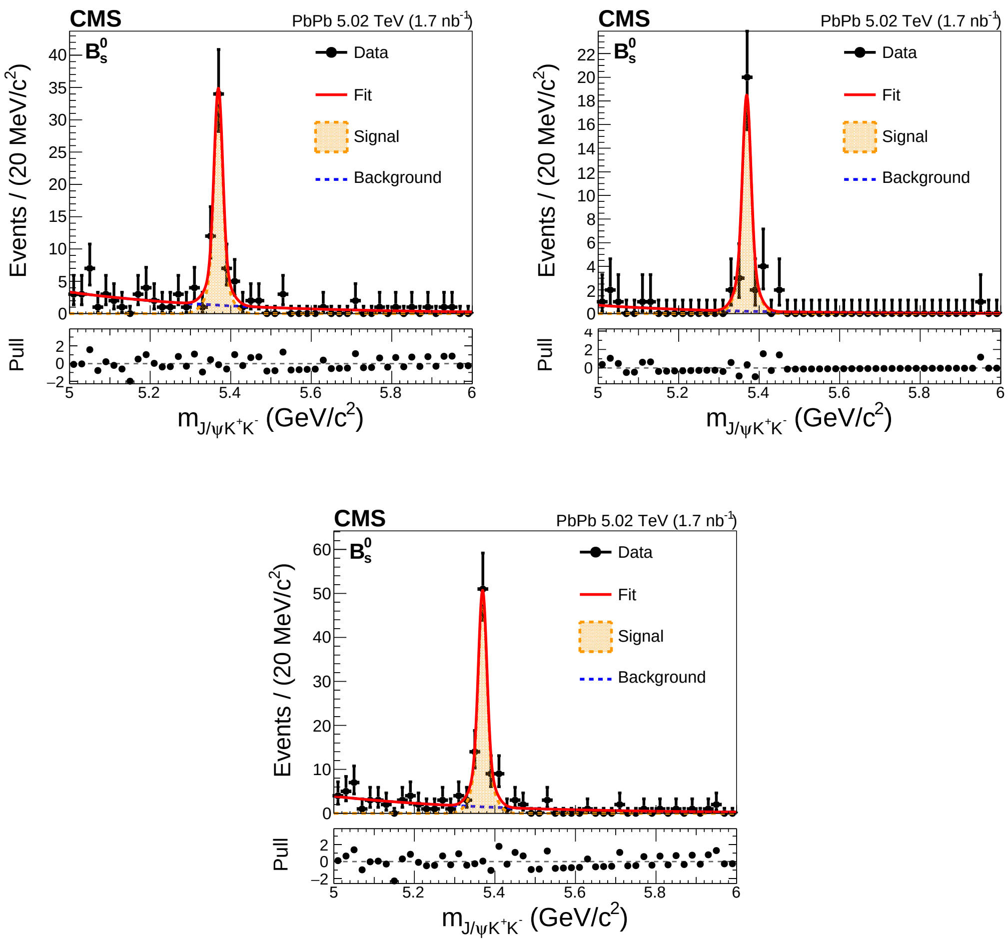


Figure 5-34: The B_s^0 invariant mass distributions as well as the fits to extract the signal raw yield N_S in different centrality bins are shown above.

$$Z = \sqrt{2 \log \frac{L_{S+B}}{L_B}} \tag{5.9}$$

Here, L_{S+B} is the likelihood of each fit and L_B is the likelihood when the number of signal N_S parameter is fixed to 0. Figure 5-37 show the likelihood scan of B_s^0 in 4 p_T bins and 3 centrality bins

Table 5.10 summarizes the significance of B_s^0 according to our likelihood estimation.

We can see that B_s^0 have significances of greater than 5σ for all its p_T and centrality bins. In fact, this is the first observation of fully reconstructed B_s^0 meson in nucleus-

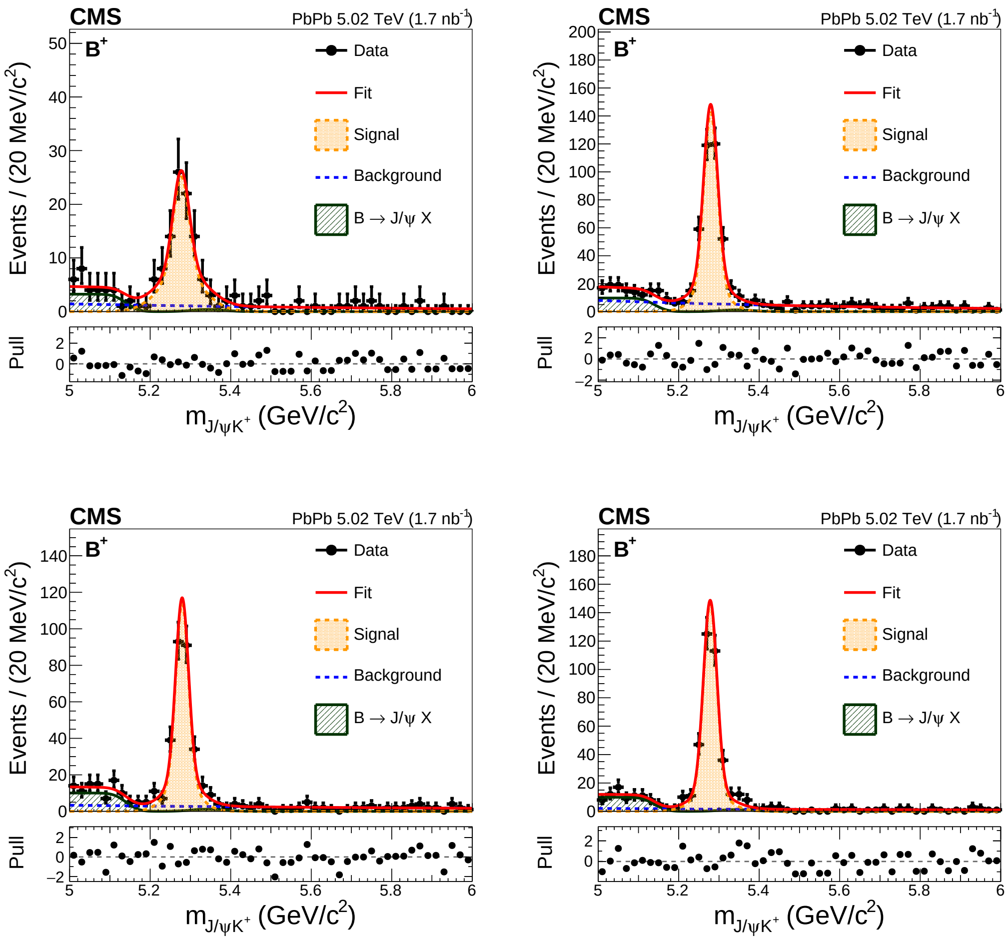


Figure 5-35: The B^+ invariant mass distributions as well as the fits to extract the signal raw yield N_S in different p_T bins are shown above.

nucleus collision with greater than 5σ significance.

5.9 B Mesons Candidates MC-Data Comparison

5.9.1 overview

Ideally, if the simulation is impeccable, the RECO distribution in the MC should match perfectly with data. Nevertheless, there is always limitation in the MC simulation because the incorrect model of physics processes in the generation side or poor model of detector conditions the in the reconstruction side. The discrepancy should

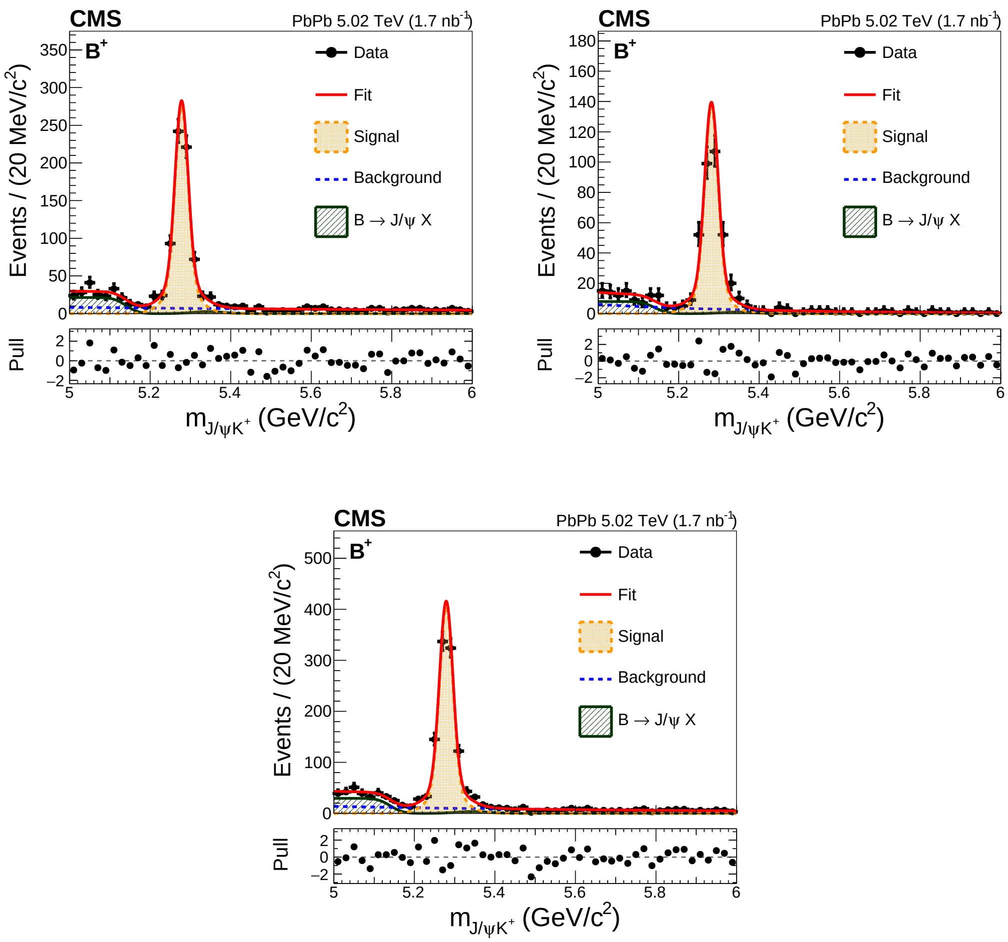


Figure 5-36: The B^+ invariant mass distributions as well as the fits to extract the signal raw yield N_S in different centrality bins are shown above.

be quantified as a source systematic uncertainties. The MC simulation plays a crucial role in the data analysis and could affect our final results significantly. In order to compare the consistency between data and MC, we need to look at the B-meson signal only candidates. In the MC, we can simply apply GEN Match selection to select all B-meson signal candidates. In the data, we need to reject the background and extract the signal. After that, we compare the normalized distributions of the data and MC to compare their shapes.

Table 5.8: The summary table of B_s^0 fits results of Gaussian mean, signal raw yield, and background raw yield as well as their uncertainties.

Centrality (%)	B_s^0 p_T (GeV/c)	Gaus Mean (GeV/c^2)	Sig Yield (N_S)	Bkgd Yield (N_B)
0 – 90	7 – 10	5.386 ± 0.007	5.94 ± 2.43	1.06 ± 1.03
0 – 90	10 – 15	5.368 ± 0.003	23.41 ± 5.28	35.59 ± 6.28
0 – 90	15 – 20	5.369 ± 0.002	26.70 ± 5.31	16.33 ± 4.20
0 – 90	20 – 50	5.371 ± 0.003	30.19 ± 5.73	16.82 ± 4.39
0 – 30	10 – 50	5.369 ± 0.002	54.64 ± 7.86	60.35 ± 8.19
30 – 90	10 – 50	5.370 ± 0.003	29.88 ± 5.58	10.13 ± 3.37
0 – 90	10 – 50	5.369 ± 0.002	80.25 ± 9.43	68.74 ± 8.79

Table 5.9: The summary table of B^+ fits results of Gaussian mean, signal raw yield, and background raw yield as well as their uncertainties.

Centrality (%)	B^+ p_T (GeV/c)	Gaus Mean (GeV/c^2)	Sig Yield (N_S)	Bkgd Yield (N_B)
0 – 90	7 – 10	5.277 ± 0.004	92.33 ± 10.88	7.90 ± 2.33
0 – 90	10 – 15	5.279 ± 0.001	354.6 ± 20.68	39.9 ± 6.28
0 – 90	15 – 20	5.279 ± 0.001	26.70 ± 5.31	16.33 ± 4.20
0 – 90	20 – 50	5.278 ± 0.001	30.19 ± 5.73	16.82 ± 4.39
0 – 30	10 – 50	5.278 ± 0.001	657.7 ± 27.7	53.46 ± 5.28
30 – 90	10 – 50	5.281 ± 0.001	327.0 ± 19.5	19.57 ± 4.69
0 – 90	10 – 50	5.279 ± 0.001	971.6 ± 33.9	74.16 ± 6.91

5.9.2 Splot Techniques

To carry out MC-Data comparison studies, the dedicated *Splot* method is used. It is a likelihood-based method by which we reweigh the data using the unbinned fit result. The weights are added to the dataset based on model and yield extraction variables. Each event has two weights: probability of belonging to the signal given its mass, probability of belonging to the background given its mass. The *Splot* class gives us the distributions of our variables for a given species (signal or background). The advantage of using this method is that we use the full dataset for the comparison in contrast to the sideband subtraction method where one should select the investigation range of signal and background. Furthermore, we use likelihood to describe events'

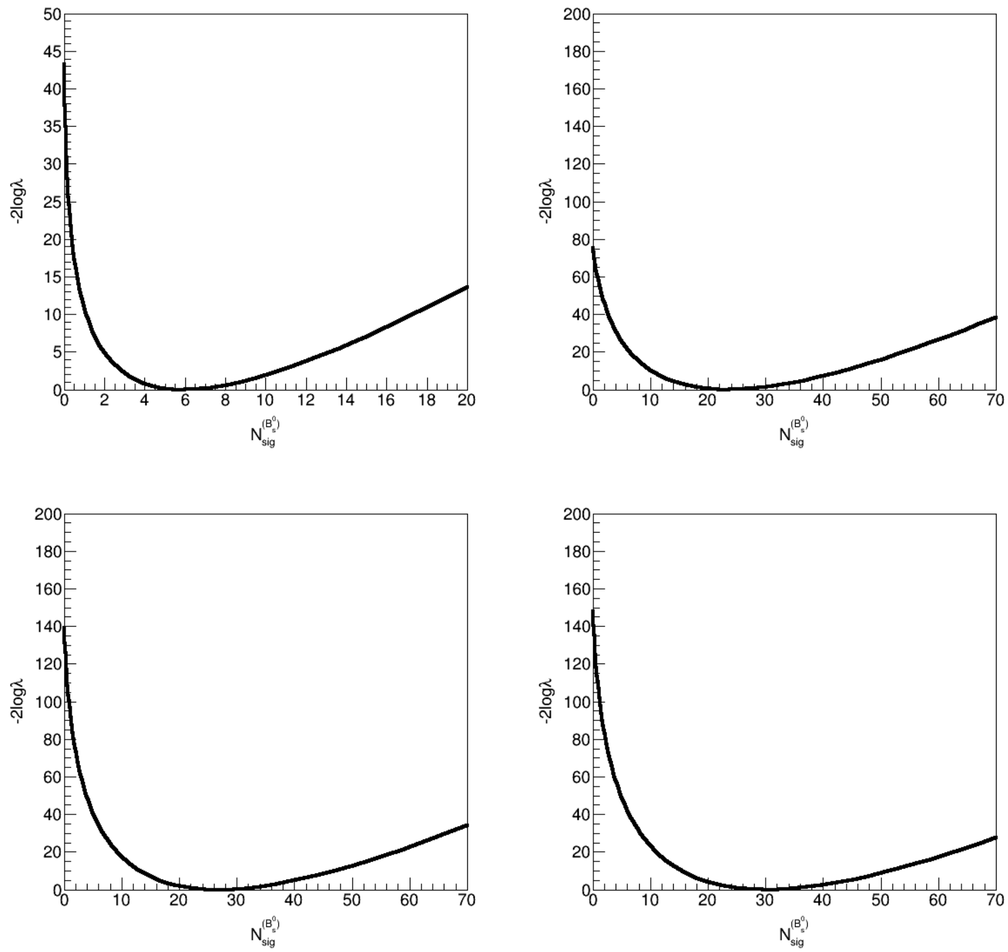


Figure 5-37: The significance vs signal yield for p_T bins at 7 – 10, 10 – 15, 15 – 20, and 20 – 50 GeV/c for 0 - 90% centrality are shown above

behavior in contrast to the potential misidentification of signal events in background region which might occur in sideband subtraction method.

In order to obtain the *Splot* weight, we first need to fit the data using a discriminating variable, namely the B meson candidates invariant mass. This method, like the one described in the previous section, assumes the discriminating variable chosen to be independent of the variables we wish to study. We then use the

fit to attribute to each event two weights: w_S, which corresponds to the probability of it belonging to the signal, and w_B, which corresponds to the probability of fit belonging to the background. The weights are qualitatively demonstrated in Figure

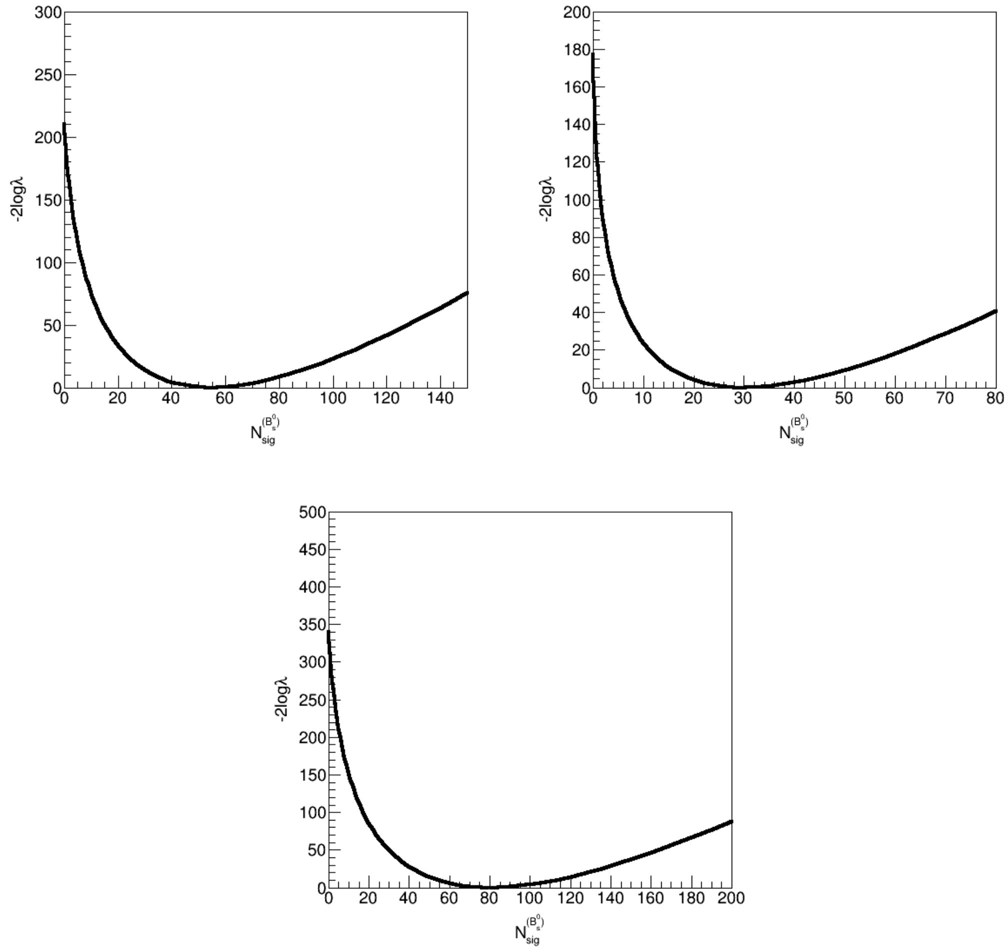


Figure 5-38: The significance vs signal yield for centrality bins in 0 - 30%, 30% - 90% 0 - 90% for p_T in 10 - 50 GeV/c are shown above

5-39.

5.9.3 Splot Variable Correlation Studies

First, in order to apply Splot to perform Data-MC comparison, we need to confirm our variables are indeed not correlated to the invariant mass. Therefore, the correlation matrices for BDT variables vs B^+ invariant mass (Bmass) for data and MC are shown in Figure 5-41:

Table 5.10: The summary table of B_s^0 likelihood significance for each p_T and centrality bin.

Centrality (%)	B_s^0 p_T (GeV/c)	Likelihood Significance (Z)
0 – 90	7 – 10	5.3
0 – 90	10 – 15	7.9
0 – 90	15 – 20	10
0 – 90	20 – 50	10
0 – 30	10 – 50	14
30 – 90	10 – 50	11
0 – 90	10 – 50	16

5.9.4 Splot Results for Data-MC Comparison

No significant correlation between the BDT variables and the invariant mass is observed in either B_s^0 and B^+, which also validate their BDT trainings. Therefore, Splot will be applicable to compare data and MC. Here, we focus on BDT values that are directly used in our signal extraction and related to MC distribution validation, rather than the variables themselves used in BDT training. In a wide ranges of BDT, the two distributions show good agreement. We only focus on the region where BDT is greater than the working point and is smaller than the maximum value that candidates have.

From Figure 5-46, the B^+ BDT variables in for all p_T bins have overall reasonably good agreement between Data and MC. Their ratio is near unity with some fluctuations due to limited statistics. In the systematic section, we will quantify the discrepancy between Data and MC as a source of systematic uncertainties. For B_s^0, unlike B^+ whose statistics is in general > 100, the total number of signal B_s^0 candidate is < 100, which results in large error bar shown in Figure 5-42. Nonetheless, the ratio of Data to MC shape is still near unity within the large uncertainties. Nevertheless, since the events used by B_s^0 and B^+ are essentially the same and their tracks are very similar, the overall good B^+ Data-MC agreement can provide indirect validation to B_s^0. The CMS analysis note AN-19-219 [**?**] documents more details on general description of *Splot* method applied to this analysis.

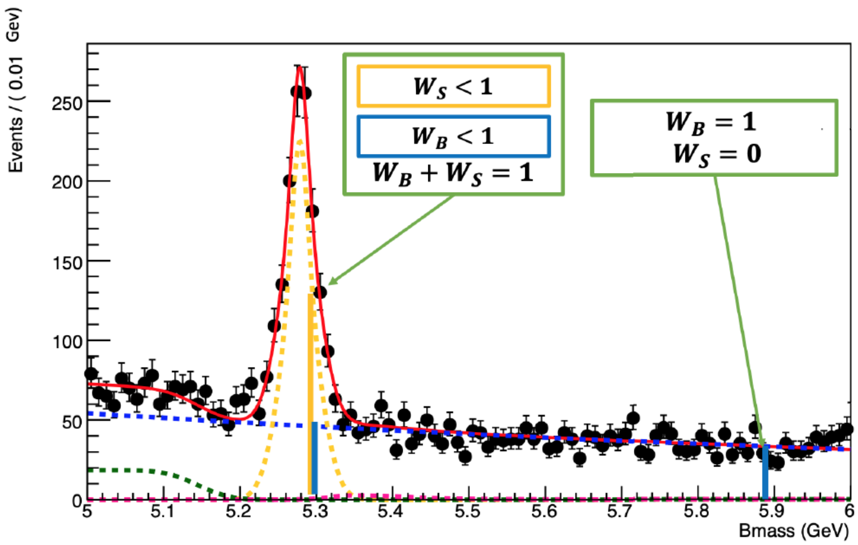


Figure 5-39: The illustration of *Splot* techniques via an unbinned fit to B^+ invariant mass distribution using our fitting model to extract the *Splot* weights w_S and w_B is shown above.

5.10 Acceptance and Efficiency Correction

5.10.1 Overview

Now, we have clearly seen B-meson signals and extract their signal raw yields N_S from their invariant distributions. In the next step, we need to correct the acceptance and selection efficiency of the B mesons in order to obtain the cross section. The following procedures define the method we use to determine B mesons acceptance and selection efficiency with B-meson MC samples:

First, We count the total number of GEN-level B meson candidates reweighed by the centrality, PVz, and $\hat{p}_T$ as **NBGen** within the given B-meson rapidity region $|y| < 2.4$.

Next, we count the number of generated B meson candidates passing the follow selection

Muon track selections:

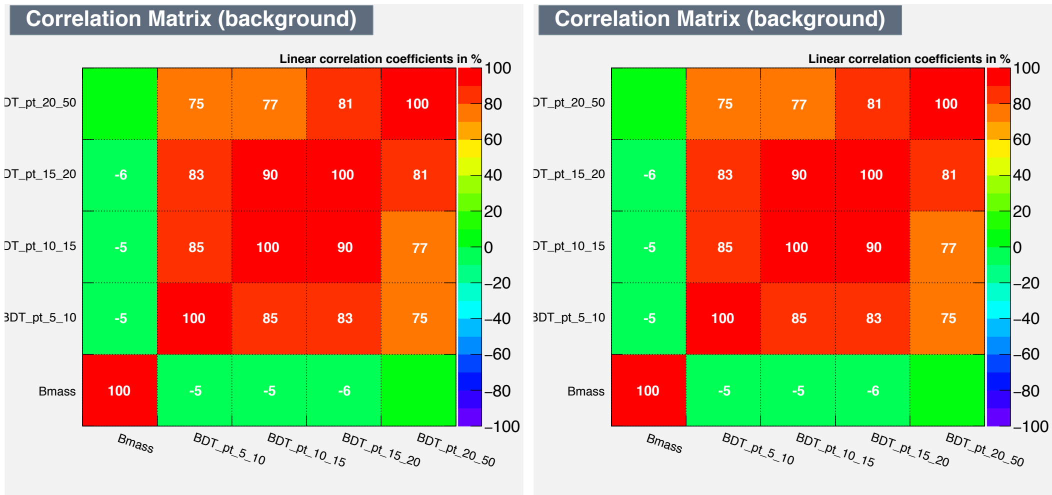


Figure 5-40: The correlation matrices in data (left) and MC (right) of B_s^0 are shown above.

$$
\begin{aligned}
p_T^{\mu} &> 3.5GeV/c && \text{for } |\eta^{\mu}| < 1.2 \\
p_T^{\mu} &> (5.47 - 1.89 \times |\eta^{\mu}|)GeV/c && \text{for } 1.2 \le |\eta^{\mu}| < 2.1 \\
p_T^{\mu} &> 1.5GeV/c && \text{for } 2.1 \le |\eta^{\mu}| < 2.4
\end{aligned} \tag{5.10}
$$

Kaon track selections:

$$
\begin{aligned}
p_T^{K} &> 0.9GeV/c \\
|\eta^{K}| &< 2.4
\end{aligned} \tag{5.11}
$$

to the generated B meson candidates denoted as **NPassAcc**.

Finally, we apply all the selections on the GEN-matched RECO B mesons. We count the reconstructed B meson candidates passing the selections mentioned in Section 4.5 “Muon and J/ψ candidates selections" and the optimal BDT selections in Table 3 in Section 4.1 and denote the number as **NSelPass**.

The acceptance is defined by: **acceptance** $=$ **NPassAcc/NBGen**.

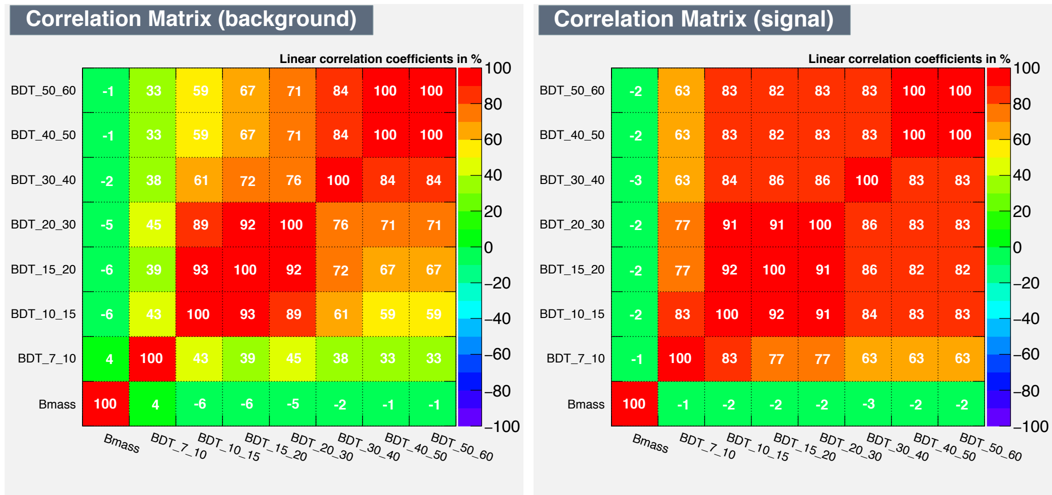


Figure 5-41: The correlation matrices in data (left) and MC (right) of B^+ are shown above.

The selection efficiency is defined by: **selection efficiency** $=$ **NSelPass/NPassAcc**.

Here, we denote α **as the acceptance** and ϵ **as the selection efficiency** andtheir product, which is $\alpha \times \epsilon =$ **NSelPass/NBGen**, is call (total) efficiency.

5.10.2 Tag & Probe Techniques

The previous subsection mentioned the efficiency correct for B-mesons using MC simulations only. However, we know that both B_s^0 and B^+ decay to J/ψ. A dedicated analysis techniques called tag & probe is developed to correct the B-meson efficiency in data-driven way. The tag-&-probe method uses the *scale factor* defined the ratio of Data/MC for L2 and L3 triggered single muon efficiency tagged from the J/ψ resonance [214]. There are three scale factors: identification (id), tracking (trk), and trigger (trg). Figure 5-44 shows workflow for tag-&-probe in B-meson analysis

The total scale factor (SF) of two muons ($SF^{\mu\mu}$) is given by the following formula

$$SF^{\mu\mu} = id^{\mu_1} \times trk^{\mu_1} \times trg^{\mu_1} \times id^{\mu_2} \times trk^{\mu_2} \times trg^{\mu_2} \tag{5.12}$$

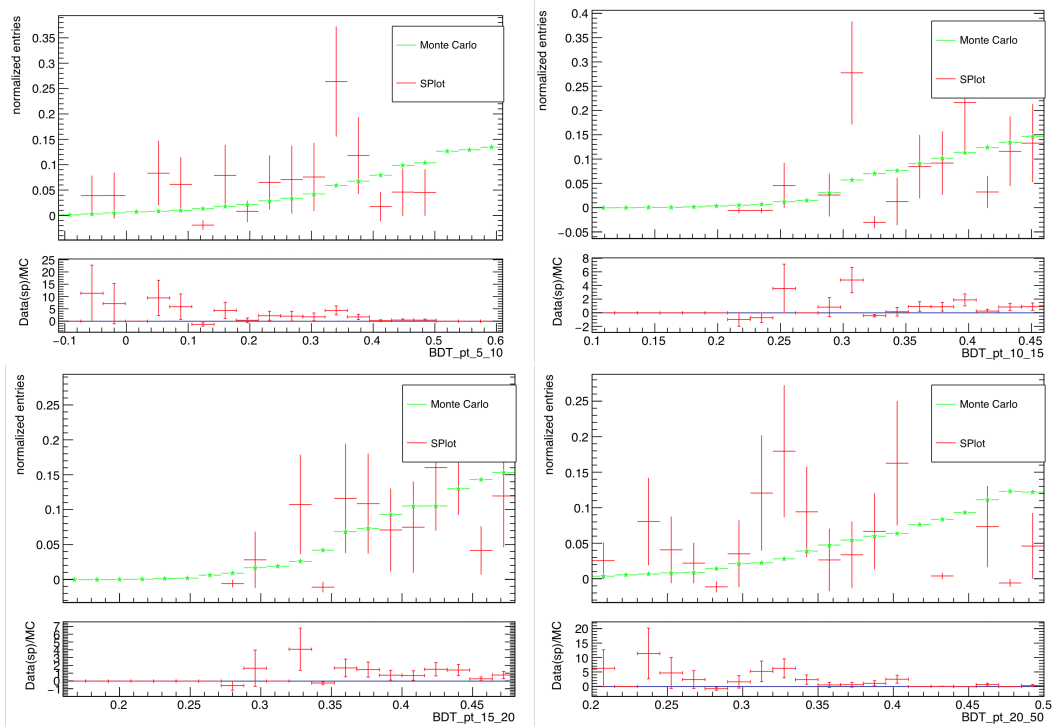


Figure 5-42: Comparison of B_s^0 BDT distribution in data (red) and MC (green) using Splot method.

It should note that, as mentioned in Section 2.2.5, the dimuon PbPb triggered datasets consistent of one L2 muon and one L3 muon. In our analysis, the two muons used to reconstruct B-meson are made of either one L2 muon and one L3 muon (L2,L3) or two L3 muons (L3,L3). The muon type L2 and L3 only affects the trg SF. In the (L2,L3) case, we simply correct the scale factor according to their individual scale factor values. However, in the (L3,L3) case, we take treat them as the combination of (L2,L3) and (L3,L2) by considering one of the muon as L2 muon. Then we compute the SF for both cases and take the average of the two SF as the SF for (L3,L3).

5.10.3 Traditional Efficiency Correction Results

Traditionally, we simply compute the efficiency as a function of binned p_T and centrality and use it to correct the signal raw yield within the p_T and centrality bins. The B_s^0 and B^+ acceptance, selection efficiency, and total efficiency as a function of

p_T and centrality are shown respectfully below in Figure **??** and Figure **??**

5.10.4 Analysis Challenges

Although B-meson p_T and y distributions are reasonably modeled by FONLL in pp collisions, the precise B-meson p_T and y distribution shapes are still unknown in PbPb, which could significantly affect the efficiency determination. In the following, we demonstrate of the generated B-meson kinematics how the unknown shape can affect the efficiency.

Since we know the efficiency is given by $\alpha \times \epsilon$ = **NSelPass/NBGen**. NBGen in the simulation could not represent the truth. A weight function $w(x)$ is defined as the ratio between the generated B-meson cross section to the cross section of actual B-meson in PbPb for some kinematic variable x as follows

$$w(x) = \frac{\sigma_B^{MC}(x)}{\sigma_B^{PbPb}(x)} \tag{5.13}$$

Therefore, this weight function needs to be applied to both GEN-level $G(x)$ and RECO-level $R(x)$ in order to obtain the correct selection efficiency as a function of B mesons p_T.

$$\epsilon_{corr}(x) = \frac{\langle R(x) \rangle}{\langle G(x) \rangle} = \frac{\int R(x)w(x)dx}{\int G(x)w(x)dx} \tag{5.14}$$

Experimentally, the data are discreet. Therefore, the measurement is done in bins. Now considering a bin of $[x_1, x_2]$, the expression above is written as

$$\epsilon_{corr}(x_1 < x < x_2) = \frac{\int_{x_1}^{x_2} R(x)w(x)dx}{\int_{x_1}^{x_2} G(x)w(x)dx} \tag{5.15}$$

When $x_1 \to x_2$, we have

$$\epsilon_{corr}(x_1 < x < x_2) = \frac{\int_{x_1}^{x_2} R(x)w(x)dx}{\int_{x_1}^{x_2} G(x)w(x)dx} \sim \frac{R(x)w(x)(x_2 - x_1)}{G(x)w(x)(x_2 - x_1)} = \frac{R(x)}{G(x)} = \epsilon(x) \tag{5.16}$$

We can see that, for any non-trivial $w(x)$, with very fine bin, the effect of $w(x)$ is gone. When we measure B-meson over a large p_T range, this effect could be significant, particular for B_s^0 where the change of efficiency between 7 – 10 GeV/c and 10 – 15 GeV/c and the slope of uncertainties on $w(x)$ both are huge. Hence, in order to eliminate effect of the uncertainties due to the B-meson kinematics $w(x)$ on efficiency correction, we can bin the efficiency finely.

5.10.5 Fiducial Measurement

In fact, to better estimate B-meson efficiency, we should use 2D efficiency maps as functions of B-meson p_T and $|y|$. Figure 5-47 shows reconstructed B_s^0 and B^+ distribution in the MC

As we can see, for B-meson p_T below 10 GeV/c, we have very little B mesons reconstructed at the rapidity region of $|y| < 1.5$. This is due to the limited acceptance of muons because the muon tracks at low p_T cannot reach the muon systems. Hence, a fiducial measurement is carried out. We only correct B-meson for $p_T < 10$ GeV/c to the $1.5 < |y| < 2.4$ instead of $|y| < 2.4$. For $p_T > 10$ GeV/c, we still correct them to $|y| < 2.4$. The fiducial B-meson measurements will be carried out throughout this thesis.

5.10.6 Finely Binned 2D Efficiency Map

After unknown B-meson kinematics effects on efficiency correction and choosing the fiducial region for our measurement, we propose to measure the inverse of the total efficiency: $\frac{1}{\alpha\times\epsilon}$ as a function of p_T and $|y|$ to correct the B-meson raw yield to production yield. We implement the following workflow in Figure 5-48 to estimate $\langle\frac{1}{\alpha\times\epsilon}\rangle$

The $\frac{1}{\alpha\times\epsilon}$ applied with the tag-&-probe SFs as functions of p_T and $|y|$ for B_s^0 and B^+ are shown on Figure 5-49 and Figure 5-50 respectfully

The p_T bin width is 0.5 GeV/c from 5 - 10 GeV/c and 1 GeV/c from 10 - 50 GeV/c. The $|y|$ binning is [0, 1.2, 1.8, 2.1, 2.4].

5.10.7 Data-Drive Efficiency Correction

Finally, we propose to correct the efficiency with a data-driven method. We correct the signal B-meson candidate efficiency. This can be done by looping B-meson data signal region candidates on the 2D $\frac{1}{\alpha\times\epsilon}$ map according their kinematics. Then, referring to techniques of the published J/ψ analysis [?], we compute the average of all signal B-meson candidate efficiency: $\langle\frac{1}{\alpha\times\epsilon}\rangle$ within the p_T and centrality bins. Here, we call $\langle\frac{1}{\alpha\times\epsilon}\rangle$ as the efficiency correction factor. Mathematically, it is written as

$$\langle\frac{1}{\alpha \times \epsilon}\rangle = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\alpha_i \times \epsilon_i} \tag{5.17}$$

It should note that we will measure $\langle\frac{1}{\alpha\times\epsilon}\rangle$ instead of $\frac{1}{\langle\alpha\times\epsilon\rangle}$ is because $\langle\frac{1}{\alpha\times\epsilon}\rangle$ has better closure. A dedicated closure test on this approach will be carried out and the results discussed the Section 4.11.

Also, there are background contaminations even within the signal region. This could be improve by Splot method. However, it turns out that the difference in the efficiency between applying Splot and not applying Splot is very small. We do not consider using the Splot method in this measurement.

Finally, this method only gives the nominal results. The statistical uncertainties turn out to be correlated with the both the data and MC statistics. A dedicated data bootstrapping approach to estimate the statistical uncertainties of the efficiency correction factor $\frac{1}{\langle\alpha\times\epsilon\rangle}$ will be carried out in Section 4.12.

5.10.8 Results

After applying data-drive method to compute efficiency correction factor, we obtain Figure 5-51 and Figure 5-52 showing $\langle\frac{1}{\alpha\times\epsilon}\rangle$ as a function p_T and centrality for both B_s^0 and B^+.

Table 5.11 and 5.12 summarize the nominal efficiency of B_s^0 and B^+ respectfully and their upper and lower asymmetric statistical uncertainties obtained from Section 4.12.

Table 5.11: The summary table of B_s^0 efficiency correction factor $\langle \frac{1}{\alpha \times \epsilon} \rangle$ for each p_T and centrality bin.

Centrality (%)	B_s^0 p_T (GeV/c)	$\langle \frac{1}{\alpha \times \epsilon} \rangle$	$\langle \frac{1}{\alpha \times \epsilon} \rangle$ Error Up +)	$\langle \frac{1}{\alpha \times \epsilon} \rangle$ Error Down -)
0 – 90	7 – 10	381.5	19.3%	20.3%
0 – 90	10 – 15	75.92	12.4%	12.3%
0 – 90	15 – 20	22.35	6.41%	6.47%
0 – 90	20 – 50	10.63	5.90%	6.20%
0 – 30	10 – 50	45.90	14.9%	14.6%
30 – 90	10 – 50	14.19	10.3%	9.78%
0 – 90	10 – 50	34.90	17.9%	16.3%

Table 5.12: The summary table of B^+ efficiency correction factor $\frac{1}{\langle \alpha \times \epsilon \rangle}$ for each p_T and centrality bin.

Centrality (%)	B^+ p_T (GeV/c)	$\langle \frac{1}{\alpha \times \epsilon} \rangle$	$\langle \frac{1}{\alpha \times \epsilon} \rangle$ Error Up +)	$\langle \frac{1}{\alpha \times \epsilon} \rangle$ Error Down -)
0 – 90	7 – 10	105.9	15.8%	15.1%
0 – 90	10 – 15	37.55	4.10%	7.95%
0 – 90	15 – 20	10.94	6.54%	6.50%
0 – 90	20 – 50	5.932	6.90%	5.26%
0 – 30	10 – 50	21.67	5.81%	5.54%
30 – 90	10 – 50	12.28	6.71%	7.06%
0 – 90	10 – 50	19.23	5.06%	4.54%

At this point, we have all the ingredients to measure the B-meson cross section and study the physics.

5.11 Cross Section Measurement

The goal is this thesis is to measure the cross section of B-meson in PbPb collisions as functions of B mesons p_T and PbPb collision event centrality. Here, in high energy heavy-ion physics, cross section really means production yield.

In terms of B mesons p_T, the cross section is defined as mathematically as follows:

$$\frac{1}{T_{AA}}\frac{dN}{dp_T} = \frac{1}{T_{AA}}\frac{1}{2}\frac{1}{\Delta p_T}\frac{1}{N_{MB}BR}N_S\langle\frac{1}{\alpha \times \epsilon}\rangle \tag{5.18}$$

In terms of PbPb collision event centrality, the cross section is defined as mathematically as follows:

$$\frac{1}{T_{AA}}N = \frac{1}{T_{AA}}\frac{1}{2}\frac{1}{N_{MB}BR}N_S\langle\frac{1}{\alpha \times \epsilon}\rangle \tag{5.19}$$

The definitions of the variables above are shown below:

- N: B-meson production yield
- T_{AA}: Nuclear overlapping function - here is for spherical nucleus $^{208}_{82}$Pb
- N_{MB}: Number of minimum biased events corresponding to the dimuon PbPb datasets
- $\frac{1}{2}$: Divide by 2 for the purpose of normalizing particle and antiparticle
- BR: B-meson decay branching ratio
- Δp_T: B-meson transverse momentum bin width
- N_S: Signal raw yield
- $\langle\frac{1}{\alpha\times\epsilon}\rangle$: Data-drive efficiency correction factor

It should be pointed out that we aim at measuring the p_T differential cross section as a function B mesons p_T while as we plan to present the p_T integrated cross section a function of event centrality in this thesis. Before presenting our final results of the cross section measurement, we need to validate our signal raw yield N_S and efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ measurements and estimate the correct statistical and systematic uncertainties of the B-meson measurements.

5.12 Validation Tests

Before presenting the final results, some more validation tests need to complete to ensure our results are unbiased.

5.12.1 Mass Scraping Test

From Figure 5-40, we see that the BDT variables are essentially uncorrelated to the B_s^0 invariant mass. However, uncorrelation is not equivalent to independence. We need to make sure that in the BDT training procedure, no significant invariant mass dependence is introduced. This effect is called the mass scraping effect. Therefore, we propose the following tests to explicitly quantify the B_s^0 BDT variables dependence on the B_s^0 invariant mass.

- We look at the BDT shape in different B_s^0 data invariant mass side band regions (both left and right hand side) that is far away from the B_s^0 PDG mass.

- Then we take the ratio from other BDT shape with respect to one side band to produce the ratio as a function of the BDT in different mass region

- We fit the BDT shapes with a linear function, which is treated as a weight as a function of BDT: $w(BDT)$

- We apply the weight $w(BDT)$ to B_s^0 candidate and plot their invariant mass distribution with the weight

- We fit the $w(BDT)$ weighed invariant mass distribution and extracted the $w(BDT)$ weighed yield N_S^w

- We count the total number of candidates in both $w(BDT)$ unweighed and weighted case N and N^w

- We compute the percent difference δ from the rescaled $w(BDT)$ B_s^0 signal raw yield $\frac{N}{N^w} N_S^w$ to the unweighed nominal B_s^0 signal raw yield N_S

$$\delta = \frac{\frac{N}{N^w}N_S^w - N_S}{N_S} \times 100\% \tag{5.20}$$

Table 5.13 shows the nominal and linearly weighted rescaled signal raw yield and the percent deviation of the variated signal raw yield from nominal raw yield for each p_T and centrality bin.

Table 5.13: The summary table of BDT vs mass dependence systematics on the background yield from the fits in different p_T and centrality bins are shown below.

Centrality	p_T (GeV/c)	$\frac{N}{N^w}N_S^w$	N_s	Percent Deviation (δ)
0 - 90%	7 - 10	6.831	6.787	**0.64%**
0 - 90%	10 - 15	27.14	27.23	**0.33%**
0 - 90%	15 - 20	26.69	26.78	**0.34%**
0 - 90%	20 - 50	31.26	31.41	**0.48%**
0 - 30%	10 - 50	59.25	60.19	**1.59%**
30 - 90%	10 - 50	29.88	30.04	**0.20%**
0 - 90%	10 - 50	84.95	86.07	**0.14%**

We can see the mass scraping effect of B_s^0 is less than 2%, which is negligible. This fully validates that our B-meson BDT variables have little to no dependence on B-meson invariant mass.

5.12.2 Raw Yield Closure

In addition, we test the closure of the unbinned fits to extract signal raw yields and make sure they are good. 5000 MC toy samples are generated according to mean and uncertainties of the B-meson invariant distributions. Then, each sample is fitted with our models and produces a signal raw yield value, error, and pull. The pull of a fit parameter is related to its value p_i and the error σ_{p_i} and the mean of the total dataset $\bar{p}$ as follows:

$$Pull = \frac{p_i - \bar{p}}{\sigma_{p_i}} \tag{5.21}$$

Finally, we plot pull distribution of all 5000 samples and perform the Gaussian fits to the pull distribution to obtain the mean and width. We expect to the Gaussian to have a mean $\mu = 0$ and the width $\sigma = 1$, which is called a unit pull. Figure 5-53, 5-54, **??** and 5-56 show the pull distributions and the Gaussian fits to them for B_s^0 and B^+ signal raw yield parameter N_S for each p_T and centrality bin respectfully

According to the fits results, all pulls appear to be unit pull. This validates the closure of our fits and confirms our signal raw yield N_S has the correct mean and error yield.

5.12.3 Efficiency Closure

Next, after validating the fit closure for B-meson signal raw yield extraction, we also need to validate the efficiency correction approach to explicitly show that the efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ are indeed consistent to the truth efficiency to avoid potential bias. To make the analysis simple, we use $\hat{p}_T = 5$ the MC sample. Here, we consider both cases. The high statistics case: one whole $\hat{p}_T = 5$ MC sample. The low statistics case: 2000 spilt data-like $\hat{p}_T = 5$ MC samples same as the raw yield fit studies. In both cases, we use the same 2D $\frac{1}{\alpha\times\epsilon}$ efficiency correction map without tag-&-probe weights from the input MC sample. We compare the $\langle\frac{1}{\alpha\times\epsilon}\rangle$ to the expected value, which is the total number of generated B_s/total number of reconstructed B mesons and compute the

$$\%Dev = \frac{\langle\frac{1}{\alpha\times\epsilon}\rangle - GEN/RECO_{truth}}{GEN/RECO_{truth}} \tag{5.22}$$

to quantify the potential bias of the efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ in the analysis.

High Statistics Limit

The efficiency closure test results in high statistics limit of B_s^0 and B^+ are summarized in table 5.14 and 5.15 respectfully below

In conclusion, we can see that the biases of the $\langle\frac{1}{\alpha\times\epsilon}\rangle$, which are all below 3.5%, are negligible compared other sources of uncertainties. On the contrary, the efficiency correction factor $\frac{1}{\langle\alpha\times\epsilon\rangle}$ has a large bias. That also explains why we use $\langle\frac{1}{\alpha\times\epsilon}\rangle$ instead

Table 5.14: The results of B_s^0 efficiency factors $\langle\frac{1}{\alpha\times\epsilon}\rangle$ and $\frac{1}{\langle\alpha\times\epsilon\rangle}$ and their percent deviation are shown above.

Centrality	B_s^0 p_T (GeV/c)	NRECO	GEN/RECO	$\langle\frac{1}{\alpha\times\epsilon}\rangle$	% Dev	$\frac{1}{\langle\alpha\times\epsilon\rangle}$	% Dev
0 - 90%	7 - 10	12102	117.495	112.958	**-3.4%**	80.5755	**-30.6%**
0 - 90%	10 - 15	1788	43.3865	43.1728	**-0.493%**	27.6409	**-36.3%**
0 - 90%	15 - 20	4577	13.7835	13.639	**-1.05%**	12.6465	**-8.25%**
0 - 90%	20 - 50	35980	6.38043	6.36748	**-0.203%**	5.9658	**-6.50%**
0 - 90%	10 - 50	9522	12.2694	12.2668	**-0.0212%**	6.5642	**-28.3%**
0 - 30%	10 - 50	33143	7.70383	7.70087	**-0.0384%**	8.79954	**-24.2%**
30 - 90%	10 - 50	42453	8.72094	8.71793	**-0.0345%**	5.8419	**-24.7%**

Table 5.15: The results of B^+ efficiency factors $\langle\frac{1}{\alpha\times\epsilon}\rangle$ and $\frac{1}{\langle\alpha\times\epsilon\rangle}$ and their percent deviation are shown above.

Centrality	B^+ p_T (GeV/c)	NRECO	GEN/RECO	$\langle\frac{1}{\alpha\times\epsilon}\rangle$	% Dev	$\frac{1}{\langle\alpha\times\epsilon\rangle}$	% Dev
0 - 90%	7 - 10	12102	117.495	112.958	**-3.4%**	80.5755	**-30.6%**
0 - 90%	10 - 15	1788	43.3865	43.1728	**-0.493%**	27.6409	**-36.3%**
0 - 90%	15 - 20	4577	13.7835	13.639	**-1.05%**	12.6465	**-8.25%**
0 - 90%	20 - 50	35980	6.38043	6.36748	**-0.203%**	5.9658	**-6.50%**
0 - 90%	10 - 50	9522	12.2694	12.2668	**-0.0212%**	6.5642	**-28.3%**
0 - 30%	10 - 50	33143	7.70383	7.70087	**-0.0384%**	8.79954	**-24.2%**
30 - 90%	10 - 50	42453	8.72094	8.71793	**-0.0345%**	5.8419	**-24.7%**

of $\frac{1}{\langle\alpha\times\epsilon\rangle}$ in our B-meson data analysis.

Low Statistics Limit

We compute the efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ on each of the 2000 data-like MC samples. Then, we plot their percent deviation for all p_T and centrality bins and compute their mean values. Our results are shown on Figure **??**

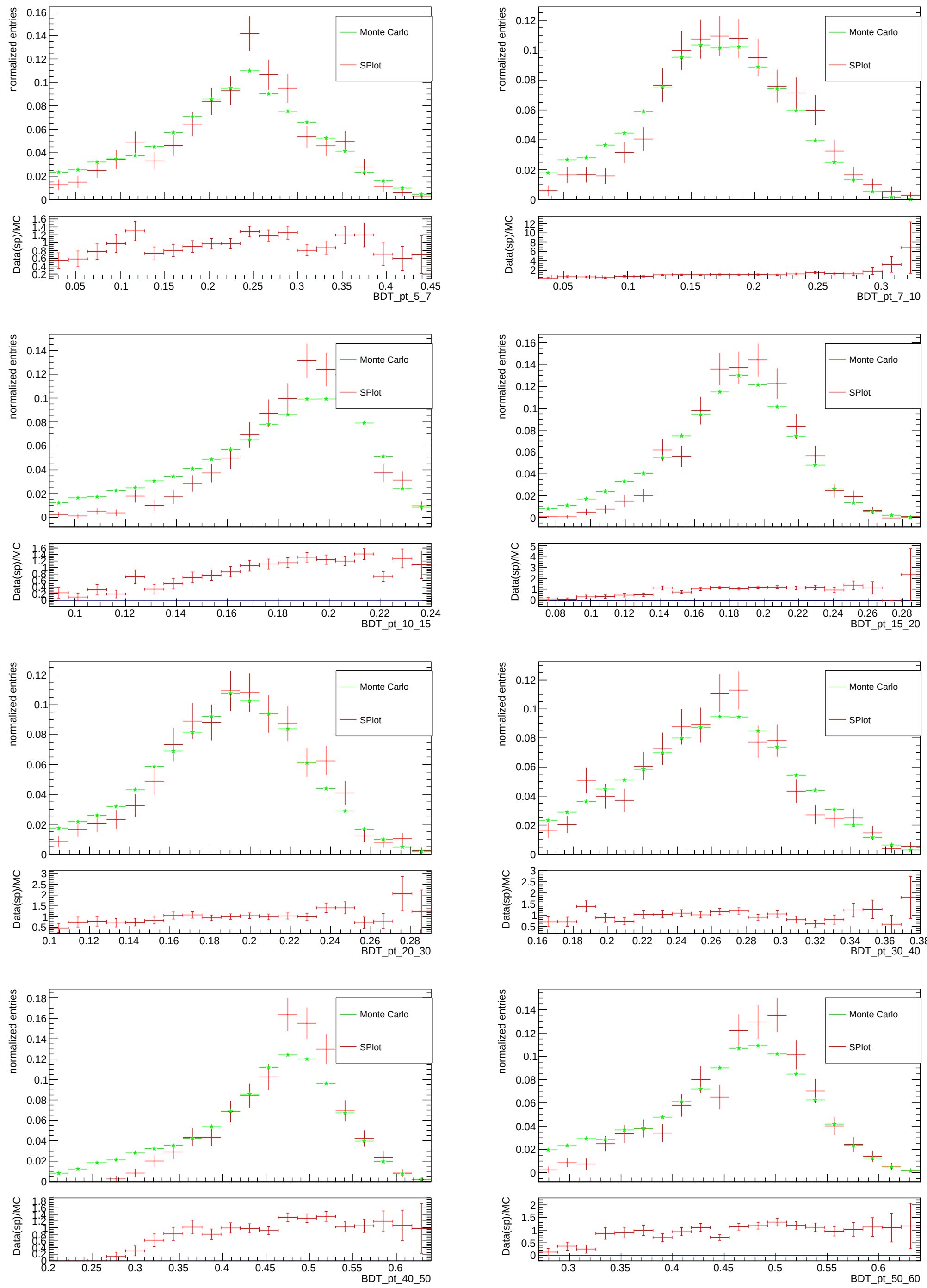


Figure 5-43: Comparison of B^+ BDT distribution in data (red) and MC (green) using Splot method.

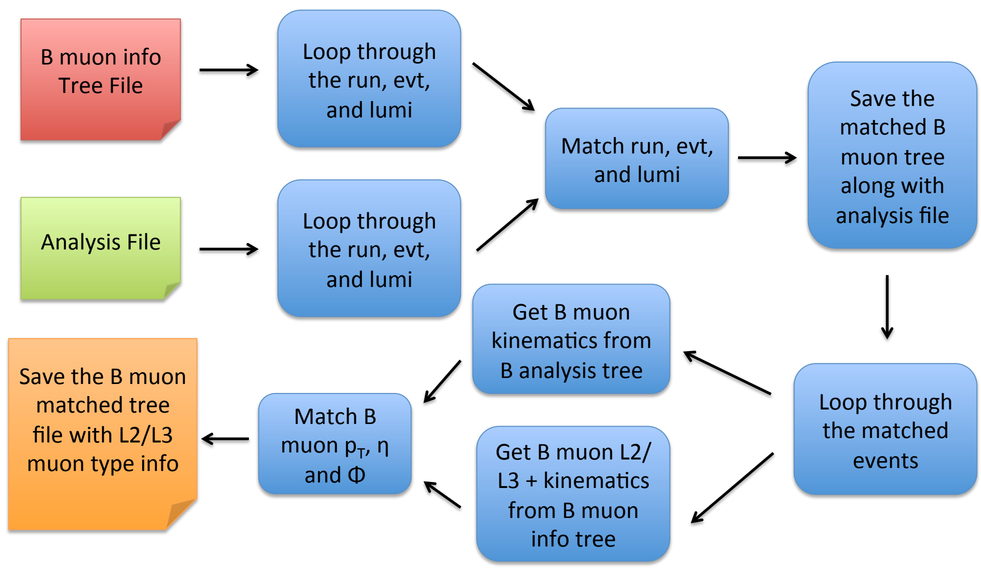


Figure 5-44: The workflow to obtain L2 and L3 muons in order to apply tag-&-probe method.

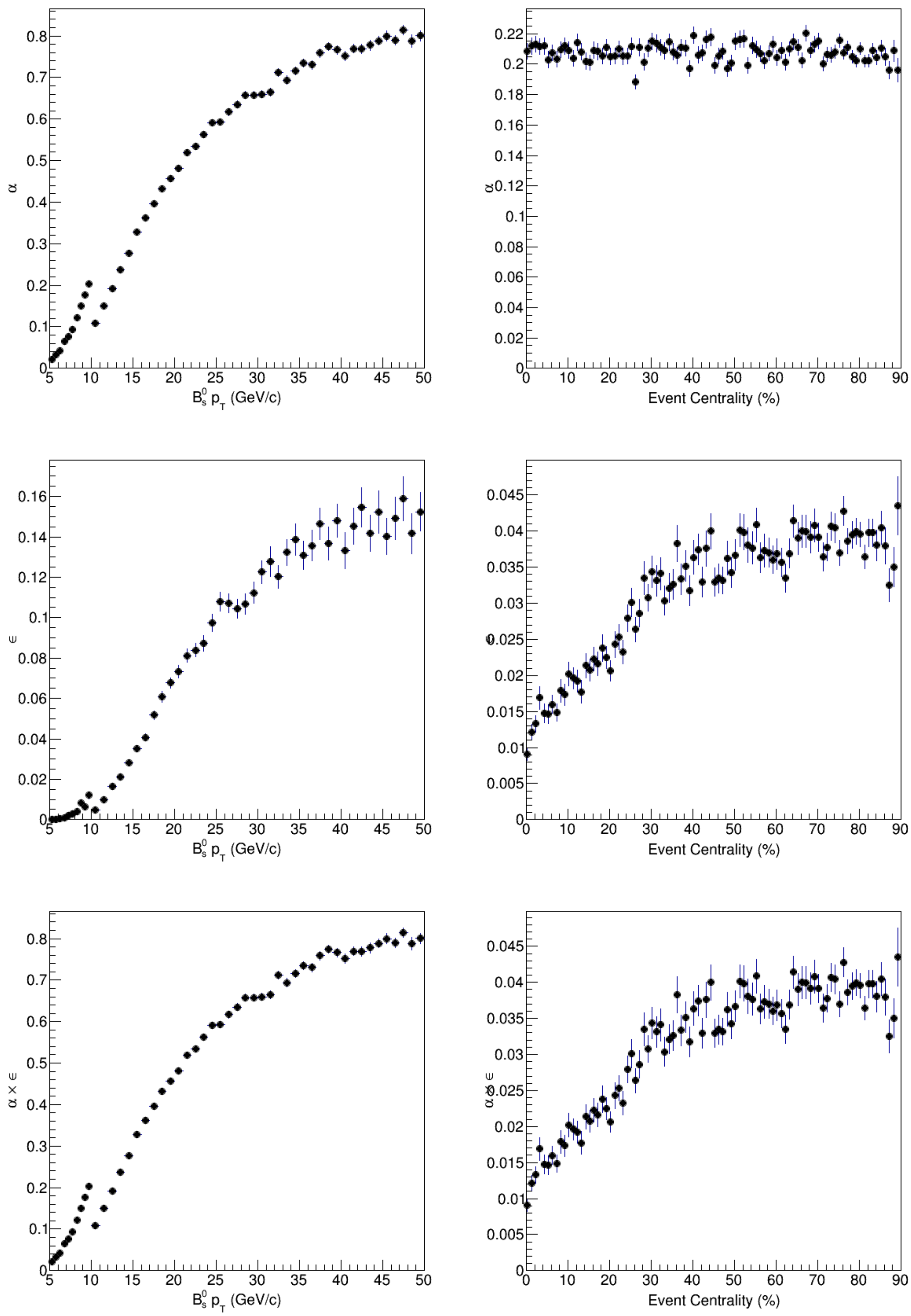


Figure 5-45: The B_s^0 acceptance (top), selection efficiency (middle), and efficiency (bottom) as a function of p_T (left) and event centrality (right) are shown respectfully above. We should note that there is no significant centrality dependence on the B_s^0 acceptance, which makes sense.

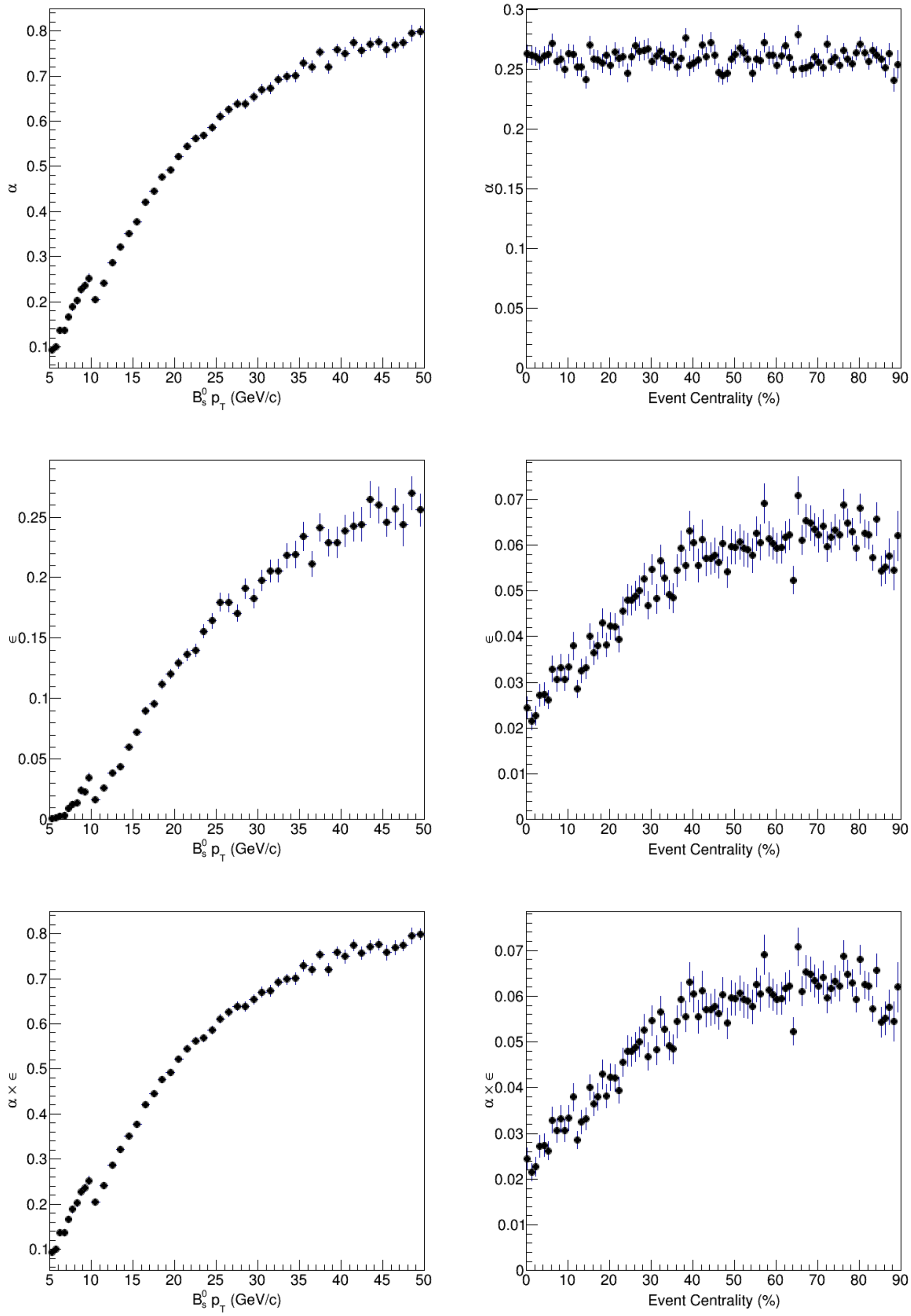


Figure 5-46: The B^+ acceptance (top), selection efficiency (middle), and efficiency (bottom) as a function of p_T (left) and event centrality (right) are shown respectfully above. We should note that there is no significant centrality dependence on the B^+ acceptance, which makes sense.

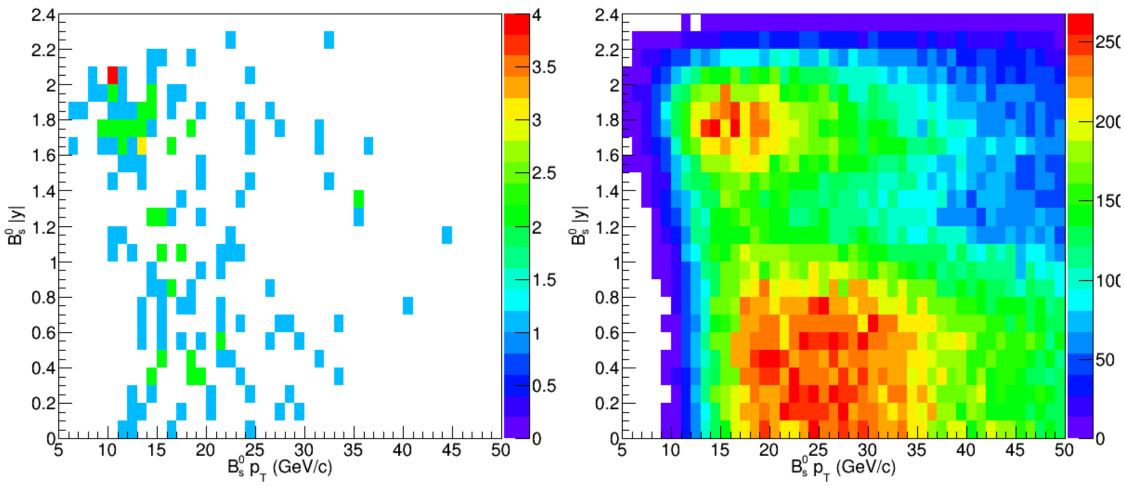


Figure 5-47: The finely binned 2D candidates distribution vs B p_T and B $|y|$ for data and MC at centrality 0 - 90% are shown respectfully above.

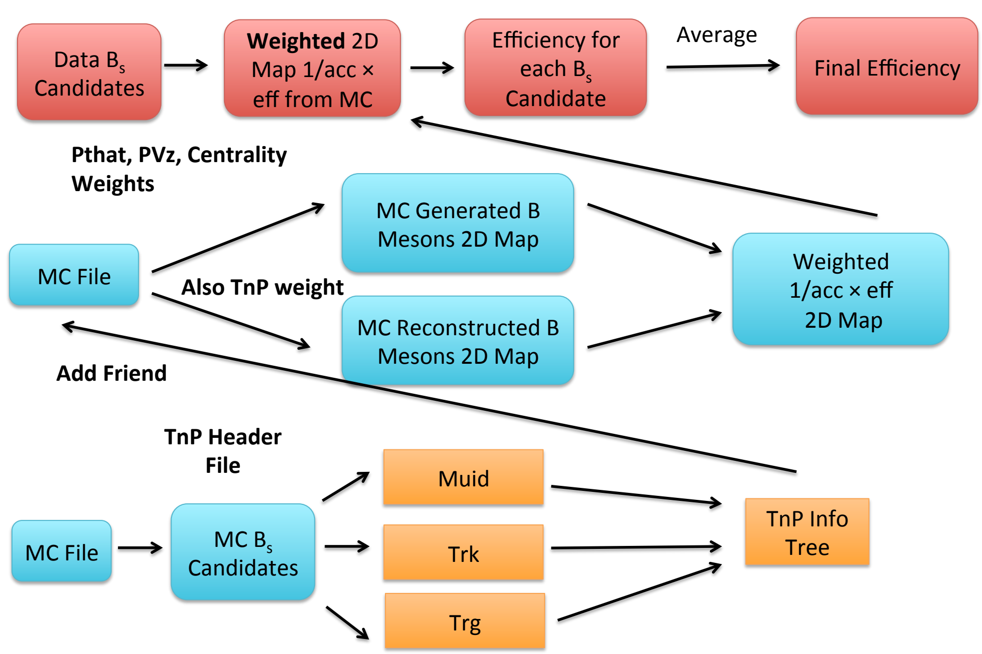


Figure 5-48: The workflow for the efficiency correction including the data driven tag-and-probe approach in B-meson analysis is shown above.

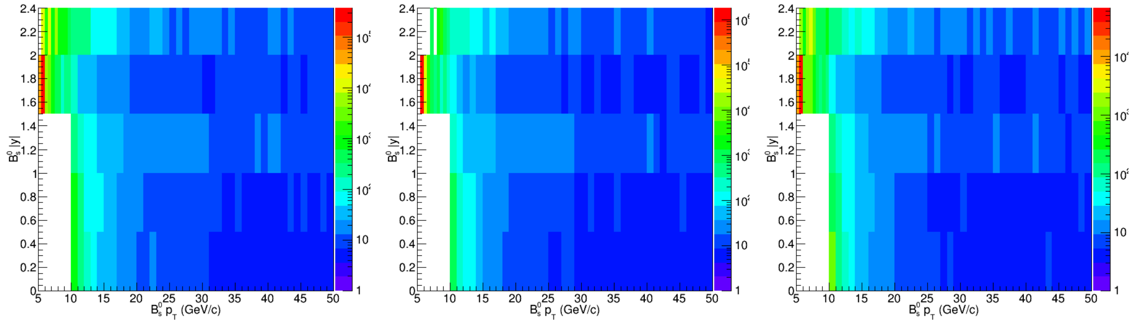


Figure 5-49: The finely binned 2D $\frac{1}{\alpha \times \epsilon}$ vs B_s^0 p_T and B $|y|$ for 0 - 90% (top), 0 - 30% (middle), and 30% - 90% (bottom) centrality are shown respectfully above.

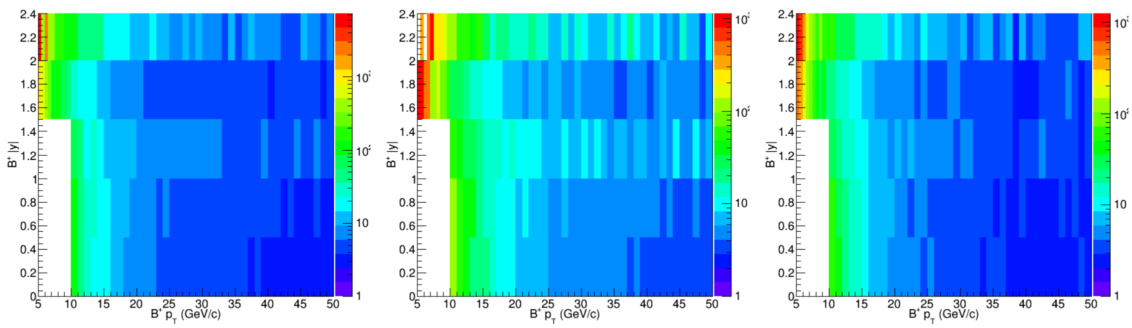


Figure 5-50: The finely binned 2D $\frac{1}{\alpha \times \epsilon}$ vs B^+ p_T and B $|y|$ for 0 - 90% (top), 0 - 30% (middle), and 30% - 90% (bottom) centrality are shown respectfully above.

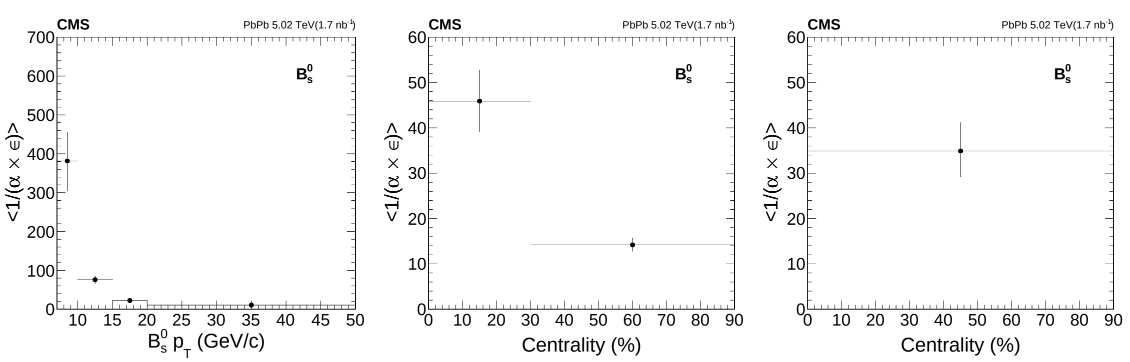


Figure 5-51: The B_s^0 efficiency correction factor $\langle\frac{1}{\alpha \times \epsilon}\rangle$ vs p_T (left) and $0 - 30$ % and $30 - 90\%$ centrality (middle) and the inclusive $0 - 90\%$ centrality (right) are shown above.

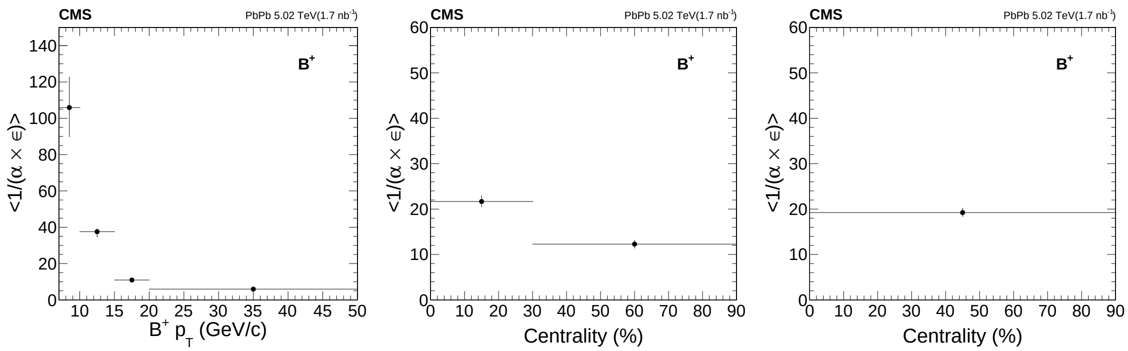


Figure 5-52: The B^+ efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ vs p_T (left) and $0-30$ % and $30-90\%$ centrality (middle) and the inclusive $0-90\%$ centrality (right) are shown above.

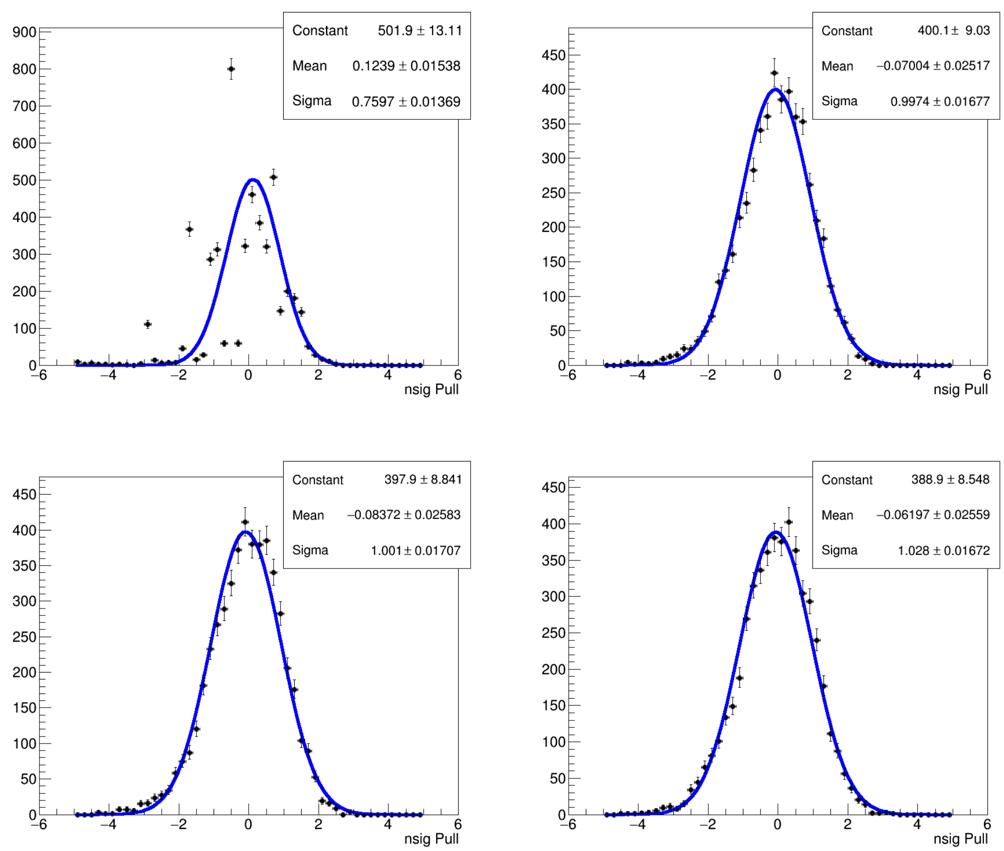


Figure 5-53: The B_s^0 pull distribution and the Gaussian fits for 0 - 90% at p_T 7 - 10, 10 - 15, 15 - 20, 20 - 50 GeV/c are shown respectfully above.

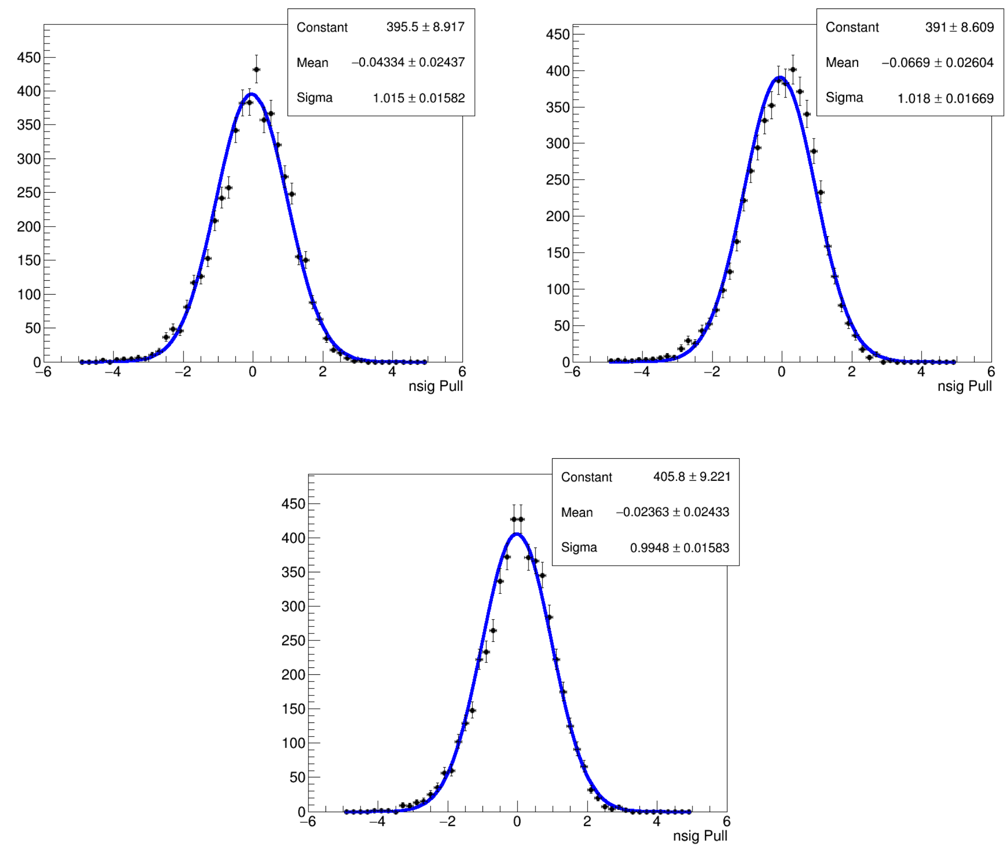


Figure 5-54: The B_s^0 pull distribution and the Gaussian fits for 0 - 30%, 30 - 90%, and 0 - 90% event centrality are shown respectfully above.

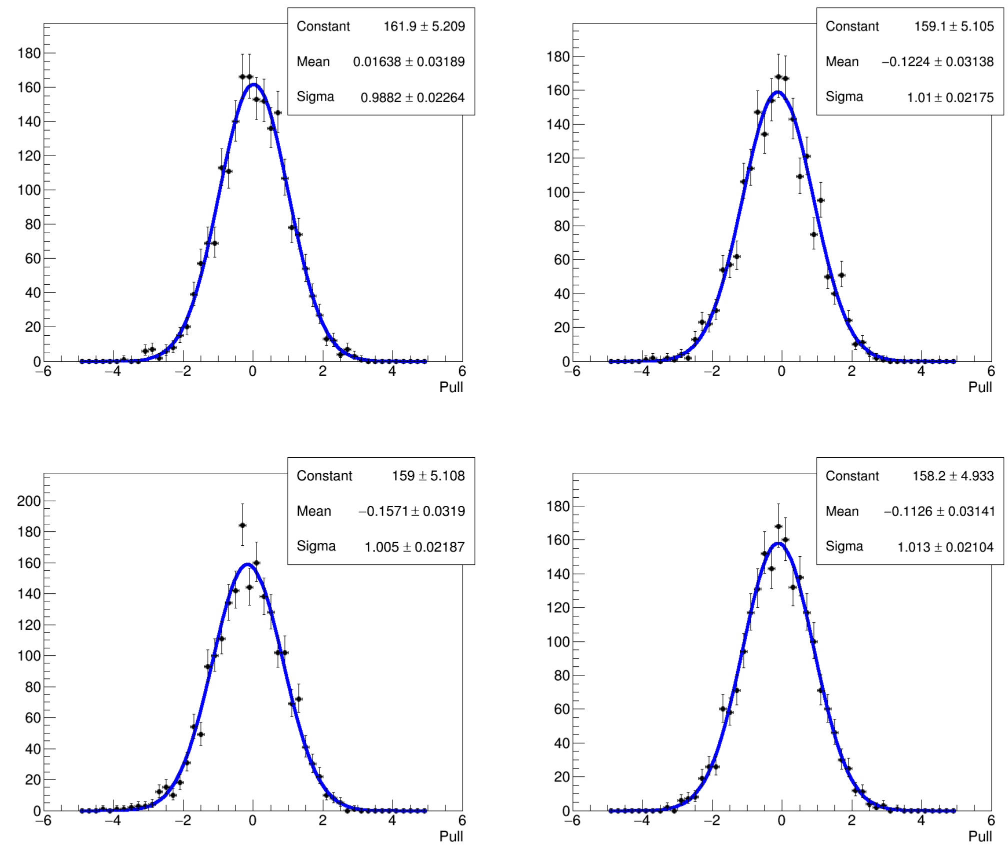


Figure 5-55: The B^{+} pull distribution and the Gaussian fits for 0 - 90% at p_T 7 - 10, 10 - 15, 15 - 20, 20 - 50 GeV/c are shown respectfully above.

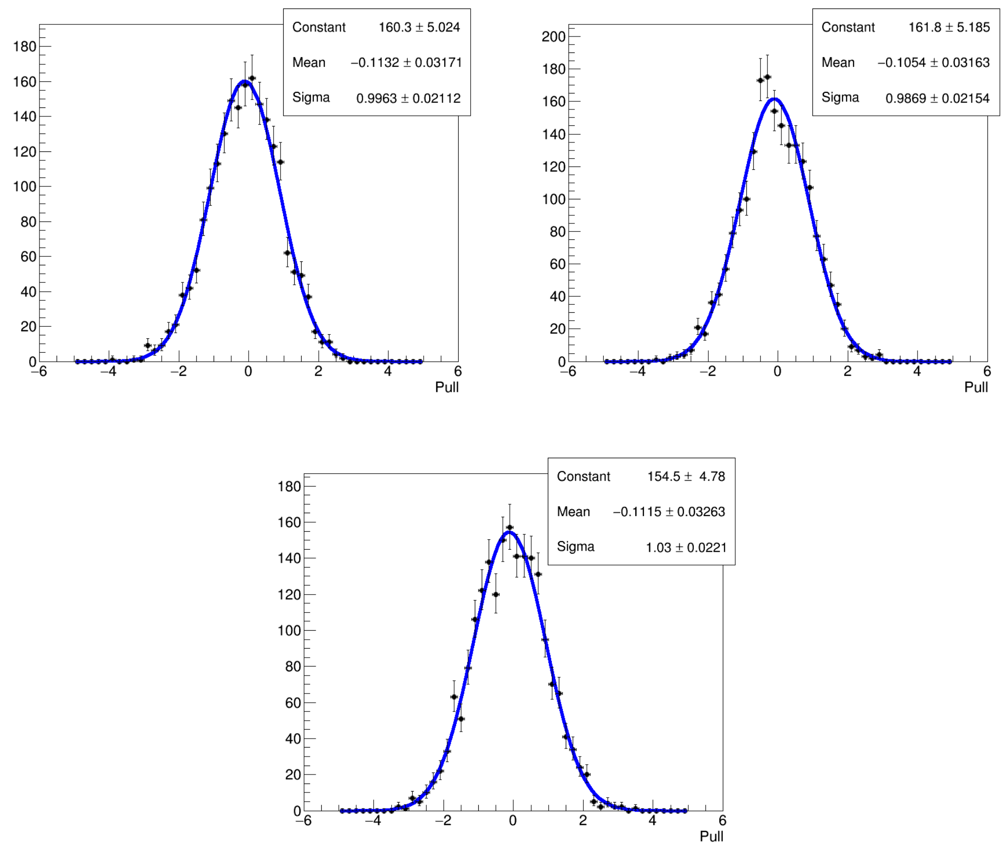


Figure 5-56: The B^+ pull distribution and the Gaussian fits for 0 - 30%, 30 - 90%, and 0 - 90% event centrality are shown respectfully above.

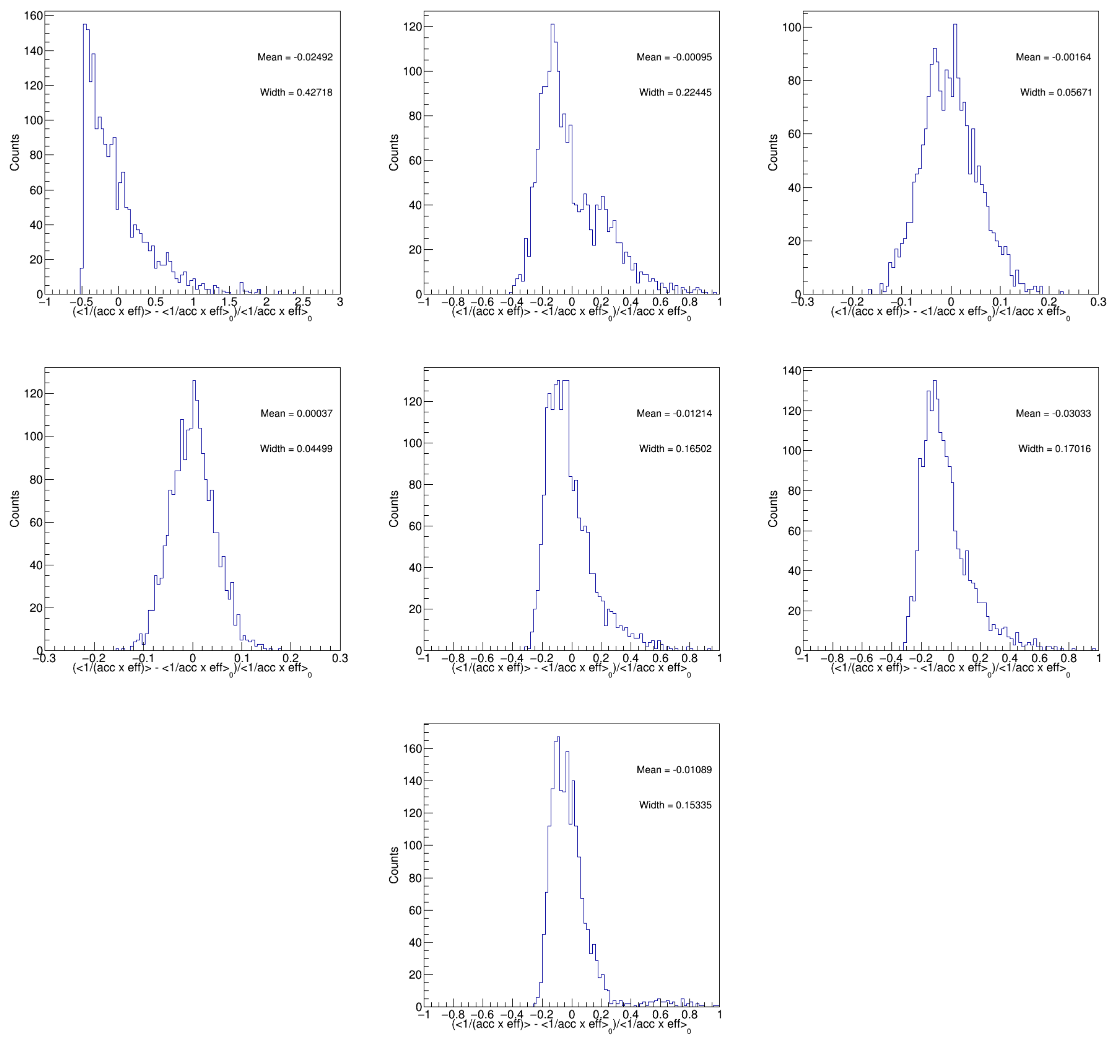


Figure 5-57: The percent deviation distributions of B_s^0 $\langle\frac{1}{\alpha\times\epsilon}\rangle$ to RECO/GEN for the data-like randomly sampled MC samples for 0 - 90% at 7 - 10, 10 - 15, 15 - 20, and 20 - 50 GeV/c as well as 0 - 90%, 0 - 30%, and 30 - 90% event centrality are shown respectfully above.

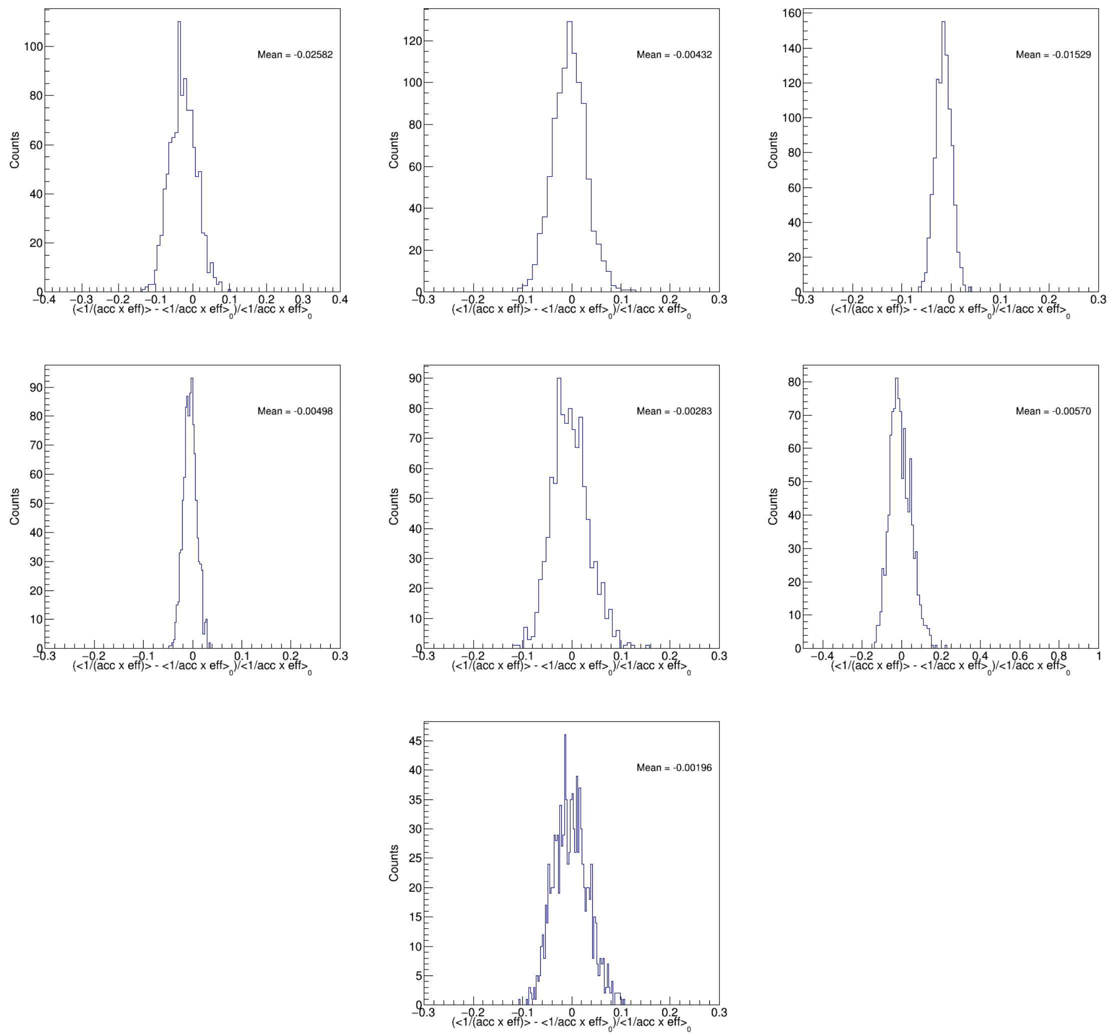


Figure 5-58: The percent deviation distributions of B^+ $\langle\frac{1}{\alpha\times\epsilon}\rangle$ to RECO/GEN for the data-like randomly sampled MC samples for 0 - 90% at 7 - 10, 10 - 15, 15 - 20, and 20 - 50 GeV/c as well as 0 - 90%, 0 - 30%, and 30 - 90% event centrality are shown respectfully above.

The percent deviation of B_s^0 and B^+ efficiency correct factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ are summarize in thed Table 5.16 and Table 5.17 respectfully

In conclusion, we can see that, even in the limit of low statistics, which is similar to the statistics in our data analysis, the $\langle\frac{1}{\alpha\times\epsilon}\rangle$ correction method still gives us satisfying closure with bias within 3%. All these show that the efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ will not introduce significant bias in both B_s^0 and B^+ analyses.

Table 5.16: The percent deviation of the efficiency factors from the expected value in the statistics similar to the data analysis are shown above.

Centrality	B_s^0 p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**-2.49%**
0 - 90%	10 - 15	**-0.10%**
0 - 90%	15 - 20	**-0.16%**
0 - 90%	20 - 50	**+0.03%**
0 - 90%	10 - 50	**-1.09%**
0 - 30%	10 - 50	**-1.21%**
30 - 90%	10 - 50	**-3.03%**

Table 5.17: The percent deviation of the efficiency factors from the expected value in the statistics similar to the data analysis are shown above.

Centrality	B^+ p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**-2.49%**
0 - 90%	10 - 15	**-0.10%**
0 - 90%	15 - 20	**-0.16%**
0 - 90%	20 - 50	**+0.03%**
0 - 90%	10 - 50	**-1.09%**
0 - 30%	10 - 50	**-1.21%**
30 - 90%	10 - 50	**-3.03%**

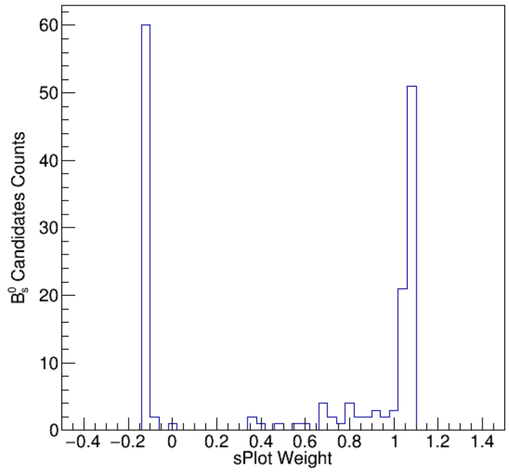
5.12.4 *Splot* Closure on Efficiency

Finally, we will validate the efficiency correction factor $\langle \frac{1}{\alpha \times \epsilon} \rangle$ using all signal region data candidates. We compare our nominal $\langle \frac{1}{\alpha \times \epsilon} \rangle$ with the results with an *Splot* weights w^S, which is obtain from the Splot analysis techniques shown in Section 4.9. Figure **??** shows the *Splot* weight distribution for B_s^0 and B^+ candidates

We can see that most of the candidate are either background-like (near 0) and signal-like (near 1). But some candidate are in between. Hence, the *Splot* weighed efficiency correction factor $\langle \frac{1}{\alpha \times \epsilon} \rangle'$ is given by

$$\langle \frac{1}{\alpha \times \epsilon} \rangle' = \frac{\sum_{i=1}^{N} \frac{w_i^S}{\alpha_i \times \epsilon_i}}{\sum_{i=1}^{N} w_i^S} \tag{5.23}$$

We can compute the the *Splot* weighed efficiency correction factor $\langle \frac{1}{\alpha \times \epsilon} \rangle'$ and look



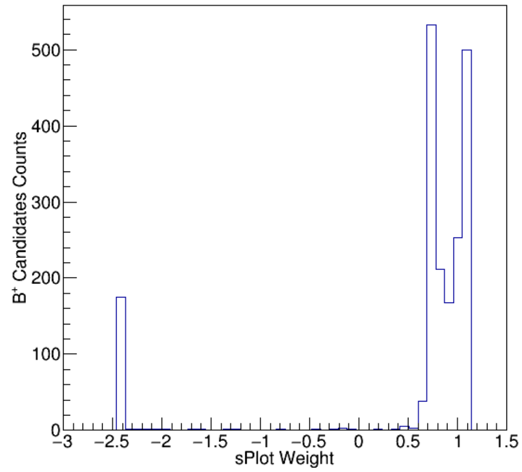


Figure 5-59: The sPlot weight distributions for B_s^0 (left) and B^+ (right) candidates are shown above.

at its deviation from the nominal efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$.

$$\%Dev = \frac{\langle\frac{1}{\alpha\times\epsilon}\rangle' - \langle\frac{1}{\alpha\times\epsilon}\rangle}{\langle\frac{1}{\alpha\times\epsilon}\rangle} \tag{5.24}$$

The percent deviation of B_s^0 and B^+ *Splot* weighed efficiency correct factor $\langle\frac{1}{\alpha\times\epsilon}\rangle'$ to the nominal efficiency correct factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ are summarize in Table 5.18 and Table 5.19 respectfully

Table 5.18: The B_s^0 *Splot* weighed efficiency correct factor $\langle\frac{1}{\alpha\times\epsilon}\rangle'$, nominal unweighed $\langle\frac{1}{\alpha\times\epsilon}\rangle$, and their percent deviation for each p_T and centrality bin are summarized below.

Centrality	B_s^0 p_T (GeV/c)	$\langle\frac{1}{\alpha\times\epsilon}\rangle'$ (Weighed)	$\langle\frac{1}{\alpha\times\epsilon}\rangle$ (Nominal)	Percent Deviation (% Dev)
0 - 90%	7 - 10	386.5	381.5	1.31%
0 - 90%	10 - 15	77.34	75.92	1.88%
0 - 90%	15 - 20	22.24	22.35	0.49%
0 - 90%	20 - 50	10.71	10.63	0.75%
0 - 30%	10 - 50	46.20	45.90	0.63%
30 - 90%	10 - 50	14.27	14.19	0.56%
0 - 90%	10 - 50	35.15	34.90	0.72%

Table 5.19: The B^+ *Splot* weighed efficiency correct factor $\langle\frac{1}{\alpha\times\epsilon}\rangle'$, nominal unweighed $\langle\frac{1}{\alpha\times\epsilon}\rangle$, and their percent deviation for each p_T and centrality bin are summarized below.

Centrality	B^+ p_T (GeV/c)	$\langle\frac{1}{\alpha\times\epsilon}\rangle'$ (Weighed)	$\langle\frac{1}{\alpha\times\epsilon}\rangle$ (Nominal)	Percent Deviation (% Dev)
0 - 90%	7 - 10	104.1	105.9	1.72%
0 - 90%	10 - 15	37.51	37.55	0.11%
0 - 90%	15 - 20	10.95	10.94	0.09%
0 - 90%	20 - 50	5.932	5.932	0.01%
0 - 30%	10 - 50	21.65	21.67	0.09%
30 - 90%	10 - 50	12.15	12.28	1.06%
0 - 90%	10 - 50	19.17	19.23	0.31%

Since the percent deviation are within 2%, there is no need to implement *Splot* weight to compute the efficiency correct factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$. Up to here, our analysis procedure have been fully validated.

5.13 Statistical Uncertainties Determination

5.13.1 Data Bootstrapping

After showing that there is no explicit bias throughout the entire analysis, next step is to evaluate statistical uncertainties of the analysis. To estimate the statistical uncertainties of a measurement, we can simply follow the toy approach. We can repeat the counting experiment with identical conditions and randomly sample the statistics according to the Poisson distribution defined as follows:

$$P(\mu, x) = \frac{\mu^N e^{-\mu}}{N!} \tag{5.25}$$

In order to estimate the statistical uncertainties, we develop the following procedures named as "Data Bootstrapping", to quantify the statistical uncertainties of the final cross section measurement, which is defined in Section 4.11:

- Randomly resample the data after passing the selections for each p_T and centrality bin. Each resampled data has the number of B mesons according to Poisson distribution with the same mean as the signal raw yield in data analysis.

- Construct 1000 randomly resampled datasets. Here, we allow repeated events in the resampled dataset.

- Carry out the same workflow on each of the resampled datasets as the data analysis and compute the corrected yield, which is defined as the product of the signal raw yield N_S and efficiency correction factor $\langle \frac{1}{\alpha \times \epsilon} \rangle$

- Plot the B-meson corrected yield distribution and compute the mean and RMS for each p_T and centrality bin

- Quote the RMS/Mean values of the distributions as the percent statistical uncertainties

Figure 5-60 and 5-61 show the corrected yield distribution of the 1000 B_s^0 and B^+ resampled datasets respectfully

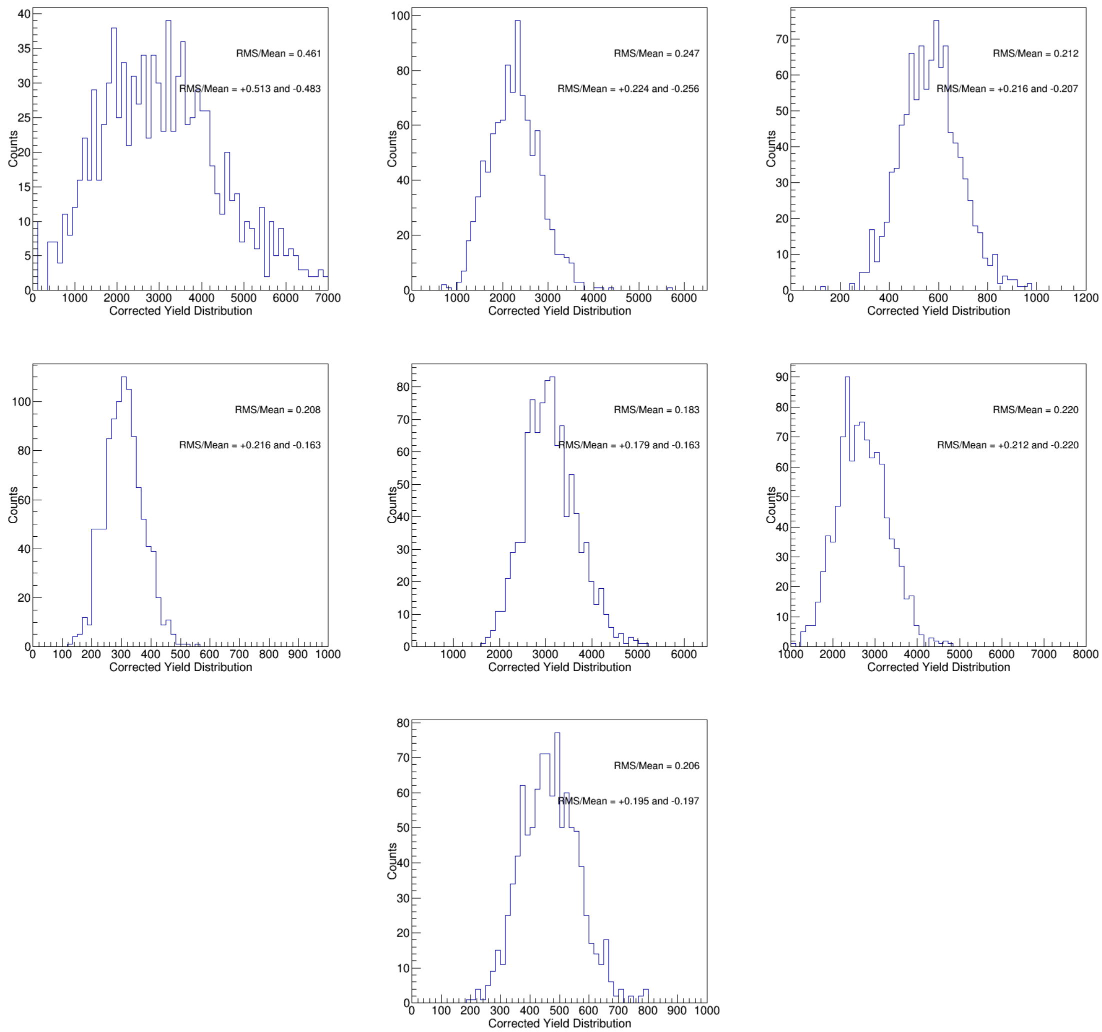


Figure 5-60: The B_s^0 corrected yield distributions of the 1000 data-like randomly resampled datasets for centrality in 0 - 90% in the p_T range of 7 - 10, 10 - 15, 15 - 20, and 20 - 50 GeV/c as well as 0 - 90%, 0 - 30%, and 30 - 90% in the p_T range of 10 - 50 GeV/c are shown above.

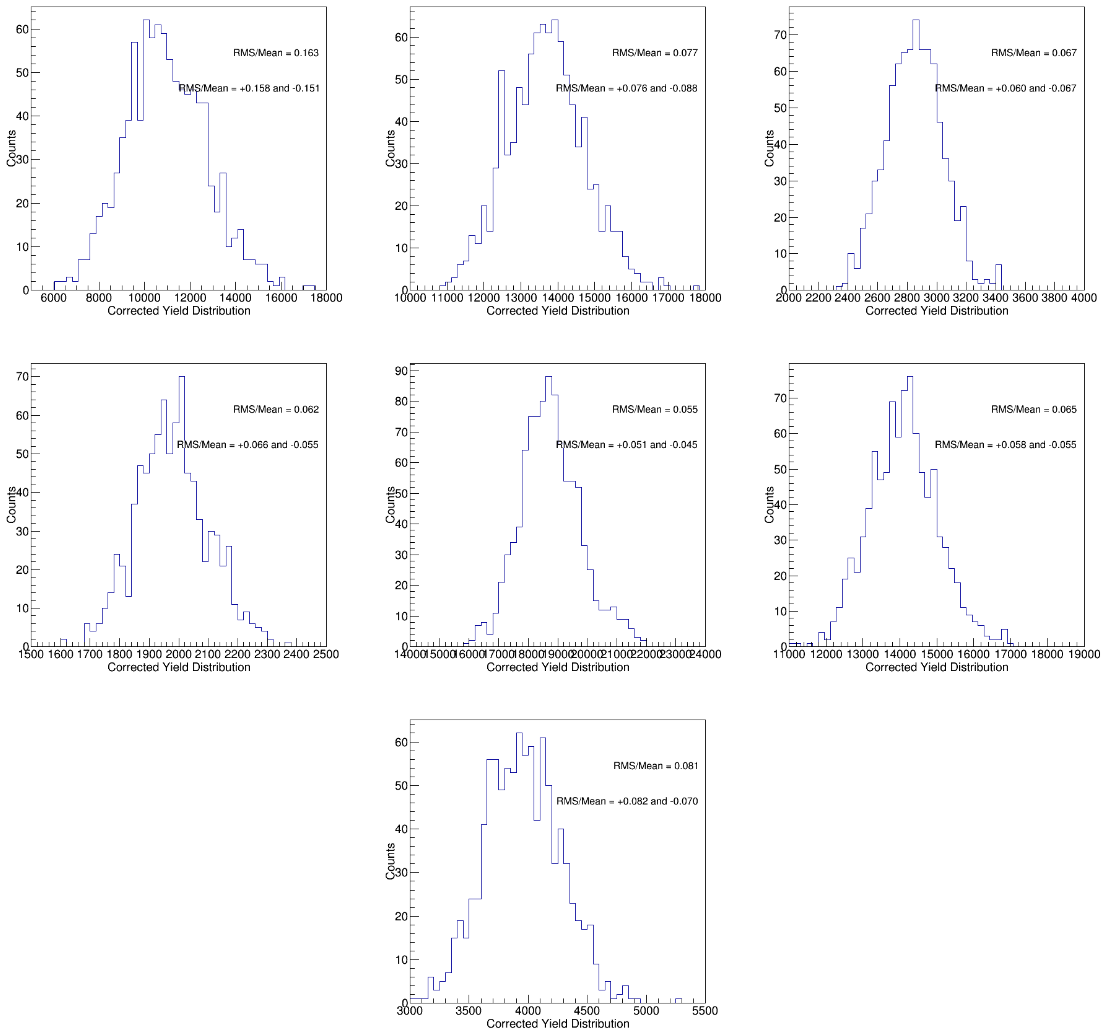


Figure 5-61: The B^+ corrected yield distributions of the 1000 data-like randomly resampled datasets for centrality in 0 - 90% in the p_T range of 7 - 10, 10 - 15, 15 - 20, and 20 - 50 GeV/c as well as 0 - 90%, 0 - 30%, and 30 - 90% in the p_T range of 10 - 50 GeV/c are shown above.

5.13.2 Statistical Uncertainties Interpretation

A refinement is needed as we see from Figure 5-60 and Figure 5-61 that the corrected yield distributions are indeed asymmetric. Hence, asymmetric statistical uncertainties will be introduced in our B-meson cross section measurements. To quantify the asymmetric statistical uncertainties, we do the following

- Find the location of the bin corresponding to the mean of corrected yield distribution

- Integrate up/down (+/−) side by increasing/decreasing the bin number from the mean bin

- Take integral count from the mean bin ratio to the total up/down side counts from the mean so that they reach one σ, which is $34.14\% \times 2 = 68.28\%$. Mathematically, the method is shown below:

$$\frac{\int_{\langle Y \rangle}^{Y^+} f(x)dx}{\int_{\langle Y \rangle}^{\infty} f(x)dx} = 68.28\% \tag{5.26}$$

- Read out the corresponding up/down corrected yield values, for instance Y^+ as shown on the equation above

- Compute the percent deviations (% Dev) from the mean values for both up/down corrected yields and quote the deviations as up/down statistical uncertainties

$$\%Dev^{\pm} = \frac{Y^{\pm} - \langle Y \rangle}{\langle Y \rangle} \tag{5.27}$$

Table 5.20 and Table 5.21 summarize the estimations of statistical uncertainties for B_s^0 and B^+ cross sections in each p_T and centrality bin

Table 5.20: B_s^0 RMS/Mean and their asymmetric up and down statistical uncertainties of the corrected yield distribution are summarized below.

Centrality	B_s^0 p_T (GeV/c)	RMS/Mean	Stat. Up (+)	Stat Down (-)
0 - 90%	7 - 10	46.1%	**51.3%**	**48.3%**
0 - 90%	10 - 15	24.7%	**22.4%**	**25.6%**
0 - 90%	15 - 20	21.2%	**21.6%**	**20.7%**
0 - 90%	20 - 50	20.8%	**21.6%**	**16.3%**
0 - 30%	10 - 50	22.0%	**21.2%**	**22.0%**
30 - 90%	10 - 50	20.6%	**19.5%**	**19.7%**
0 - 90%	10 - 50	18.3%	**17.9%**	**16.3%**

Table 5.21: B^+ RMS/Mean and their asymmetric up and down statistical uncertainties of the corrected yield distribution are summarized below.

Centrality	B^+ p_T (GeV/c)	RMS/Mean	Stat. Up (+)	Stat Down (-)
0 - 90%	7 - 10	16.4%	**15.8%**	**15.1%**
0 - 90%	10 - 15	7.70%	**7.60%**	**8.77%**
0 - 90%	15 - 20	6.72%	**6.00%**	**6.66%**
0 - 90%	20 - 50	6.19%	**6.64%**	**5.49%**
0 - 30%	10 - 50	6.52%	**5.81%**	**5.54%**
30 - 90%	10 - 50	8.10%	**8.22%**	**6.97%**
0 - 90%	10 - 50	5.49%	**5.06%**	**4.54%**

5.14 Systematic Uncertainties Estimation

The final step is to estimate the systematic uncertainties of the B-meson measurements. We know that systematic uncertainties always exist in experiments due to the imperfection of conditions, unknown variation in the experiments, and limitation of the analysis techniques, which cannot be cure even with infinite statistics. They mainly come from 3 parts: the common constant scale factors for both B_s^0 and B^+ mesons, the signal raw yield extraction, and the efficiency correction. Based on the formula to obtain the cross section (Equation 4.18 and 4.19), we identify the main sources of systematic uncertainties in term of percentage and the method to estimate each of them in the follow subsections.

5.14.1 Global Observables

The uncertainties of global observables T_{AA} and N_{MB} are summarized at Table 5.3 and Table 5.2. We simply quote the uncertainties from them in our cross section measurements.

5.14.2 Branching Ratios

According to PDG 2018 [215], the relevant decay branching ratio and their uncertainties are listed as follows

- $BR(B_s^0 \to J/\psi\phi) = (1.08 \pm 0.08) \times 10^{-3}$
- $BR(B^+ \to J/\psi K^+) = (1.010 \pm 0.029) \times 10^{-3}$
- $BR(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033)\%$
- $BR(\phi \to K^+K^-) = (49.2 \pm 0.5)\%$

Hence, we can compute total branching ratio of the B mesons decay chain by multiply the partial decays. The uncertainties on the BR will propagate in an uncorrelated manner.

- $BR(B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-) = (3.17 \pm 0.24) \times 10^{-5}$
- $BR(B^+ \to J/\psi K^+ \to \mu^+\mu^-K^+) = (6.03 \pm 0.18) \times 10^{-5}$

Hence, we can compute percent uncertainties of the B_s^0 and B^+ branching ratios. The systematic uncertainties due to decay branching ratio of B_s^0 is **7.50%** and B^+ is **2.92%**.

5.14.3 Tracking Efficiency

The difference in the track reconstruction efficiency in data and simulation is estimated by comparing 3-prong D^* decay $D^* \to K\pi\pi$ and 5-prong D^* decay $D^* \to K\pi\pi\pi\pi$, **??**. According to CMS tracking group studies, this results in **5%** for each track. Hence, the systematic uncertainties due to tracking efficiency is **5%** for B^+ since it has only one kaon track and **10%** for B_s^0 since it has two kaon tracks. These apply to all p_T and centrality bins.

5.14.4 Muon Efficiency

The systematic uncertainties due to muon efficiency can be quantified by taking the up and down cases of the tag-&-probe scale factors determined by the CMS dilepton group. Figure 5-62 shows the workflow to carry out the systematic uncertainties studies using tag-&-probe method in this analysis:

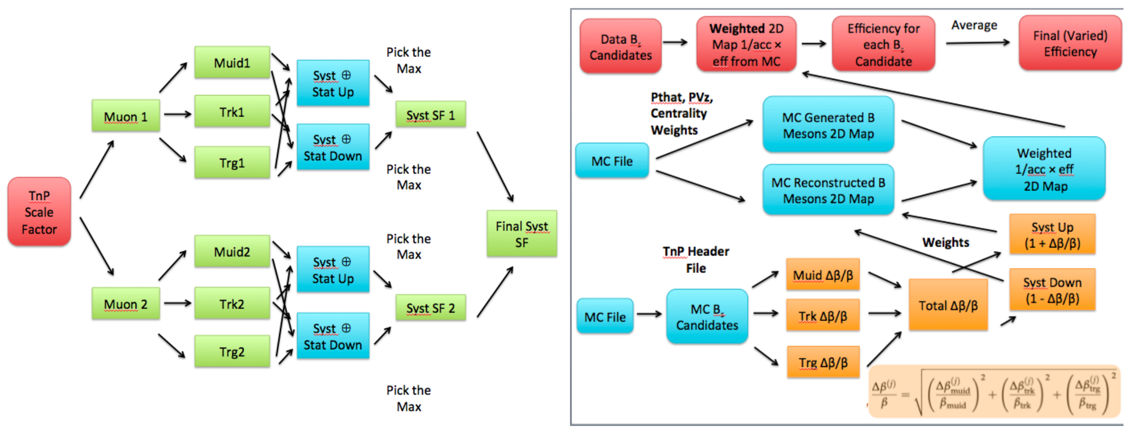


Figure 5-62: The schematic block diagrams demonstrating the calculations of the uncertainties on tag-&-probe scale factors and the asymmetric systematic uncertainties due to muon efficiency on the efficiency correction factor (right) are shown above.

Table 5.22 and Table 5.23 summarize the systematic uncertainties due to muon efficiency on B_s^0 and B^+ respectfully

Table 5.22: The B_s^0 systematic uncertainty due to muon efficiency for each p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	Uncertainty Up (+)	Uncertainty Down (-)
0 - 90%	7 - 10	**8.88%**	**7.49%**
0 - 90%	10 - 15	**5.97%**	**5.24%**
0 - 90%	15 - 20	**3.73%**	**3.46%**
0 - 90%	20 - 50	**3.91%**	**3.60%**
0 - 30%	10 - 50	**5.52%**	**4.85%**
30 - 90%	10 - 50	**4.63%**	**4.19%**
0 - 90%	10 - 50	**5.29%**	**4.70%**

Table 5.23: The B^+ systematic uncertainty due to muon efficiency for each p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	Uncertainty Up (+)	Uncertainty Down (-)
0 - 90%	7 - 10	**7.21%**	**6.28%**
0 - 90%	10 - 15	**4.29%**	**3.92%**
0 - 90%	15 - 20	**3.83%**	**3.53%**
0 - 90%	20 - 50	**3.87%**	**3.56%**
0 - 30%	10 - 50	**4.18%**	**3.83%**
30 - 90%	10 - 50	**4.14%**	**3.80%**
0 - 90%	10 - 50	**4.16%**	**3.81%**

5.14.5 Selection Efficiency

Next, we will estimate uncertainties due to the B-meson selection efficiency. The efficiency correction heavily relies on the MC performance. The poor descriptions of detector performance in the MC simulation, the limited MC statistics, and poor underlying physics processes modeling of the MC B-meson spectra will contribute to the uncertainties in efficiency correction. In fact, this is the major systematic uncertainty throughout the analysis.

Data-MC Discrepancy

We have previous compared the data and MC distributions BDT scores of B_s^0 and B^+ mesons. To quantify the systematic uncertainties due to Data-MC discrepancy, we can simply apply the *Splot* weights w obtained using *Splot* techniques in the BDT distributions to obtain a weighed 2D efficiency correction map $\frac{w}{\alpha\times\epsilon}$. Then, we compute the weighed efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle_w$.Finally, we take the percent deviation (% Dev) of the BDT *Splot* weighed $\langle\frac{1}{\alpha\times\epsilon}\rangle_w$ to the nominal efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$ as the systematic uncertainties.

$$\%Dev = \frac{\langle\frac{1}{\alpha\times\epsilon}\rangle_w - \langle\frac{1}{\alpha\times\epsilon}\rangle}{\langle\frac{1}{\alpha\times\epsilon}\rangle} \tag{5.28}$$

This method works for B^+ since it has sufficient statistics. However, for B_s^0, due to its limited statistics, we decide to apply the *Splot* weights obtained from kinematic variables of the kaon track in B^+ to the efficiency correction factor and quote the largest one as the systematic uncertainties. The list of kinematic variables are shown below:

- The transverse distance of closest approach the kaon track to the primary vertex
- The error on the transverse distance of closest approach the kaon track to the primary vertex
- The longitudinal distance of closest approach the kaon track to the primary vertex

- The error on the longitudinal distance of closest approach the kaon track to the primary vertex

- The kaon track pseudorapidity

- The kaon track rapidity

- The kaon track transverse momentum

Table 5.24 and Table 5.25 summarize the systematic uncertainties due to MC-Data discrepancy on B_s^0 and B^+ respectfully

Table 5.24: The B_s^0 systematic uncertainty due to Data-MC discrepancy for each p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**34.65%**
0 - 90%	10 - 15	**5.64%**
0 - 90%	15 - 20	**4.66%**
0 - 90%	20 - 50	**10.24%**
0 - 30%	10 - 50	**3.09%**
30 - 90%	10 - 50	**3.66%**
0 - 90%	10 - 50	**3.19%**

Table 5.25: The B^+ systematic uncertainty due to Data-MC discrepancy for each p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**4.17%**
0 - 90%	10 - 15	**15.25%**
0 - 90%	15 - 20	**3.01%**
0 - 90%	20 - 50	**1.65%**
0 - 30%	10 - 50	**13.28%**
30 - 90%	10 - 50	**8.49%**
0 - 90%	10 - 50	**11.51%**

Finite MC Statistics

Another source of uncertainties on the selection efficiency is the statistics of the MC samples. Ideally, this uncertainties should be 0 because we can in principle simulate as many MC events as we want. However, in reality, we only generated about 2.5 million MC events, which is finite. Particularly, when we create the finely binned 2D map of efficiency vs B-meson p_T and $|y|$, due to the limited acceptance and fine binning, very low p_T B meson will have very few candidates.

To quantify the systematic uncertainties due to limited B meson MC events, we the following procedures are carried out

- Generate 10000 2D map of efficiency vs B-meson p_T and $|y|$ according to the smear with Gaussian distribution according to the mean and error of the nominal 2D map

- Carry out the analysis workflow to obtain the efficiency correct factors for $\langle \frac{1}{\alpha \times \epsilon} \rangle$ all 10000 2D maps using the same B-meson experimental data

- Plot the distribution of the efficiency correct factors for $\langle \frac{1}{\alpha \times \epsilon} \rangle$

- Compute the RMS/Mean and quote it as the systematic uncertainties due to finite MC statistics

Figure 5-63 and Figure 5-64 show the $\langle \frac{1}{\alpha \times \epsilon} \rangle$ distributions for B_s^0 and B^+ respectfully

As expected, the distributions have symmetric Gaussian shapes. Their means of the distributions agree with the nominal value of $\langle \frac{1}{\alpha \times \epsilon} \rangle$ in the analysis, which makes sense and thus validates the procedures. Table 5.26 and Table 5.27 summarize the RMS/Mean results of the distribution for B_s^0 and B^+ respectfully

The systematic uncertainties due to finite MC simulation events is particularly large at low p_T due to the limited statistics.

Table 5.26: The B_s^0 systematic uncertainty due to limited MC sample statistics for each p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**26.5%**
0 - 90%	10 - 15	**6.28%**
0 - 90%	15 - 20	**3.08%**
0 - 90%	20 - 50	**3.21%**
0 - 30%	10 - 50	**6.59%**
30 - 90%	10 - 50	**2.27%**
0 - 90%	10 - 50	**4.37%**

Table 5.27: The B^+ systematic uncertainty due to MC sample statistics for each p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**9.22%**
0 - 90%	10 - 15	**3.36%**
0 - 90%	15 - 20	**1.92%**
0 - 90%	20 - 50	**1.35%**
0 - 30%	10 - 50	**3.37%**
30 - 90%	10 - 50	**2.26%**
0 - 90%	10 - 50	**2.49%**

Residue p_T Shape Effect

According to the B-meson measurement in pp collisions from LHCb in Figure 1-33, we do not see significantly rapidity dependence of the B-meson. Assuming similar insignificant rapidity depend also holds in PbPb collisions, the rapidity does not contribute to the weigh function $w(x)$. Hence, only p_T shape plays an important role in creating systematic uncertainties to selection efficiency.

As demonstrated in Section 4.10.4, in the limit of zero bin width, the uncertainties due to the unknown kinematic of B-mesons will be completely eliminated. However, since we have a very fine but still finite p_T binning, there will still be residue effects on systematic uncertainties p_T shape. To quantify the uncertainties, we simply apply the B p_T weights obtained in Section 4.3.4 according to the fits functions obtained in

Figure 5-4 and Figure 5-5.

We apply the B p_T weights to the 2D $\frac{1}{\alpha\times\epsilon}$ map. Then, we compute $\langle\frac{1}{\alpha\times\epsilon}\rangle^{w_{p_T}}$ and take the percent deviation to the nominal efficiency correction factor $\langle\frac{1}{\alpha\times\epsilon}\rangle$. Then, we choose the largest percents deviated from 0 and quote them as the systematic uncertainties due to the B-meson p_T shape. Table 5.35 and Table 5.29 summarize the results

Table 5.28: The B_s^0 systematic uncertainty due to unknown B-meson p_T shape in PbPb collisions for each p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**0.015%**
0 - 90%	10 - 15	**0.050%**
0 - 90%	15 - 20	**0.010%**
0 - 90%	20 - 50	**0.022%**
0 - 30%	10 - 50	**0.067%**
30 - 90%	10 - 50	**0.015%**
0 - 90%	10 - 50	**0.037%**

Table 5.29: The B^+ systematic uncertainty due to unknown B-meson p_T shape in PbPb collisions for each p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	% Dev
0 - 90%	7 - 10	**0.166%**
0 - 90%	10 - 15	**0.198%**
0 - 90%	15 - 20	**0.009%**
0 - 90%	20 - 50	**0.005%**
0 - 30%	10 - 50	**0.158%**
30 - 90%	10 - 50	**0.093%**
0 - 90%	10 - 50	**0.140%**

We can see that the remaining p_T systematic uncertainties are essentially negligible, which demonstrates the success of $\langle\frac{1}{\alpha\times\epsilon}\rangle$ efficiency correction approach.

Total Uncertainties on Selection Efficiency

Hence, in summary, the total uncertainties due to selection efficiency are simply the sum in of uncertainties due to data-MC discrepancy, finite MC events, and the remaining B-meson p_T shape in quadrature: efficiency = data-MC discrepancy $\oplus$ finite MC events $\oplus$ remaining B-meson p_T shape. Table **??** and Table **??** summarize the selection efficiency of B_s^0 and B^+ respectfully

Table 5.30: The B_s^0 selection efficiency uncertainty in p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	MC-Data	MC Stat.	p_T Shape	Total Uncertainties
0 - 90%	7 - 10	34.65%	26.5%	0.015%	**43.62%**
0 - 90%	10 - 15	5.64%	6.28%	0.050%	**8.44%**
0 - 90%	15 - 20	4.66%	3.08%	0.010%	**5.85%**
0 - 90%	20 - 50	10.24%	3.21%	0.022%	**10.73%**
0 - 30%	10 - 50	3.09%	6.59%	0.067%	**7.28%**
30 - 90%	10 - 50	3.66%	2.27%	0.015%	**4.31%**
0 - 90%	10 - 50	3.19%	4.37%	0.037%	**5.41%**

Table 5.31: The B^+ selection efficiency uncertainty in p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	MC-Data	MC Stat.	p_T Shape	Total Uncertainties
0 - 90%	7 - 10	4.17%	9.22%	0.166%	**10.12%**
0 - 90%	10 - 15	15.25%	3.36%	0.198%	**15.62%**
0 - 90%	15 - 20	3.01%	1.92%	0.009%	**3.57%**
0 - 90%	20 - 50	1.65%	1.35%	0.005%	**2.13%**
0 - 30%	10 - 50	13.28%	3.37%	0.158%	**13.70%**
30 - 90%	10 - 50	8.49%	2.26%	0.093%	**8.79%**
0 - 90%	10 - 50	11.51%	2.49%	0.140%	**11.78%**

5.14.6 Signal Extraction

Finally, there is also systematic uncertainties due to signal extraction. To evaluate the uncertainties, we try different models for the signal and background components in the unbinned fit to the B-meson invariant mass distributions. Then, we quote the percent deviation (% Dev) of variated signal raw yield N_S' from the nominal signal raw yield N_S as the systematic uncertainties.

Signal PDF Variation

For the signal component, we consider several variations listed as follows. First, we consider switching the nominal double Gaussian to triple Gaussian PDF.

- Switch double Gaussian to triple Gaussian
- Fix the mean of the double Gaussian to be the mean from the MC fits
- Increase/decrease the widths nominal double Gaussian function by 10%: $\sigma_{1,2} \rightarrow (1 + \pm 10\%)\sigma_{1,2}$

Figure 5-65 and Figure 5-66 show examples of signal PDF variation done in B_s^0 and B^+ invariant mass fits to estimate the systematic uncertainties due to signal extraction

Next, we simply take the largest percent variations of the variated signal raw yields compared to the nominal ones and quote them as the systematic uncertainties. Table **??** and Table **??** summarize the signal PDF systematic uncertainties

Background PDF Variation

The variation of background is simpler. We replace the nominal exponential decay function by first, second, and third order polynomials. Again, same as the signal PDF variation, we quote the background PDF systematic uncertainties as the percent deviation from the variated signal raw yield to the nominal signal raw yield.

Table 5.32: The B_s^0 systematic uncertainty due to signal PDF variation in p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	Triple Gaussian	Fixed Mean	+10% Width	-10% Width
0 - 90%	7 - 10	**0.862%**	0.762%	0.066%	0.056%
0 - 90%	10 - 15	0.144%	0.137%	1.83%	**2.34%**
0 - 90%	15 - 20	**0.110%**	0.066%	**1.02%**	1.00%
0 - 90%	20 - 50	**0.629%**	0.056%	1.40%	**1.88%**
0 - 30%	10 - 50	0.241%	0.192%	1.47%	**1.74%**
30 - 90%	10 - 50	**2.96%**	0.208%	0.840%	0.767%
0 - 90%	10 - 50	1.56%	0.20%	1.33%	**1.65%**

Table 5.33: The B^+ systematic uncertainty due to signal PDF variation in p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	Triple Gaussian	Fixed Mean	+10% Width	-10% Width
0 - 90%	7 - 10	0.004%	1.13%	3.84%	**4.46%**
0 - 90%	10 - 15	0.235%	0.046%	2.27%	**2.67%**
0 - 90%	15 - 20	0.050%	0.025%	2.32%	**2.74%**
0 - 90%	20 - 50	0.750%	0.010%	1.79%	**2.36%**
0 - 30%	10 - 50	0.415%	0.155%	2.12%	**2.50%**
30 - 90%	10 - 50	0.370%	0.064%	2.10%	**2.57**%
0 - 90%	10 - 50	0.494%	0.060%	2.19%	**2.60%**

Figure 5-67 and Figure 5-68 show examples of signal PDF variation done in B_s^0 and B^+ invariant mass fits to estimate the systematic uncertainties due to signal extraction

Again, we simply take the largest percent variations of the variated signal raw yields compared to the nominal ones and quote them as the systematic uncertainties. Table **??** and Table **??** summarize the background PDF systematic uncertainties

Total PDF Variation

To calculate the total PDF variation, we simply add the PDF variation of the signal and background into quadrature: signal extraction uncertainties = PDF variation

Table 5.34: The B_s^0 systematic uncertainty due to background PDF variation in p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	First Order	Second Order	Third Order
0 - 90%	7 - 10	0.02%	0.92%	**0.96%**
0 - 90%	10 - 15	1.45%	**2.69%**	2.64%
0 - 90%	15 - 20	**1.45%**	1.36%	0.28%
0 - 90%	20 - 50	3.30%	3.54%	**6.11%**
0 - 30%	10 - 50	**1.83%**	0.495%	0.029%
30 - 90%	10 - 50	1.14%	**1.33%**	0.912%
0 - 90%	10 - 50	**1.58%**	0.773%	0.592%

Table 5.35: The B^+ systematic uncertainty due to background PDF variation in p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	First Order	Second Order	Third Order
0 - 90%	7 - 10	0.021%	0.117%	**0.093%**
0 - 90%	10 - 15	0.312%	0.380%	**0.546%**
0 - 90%	15 - 20	0.386%	0.432%	**0.576%**
0 - 90%	20 - 50	**0.196%**	0.238%	**1.03%**
0 - 30%	10 - 50	0.157%	0.245%	**0.412%**
30 - 90%	10 - 50	**1.13%**	0.065%	0.102%
0 - 90%	10 - 50	0.384%	**0.427%**	0.422%

of signal $\oplus$ background. Table 5.36 and Table 5.37 show the total PDF systematic uncertainties contributing to B_s^0 and B^+ signal extraction respectfully.

5.14.7 Summary

Finally, we collect all studies above and all all sources of systematic uncertainties into quadrature to obtain the total systematic uncertainties. We compile the summary for the systematic uncertainties in measurements shown in Table 5.40 and Table 5.41 for B_s^0 p_T and centrality cross section measurements

Table **??** and Table **??** show B^+ p_T and centrality cross section measurements respectfully

Finally, Figure 5-69 and Figure 5-70 below illustrate the plot of all sources of B_s^0

Table 5.36: The B_s^0 signal extraction systematic uncertainty due to PDF variation in p_T and centrality bin is summarized below.

Centrality	B_s^0 p_T (GeV/c)	Signal PDF	Background PDF	Total Uncertainties
0 - 90%	7 - 10	0.762%	0.96%	**1.23%**
0 - 90%	10 - 15	2.34%	2.69%	**3.57%**
0 - 90%	15 - 20	1.02%	1.45%	**1.77%**
0 - 90%	20 - 50	1.88%	6.11%	**6.39%**
0 - 30%	10 - 50	1.74%	1.83%	**2.53%**
30 - 90%	10 - 50	2.96%	1.33%	**3.25%**
0 - 90%	10 - 50	1.65%	1.58%	**2.28%**

Table 5.37: The B^+ signal extraction systematic uncertainty due to PDF variation in p_T and centrality bin is summarized below.

Centrality	B^+ p_T (GeV/c)	Signal PDF	Background PDF	Total Uncertainties
0 - 90%	7 - 10	4.46%	0.117%	**4.46%**
0 - 90%	10 - 15	2.67%	0.546%	**2.73%**
0 - 90%	15 - 20	2.74%	0.576%	**2.80%**
0 - 90%	20 - 50	2.36%	1.03%	**2.57%**
0 - 30%	10 - 50	2.50%	0.412%	**2.53%**
30 - 90%	10 - 50	2.57%	1.13%	**2.81%**
0 - 90%	10 - 50	2.60%	0.427%	**2.64%**

and B^+ systematic uncertainties respectfully for each p_T and centrality bin.

Table 5.38: Summary of systematic uncertainties from each B_s^0 p_T bin. All the values are shown in percentage.

B_s^0 p_T (GeV/c)	7 - 10	10 - 15	15 - 20	20 - 50
Tracking Efficiency	10%	10%	10%	10%
Muon Efficiency	+8.88%	+5.97%	+3.73%	+3.91%
	−7.49%	−5.24%	−3.46%	−3.60%
Selection Efficiency	43.62%	8.44%	5.85%	10.73%
Signal Extraction	1.23%	3.57%	1.77%	6.39%
Total	+45.64%	+14.82%	+12.18%	+16.47%
	−45.39%	−14.54%	−12.10%	−16.40%
N_{MB}	1.26%	1.26%	1.26%	1.26%
T_{AA}	2.2%	2.2%	2.2%	2.2%
Branching Ratio	7.5%	7.5%	7.5%	7.5%
Global Systematics	7.92%	7.92%	7.92%	7.92%

Table 5.39: Summary of systematic uncertainties from each B_s^0 centrality bin. All the values are shown in percentage.

PbPb Collision Centrality	0 - 30%	30 - 90%	0 - 90%
Tracking Efficiency	10%	10%	10%
Muon Efficiency	+5.52%	+4.63%	+5.29%
	−4.85%	−4.19%	−4.70%
Selection Efficiency	7.28%	4.31 %	5.41%
Signal Extraction	2.53%	3.25%	2.28%
T_{AA}	2%	3.6%	2.2%
N_{MB}	1.26%	1.26%	1.26%
Total	+13.68%	+12.71%	+13.00%
	-13.52%	-12.60%	-12.77%
Branching fractions	7.5%	7.5%	7.5%
Global Systematics	7.5%	7.5%	7.5%

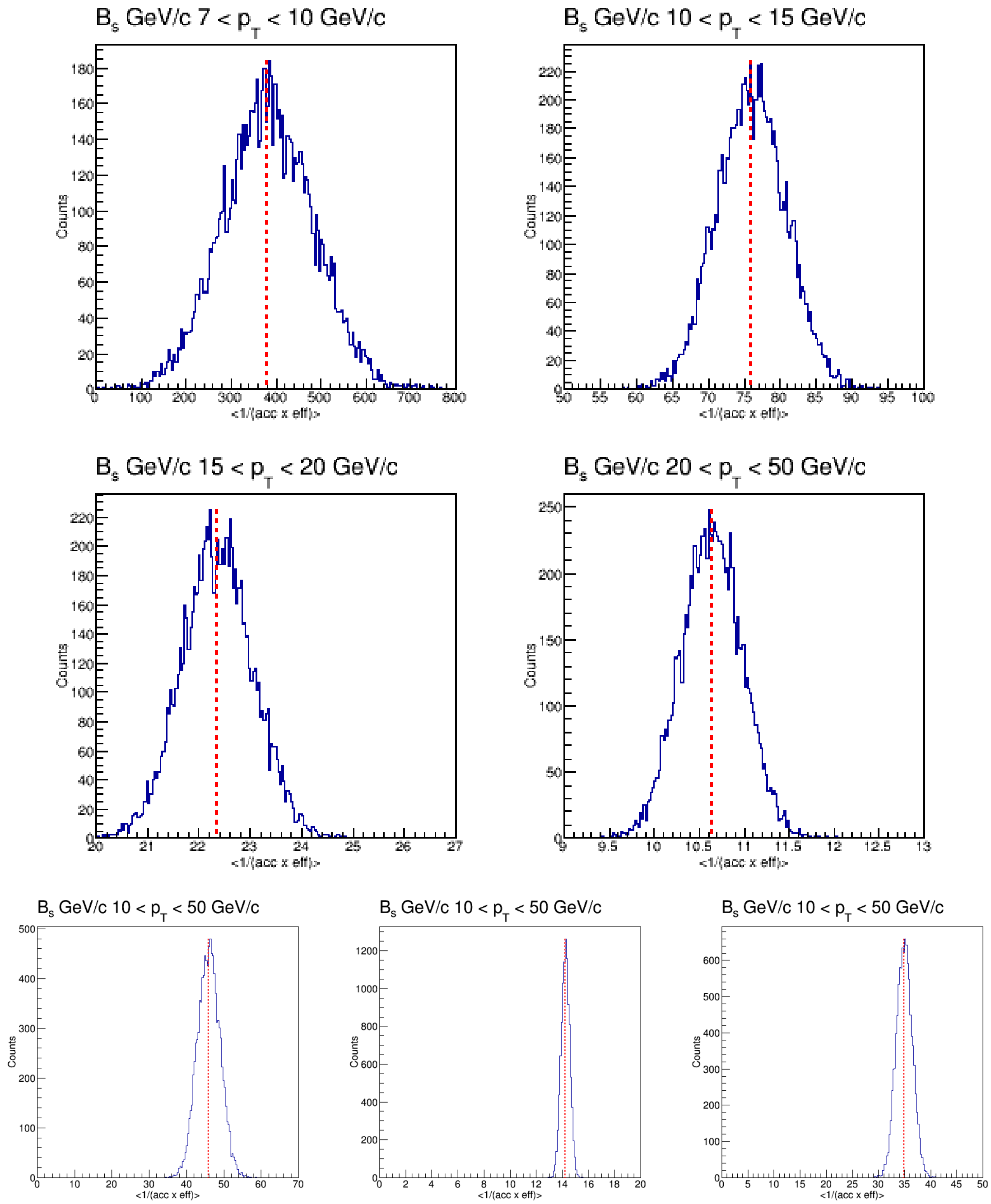


Figure 5-63: The distribution of B_s^0 $\langle \frac{1}{\alpha \times \epsilon} \rangle$ for centrality in 0 - 90% in the p_T range of 7 - 10, 10 - 15, 15 - 20, and 20 - 50 GeV/c as well as 0 - 90%, 0 - 30%, and 30 - 90% in the p_T range of 10 - 50 GeV/c are shown above. The red dash lines are our nominal value for efficiency correction.

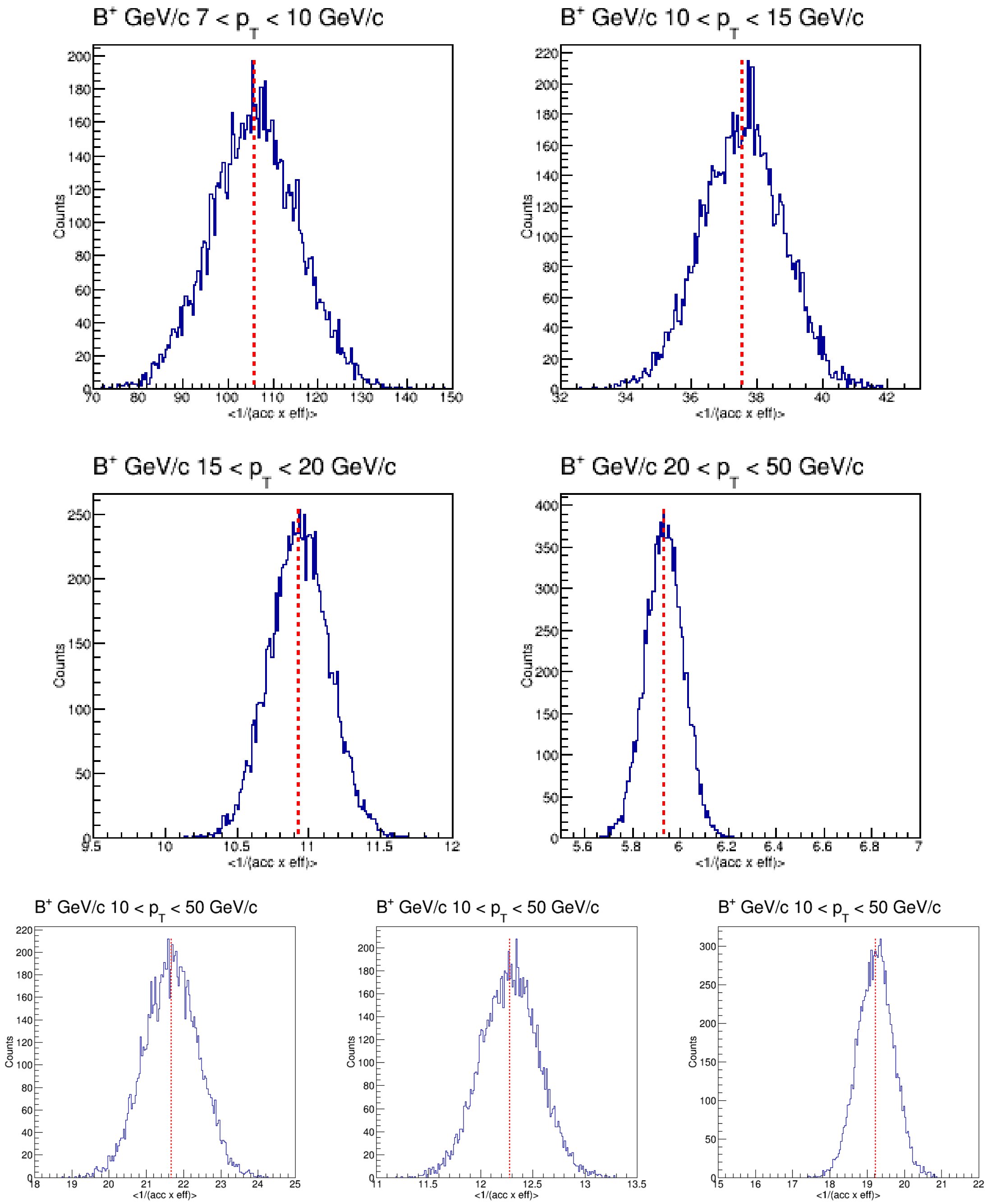


Figure 5-64: The distribution of B^+ $\langle\frac{1}{\alpha\times\epsilon}\rangle$ for centrality in 0 - 90% in the p_T range of 7 - 10, 10 - 15, 15 - 20, and 20 - 50 GeV/c as well as 0 - 90%, 0 - 30%, and 30 - 90% in the p_T range of 10 - 50 GeV/c are shown above. The red dash lines are our nominal value for efficiency correction.

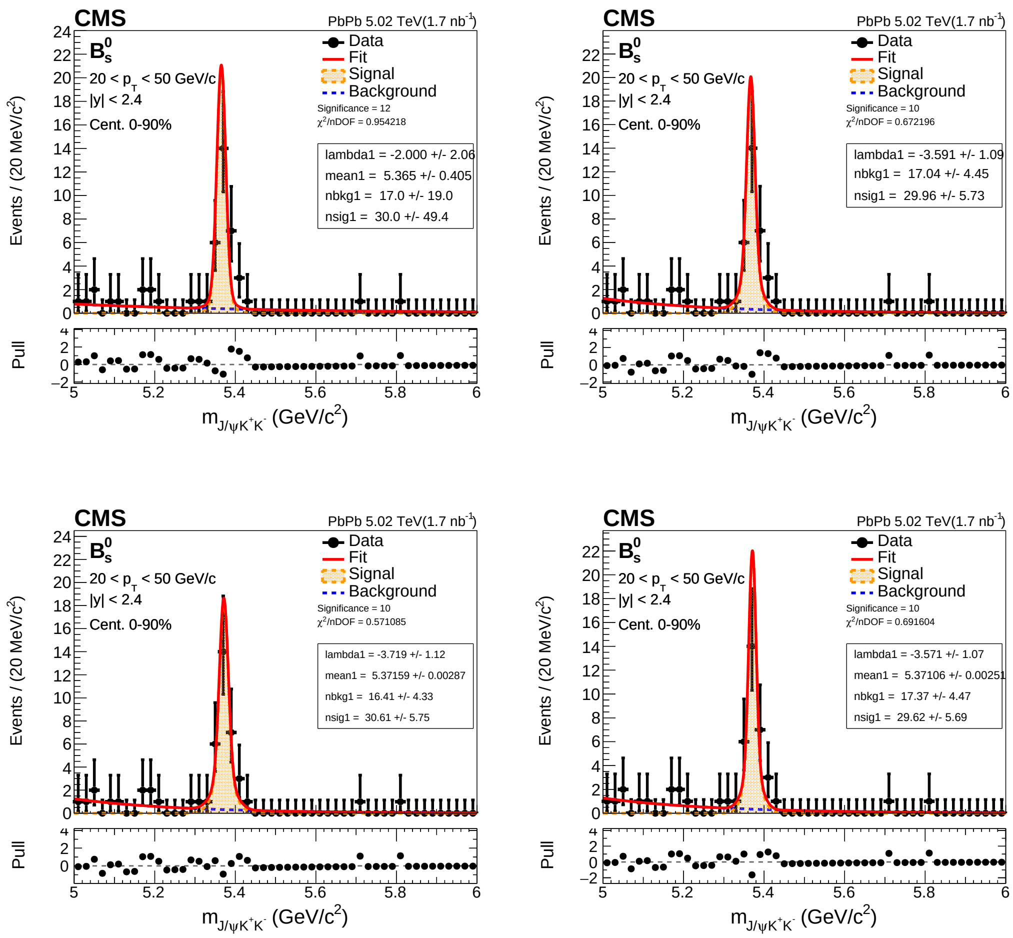


Figure 5-65: Invariant mass fit of B_s^0 candidates for B_s^0 p_T from 20 - 50 GeV/c and centrality from 0 to 90% in 5.02 TeV PbPb. The signal pdf from left to right is triple gaussian (with widths and relative proportions fixed from MC), double gaussian with all the parameters fixed (including the mean), increased width (a=1.1), and decreased width (a=0.9).

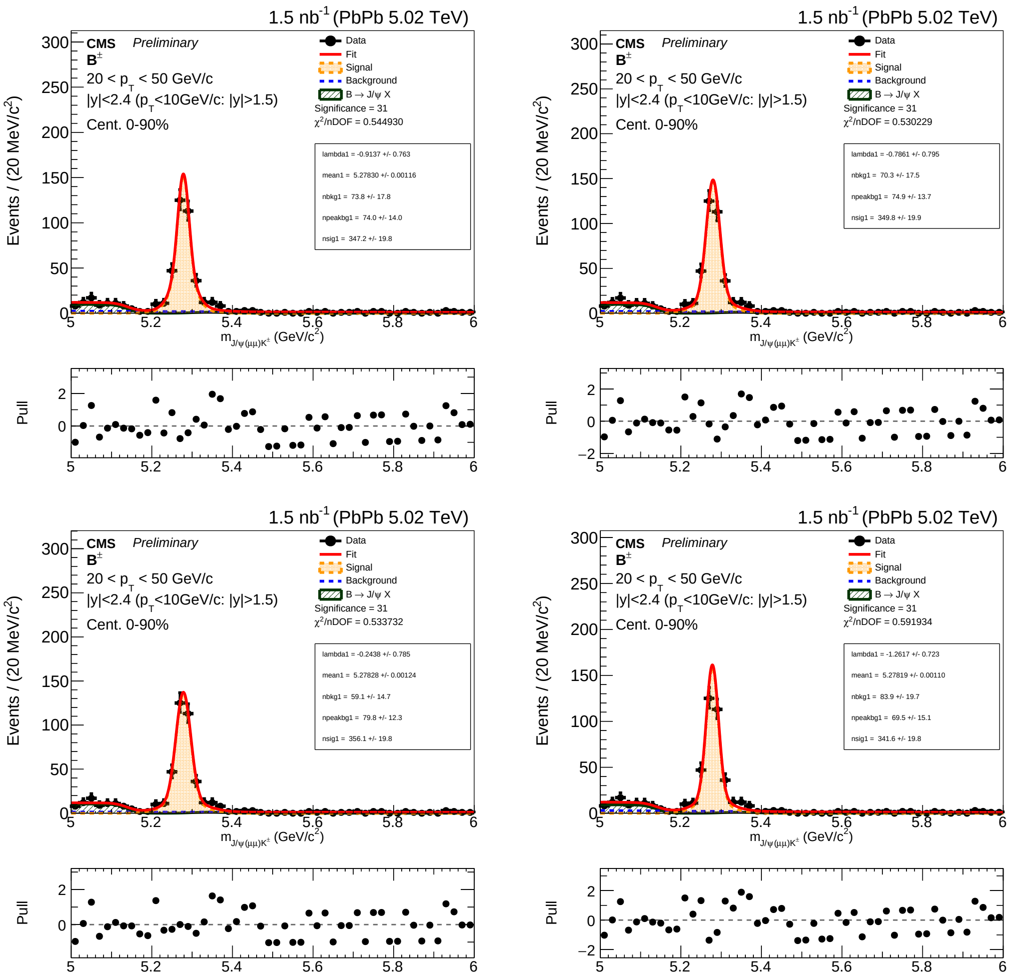


Figure 5-66: Invariant mass fit of B^+ candidates for B^+ p_T from 20 - 50 GeV/c and centrality from 0 to 90% in 5.02 TeV PbPb. The signal pdf from left to right is triple gaussian (with widths and relative proportions fixed from MC), double gaussian with all the parameters fixed (including the mean), increased width (a=1.1), and decreased width (a=0.9).

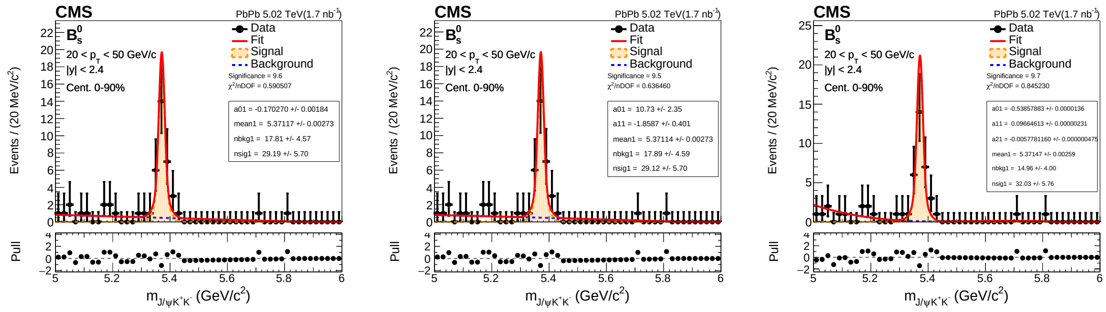


Figure 5-67: Invariant mass fit of B_s^0 candidates for B_s^0 p_T from 20 - 50 GeV/c and centrality from 0 to 90% in 5.02 TeV PbPb. The background PDFs from left to right are first, second, and third order polynomials.

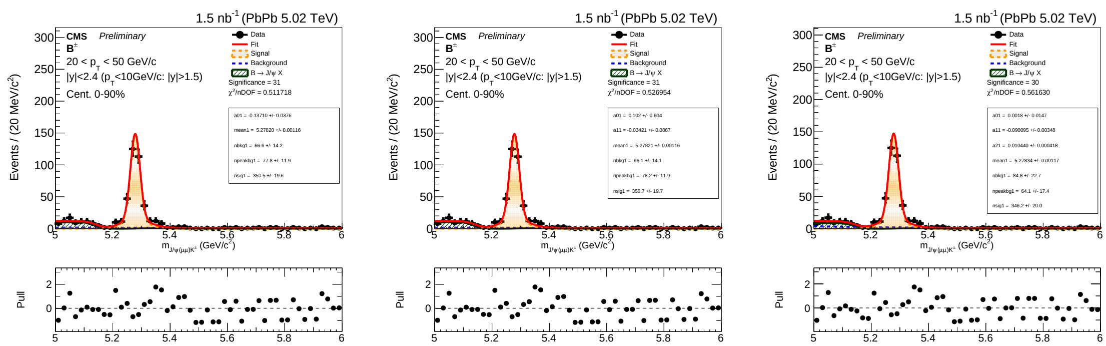


Figure 5-68: Invariant mass fit of B^+ candidates for B^+ p_T from 20 - 50 GeV/c and centrality from 0 to 90% in 5.02 TeV PbPb. The background PDFs from left to right are first, second, and third order polynomials.

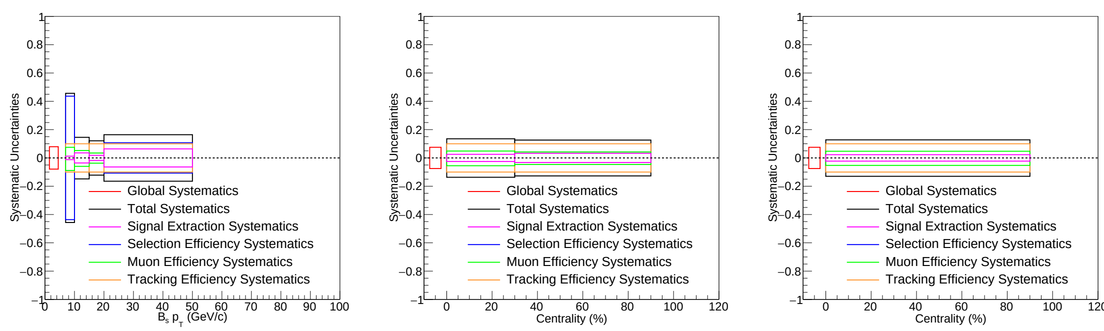


Figure 5-69: The summary of B_s^0 systematic uncertainties plotted as a function of p_T (left), centrality in 0 - 30%, and 30 - 90% (middle), and the inclusive centrality bin 0 - 90% (right) are shown above.

Table 5.40: Summary of systematic uncertainties from each B^+ p_T bin. All the values are shown in percentage.

B^+ p_T (GeV/c)	7 - 10	10 - 15	15 - 20	20 - 50
Tracking Efficiency	5%	5%	5%	5%
Muon Efficiency	$+7.21\%$	$+4.29\%$	$+3.83\%$	$+3.87\%$
	-6.28%	-3.92%	-3.53%	-3.56%
Selection Efficiency	10.12%	15.62%	3.57%	2.13%
Signal Extraction	4.46%	2.73%	2.80%	2.57%
Total	$+14.04\%$	$+17.14\%$	$+7.75\%$	$+7.15\%$
	-13.59%	-17.05%	-7.61%	-6.99%
$\mathrm{N_{MB}}$	1.26%	1.26%	1.26%	1.26%
T_{AA}	2.2%	2.2%	2.2%	2.2%
Branching Ratio	2.9%	2.9%	2.9%	2.9%
Global Systematics	3.85%	3.85%	3.85%	3.85%

Table 5.41: Summary of systematic uncertainties from each B^+ centrality bin. All the values are shown in percentage.

PbPb Collision Centrality	0 - 30%	30 - 90%	0 - 90%
Tracking Efficiency	10%	10%	10%
Muon Efficiency	$+4.18\%$	$+4.14\%$	$+4.16\%$
	-3.83%	-3.80%	-3.81%
Selection Efficiency	13.70%	8.79%	11.78%
Signal Extraction	2.53%	2.81%	2.64%
T_{AA}	2%	3.6%	2.2%
$\mathrm{N_{MB}}$	1.26%	1.26%	1.26%
Total	$+15.53\%$	$+11.89\%$	$+13.92\%$
	-15.43%	-11.77%	-13.82%
Branching fractions	2.92%	2.92%	2.92%
Global Systematics	2.92%	2.92%	2.92%

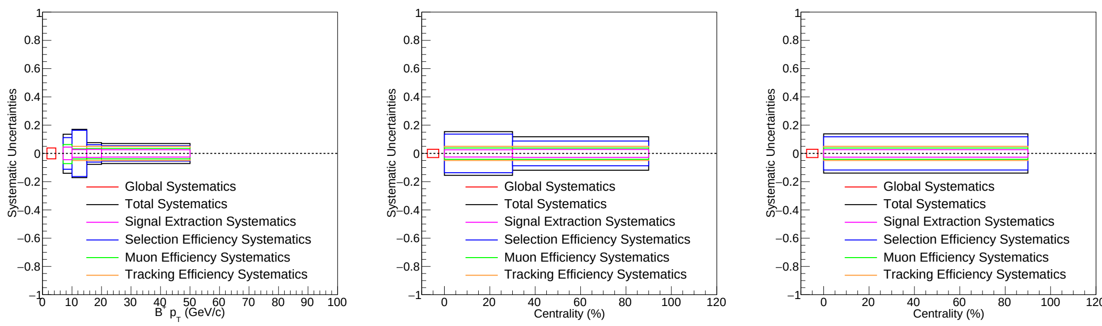


Figure 5-70: The summary of B^+ systematic uncertainties plotted as a function of p_T (left), centrality in 0 - 30%, and 30 - 90% (middle), and the inclusive centrality bin 0 - 90% (right) are shown above.

5.15 Final Results

5.15.1 Overview

At this point, we have fully validated the analysis with detail evaluation of both statistical and systematic uncertainties. We are ready to report the experimental measurements of B_s^0 and B^+ cross section and the $B_s^0/B+$ ratio as function of p_T and PbPb collision centrality.

5.15.2 B_s^0 and B^+ Cross Section

When B-meson measurement as a function of p_T plots are created, the abscissa of each data point is set to the mean value of the p_T distribution, after background subtraction via *Splot*. Figure 5-71 shows B_s^0 and B^+ p_T differential cross section $\frac{1}{T_{AA}}\frac{dN}{d_pT}$ as a function p_T in PbPb collisions at $\sqrt{s_{NN}}$ =5.02 TeV with the CMS detector using the 2018 dimuon PbPb dataset.

Table 5.42 and Table 5.43 summarize the B_s^0 and B^+ cross section as functions of p_T:

Table 5.42: The numerical values and uncertainties of B_s^0 cross section as a function of p_T are summarized below.

p_T (GeV/c)	$\frac{1}{T_{AA}}\frac{dN}{d_pT}$ (pb c/GeV)	Stat. Up (+)	Stat. Down (-)	Syst. Up (+)	Syst. Down (-)
7 - 10	160432	51.3%	48.3%	45.6%	45.4%
10 - 15	75523.7	22.4%	25.6%	14.8%	14.5%
15 - 20	25355	21.6%	20.7%	12.2%	12.1%
20 - 50	2272.18	21.6%	16.3%	16.5%	16.4%

It should note that for the p_T bin 7 - 10 GeV/c, the measurement is a fiducial measurement correcting to B mesons rapidity only up to $1.5 < |y| < 2.4$. The measurement has a p_T range of 7 - 50 GeV/c. The uncertainties are large for B_s^0 p_T from 7 to 10 GeV/c due to the limited statistics.

Figure 5-72 shows B_s^0 and B^+ p_T integrated cross section as function of average

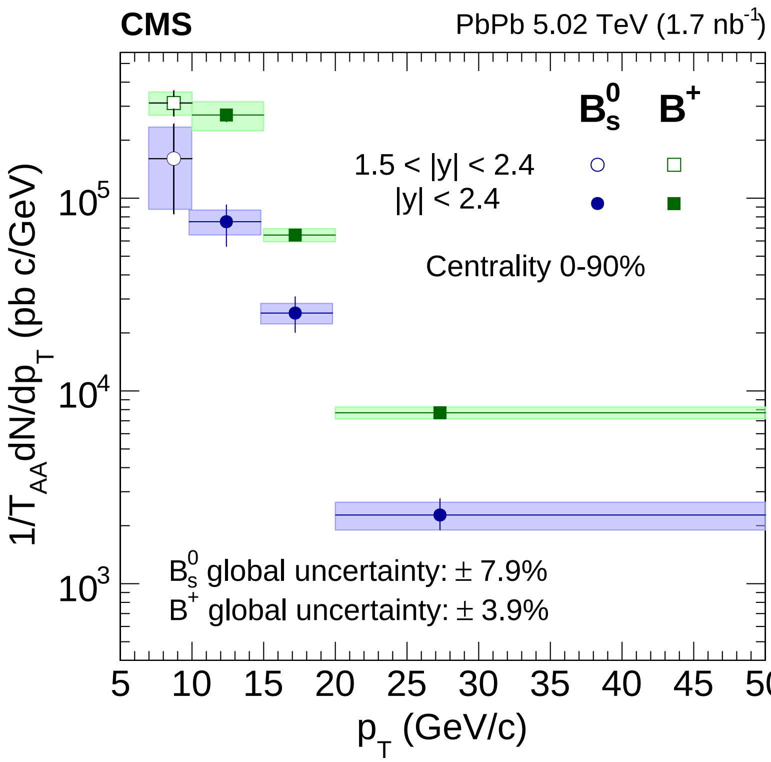


Figure 5-71: The measurement of B_s^0 and B^+ p_T differential cross section $\frac{1}{T_{AA}}\frac{dN}{dp_T}$ as a function of B-meson p_T within 0 - 90% centrality is shown above. It should be pointed out that the plot is plot in the unit of pb c/GeV since T_{AA} is in the unit of pb^{-1} while the p_T is in the unit of GeV/c. The open markers from 7 - 10 GeV/c stands for the fiducial measurement at the B-meson rapidity of $1.5 < |y| < 2.4$.

Table 5.43: The numerical values and uncertainties of B^+ cross section as a function of p_T are summarized below.

p_T (GeV/c)	$\frac{1}{T_{AA}} \frac{dN}{d_pT}$ (pb c/GeV)	Stat. Up (+)	Stat. Down (-)	Syst. Up (+)	Syst. Down (-)
7 - 10	311668	15.9%	14.3%	14.0%	13.6%
10 - 15	270167	6.63%	7.95%	17.1%	17.1%
15 - 20	64384.4	6.54%	6.50%	7.75%	7.61%
20 - 50	7704.11	6.90%	5.26%	7.15%	6.99%

number of participant nucleons $\langle N_{part} \rangle$, with the correspondence to PbPb collision centrality labelled in the plot

Table 5.44 and Table 5.45 summarize the B_s^0 and B^+ cross section as functions of centrality and their statistical and systematic uncertainties.

Table 5.44: The numerical values and uncertainties of B_s^0 cross section as a function of centrality bin are summarized below.

Centrality	$\frac{1}{T_{AA}} N$ (pb)	Stat. Up (+)	Stat. Down (-)	Syst. Up (+)	Syst. Down (-)
0 - 30%	650790	21.2%	22.0%	14.0%	13.7%
30 - 90%	497359	19.5%	19.7%	12.8%	12.7%
0 - 90%	595064	17.9%	16.3%	13.0%	12.8%

Table 5.45: The numerical values and uncertainties of B^+ cross section as a function of centrality bin are summarized below.

Centrality	$\frac{1}{T_{AA}} N$ (pb)	Stat. Up (+)	Stat. Down (-)	Syst. Up (+)	Syst. Down (-)
0 - 30%	1780710	5.52%	6.72%	15.5%	15.4%
30 - 90%	2286890	6.71%	7.06%	11.9%	11.8%
0 - 90%	1936560	4.47%	4.76%	13.9%	13.8%

It also should note that this is the first centrality differential fully measurement of reconstructed B_s^0 with the CMS experiment. From the cross section above, we can see that the nominal value of the cross section is lower for both B_s^0 and B^+ in more

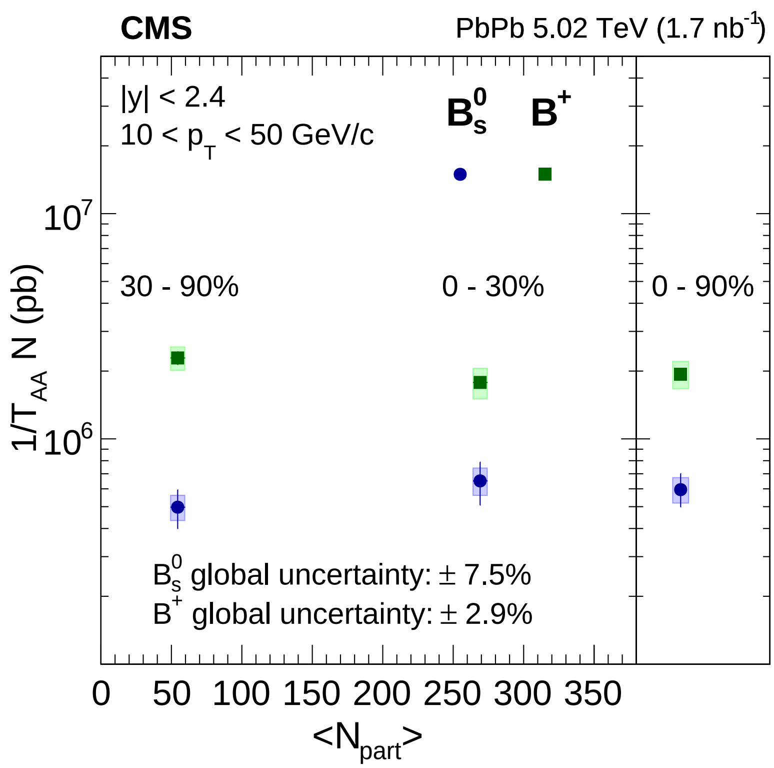


Figure 5-72: The measurement of B_s^0 and B^+ p_T differential cross section $\frac{1}{T_{AA}}\frac{dN}{d_pT}$ measurement as function of $\langle N_{part}\rangle$ at different PbPb collision centrality within B-meson p_T from 10 - 50 GeV/c is shown above.

central collision (0 - 30%) compare to the peripheral collision (30 - 90%). However, due to the large uncertainties, we can not draw a conclusion that this suppression due to the QGP medium effect.

5.15.3 B_s^0/B^+ Ratio

Since we have measured the B_s^0 and B^+ cross section, the next step is to obtain their ratio, which is an observable to study beauty hadronization mechanism as mention in Section 1.8.4. We should note that when we take the ratio between B_s^0 and B^+, the systematic uncertainties of T_{AA} and N_{MB} cancel. Also, the systematic uncertainties of tracking efficiency is reduced to 5% instead of adding into quadrature. Finally, the systematic uncertainties of muon efficiency are treated as perfectly correlated. This means we both vary the B_s^0 and B^+ cross section with the muon efficiency systematics up and down and compute the ratio. Then, we calculate the percent deviation of up and down variated B_s^0/B^+ to the nominal B_s^0/B^+ and quote those numbers as the systematic uncertainties. Everything other sources of systematic uncertainties of B_s^0 and B^+ cross section are added into quadrature for B_s^0/B^+ ratio.

Figure 5-73 shows the B_s^0/B^+ ratio as a function of B-meson p_T

Table 5.47 summarizes B_s^0/B^+ ratio as a function of B-meson p_T

Table 5.46: The numerical values and uncertainties of B_s^0/B^+ cross section ratio as a function of p_T are summarized below.

p_T (GeV/c)	Abc. (GeV/c)	B_s^0/B^+	stat. up (+)	stat. down (-)	syst. up (+)	syst. down (-)
(7,10)	8.75	0.5148	53.7%	50.4%	46.0%	46.1%
(10,15)	12.6	0.2795	22.3%	26.8%	19.3%	19.3%
(15,20)	17.4	0.3938	22.6%	21.7%	9.10%	9.11%
(20,50)	27.3	0.2949	22.7%	17.1%	13.9%	13.9%

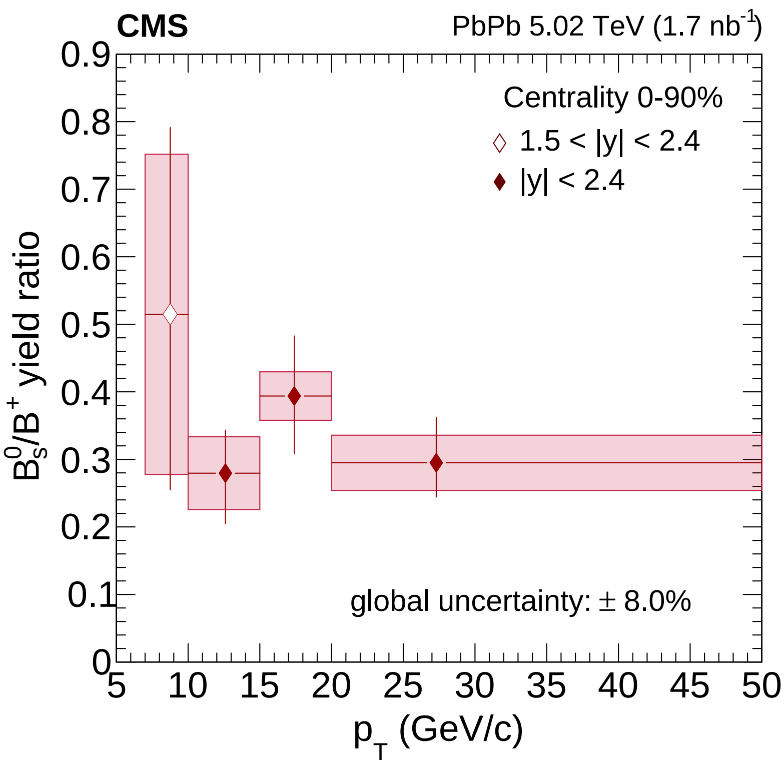


Figure 5-73: The measurement of B_s^0/B^+ as a function of p_T is shown above. The open markers from 7 - 10 GeV/c stands for the fiducial measurement at the B-meson rapidity of $1.5 < |y| < 2.4$.

Figure 5-74 shows the B_s^0/B^+ ratio as a function of PbPb collision centrality

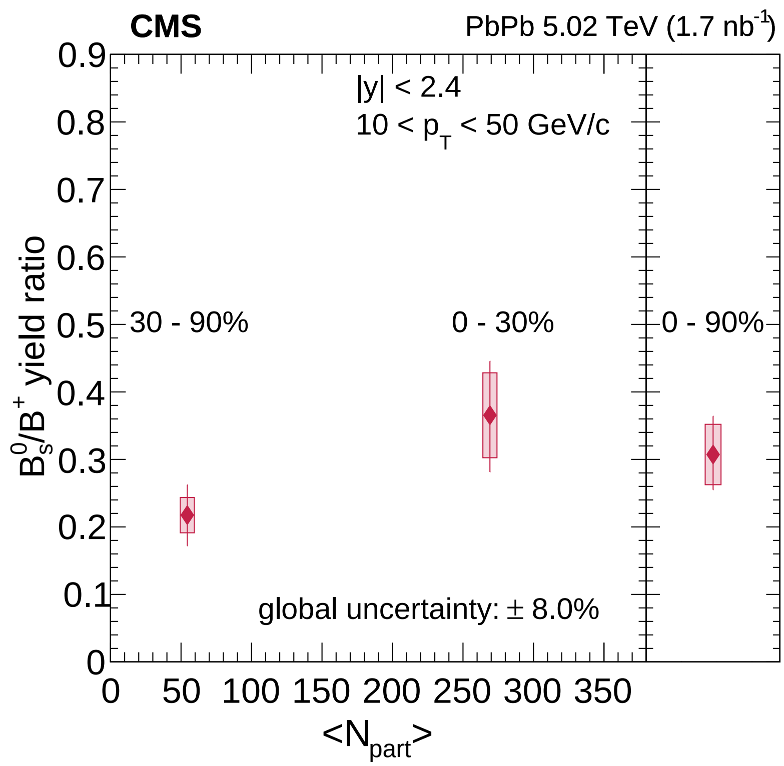


Figure 5-74: The measurement of B_s^0/B^+ as a function of p_T is shown above.

Table 5.47 summarizes B_s^0/B^+ ratio as a function of PbPb collision centrality

Table 5.47: The numerical values and uncertainties of B_s^0/B^+ cross section ratio as a function of centrality are summarized below.

Centrality	B_s^0/B^+	stat. up (+)	stat. down (-)	syst. up (+)	syst. down (-)
0 - 30%	0.3655	21.9%	23.0%	17.2%	17.2%
30 - 90%	0.2175	20.6%	20.9%	12.0%	12.0%
0 - 90%	0.3073	18.4%	17.0%	14.5%	14.6%

Finally, we can also cross check our measurement with the 2015 published B_s^0/B^+ ratio. Figure **??** show the direct comparison of our data .

Within uncertainties, we see that our results are consistent to the 2015 published results but provide much more information since our measurement is more precise and differential.

Hence, We have reported the measurement of B_s^0 and B^+ mesons cross section and the B_s^0/B^+ ratio in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV with the CMS experiment. In the next chapter, we will answer the questions raised in Section 2.4 through the comparison our data with other experimental references and theoretical predictions.

Chapter 6

Conclusions

With B_s^0 and B^+ measurements in PbPb collision, we can study beauty hadrochemistry and answer questions raised in Section 2.7.

6.1 pp Reference and Theoretical Models

Because our fully reconstructed B-meson analysis in pp as the reference is still ongoing, in order to understand our PbPb data, we need to add the B-meson pp measurements from other experiments at the LHC. The follow list the pp references we use to compare our PbPb measurement

LHCb 7 TeV pp result at $2 < |y| < 5$**:** This reference is chosen because it is one of the most precise B_s^0/B^+ measurements with energy is closest to the 5.02 TeV in our analysis **??** . The original results is the efficiency corrected yield ratio $\mathcal{R} = \frac{N(B_s^0 \to J/\psi\phi)}{N(B^+ \to J/\psi K^+)} \cdot \frac{\epsilon(B^+ \to J/\psi K^+)}{\epsilon(B_s^0 \to J/\psi\phi)}$ [134]. We scale $\mathcal{R}$ by the branching ratio of $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$ to $B^+ \to J/\psi K^+ \to \mu^+\mu^-K^+$ and make them to be the same quantity as our B_s^0/B^+ measurement.

ATLAS 7 TeV pp results at $|y| < 2.5$**:** This reference is chosen because it is measured over a rapidity range similar to our measurement range **??** . The original results is the ratio of the fragmentation fraction f_s/f_d. Using the isospin symmetry, we $f_d = f_u$. In addition, the ATLAS paper uses the QCD calculation $BF(QCD) = \frac{BR(B_s^0 \to J/\psi\phi)}{BR(B^0 \to J/\psi K^{*0})} = 0.83$ instead of quote the PDG the branching ratios

$BF(PDG) = \frac{BR(B_s^0 \to J/\psi\phi)}{BR(B^0 \to J/\psi K^{*0})} = 0.85$. Hence, we relate ATLAS f_s/f_d to our B_s^0/B^+ via: $B_s^0/B^+ = BF(PDG)/BF(QCD) f_s/f_d$ and compare the ATLAS scaled pp data to our data.

LHCb and ATLAS at different rapidity ranges. Since the rapidity dependence is not significant in B_s^0/B^+ ratio as demonstrated in Figure **??** from the LHCb publication [134], assuming it is also insignificant in PbPb, we can use the pp reference at different rapidity ranges as references in our B_s^0/B^+ measurement

In addition to the pp reference, we also include the prediction from TAMU (labelled as "PbPb: TAMU" in orange), Cao, Sun, Ko (labelled as "PbPb: Langevin" in green), and (labelled as) which have been introduced in Section 1. Here we summarize

Figure 6-1 show the comparison between our B_s^0/B^+ measurement with pp references and theoretical model calculations.

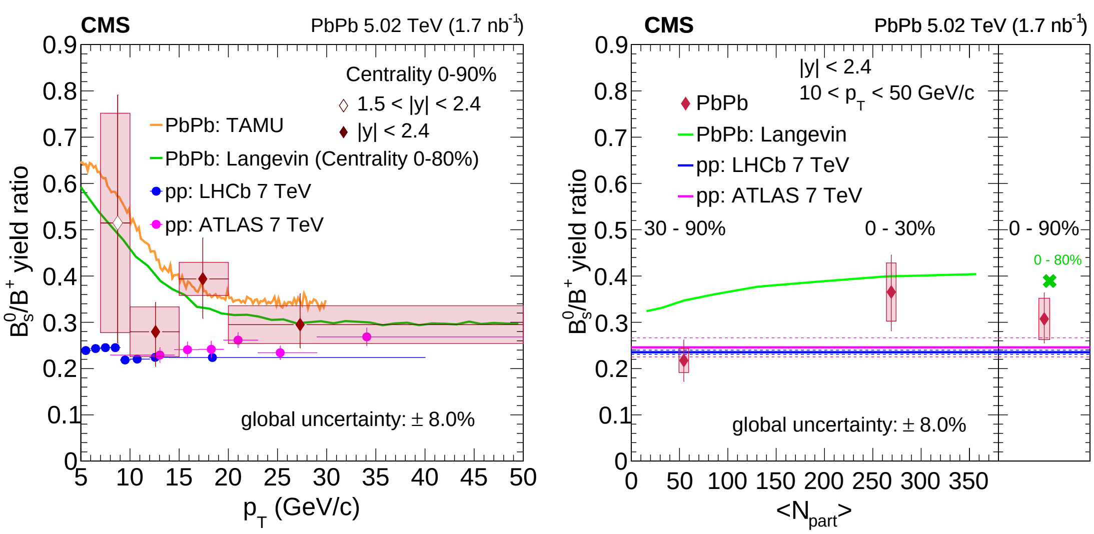


Figure 6-1: The fully reconstructed B_s^0/B^+ (left) and B_s^0/B^+ R_{AA} ratio (right) as a function of p_T using the 2015 CMS pp and PbPb datasets are shown above. Both plot include the ATLAS (magenta) and LHCb (blue) 7 TeV reference. The TAMU model (orange) has only p_T dependent prediction plot on the left figure while the Cao, Sun, Ko model (green) has both p_T and centrality predictions plotted on both figures.

6.2 Physics Messages Discussion

Figure 6-1 conveys a lot of information. We will discuss the physics message of our measurement through the comparison of our data with the pp references and theoretical model prediction:

Substantial Uncertainties at Low p_T: Both statistical and systematic uncertainties of B_s^0/B^+ ratio is large in $7 < p_T < 10$ GeV/c. They come B_s^0. However, we know that the statistics of B_s^0 in the p_T range 7 - 10 GeV/c is indeed very small. In fact, from FONLL calculation, we expect to get only about B_s^0 candidates. From our estimation, we expect only about 100. Unfortunately, some of the systematic uncertainties, for instance, the one due to finite MC simulation statistics, which contribute a lot (25%) to the total systematic uncertainties (46%) can be in principle further reduce.

No significant p_T dependence: According to the B^0/B^+ ratio as a function of B-meson p_T, apparently, there is no significant change of the central values above 10 GeV/c. For 7 - 10 GeV/c, the central value jumps from 0.28 up to 0.51. However, the uncertainties of the measurement is also large. Considering all the uncertainties, we do not observe significant p_T dependence on the B_s^0/B^+ ratio.

Good Agreement with theoretical models: Comparing our data to the TAMU and Cao, Sun Ko model calculations, the B_s^0/B^+ vs p_T data agree well with these two models. They both predict the trend of the central values of our data, which decreases and then approaches to flat values as p_T increases. The TAMU model always lies above the Cao, Sun, Ko model because it only employs quark coalescence model in hadronization. In Cao, Sun, Ko model, fragmentation hadronization is also considered.

However, we know that in the limit $p_T \to \infty$, the B_s^0/B^+ in PbPb collision will be very similar to pp since the fast moving beauty quark traverse through the medium within a very short time and is not likely to combine with any quarks in the medium. Hence, fragmentation haronization dominates for fast moving beauty quarks.

In the centrality measurement, the Cao, Sun, Ko model prediction are also reasonably consistent to our for 0 - 30% and 0 - 90%. However, in the peripheral 30 - 90% collisions, the Cao, Sun, Ko model has a lager $B_s^0/B+$ ratio compared to our data, which coincides with the pp reference.

Compatible to pp references: While the center B_s^0/B^+ data systematically lies above the pp reference, taking into account all the uncertainties, they are within about 1 σ except the peripheral 30 - 90% bin with a small number of participant nucleons. However, it should note that the energy of pp reference is higher than the PbPb data. LHCb has reported that B_s^0/B^+ ratio increases with increasing energy [134]. Therefore, it would make the comparison much better if we could also perform B_s^0/B^+ measurement in pp collisions with CMS and compare to the PbPb results

6.3 Conclusions

With the physics message obtained from the discussion, we are prepared to answer the question raised in Section 2.3 and draw conclusions of our studies in this thesis below:

First Observation of B_s^0 in Nucleus-Nucleus Collisions: In the analysis, we have fully reconstructed B_s^0 with greater 5σ significance in all p_T and centrality bins. Therefore, we have improved our measurement compared to the 2015 published results and first observed fully reconstructed B_s^0 in heavy-ion collisions.

Significant Improvement of the Previous Results: While we have successfully reproduced the published results using 2015 datasets, our new results measure B_s^0/B^+ as a function of centrality for the first time. In addition, thanks to the higher statistics of the dataset, we are able to measure four p_T bin with smaller uncertainties, provide more information about the p_T dependence of the B_s^0/B^+ ratio. In our measurement, we find no significant p_T dependence down to at least 10 GeV/c. In addition, there is a hint of suppression of B-meson cross section in central collision compare to peripheral collisions to be confirmed with larger statistics.

Inconclusive about Strangeness Enhancement: There is a low hint of potential strangeness enhancement in PbPb collisions, particularly at low p_T, compare to pp collisions since the B_s^0/B^+ ratio in PbPb are systematically higher than pp with about 1 - 1.5 σ. However, the hint is not strong enough. We will need more statistics to confirm this hint.

Fragmentation mechanism alone is not enough to describe our data: We can see that quark coalescence effect must be there because our data lies systematically above the pp reference. Looking at the most central collision from 0 - 30 %, the B_s^0/B^+ ratio is about 1.25σ, which corresponds to about 80% confidence, above the LHCb pp reference. The explicit computation is shown as follows:

$$\%Dev = (0.3655 - 0.2353)/(0.3655 * \sqrt{0.23^2 + 0.172^2}) = 1.25 \tag{6.1}$$

Not Enough Precision to Constrain Theoretical Models: Base on the uncertainties of our data, we find that the theoretical models using quark coalescence as hadronization model, for instance, the TAMU, Cao, Sun, Ko, and models, are all in reasonable with the PbPb data, both in terms of central values and the decreasing trends as p_T increases.

Missing B-meson Measurement in the pp: Currently, the B-meson pp analysis is still ongoing. More results will be coming in the future to answer the questions such as the beauty energy loss mechanism in the QGP, and hadornization mechanism in small systems, constrain heavy quark diffusion coefficient, and the jet transport parameter to probe the inner workings of the QGP.

In conclusion, the larger PbPb data sets that should be accumulated in upcoming LHC Run 3 and high-luminosity(HL) LHC heavy ion runs will provide greater precision and allow more differential B-meson measurement not only about traditional observables with but also modern observables such as $B-\bar{B}$ angular correlations with more fully reconstructed b-hadron species such as Λ_b, B_c^0 and Ω_b. In addition, the MTD upgrade in CMS will allow us to perform hadronic PID. We will be able fully reconstruct b-hadron down very low p_T and carry out measurements with high precision. These future beauty measurements could help to further characterize beauty hadronization in vacuum and QGP.

6.4 Future Outlooks

As mentioned previously, our pp data analysis is still working in progress. Figure 6-2, Figure 6-3, and Figure 6-4 show our ongoing analysis of fully reconstructed B_s^0, B^+, and B^0 at very low p_T respectfully.

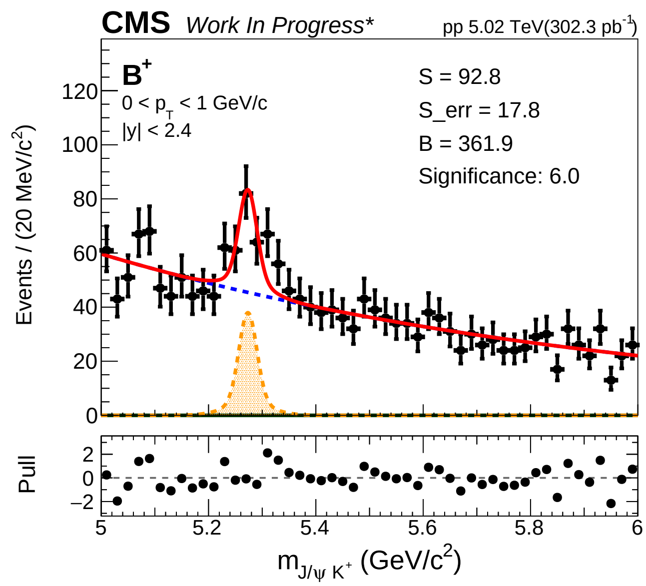


Figure 6-2: The fully reconstructed B^+ via the decay channel of $B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+$ in the p_T range of 0 - 1 GeV/c using the full CMS 2017 pp dataset is shown above. The statistical significance is about 6. The selection is optimized with BDT algorithm with a subset of topological variables used in PbPb B^+ studies.

Thanks to the powerful machine learning algorithms, even without hadron PID, very clear B-meson signals have still been observed down to $p_T = 0$. The estimated significance are all greater 4. With these signals, we can perform precise measurement on B^+ cross section in pp collisions down to 0 p_T, which allows us to perform measurement of inclusive beauty production cross section. In addition, we will also be able to measure B_s^0/B^+ ratio down to 3 GeV/c. Also, according to the multiplicity studies, we can also B_s^0/B^+ as a function of multiplicity up to 150, which also help us answer questions raised in Chapter 2.

In the era of LHC Run 3 and HL-LHC, much more statistics will be collected to

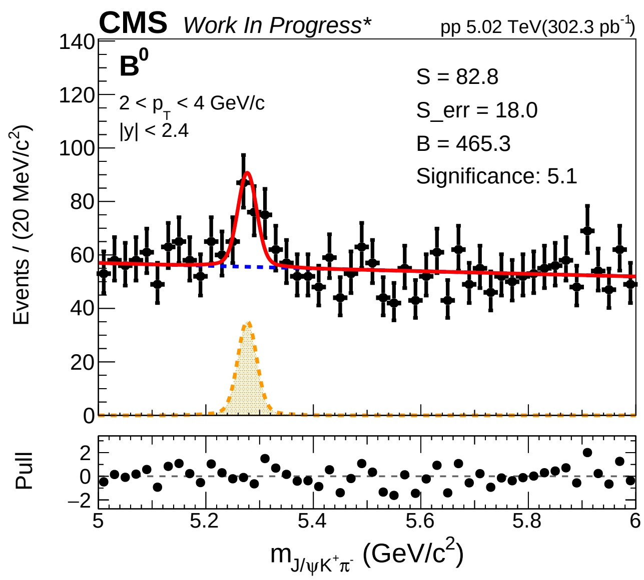


Figure 6-3: The fully reconstructed B^0 via the decay channel of $B^0 \rightarrow J/\psi K^{0*} \rightarrow \mu^+\mu^- K\pi$ in the p_T range of 2 - 4 GeV/c using the full CMS 2017 pp dataset is shown above. The statistical significance is about 5.1. The selection is optimized with BDT algorithm with a subset of topological variables used in PbPb B_s^0 studies.

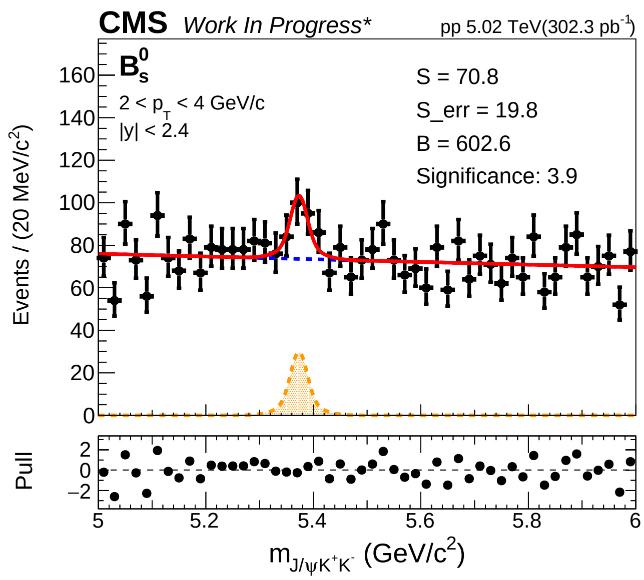


Figure 6-4: The fully reconstructed B_s^0 via the decay channel of $B_s^0 \rightarrow J/\psi\phi \rightarrow \mu^+\mu^-K^+K^-$ in the p_T range of 2 - 4 GeV/c using the full CMS 2017 pp dataset is shown above. The statistical significance is about 3.9. The selection is optimized with BDT algorithm with a subset of topological variables used in PbPb B_s^0 studies.

perform measurement on fully reconstructed B_c^+ and Λ_b. With the potential update of CMS MTD, we can good performance of PID for particle momentum from to GeV/c, which allow us to measure down to very low p_T.

Finally, at RHIC, with the sPHENIX experiment taking data in 2023, we can also fully reconstruct b-hadrons at RHIC energy to study a QGP medium at a lower temperature and higher baryon chemical potential. The beauty measurement at RHIC will be complementary to the measurements at the LHC. Together, these will help determine heavy quark diffusion coefficient with different temperature, constrain the fundamental property of QGP η/s, and probe the inner workings of QGP. To fully explore open heavy flavor physics and understand the inner workings of QGP, lots of challenges and opportunities are awaiting. A bright and exciting future of heavy flavor physics in heavy-ion collision will be forthcoming.

References

[1] S. Weinberg, "A Theory of Leptons", Phys. Rev. Lett. 19 1264 - 1266 (1967)

[2] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, "The Standard Model of Particle Physics", Rev. Mod. Phys. 71 (1999)

[3] C. D. Roberts, "Nonperturbative effects in QCD at Finite Temperature and Density", Phys. Part. Nucl. 30 (1999)

[4] P.A. Zyla et al. (Particle Data Group), "Review of Particle Physics", Prog. Theor. Exp. Phys. 2020, 083 C01 (2020)

[5] J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories", Phys. Rev. Lett. 30, 1343 (1973)

[6] S. Dürr et al. "Ab Initio Determination of Light Hadron Masses", Science. 322 (5905): 1224 7 (2008)

[7] N. Fettes, U.-G. Meißner, and S. Steininger, "Pion-nucleon scattering in chiral perturbation theory I: Isospin-symmetric case", Nucl. Phys. A 640 (1998)

[8] P. Navrtil and E. Caurier, "Nuclear structure with accurate chiral perturbation theory nucleon nucleon potential: Application to ${}^{6}_{3}Li$ and ${}^{10}_{5}B$", Phys. Rev. C 69, 014311 (2004)

[9] C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou, and A. V. Avilés-Casco, "Nucleon spin structure from lattice QCD", PoS DIS2018, 148 (2018)

[10] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization of Hard Processes in QCD", Adv. Ser. Direct. High Energy Phys. 5 (1989)

[11] D. E. Soper, "Parton distribution functions", Nucl. Phys. B Proc. Suppl. 53, 69-80 (1997)

[12] A. Metz and A. Vossen, "Parton Fragmentation Functions", Prog. Part. Nucl. Phys. 91, 136-202 (2016)

[13] J. J. Ethier and E. R. Nocera, "Parton Distributions in Nucleons and Nuclei", Ann. Rev. Nucl. Part. Sci. 70, 43-76 (2020)

[14] Francesco Becattini, "What is the meaning of the statistical hadronization model?", J. Phys. Conf. Ser. 5 (2005)

[15] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics", Phys. Rep. 97 (1983)

[16] B. R. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference", Nucl. Phys. B 238 492-528 (1984)

[17] R. J. Fries, V. Greco, and P. Sorensen "Coalescence Models For Hadron Formation From Quark Gluon Plasma", Ann. Rev. Nucl. Part. Sci. 58 (2008)

[18] F. Wilczek, "QCD In Extreme Conditions", Contribution to: 9th CRM Summer School: Theoretical Physics at the End of the 20th Century, 567-636

[19] E. d'Enterria, David G., et al., "CMS physics technical design report: Addendum on high density QCD with heavy ions", J. Phys.G 34 (2007)

[20] E. Altman, "Many-body localization and quantum thermalization", Nat. Phys. 14, 979 - 983 (2018).

[21] M. P. Heller, R. A. Janik, and P. Witaszczyk, "'Characteristics of Thermalization of Boost-Invariant Plasma from Holography", Phys. Rev. Lett. 108, 201602 (2012)

[22] G. Parisi, “Some considerations on the Quark-Gluon Plasma”, Quark Matter 2018 Conference (2018)

[23]

[24] H.C. Chandola, G. Punetha, and H. Dehnen, “Dual QCD thermodynamics and quark-gluon plasma”, Nucl. Phys. A 945 (2016)

[25] R. Stock, “Relativistic Nucleus-Nucleus Collisions and the QCD Matter Phase Diagram”, In *Landolt-Boernstein I 21A: Elementary particles* 7

[26] T.D. Lee and G.C. Wick, “Vacuum stability and vacuum excitation in a spin-0 field theory”, Phys. Rev. D9 2291(1974)

[27] J.O. Andersen and T. Brauner, “Linear sigma model at finite density in the 1/N expansion to next-to-leading order”, Phys .Rev. D 78:014030 (2008)

[28] M. Asakawa amd K. Yazaki, “Chiral Restoration at Finite Density and Temperature", Nucl. Phys. A 504 (1989)

[29] K. Fukushima, D.E. Kharzeev, and H.J. Warringa, “The Chiral Magnetic Effect”, Phys. Rev. D 78 074033 (2008)

[30] S. Shi, H. Zhang, D. Hou, and J. Liao, “Signatures of Chiral Magnetic Effect in the Collisions of Isobars”, Phys. Rev. Lett. 125 (2020)

[31] J. Zhao and F-Q. Wang, “Experimental searches for the chiral magnetic effect in heavy-ion collisions”, Prog. Part. Nucl. Phys.107 (2019)

[32] D.E. Kharzeev, J. Liao, S. A. Voloshin, and G. Wang, “Chiral Magnetic and Vortical Effects in High-Energy Nuclear Collisions — A Status Report”, Prog. Part. Nucl. Phys. 88 (2016)

[33] S. Choudhury, G. Wang, W. He, Y. Hu, and H.Z. Huang, “Background evaluations for the chiral magnetic effect with normalized correlators using a multiphase transport model”, Eur. Phys. J. C 80 (2020)

[34] H. S. Chung, J. Lee, and D. Kang, “Cornell potential parameters for S-wave heavy quarkonia”, J. Korean Phys. Soc. 52 (2018)

[35]

[36] J. Harris and B. Muller, “The Search for the quark-gluon plasma”, Ann. Rev. Nucl. Part. Sci. 46 (1996)

[37] A. Dumitru, Y. Guo, A. Mócsy, and M. Strickland, “Quarkonium states in an anisotropic QCD plasma”, Phys. Rev. D 79 (2009)

[38] R. Hagedorn, “Statistical thermodynamics of strong interactions at high energies”, Nuovo Cim. , Suppl. 3 (1965)

[39] J. Rafelski, “Melting Hadrons, Boiling Quarks”, from Hagedorn Temperature to Ultra-Relativistic Heavy-Ion Collisions at CERN. Springer, Cham.

[40] C.A. Dominguez, “Color Deconfinement in QCD at Finite Temperature”, Nucl. Phys. B Proc. Suppl.15 (1990)

[41] S.A. Chin, “Transition to Hot Quark Matter in Relativistic Heavy Ion Collision”, Phys. Lett. B 78, 552-555 (1978)

[42] F. Karsch, “Lattice QCD at High Temperature and Density”, Lect. Notes Phys. 583, 209-249 (2002)

[43] P. Braun-Munzinger and J. Wambach, “The Phase Diagram of Strongly-Interacting Matter”, Rev. Mod. Phys. Vol 81, 1031-1050 (2009)

[44] S. Gupta, X. Luo, B. Mohanty, H. G. Ritter, N. Xu, “Scale for the Phase Diagram of Quantum Chromodynamics”, Science 332 (2011)

[45] U. Heinz, “The Little Bang: Searching for quark-gluon matter in relativistic heavy-ion collisions”, Nucl. Phys. A 685, 414-431, 2001

[46] B.Alver and G.Roland, “Collision geometry fluctuations and triangular flow in heavy-ion collisions”, Phys. Rev. C 81, 054905 (2010)

[47] Jamal. Jalilian-Marian and X.N. Wang, “Shadowing of gluons in perturbative QCD: A comparison of different models”, Phys. Rev. D63, 096001 (2001)

[48] European Muon Collaboration, “The ratio of the nucleon structure functions F_2^N for iron and deuterium”, Phys. Lett. B 123, 275-278 (1983)

[49] E. Wang and X.-N. Wang, “Jet Tomography of Hot and Cold Nuclear Matter”, Phys. Rev. Lett. 89, 162301 (2002)

[50] R.X. Xu, “Strange quark stars - A Review”, IAU Symp. 214 (2003)

[51] Y.-Z. Fan, Y.-W. Yu, D. Xu, Z.-P. Jin, X.-F. Wu, D.-M. Wei, and B. Zhang, “A supra-massive magnetar central engine for short GRB 130603B”, Astrophys. J. Lett. 779 (2013)

[52] Z. G. Dai, S. Q. Wang, J. S. Wang, L. J. Wang, and Y. W. Yu, “The Most Luminous Supernova ASASSN-15lh: Signature of a Newborn Rapidly-Rotating Strange Quark Star”, Astrophys. J. 817 (2016)

[53]

[54] D. Trbojevic and S. Peggs, “Required Accuracy of the RHIC Circumference”, United States: N. p., Web. doi:10.2172/1119398 (1993)

[55] M. J. Rhoades-Brown, “The Heavy Ion Injection Scheme for RHIC”, Proc. of the Workshop on the RHIC Performance (1988)

[56] D. B. Steski, J. Alessi, J. Benjamin, C. Carlson, M. Manni, P. Thieberger, and M. Wiplich, “Operation of the Relativistic Heavy Ion Collider Au^- ion source”, Review of Scientific Instruments 73, 797 (2002)

[57] D.B. Steski and P. Thieberger, “Stripping foils at RHIC”, Nucl. Instrum. Meth. A 613 (2010)

[58] P. Thieberger, L. Ahrens, J. Alessi, J. Benjamin, M. Blaskiewicz, J. M. Brennan, K. Brown, C. Carlson, C. Gardner, W. Fischer, D. Gassner, J. Glenn, W.

Mac Kay, G. Marr, T. Roser, K. Smith, L. Snydstrup, D. Steski, D. Trbojevic, N. Tsoupas, V. Zajic, and K. Zeno, “Improved gold ion stripping at 0.1 and 10 GeV/nucleon for the Relativistic Heavy Ion Collider”, Phys. Rev. ST Accel. Beams 11, 011001 (2008)

[59] L. Evans, “The Large Hadron Collider”, Phil. Trans. R. Soc. A 370 (2012)

[60] J. Brewer, A. Mazeliauskas, and W. van der Schee, “Opportunities of OO and pO collisions at the LHC”, CERN Theory Report: CERN-TH-2021-028 (2021)

[61] M. Schaumann, R. Alemany-Fernandez, H. Bartosik, T. Bohl, R. Bruce, G-H Hemelsoet, S. Hirlaender, J. Jowett, V. Kain, M. Krasny, J. Molson, G. Papotti, M.S. Camillocci, H. Timko, and J. Wenninger, “First partially stripped ions in the LHC ($^{208}Pb^{81+}$)” J. Phys. Conf. Ser. 1350, 012071 (2019)

[62] C.Y. Wong, “Introduction to high-energy heavy ion collisions”, Singapore, Singapore: World Scientific (1994) 516 p

[63] Z.-T. Liang and X.-N. Wang , “Globally Polarized Quark-Gluon Plasma in Non-central A + A Collisions”, Phys.Rev.Lett. 96, 039901 (2006)

[64] B.Alver and G.Roland, “Collision geometry fluctuations and triangular flow in heavy-ion collisions”, Phys. Rev. C 81, 054905 (2010)

[65] M. L. Miller, K. Reygers, S. J. Sanders and P. Steinberg, “Glauber modeling in high energy nuclear collisions”, Ann. Rev. Nucl. Part. Sci. 57, 205 (2007)

[66] I. Altsybeev and V. Kovalenko, “Classifiers for centrality determination in proton-nucleus and nucleus-nucleus collisions”, EPJ Web Conf. 137, 11001

[67] P. Cortese, “Performance of the ALICE Zero Degree Calorimeters and upgrade strategy”, J. Phys. Conf. Ser. 1162, 012006 (2019)

[68] Oliver Suranyi, “Study of Very Forward Neutrons with the CMS Zero Degree Calorimeter”, Universe 5 10, 210 (2019)

[69] P. Dmitrieva and I. Pshenichnov, “On the performance of Zero Degree Calorimeters in detecting multinucleon events”, Nucl. Instrum. Meth. A 906 (2018)

[70] Star Collaboration, `https://www.star.bnl.gov/public/tpc/tpc.html`

[71] R. J. Glauber, “Quantum Optics and Heavy Ion Physics”, Nucl. Phys. A 774 (2006)

[72] J. Chauvin, D. Bebrun, A. Lounis, and M. Buenerd, “Low and intermediate energy nucleus-nucleus elastic scattering and the optical limit of Glauber theory”, Phys. Rev. C. 28, 1970 (1983)

[73] T. Wibig and D. Sobczynska, “Proton-nucleus cross section at high energies”, J. Phys. G: Nucl. Part. Phys. 24, 2037 (1998)

[74] B. B. Back, “Studies of multiplicity in relativistic heavy-ion collisions”, J.Phys.Conf.Ser. 5 (2000)

[75] A. Milov, “Electroweak probes with ATLAS”, PoS High-pT2017 016 (2019)

[76] L. Van Hove, “Theoretical prediction of a new state of matter, the "quark-gluon plasma" (also called "quark matter")”, Part of Multipartcle Dynamics. Proceedings, 17th International Symposium, Seewinkel, Austria, June 16-20, 801-818 (1986)

[77] S. A. Bass, M. Gyulassy, H. Stoecker, and W. Greiner, “Signatures of Quark-Gluon-Plasma formation in high energy heavy-ion collisions: A critical review”, J. Phys. G 25 R1-R57 (1999)

[78] C. Quigg and J. L. Rosner, “Quantum Mechanics with Applications to Quarkonium”, Phys. Rept. 56 167-235 (1979)

[79] P. Petreczky, “Quarkonium in Hot Medium”, J. Phys. G 37, 094009 (2010)

[80] G. S. Bali, H. Neff, T. Duessel, T. Lippert, K. Schilling, “Observation of string breaking in QCD”, Phys. Rev. D 71, 114513 (2005)

[81] P. Petreczky, “Quarkonium in Hot Medium”, J. Phys. G 37, 094009 (2010)

[82] STAR Collaboration, “Measurement of inclusive J/ψ suppression in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV through the dimuon channel at STAR”, Phys. Lett. B 797, 134917 (2019)

[83] A. Andronic et. al., “Heavy-flavour and quarkonium production in the LHC era: from proton-proton to heavy-ion collisions”, Eur.Phys.J.C 76, 107 (2016)

[84] STAR Collaboration, “Suppression of Υ production in d+Au and Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV”, Phys. Lett. B 735, 127-137 (2014)

[85] CMS Collaboration, “Suppression of $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ production in PbPb collisions at $\sqrt{s_{NN}} = 200$ GeV”, Phys. Lett. B 770 357-379 (2017)

[86] ZEUS Collaboration, “Forward jet production in deep inelastic ep scattering and low-x parton dynamics at HERA”, Phys. Lett. B 632 13-26 (2006)

[87] R. A. Gerwin, “Energy loss of a relativistic electron beam in a plasma”, the Physics of Fluids 18, 614 (1975)

[88] STAR Collaboration, “Disappearance of back-to-back high p_T hadron correlations in central Au+Au collisions at ’$\sqrt{s_{NN}} = 200$ GeV’, Phys.Rev.Lett. 90, 082302 (2003)

[89] ALICE Collaboration, “Measurements of inclusive jet spectra in pp and central Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, Phys. Rev. C 101, 034911 (2020)

[90] C. McGinn, “Mapping the redistribution of jet energy in PbPb collisions at the LHC with CMS”, MIT PHD Thesis (2019)

[91] P. Bozek and I. Wyskiel, “Directed flow in ultrarelativistic heavy-ion collisions”, Phys. Rev. C 81, 054902 (2010)

[92] CMS Collaboration, “Constraints on the chiral magnetic effect using charge-dependent azimuthal correlations in pPb and PbPb collisions at the CERN Large Hadron Collider”, Phys. Rev. C 97, 044912 (2018)

[93] A. M. Poskanzer and S.A. Voloshin, "Methods for analyzing anisotropic flow in relativistic nuclear collisions", Phys.Rev.C 58 1671-1678 (1998)

[94] R. S. Bhalerao, J.-Y. Ollitrault,"Eccentricity fluctuations and elliptic flow at RHIC", Phys. Lett. B 641, 260-264 (2006)

[95] P. F. Kolb and U. Heinz, "Hydrodynamic description of ultra relativistic heavy-ion collisions", Part of Quark-gluon plasma 4, 634-714 (2003)

[96] STAR Collaboration, "Elliptic flow in Au+Au collisions at $\sqrt{s_{NN}} = 130$ GeV", Phys. Rev. Lett. 86, 402-407 (2001)

[97] ALICE Collaboration, "Elliptic flow of charged particles in Pb-Pb collisions at 2.76 TeV", Phys. Rev. Lett. 105, 252302 (2010)

[98]

[99] J. Rafelski and B. Muller, "Strangeness Production in the Quark-Gluon Plasma", Phys. Rev. Lett. 48, 1066 (1982)

[100] STAR Collaboration, "Measurements of ϕ meson production in relativistic heavy-ion collisions at RHIC", Phys. Rev. C 79, 064903 (2009)

[101] ALICE Collaboration, "Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions", Nature Phys. 13, 535-539 (2017)

[102] U. W. Heinz and M. Jacob, "Evidence for a new state of matter: An Assessment of the results from the CERN lead beam program", CERN Special Seminar Report, (2000)

[103] "Quark Gluon Plasma an Color Glass Condensate at RHIC? The perspective from the BRAHMS experiment" , Nucl. Phys. A 757, 1-27 (2005)

[104] PHOBOS Collaboration, "The PHOBOS Perspective on Discoveries at RHIC", Nucl. Phys. A 757, 28-101 (2005)

[105] STAR Collaboration, “Experimental and theoretical challenges in the search for the quark?gluon plasma: The STAR Collaboration’s critical assessment of the evidence from RHIC collisions”, Nucl. Phys. A 757, 102-183 (2005)

[106] PHENIX Collaboration, “Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration”, Nucl. Phys. A 757, 184-283 (2005)

[107] J. Rafelski, “Discovery of Quark-Gluon-Plasma: Strangeness Diaries”, Eur. Phys. J. ST 229, 1-140 (2020)

[108] C. Markert, R. Bellwied, and I. Vitev, “Formation and decay of hadronic resonances in the QGP”, Phys. Lett. B 669, 92-97 (2008)

[109] T. Kodama, “Hunt for the quark-gluon plasma: 20 years later”, Braz. J. Phys. 34, 205-210 (2004)

[110] A. Kurkela and A. Mazeliauskas, “Kinetic and Chemical Equilibration of Quark-Gluon Plasma”, Springer Proc. Phys. 250, 177-181 (2020)

[111] J. L. Nagle, “The Letter S (and the sQGP)”, Eur. Phys. J. C 49, 275-279 (2007)

[112] S.M. Sanches, F.S. Navarra, and D.A. Fogaca, “The quark gluon plasma equation of state and the expansion of the early Universe”, Nucl. Phys. A 937, 1-16 (2015)

[113] U. Heinz, C. Shen, and H. Song, “The viscosity of quark-gluon plasma at RHIC and the LHC”, AIP Conf. Proc. 1441, 766-770 (2012)

[114] G. Policastro, D.T. Son, and A.O. Starinets, “Shear viscosity of strongly coupled N=4 supersymmetric Yang-Mills plasma”, Phys. Rev. Lett. 87, 081601 (2001)

[115] J. Bjorken, “Highly Relativistic Nucleus-Nucleus Collisions: The Central Rapidity Region”, Phys.Rev.D 27 140-151 (1983)

[116] B. V. Jacak and B. Müller, “The Exploration of Hot Nuclear Matter”, Science 337, 310-314 (2012)

[117] A. Adil and M. Gyulassy, “Energy systematics of jet tomography at RHIC: $\sqrt{s_{NN}} = 62.4$ vs. 200 AGeV”,

[118] Heavy Ion Collisions: The Big Picture, and the Big Questions, “Heavy Ion Collisions: The Big Picture, and the Big Questions”, Ann. Rev. Nucl. Part. Sci. 68, 339-376 (2018)

[119] E. Rutherford, “The Scattering of α and β Particles by Matter and the Structure of the Atom”, Philos. Mag, 6, 21 (1911)

[120] H. W. Kandall, “Deep inelastic scattering: Experiments on the proton and the observation of scaling”, Rev. Mod. Phys. 63, 597-614 (1991)

[121] R. E. Taylor, “Deep inelastic scattering: The Early years”, Rev. Mod. Phys. 63, 573-595 (1991)

[122] J. I. Friedman, “Deep inelastic scattering: Comparisons with the quark model”, Rev. Mod. Phys. 63, 615-629 (1991)

[123]

[124] C. Lourenco and H. Satz, “Proceedings, 1st International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions (Hard Probes 2004) : Ericeira, Portugal, November 4-10, 2004”, Eur. Phys. J. C 43 1-4 (2005)

[125] B. Betz and M. Gyulassy, “Constraints on the Path-Length Dependence of Jet Quenching in Nuclear Collisions at RHIC and LHC”, JHEP 08, 090 (2014)

[126] CMS Collaboration, “Event displays and some infographics of jets in heavy ion collisions”, CMS-PHO-EVENTS-2021-007, `http://cds.cern.ch/record/2757389`

[127] J. Casalderrey-Solana, G. Milhano, D. Pablos, and K. Rajagopal, “Modification of Jet Substructure in Heavy Ion Collisions as a Probe of the Resolution Length of Quark-Gluon Plasma”, JHEP 01, 044 (2020)

[128] ALICE Collaboration, “Exploring jet substructure with jet shapes in ALICE”, Nucl. Phys. A 967 528-531(2017)

[129] CMS Collaboration, “Measurement of isolated photon production in pp and PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, Phys. Lett. B 710, 256 (2012)

[130] CMS Collaboration, “Study of Z production in PbPb and pp collisions at $\sqrt{s_{NN}}$ $= 2.76$ TeV in the dimuon and dielectron decay channels”, JHEP 03, 022 (2015)

[131] CMS Collaboration, “Study of Jet Quenching with Z + jet Correlations in Pb-Pb and pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, Phys. Rev. Lett. 119, 082301 (2017)

[132]

[133] S.J. Brodsky, H.J. Pirner, and J. Raufeisen “Scaling properties of high p_T inclusive hadron production”, Phys. Lett. B 637, 58-63 (2006)

[134] LHCb Collaboration, “Measurement of f_s/f_u variation with proton-proton collision energy and B-meson kinematics”, Phys. Rev. Lett. 124, 122002 (2020)

[135] ALICE Collaboration “Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC”, CERN-EP-2021-088

[136] X. Dong, Y.-J. Lee, and R. Ralf, “Open Heavy-Flavor Production in Heavy-Ion Collisions”, Ann. Rev. Nucl. Part. Sci. 69, 417-445 (2019)

[137] Y. Liu, C. M. Ko, and F. Li, “Heavy quark correlations and the effective volume for quarkonia production”, Phys. Rev. C 93, 034901 (2016)

[138] F. Prino and R. Rapp, “Open Heavy Flavor in QCD Matter and in Nuclear Collisions”, J. Phys. G 43, 093002 (2016)

[139] A. M. Adare, M. P. McCumber, J. L. Nagle, and P. Romatschke, “Tests of the Quark-Gluon Plasma Coupling Strength at Early Times with Heavy Quarks”, Phys. Rev. C 90, 024911 (2014)

[140] Y. Akiba, “Quest for the quark-gluon plasma - hard and electromagnetic probes” Prog. Theor. Exp. Phys. 2015, 03A105 (2015).

[141] K. M. Burke et al. (JET Collaboration), “Extracting the jet transport coefficient from jet quenching in high-energy heavy-ion collisions”, Phys. Rev. C 90, 014909 (2014)

[142] J. Casalderrey-Solana, D. C. Gulhan, J. G. Milhano, D. Pablos, and K. Rajagopal, “Predictions for boson-jet observables and fragmentation function ratios from a hybrid strong/weak coupling model for jet quenching”, JHEP 053 (2016)

[143] S. S. Gubser, “Drag force in AdS/CFT”, Phys. Rev. D 74, 126005 (2006)

[144] A. Ficnar, J. Noronha, and M. Gyulassy, “Non-conformal Holography of Heavy Quark Quenching”, Nucl. Phys. A 855 (2011)

[145] H. Bichsel and H. Schindler, “The Interaction of Radiation with Matter”, In: Fabjan C., Schopper H. (eds) Particle Physics Reference Library. Springer, Cham. (2020)

[146] Y. L. Dokshitzer and D. E. Kharzeev, “Heavy quark colorimetry of QCD matter”, Phys. Lett. B 519, 199 (2001)

[147] JET Collaboration, “Extracting the jet transport coefficient from jet quenching in high-energy heavy-ion collisions”, Phys. Rev. C 90, 014909 (2014)

[148] S.-Q.Li, W.-J. Xing, F.-L. Liu, S. Cao, and C.-Y. Qin, “Heavy flavor quenching and flow: the roles of initial condition, pre-equilibrium evolution, and in-medium interaction”, Chin. Phys. C 44, 114101 (2020)

[149] M. He, R. J. Fries, and R. Rapp, “Heavy flavor at the large hadron collider in a strong coupling approach", Phys. Lett. B 735, 445 - 450 (2014)

[150] F. Riek and R. Rapp, “Quarkonia and heavy-quark relaxation times in the quark-gluon plasma”, Phys. Rev. C 82, 035201 (2010).

[151] Min He, Rainer J. Fries, and Ralf Rapp, “Heavy-quark diffusion and hadronization in quark-gluon plasma”, Phys. Rev. C 86, 014903 (2012)

[152] S. Cao et. al., “Charmed hadron chemistry in relativistic heavy-ion collisions", Phys. Lett. B 807, 135561 (2020)

[153] J. Song, H-H Li, F. Shao, “New feature of low p_T charm quark hadronization in pp collisions at $\sqrt{s}$ = 7 TeV”, Eur. Phys. J. C 78 344 (2018)

[154] S. Cao, G.-Y. Qin, and S. A. Bass, Phys. Rev. C 88, 044907 (2013)

[155] S. Cao, G.-Y. Qin, and S. A. Bass, Phys. Rev. C 92, 024907 (2015)

[156] Min He and Ralf Rapp, “Hadronization and Charm-Hadron Ratios in Heavy-Ion Collisions”, Phys. Rev. Lett. 124, 042301 (2020)

[157] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, JHEP 0605, 026 (2006)

[158] STAR Collaboration, “Measurement of D^0 Azimuthal Anisotropy at Midrapidity in Au + Au Collisions at $\sqrt{s}_{NN}$ = 200 GeV”, Phys. Rev. Lett. 118, 212301 (2017)

[159] CMS Collaboration, “Elliptic Flow of Charm and Strange Hadrons in High-Multiplicity $p+Pb$ Collisions at $\sqrt{s}_{NN} = 8.16$ TeV”, Phys. Rev. Lett. 121, 082301 (2018)

[160] ALICE Collaboration, “Elliptic Flow of Electrons from Beauty-Hadron Decays in Pb-Pb Collisions at $\sqrt{s}_{NN} = 5.02$ TeV”, Phys. Rev. Lett. 126, 162001 (2021)

[161] ATLAS Collaboration, “Measurement of azimuthal anisotropy of muons from charm and bottom hadrons in Pb+Pb collisions at $\sqrt{s}_{NN} = 5.02$ TeV with the ATLAS detector”, Phys. Lett. B 807, 135595 (2020)

[162] ALICE Collaboration, “D-meson production in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and in pp collisions at $\sqrt{s}$ =7 TeV”, Phys. Rev. C 94, 054908 (2016)

[163] CMS Collaboration, “Study of B Meson Production in p + Pb Collisions at $\sqrt{s_{NN}}$ = 5.02 TeV Using Exclusive Hadronic Decays”, Phys. Rev. Lett. 116 032301 (2016)

[164] STAR Collaboration, “Observation of D^0 Meson Nuclear Modifications in Au+Au Collisions at $\sqrt{s}_{NN}$ = 200 GeV”, Phys. Rev. Lett. 113, 142301 (2014)

[165] CMS Collaboration, “Nuclear modification factor of D^0 mesons in PbPb collisions at $\sqrt{s}_{NN}$ = 5.02 TeV”, Phys. Lett. B 782, 474-496 (2018)

[166] CMS Collaboration, “Measurement of B^0_s meson production in pp and PbPb collisions at $\sqrt{s}_{NN}$ = 5.02 TeV”, Phys. Lett. B 796 168-190 (2019)

[167] I. Kuznetsova and J. Rafelski, “Heavy flavor hadrons in statistical hadronization of strangeness-rich QGP”, Eur. Phys. J. C 51, 113-133 (2007)

[168] Y. Oh, C. M. Ko, S. H. Lee, and S. Yasui, “Heavy baryon/meson ratios in relativistic heavy-ion collisions”, Phys. Rev. C 79, 044905 (2009)

[169] STAR Collaboration, “First measurement of Λ_C baryon production in AuAu collisions at $\sqrt{s_{NN}}$ = 200 GeV”, Phys. Rev. Lett. 124, 172301 (2020)

[170] ALICE Collaboration, “\$$\Lambda^+_C$ production in Pb - Pb collisions at $\sqrt{s}_{NN}$ = 5.02 TeV”, Phys.Lett.B 793, 212-223 (2019)

[171] ALICE QM Conference proceedings

[172] LHCb Collaboration, “Measurement of B^+, B^0 and Λ^0_b production in pPb collisions at $\sqrt{s_{NN}}$ = 8.16 TeV”, Phys. Rev. D 99, 052011 (2019)

[173] Z. Tang, L. Yi, L. Ruan, M. Shao, H. Chen, C. Li, B. Mohanty, P. Sorensen, A. Tang, and Z. Xu, “Statistical Origin of Constituent-Quark Scaling in the QGP hadronization”, Chin. Phys. Lett. 30 031201(2013)

[174] CMS Collaboration, “Studies of charm quark diffusion inside jets using PbPb and pp collisions at $\sqrt{s}_{NN}$ = 5.02 TeV”, Phys. Rev. Lett. 125, 102001 (2020)

[175] F. Colamaria, "Study of azimuthal correlations between D mesons and charged particles with the ALICE experiment", J. Phys. Conf. Ser. 612, 012024 (2015)

[176] STAR Collaboration, "Measurement of D^0 meson + hadron two-dimensional angular correlations in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV", Phys. Rev. C 102, 014905 (2020)

[177] LHCb Collaboration, "

[178] H, van H. and R. Rapp, "Thermalization of heavy quarks in the quark-gluon plasma", Phys. Rev. C 71, 034907 (2005)

[179] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST 3 S08004 (2008).

[180] CMS Collaboration, "A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider", Science 338, 1569-1575 (2012)

[181] CMS Collaboration, "The CMS trigger system", JINST 12 01, P01020 (2017)

[182] Y. Chao, "Minimum-Bias and Underlying Event Studies at CMS", Proceedings, 28th International Conference on Physics in Collision (PIC 2008) : Perugia, Italy, June 25-28, (2008)

[183] Zhaozhong Shi, "Measurement of B_s^0 and B^+ meson yields in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV", CMS-PAS-HIN-19-011

[184] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker", CERN-LHCC-2017-009

[185] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", JINST 9 10, P10009 (2014)

[186] P. Billoir, "Progressive track recognition with a Kalman like fitting procedure", Comput. Phys. Commun. 57, 390 (1989)

[187] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method", Nucl. Instrum. Meth. A 294, 219 (1990)

[188] R. Mankel, "A Concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system", Nucl. Instrum. Meth. A 395, 169 (1997)

[189] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", Nucl. Instrum. Meth. A 262, 444 (1987)

[190] K. Rose, "Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems", Proceedings of the IEEE 86 (1998)

[191] CMS Collaboration, "The CMS ECAL performance with examples", JINST 9 C02008 (2014)

[192] CMS Collaboration, "The CMS hadron calorimeter project : Technical Design Report", CERN-LHCC-97-031 (1997)

[193] A. Penzo and Y. Onel, "The CMS-HF quartz fiber calorimeters", J.Phys.Conf.Ser. 160, 012014 (2009)

[194] CMS Collaboration, "Observation and studies of jet quenching in PbPb collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV", Phys. Rev. C 84, 024906 (2011)

[195] B. Roland, "Forward Physics Capabilities of CMS with the CASTOR and ZDC detectors", Part of Proceedings, 17th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2019), Madrid, Spain, April 26-30 (2009)

[196] T. Bawej el. at., "The New CMS DAQ System for Run-2 of the LHC", IEEE Trans. Nucl. Sci. 62, 1099-1103 (2015)

[197] CMS Collaboration, "Commissioning and performance of the CMS pixel tracker with cosmic ray muons", JINST 5 T03007 (2010)

[198] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC", J. Phys. G 31 N9 (2005)

[199] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", J. Phys. G 34 N343 (2007)

[200] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", in Hadron collider physics. Proceedings, 22nd Conference, HCP 2010, Toronto, Canada, August 23-27 (2010)

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

[202] E. Barberio, B. van Eijk, and Z.Was, "PHOTOS: A Universal Monte Carlo for QED radiative corrections in decays", Comput. Phys. Commun. 66 115 -128 (1991)

[203] I. P. Lokhtin and A. M. Snigirev, "A model of jet quenching in ultrarelativistic heavy ion collisions and high-pT hadron spectra at RHIC", Eur. Phys. J. C45 211(2006)

[204] CMS Collaboration, "Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.76 TeV", JHEP 02 041 (2010)

[205] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", JINST 7 P10002 (2012)

[206] CMS Collaboration, "Muon performance studies in 2017 pp and 2018 PbPb 5.02TeV data", CMS Analysis Note: AN-18-316 (2020)

[207]

[208]

[209] V. W. Berger and Y. Zhou, "Kolmogorov?Smirnov test: Overview", Wiley Statsref: Statistics Reference Online, (2014)

[210] M. Kearns and L. Valiant (1989). Crytographic limitations on learning Boolean formulae and finite automata. Symposium on Theory of Computing. 21. ACM. 433 - 444. (1989)

[211] J. Therhaag, "TMVA: Toolkit for multivariate data analysis", AIP Conf.Proc. 1504, 1013-1016 (2012)

[212] CMS Collaboration "Measurement of the B^+ hadron production cross section in pp collisions at 13 TeV", CMS Physics Analysis Summary, CMS-PAS-BPH-15-004 (2015)

[213] Wouter Verkerke and David P. Kirkby, "The RooFit toolkit for data modeling", eConf. C0303241, MOLT007 (2003)

[214] CMS Collaboration, "Upsilon production cross-section in pp collisions at $\sqrt{s}$ = 7 TeV", Phys. Rev. D 83, 112004 (2011)

[215] M. Tanabashi et al. (Particle Data Group), "Review of Particle Physics", Phys. Rev. D 98, 030001 (2018)

[216] CMS Collaboration, "Charged-particle nuclear modification factors in PbPb and pPb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV", JHEP 04, 039 (2017)

[217] ATLAS Collaboration, "Determination of the ratio of bb-quark fragmentation fractions f_s/f_d in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector", Phys. Rev. Lett. 115, 262001 (2015)