

RICE UNIVERSITY

A Systematic Measurement of  $\mu^+\mu^-$  Production in p+p  
and p+Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV with the  
STAR Detector

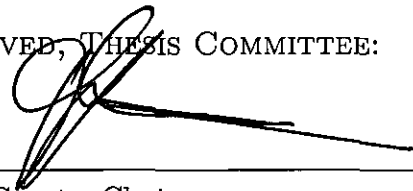
by

James Daniel Brandenburg

A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:




---

Frank Geurts, Chair  
Associate Professor of Physics and Astronomy



---

Wei Li  
Associate Professor of Physics and Astronomy



---

David Scott  
Noah Harding Professor of Statistics

Houston, Texas

August, 2018

## ABSTRACT

A Systematic Measurement of  $\mu^+\mu^-$  Production in p+p and p+Au Collisions at  
 $\sqrt{s_{NN}} = 200$  GeV with the STAR Detector

by

James Daniel Brandenburg

Most of the matter around us today is made up of protons and neutrons, but in the first few moments of the universe the temperature and density were too high for tightly bound protons and neutrons to form. Instead of being bound inside protons and neutrons, quarks and gluons existed in a plasma-like fluid called the Quark Gluon Plasma (QGP). As the universe cooled and expanded, the matter that we have today began to form. An understanding of nuclear matter and the transition from QGP to normal matter (and vice versa) can in principle be ascertained from the fundamental theory of the strong interaction, Quantum Chromodynamics (QCD). In practice though, the current state-of-the-art calculations provide only limited information about the properties of QCD matter. The transition from normal matter to QGP can be studied in the laboratory using relativistic heavy-ion collisions like those produced by the Relativistic Heavy-Ion Collider (RHIC) at Brookhaven National Laboratory.

Studying the QGP through heavy-ion collisions has its challenges though, since the created matter evolves through many stages before the final state particles can be detected. Learning about the earliest stages of the system requires penetrating probes, capable of carrying information from inside the medium out to the final state. Electromagnetic probes, such as leptons, are inert to the strong force. For this reason, they carry pristine information from all stages of the created medium. Dileptons ( $l^+l^-$ ) are even more valuable, since the various production mechanisms and time periods of the system can be distinguished through the invariant mass of the pair. For instance, the suppression in production of heavy quark

(charm and bottom) bound states, which can be identified through dileptons, has long been considered a direct probe of the QGP [1]. At lower masses dileptons can be used to measure the spectrum of thermal radiation of the medium, acting as a “fireball thermometer” [2]. Dileptons are also linked to the phenomena of spontaneous chiral symmetry breaking (and its expected restoration inside the QGP) through the  $\rho$ -meson which decays into dileptons [3].

In this thesis, the first measurements of the dilepton invariant mass spectra through the dimuon ( $\mu^+\mu^-$ ) channel with the Solenoidal Tracker at RHIC (STAR) are presented. The  $\mu^+\mu^-$  invariant mass spectra is measured in data from  $p + p$  collisions at  $\sqrt{s} = 200$  GeV and  $p+\text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV. The first measurement of the  $\phi \rightarrow \mu^+\mu^-$  spectra at STAR is also measured in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV. For these analyses novel muon identification techniques were developed to combat the contamination from hadrons and secondary muons resulting from weak decays. Techniques are presented for training and employing deep neural networks for the identification of muons and for the rejection of backgrounds. Data-driven techniques are presented for the measurement of muon-purity and for the estimation of physical backgrounds to the  $\mu^+\mu^-$  invariant mass spectra. The measurement of the  $\mu^+\mu^-$  invariant mass spectra in  $p+p$  and  $p+\text{Au}$  collisions is also compared with the expected dimuon yields from light hadron decays, open heavy flavor decays, and the Drell-Yan process. Finally, the potential for future dilepton measurements at STAR is discussed in light of the new datasets collected in the recent years.

## Acknowledgements

First and foremost I would like to thank my wife for supporting me throughout my time as a Ph.D. student. Without her support this would not have been possible for me. Her company has allowed me to enjoy every part of this even despite the stressful days.

I am deeply grateful to my advisor, Frank Geurts, who has guided me during my time as a student. His direction and support have been invaluable to me, helping me to succeed as a student at Rice. Frank always supported me and helped me to gain opportunities to speak and travel to conferences for which I am very appreciative. I believe that I have grown significantly as a scientist during my studies thanks to his guidance.

I would also like to thank Geary Eppley for allowing me to work with him and be a part of many projects with him. I greatly enjoyed the opportunities to work together and I always benefitted from his input. I want to thank Bill Llope as well. I benefitted greatly from discussion with Bill while he was at Rice and after. I am also grateful to Joey Butterworth, who helped me in countless ways when I was first starting at Rice and in STAR. I would also like to thank David Tlusty for always being willing to help me with my research through discussion and feedback. I have enjoyed the chance to work together. I am grateful to Wei Li as well. I have greatly benefited from his advice and from the chances to discuss physics with him.

Finally, I would like to thank my parents for their encouragement throughout the years. Thank you for always making the time to listen. Thank you especially to my father for encouraging my curiosity and for constantly teaching me.

# Contents

Abstract	ii
Acknowledgements	iv
List of Illustrations	viii
List of Tables	xxv
<b>1 Introduction</b>	<b>1</b>
1.1 The Standard Model	1
1.2 High Energy Nuclear Physics and QCD	4
1.2.1 Historical Origins	4
1.2.2 Quantum Chromodynamics	6
1.2.3 The Quark Gluon Plasma	10
1.2.4 Chiral Symmetry	13
1.3 Heavy-Ion Collisions	17
<b>2 Dileptons</b>	<b>21</b>
2.1 Dilepton Production in Heavy Ion Collisions	21
2.2 Survey of Past Dilepton Measurements	25
2.3 Motivations for $\mu^+\mu^-$ Measurements	38
<b>3 Experimental Apparatus</b>	<b>40</b>
3.1 Relativistic Heavy Ion Collider (RHIC)	40
3.2 Solenoidal Tracker at RHIC	42
3.2.1 Time Projection Chamber	44
3.2.2 Vertex Position Detector	47
3.2.3 Time-of-Flight Detector	50
3.2.4 Muon Telescope Detector	55
3.3 Data Sets and Event Selection	59

<b>4</b>	<b>Muon Identification</b>	<b>62</b>
4.1	Identifying Low Momentum Muons with the TPC and TOF . . . . .	62
4.2	Muon Identification Information from the MTD . . . . .	66
4.3	Preparation of Training Datasets . . . . .	67
4.3.1	Simulation of Muon Telescope Detector . . . . .	67
4.3.2	Extracting $\Delta$ TOF distributions from Data . . . . .	71
4.3.3	Background MC Closure Test Using Identified $K_S^0 \rightarrow \pi^+\pi^-$ and $\phi \rightarrow K^+K^-$ Decays . . . . .	73
4.4	Training and Evaluation of Artificial Neural Networks . . . . .	75
4.4.1	Introduction to Artificial Neural Networks and Machine Learning . .	75
4.4.2	Shallow Neural Networks . . . . .	75
4.4.3	Deep Neural Networks and Hyper-parameter Optimization . . . . .	77
4.4.4	Comparison of Multivariate Classifiers . . . . .	79
4.4.5	Identification of Muon Pairs . . . . .	80
4.4.6	Data-Driven Muon Purity Measurements . . . . .	83
<b>5</b>	<b>Dimuon Analysis Details</b>	<b>85</b>
5.1	Track Selection and Muon Identification . . . . .	85
5.2	Hadron Contamination and Muon Purity . . . . .	87
5.3	Identification of Muon Pairs . . . . .	90
5.4	Sources of Background . . . . .	93
5.4.1	Uncorrelated Combinatorial Background . . . . .	95
5.4.2	Physical Background Sources . . . . .	103
5.5	Signal Extraction . . . . .	107
5.5.1	Like-Sign Ratio Background Estimation Technique . . . . .	107
5.5.2	Correlation Weighted Event Mixing Technique . . . . .	112
5.5.3	Raw Signal Extraction . . . . .	115
5.6	Contributions from Hadronic Decays . . . . .	120
5.6.1	Decay Channels, Branching Ratios, and Kinematics . . . . .	120
5.6.2	Decays of Vector Mesons and Pseudo Scalar Mesons . . . . .	122
5.6.3	Contributions from $c\bar{c}$ , $b\bar{b}$ , and Drell-Yan . . . . .	125

5.6.4	Hadronic Cocktail for the MTD Acceptance . . . . .	127
5.7	Efficiency and Acceptance Corrections . . . . .	130
5.7.1	Trigger Correction . . . . .	130
5.7.2	Tracking and MTD Efficiency . . . . .	132
5.8	Systematic Uncertainties . . . . .	140
5.8.1	Uncertainties on the $\mu^+\mu^-$ Invariant Mass Spectra . . . . .	140
5.8.2	Uncertainty on the $\phi \rightarrow \mu^+ + \mu^-$ Yield . . . . .	144
5.8.3	Uncertainties on the Hadronic Cocktail . . . . .	146
<b>6</b>	<b>Results and Discussion</b>	<b>149</b>
6.1	Results . . . . .	149
6.1.1	First Invariant $M_{\mu\mu}$ Distribution with STAR . . . . .	149
6.1.2	First Measurement of the $\phi$ Spectra via Dimuons with STAR . . . . .	153
6.2	Summary and Discussion . . . . .	156
<b>7</b>	<b>Future <math>e^+e^-</math> Measurements with STAR</b>	<b>161</b>
	<b>Appendix A Training Neural Networks</b>	<b>167</b>
	<b>Appendix B Muon Purity in Au+Au Collisions at <math>\sqrt{s_{NN}} = 200</math></b>	
	<b>GeV</b>	<b>172</b>
	<b>Bibliography</b>	<b>180</b>

# Illustrations

1.1	(a) The particles included in the Standard Model of particle physics. The quarks and leptons make up the spin 1/2 matter particles called fermions. The spin 0 or 1 bosons make up the force carrier particles. (b) A schematic of the interactions between matter particles and force carriers in the Standard Model of particle physics. . . . .	2
1.2	The masses of the 6 known flavors of quarks shown in MeV/c <sup>2</sup> . The mass generated by the Higgs mechanism is shown in blue while the mass generated by the QCD vacuum is shown in taupe [16]. . . . .	3
1.3	Step-like lines show the spectra of non-strange mesons as a function of mass. The lowest curve includes roughly the set of mesons that Hagedorn originally used. The higher two curves include measurements from later experiments. The solid and dashed lines are exponential fits of the form $dN/dM \propto M^a e^{\frac{M}{T_H}}$ to the data. [17] . . . . .	4
1.4	The nonet of mesons with spin 0 predicted by the quark model. . . . .	5
1.5	(a) The behavior of the strong coupling constant $\alpha_s$ (shown here as $\alpha_{qq}$ ) as a function of r for several different temperatures and at zero temperature (solid line). This is the result using the $qq$ calculation scheme described in Ref. [22] (b) Experimentally measured values of $\alpha_s$ for several different values of $Q$ [23].	7
1.6	Feynman diagrams for screening (left) and anti-screening (right) effects in QCD. The anti-screening interaction is only possible with gluon self-interactions. . . . .	8



- 1.7 Diagrams for various  $\gamma\gamma \rightarrow \gamma\gamma$  interactions: Delbruck scattering (left), photon splitting (middle), and elastic light-by-light scattering (right). These types of interactions are forbidden in classical electromagnetism. The ATLAS detector at the LHC has recently found evidence for these types of reactions [25]. . . . . 9
- 1.8 The chemical freeze-out temperature versus  $\mu_b$  measured in heavy-ion collisions (a). The measured yield ( $dN/dy$ ) versus the yield from a grand canonical statistical model (b) [34] . . . . . 11
- 1.9 (a) Lattice QCD calculations of the energy density in units of  $T^4$  (which scales with the number of degrees of freedom) vs. temperature [35]. The red squares show the result of the calculation with 4 temporal bins. The black circles show the result of 6 temporal bins (finer lattice spacing). (b) Normalized pressure, energy density, and entropy of the Hadron Resonance Gas (solid lines) compared to Lattice QCD (bands). . . . . 12
- 1.10 An example QCD phase diagram in the temperature versus baryon chemical potential ( $T - \mu_b$ ) plane [37]. The points of reference are the limits of  $\mu_b$  i.e.  $\mu_b \rightarrow 0$ ,  $\mu_b \rightarrow \infty$  and the  $\mu_b$  of nuclear matter all at  $T = 0$ . Lattice QCD predicts a crossover transition from a hadron gas to a QGP at a temperature  $T_c \approx 155$  MeV at  $\mu_b = 0$ . The location and validity of the remaining features is largely speculative and is the source of much activity in the field. . . . . 12
- 1.11 (a) The Goldstone “Mexican Hat” potential  $V(\phi)$  [51]. (b) Comparison of the quark masses from the Higgs mechanism and from chiral symmetry breaking in QCD. . . . . 15
- 1.12 Left: Before colliding, length contracted ions approach with an impact parameter  $b$ . Right: During the collision, participant nucleons interact while the participant nucleons continue traveling almost undeflected in the forward and backwards directions. [56] . . . . . 17

1.13 (a) An example of the correlation between the geometric impact parameter  $b$  (not directly observable) and the final-state-observable charged-particle multiplicity ( $N_{ch}$ ). [57]. (b) Schematic view of the nucleon positions as a function of the longitudinal ( $z$ ) and one transverse direction ( $x$ ) shown for two different collision energies [58]. . . . . 18

1.14 A space-time diagram of the evolution of a heavy-ion collision. [65] . . . . . 19

1.15 A schematic of the evolution of a heavy-ion collision. [66] . . . . . 20

2.1 The Feynman diagrams for flavor creation (a,b), flavor excitation (c), and gluon splitting (d) [70–72] . . . . . 22

2.2 The Feynman diagrams for various dilepton production mechanisms (shown for  $e^+e^-$ ): the direct decay of a vector meson ( $\rho, \omega, \phi$ ) through a virtual photon (a); the Dalitz decay of a vector (V), pseudo-scalar (P), or scalar (S) meson into a neutral particle and an  $e^+e^-$  pair (b); the four-body decay of a pseudo-scalar or vector meson into  $e^+e^-$  and two pseudo-scalar mesons ( $\pi$  or  $\eta$ ) through an intermediate state containing a virtual photon or vector meson (c); the decay of a nucleon or  $\Delta$  resonance into a nucleon and a vector meson which further decays into an  $e^+e^-$  pair. [74] . . . . . 23

2.3 Inclusive  $e^+e^-$  invariant mass spectra for  $p$ +Be (a) and  $p$ +Au (b) collisions at 450 GeV. The data points are shown in filled circles while the hadronic cocktail is shown in black curves [92]. The systematic uncertainty on the hadronic cocktail is shown as a shaded band. . . . . 25

2.4 Inclusive  $e^+e^-$  invariant mass spectra in S+Au collisions at 200 A GeV (a) and in Pb+Au collisions at 158 A GeV. The dashed line shows the cocktail+vacuum  $\rho$  contribution. Models of the thermal radiation are shown for two competing scenarios: the dropping  $\rho$  mass (dot-dashed line), and the broadened  $\rho$  (solid line). While the vacuum  $\rho$  is disfavored by data, within the statistical uncertainties, the two competing thermal models cannot be distinguished [92]. . . . . 26

2.5	The NA60 detector setup near the target. The addition of the high granularity silicon-based vertex tracker was crucial for the NA60 $\mu^+\mu^-$ measurements. [97]	27
2.6	(a) The $\mu^+\mu^-$ invariant mass spectra in the acceptance of NA60 for In+In collisions without centrality selection before (red circles) and after (black triangles) subtracting the hadronic cocktail (without $\rho$ ). (b) The excess $\mu^+\mu^-$ invariant mass spectra in semi-central collisions for the acceptance of NA60. The predicted excess spectra is shown for several theoretical models.	28
2.7	(a) The $\mu^+\mu^-$ invariant mass spectra in the acceptance of NA60 for In+In collisions without centrality selection before (red circles) and after (black triangles) subtracting the hadronic cocktail (without $\rho$ ). (b) The excess $\mu^+\mu^-$ invariant mass spectra in semi-central collisions for the acceptance of NA60. The predicted excess spectra is shown for several theoretical models.	29
2.8	(a) Top panel: The HADES measurement of the $e^+e^-$ mass spectra in C+C collisions at 1 A GeV incident energy compared with expected sources. (a) Bottom panel: The ratio of the measured yield in 1 and 2 A GeV collisions to the respective cocktail A at each energy. (b) The inclusive multiplicity of the excess yield in the mass range $0.15 < M_{ee} < 0.5$ GeV/ $c^2$ . The solid curves show the excitation functions for $\pi^0$ and $\eta$ . The dotted curve is the excitation function for the $\pi^0$ scaled down by an arbitrary factor while the dashed curve shows the scaled contribution of the $\eta$ excitation function [100].	30
2.9	First measurement of the $e^+e^-$ invariant mass spectra. The hadronic cocktail describes the production over the full mass range. At RHIC energies the intermediate mass region is dominated by pairs from correlated open heavy flavor decays(a). The first PHENIX measurement of the $e^+e^-$ invariant mass spectra in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV showed a very large excess in central collisions at low mass.	31
2.10	Comparison between the measured excess of $e^+e^-$ pairs and the prediction from the broadened $\rho + QGP$ model (a). The centrality dependence of the $e^+e^-$ excess yield. [106]	33

2.11	The inclusive $\mu^+\mu^-$ invariant mass distribution in $p + p$ collisions at $\sqrt{s} = 200$ GeV within the PHENIX south arm acceptance. The lower panel shows the ratio of the measured yield to the hadronic cocktail [108]. . . . .	34
2.12	Azimuthal correlations from $c\bar{c}$ (a) and $b\bar{b}$ measured through the dimuon channel. Measurements are compared with PYTHIA and POWHEG [70, 108] .	34
2.13	STAR measurement of the $e^+e^-$ invariant mass spectra within STAR acceptance from $p + p$ collisions at $\sqrt{s} = 200$ GeV (a) [109]. Top panel of (b): The measurement of $e^+e^-$ invariant mass spectra within STAR acceptance from 0 – 80% central Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. Lower panel (b): The ratio of the measured spectra to the hadronic cocktail (excluding $\rho$ ) compared to an effective many-body model and a dynamical microscopic transport model [111, 113]. . . . .	35
2.14	The $e^+e^-$ invariant mass spectra within STAR acceptance from 0 – 80% central Au+Au collisions at $\sqrt{s_{NN}} = 19.6, 27, 39, 62.4, 200$ GeV. Statistical (bars) and systematic (boxes) uncertainties are shown with the data points. The total hadronic cocktail is shown for each energy while the individual components are shown only for the $\sqrt{s_{NN}} = 62.4$ GeV data [115]. . . . .	36
2.15	The mid-rapidity charged particle multiplicity ( $dN_{ch}/dy$ ) normalized excess dielectron mass spectra in full phase-space for Au+Au collisions at $\sqrt{s_{NN}} = 19.6$ and 200 GeV from STAR and for In+In collisions at $\sqrt{s_{NN}} = 17.3$ GeV from NA60. The solid line shows a model calculation including a broadened $\rho$ spectral function in a hadron gas (HG) and thermal radiation from a QGP [111]. . . . .	37
2.16	The stopping power ( $-\langle dE/dx \rangle$ ) of $\mu^+$ in copper as a function of $\beta\gamma = p/Mc$ [116]. . . . .	38
3.1	The RHIC accelerator complex. The AGS area is shown in (a) along with the EBIS, the Linac, the Booster and the AGS-to-RHIC line. The RHIC rings are shown in (b) [119]. . . . .	41
3.2	The STAR detector system. [121] . . . . .	43
3.3	A schematic of the STAR Time Projection Chamber [122]. . . . .	44

3.4	.....	46
3.5	The VPD vertex resolution using the trigger electronics as a function of gRefMult (the number of global tracks in $\pm 0.5$ units of $\eta$ ) in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. Four different configurations are shown. The default going into the 2016 run year provided $\sigma_{vz} \approx 4$ cm. The improved version developed for the HFT provided $\sigma_{vz} \approx 1.5$ cm at high multiplicities.	48
3.6	A photo of the Vertex Position Detectors before being installed around the beam-pipe (a). A schematic drawing the the VPD detector assembly, (b) and a schematic each single detector enclosure used by the VPD (c).	49
3.7	A schematic of the MRPCs used in the STAR BTOF. [125]	50
3.8	A schematic of a TOF tray showing the layout of the MRPCs within [126].	51
3.9	The $\beta^{-1}$ vs. $p$ measured by the TOF system in STAR for Au+Au collisions at $\sqrt{s_{NN}} = 7.7$ GeV. The expected $\beta^{-1}$ is shown for each particle species. Taken from Ref. [131]	54
3.10	Schematic views of the MTD geometry. The magnet return steel is shown in (a) with numbers indicating the order of the MTD backlegs. An isometric view of the MTD modules (b) showing that some backlegs contain 5 modules and others contain only 3 modules due to space constraints.	55
3.11	A schematic of the side-view of an MTD LMRPC module. [133]	56
3.12	A schematic drawing of the MTD trigger logic. Most trigger patches consists of 5 modules while some consist of only 3. The modules in a trigger patch are always grouped from position of equal $\eta$ . The red and magenta trigger patch numbers mark the location of various electronics boxes. The violet numbers mark trigger patches with some channels disabled where there is no material guarding the module [136].	57
3.13	A map of the magnetic field of the STAR detector in the transverse plane. (a). The energy loss parameterization used in the STAR MTD track matching algorithm (b). The energy loss is parameterized as a function of track momentum by averaging the energy loss from GEANT over many events for several sub-systems [136].	58

- 3.14 The luminosity of collected dimuon triggered events as a function of time through the run. The red line shows the initial luminosity goal, the black shows what was actually collected, and the blue shows a projection based on the first several days of data collection. The p+p dataset is shown in (a) and the p+Au dataset is shown in (b). . . . . 59
- 3.15 The Event selection cuts shown for the data collected from p+p collisions at  $\sqrt{s} = 200$  GeV. The primary vertex position ( $|z_{\text{TPC}}|$ ) before cuts is shown in (a). The difference between the TPC and VPD vertices ( $|z_{\text{VPD}} - z_{\text{TPC}}|$ ) is shown in (b). The primary vertex ranking is shown in (c). The multiplicity of MTD-matched tracks is shown in (d). In (a) and (b), the red lines show the cut applied to those quantities with the region between the red lines accepted. In (c) and (d) the red lines show the cut applied, with all events to the right of the red line being accepted. Summary of events surviving each cut (e). The bin labels correspond to : All events, dimuon triggered (after day 50), pass bad run rejection, pass  $|z_{\text{TPC}}| < 100$  cm, pass  $|z_{\text{VPD}} - z_{\text{TPC}}| < 6$  cm, pass ranking  $\geq 0$ , has at least one MTD matched track, and has at least one MTD matched primary track. The events surviving all cuts in the "tree" bin are saved to a tree for further analysis. . . . . 61
- 4.1 The  $dE/dx$  (a) and  $\beta^{-1}$  (b) for charged particles. Since both of these quantities are proportional to the particle's mass, they can be used to identify charged particles. . . . . 63
- 4.2 Comparison of Gaussian (a) versus a Pearson VII (b) distribution for the pion and muon  $\beta^{-1}$  shapes. The Gaussian significantly underestimates the tails of the distributions. Since the yield of the  $\pi$  is much larger than the  $\mu$ , underestimating the  $\pi$  tails could significantly effect the fit to the  $\mu$  contribution. The Pearson VII distribution, with a free parameter related to the distribution's kurtosis, provides a much better description of the tails in the  $\beta^{-1}$  distributions. . . . . 64

4.3	Simulated MTD cell (a) and $\Delta Z$ distributions for signal (primary $\mu s$ ) and background sources. The signal and background distributions are normalized for comparison. The effect of varying amounts of steel in the $\phi$ direction can be clearly seen in the cell distribution. Hadrons are significantly more likely to punch through the steel guarding the edge cells (0 & 11) than the central cells. . . . .	68
4.4	(a) The invariant mass distribution for unlike-sign and like-sign pairs near the $J/\psi$ mass. The N-1 cut technique is used to maximize the $J/\psi$ significance by cutting on all MTD PID variables except the $\Delta TOF$ distribution. A $p_T^{\text{leading}} > 1.5$ (GeV/c) cut is applied to further improve the purity in the $J/\psi$ mass region. The $\beta^{-1}$ vs. momentum distribution for all tracks passing basic QA cuts that are matched to hits in the MTD and the BTOF detectors. (b) The $\beta^{-1}$ calculated from the BTOF information shows clear contributions from $\pi$ , K, and $p/\bar{p}$ . . . . .	72
4.5	The $\Delta TOF$ distributions for $\mu^\pm$ from $J/\psi$ , $\pi^\pm$ , $K^\pm$ , and $p/\bar{p}$ . . . . .	72
4.6	The $M_{\pi^+\pi^-}$ distribution near the $K_S^0$ mass shown for the cases in which only one track is matched to an MTD hit (a) and the $M_{K^+K^-}$ distribution near the $\phi$ mass shown for the cases in which only one track is matched to an MTD hit (b). The $\Delta Y$ and $\Delta Z$ data / simulation ratio for both $\pi^\pm$ and $K^\pm$ (c). The MTD Cell data / simulation ratio for both $\pi^\pm$ and $K^\pm$ (d). . . . .	74
4.7	An example of a dense multilayer perceptron neural network architecture. The shallow neural networks have only a single hidden layer of neurons between the input and output layers. The deep neural networks have two or more. Bias neurons in the hidden layers are marked with a “B”. . . . .	76
4.8	The signal vs. background rejection power as a function of the number of neurons ( $NN_{HL}$ ) in the hidden layer of a shallow neural network. The performance of the SNN’s are quantified using the AUC - the area under the background rejection vs. signal efficiency curve (See 4.4.4 in text). The points are the mean value of 10 models trained with different random samples. The uncertainties show the $1\sigma$ variance of the models assuming a Gaussian variance. . . . .	76

4.9	A Neural Network with two hidden layers containing N=30,20 neurons (a) and its output for simulated signal and background tracks (b). . . . .	79
4.10	The classifier response for the BDT (a) and DNN (b). These are the responses for all simulated tracks, so the distribution is mostly dominated by the high $p_T$ tracks. . . . .	80
4.11	Comparison of the performance of the DNN and BDT (nTrees=800) for all simulated tracks ( $1.0 \leq p_T \leq 15.0$ GeV/c). The DNN out-performs the BDT at mid to high signal efficiencies. . . . .	81
4.12	Raw yield extraction of the $\phi$ meson using optimized traditional 1D PID techniques (a) compared to the DNN based PID (b). . . . .	81
4.13	The raw $\mu^+\mu^-$ invariant mass spectra from optimized traditional PID and from deep neural network based PID. The distributions are scaled in the intermediate mass region to make comparison easier. . . . .	82
4.14	The top panel shows the DNN response for muon candidates in the range $1.5 < p_T < 1.55$ GeV/c. A template fit is conducted to extract the contributions from $\mu$ (red), $\pi$ (blue), K (orange), and $p$ (magenta). The lower panel shows the ratio of the data over the sum of the contributions. . . . .	84
4.15	The result of the DNN response fit for $\mu$ , $\pi$ , K, and p contributions projected back onto the $\Delta Z$ (a) and DCA (b) distributions. The ratio of fit / data is shown in the lower panels of each figure. . . . .	84
5.1	The invariant mass of muon candidate tracks shown for unlike-sign pairs and like-sign pairs. . . . .	86
5.2	The neural network response for positive and negative tracks in $p + p$ collisions at $\sqrt{s} = 200$ GeV (a) and $p + \text{Au}$ collisions at $\sqrt{s_{NN}} = 200$ GeV (b). . . . .	87
5.3	The muon purity as a function of track $p_T$ for all muon candidate tracks and for various DNN-base pid cuts. . . . .	89
5.4	The DNN response of daughter 1 ( $r_1^{NN}$ ) versus the DNN response of daughter 2 ( $r_2^{NN}$ ) for $p + p$ collisions at $\sqrt{s} = 200$ GeV. . . . .	91



5.5	The three categories of pairs: Signal pairs resulting from signal + signal (a), cross pair resulting from background + signal (b), and background pair resulting from background + background (c). . . . .	92
5.6	The neural network architecture used for the selection of signal pairs (a). The inputs are the $p_T$ and neural network response $r^{NN}$ of each daughter track in the pair. The output of the pair DNN for the three different categories of pairs (b). Note: In this figure mixed pairs=cross pairs. . . . .	93
5.7	The correlation between $r_{pair}^{NN}$ and $2.41 \times (\sqrt{(r_1^{NN})^2 + (r_2^{NN})^2} - 1)$ (a). The three types of pairs shown using the $r_{pair}$ projection (b). . . . .	94
5.8	Plots of the yield / event from a toy Monte Carlo simulations with various types of combinatorial backgrounds present. The solid black line shows the true amount of signal / event. Top: Combinatorial background resulting from an $e^+e^-$ -like case, i.e. approximately Poisson pair production. In this case the geometric mean of the like-sign reproduced the background correctly. A charge asymmetry is introduced to demonstrate that in this case the direct mean does not reproduce the background correctly. Bottom: Combinatorial background in the case with independently produced singles. The direct mean and geometric mean both underestimate the background, leading to an overestimation of the signal. The combinatorial background estimated via the mean multiplicities gives a more accurate estimate of the unlike-sign background. . . . .	102
5.9	The $M_{\pi^+\pi^-}$ distribution near the $K_S^0$ mass (a) and the $M_{K^+K^-}$ distribution near the $\phi$ mass (b). Only MTD matched tracks passing all quality cuts are used. . . . .	104
5.10	Top: The invariant mass of muon candidates in the low mass region ( $M < 1.0 \text{ GeV}/c^2$ ) with a cut on the NN response of both daughter 1 and 2 ( $r_{NN}^{1,2} < 0.8$ ) to select predominantly background-like pairs. Bottom: The simulated dimuon invariant mass distribution for hadron decay sources and combinatorial backgrounds. Scaling is arbitrary, meant to show that the distribution in data can be qualitatively described by background-only sources.	106

- 5.11 The  $FG_{+-}$  and  $F_{\pm\pm}$  distributions for two different background-only  $r_{\text{pair}}$  selection criteria. In both cases the  $FG_{\pm\pm}$  distribution fails to describe the  $FG_{+-}$  distribution. . . . . 108
- 5.12 An example of determining the like-sign ration as a function of  $r_{\text{pair}}$ , i.e.  $R = R(r_{\text{pair}})$  in the  $\omega$  mass region,  $0.7 < M_{\mu\mu} < 0.85 \text{ GeV}/c^2$  (a) and the  $J/\psi$  mass region,  $3.0 < M_{\mu\mu} < 3.2 \text{ GeV}/c^2$ . The blue points show the  $FG_{+-}$  distribution, the red points show the  $FG_{\pm\pm}$  distribution, and the black distribution shows the ratio  $R \times 10$  to make it visible on the same axis. A dotted black line is added to show where  $R = 1$  would be. . . . . 109
- 5.13 Fits to the  $r_{\text{pair}}$  (pairPid) distributions  $FG_{+-}(r_{\text{pair}})$  (a) and  $FG_{\pm\pm}(r_{\text{pair}})$  (b). 111
- 5.14  $R_{\text{pair}}$  computed from the data (blue), which includes contribution from signal pairs and  $R_{\text{pair}}$  computed from templates (red). . . . . 112
- 5.15  $R_{\text{pair}}$  computed from the templates and evaluated in the signal selection region shown as a function of  $M_{\mu\mu}$ . The effect of the  $K_S^0$  can be clearly seen around  $M_{\mu\mu} = 0.45 \text{ GeV}/c^2$ . The value at high mass stabilizes to a small enhancement due only to the statistical effect from single muon production. 112
- 5.16 The standard mixing-event technique compared with the same-event like-sign distribution (a). The mixed-event distribution is able to match the same-event shape only at very high  $M_{\mu\mu}$ . The opening angle versus pair  $p_T$  used to weight the mixed-event distribution in order to produce a mixed-event that reproduces the shape of the same-event distribution (b). . 113
- 5.17 The mixed-event distribution weighted by the pair opening angle as a function of pair  $p_T$ . . . . . 114
- 5.18 The maximum likelihood fits used to extract the raw  $\phi \rightarrow \mu^+ + \mu^-$  yields in four momentum bins. Each fit used a Gaussian for the  $\phi$  peak and a fourth-order polynomial for the background shape. The nominal fit range was  $0.85 < M_{\mu\mu} < 1.2 \text{ GeV}/c^2$ . . . . . 116
- 5.19 The raw  $ME_{\pm\pm}$  (red) and the correlated background (blue) for  $p + p$  (a) and  $p + \text{Au}$  collisions at  $\sqrt{s_{NN}} = 200 \text{ GeV}$ . . . . . 117

5.20	The $FG_{+-}$ distribution (black) along with the $ME_{\pm\pm}$ distribution (red) and the correlated background estimation (blue). The signal dimuons are shown in magenta. . . . .	118
5.21	The signal over background ratio as a function of invariant mass for $p + p$ (top) and $p+\text{Au}$ (bottom). . . . .	119
5.22	The measured transverse momentum spectra of a large number of mesons in $p + p$ collisions at $\sqrt{s} = 200$ GeV. The solid lines are Tsallis Blast-Wave fits to each particle species [147]. . . . .	121
5.23	The Feynman diagrams for various dilepton production mechanisms (shown for $e^+e^-$ ): the direct decay of a vector meson ( $\rho, \omega, \phi$ ) through a virtual photon (a); the Dalitz decay of a vector (V), pseudo-scalar (P), or scalar (S) meson into a neutral particle and an $e^+e^-$ pair (b); the four-body decay of a pseudo-scalar or vector meson into $e^+e^-$ and two pseudo-scalar mesons ( $\pi$ or $\eta$ ) through an intermediate state containing a virtual photon or vector meson (c); the decay of a nucleon or $\Delta$ resonance into a nucleon and a vector meson which further decays into an $e^+e^-$ pair. [74] . . . . .	123
5.24	The total hadronic cocktail for $p + p$ collisions at $\sqrt{s} = 200$ GeV for a progression of kinematic cuts. Requiring only that $ y^{\mu\mu}  < 0.5$ results in the solid red curve. The addition of the $ \eta^\mu  < 0.5$ cut on each daughter $\mu$ is shown in the solid blue curve. Finally, the addition of the $p_T^\mu > 1.1$ GeV/ $c$ cut is shown in the solid black curve. . . . .	128
5.25	The hadronic cocktail for $p + p$ collisions at $\sqrt{s} = 200$ GeV. Each component is drawn separately. The sum of all contributions is drawn in the black solid curve. The hadronic cocktail for $p+\text{Au}$ collisions is identical, except that the entire cocktail is scaled by the number of binary collisions, $N_{col}$ . . . . .	129
5.26	The pair efficiency as a function of $M_{\mu\mu}$ for two scenarios: 1) A $\gamma^*$ decay with flat mass and $p_T$ and 2) using the $M_{\mu\mu}$ and $p_T^{\mu\mu}$ distributions from the total hadronic cocktail. Note, this includes only the efficiency and acceptance affects for tracks that fall within the gross kinematic acceptance of the MTD (i.e. $ \eta  < 0.5$ and $p_T > 1.1$ GeV/ $c$ ). . . . .	135

5.27	The total efficiency $\times$ acceptance for the $\phi$ (a) and the $J/\psi$ (b). Each component is shown in addition to the total. The total efficiency $\times$ acceptance is shown in small bins and in the larger bins used to extract signal.	136
5.28	The PID efficiency for $\mu^+\mu^-$ pairs in $p + p$ (a) and $p+\text{Au}$ collisions at $\sqrt{s_{NN}} = 200$ GeV selected with $r_{\text{pair}} > 1.36$ and $r_{\text{pair}} > 1.34$ for $p + p$ and $p+\text{Au}$ respectively.	138
5.29	The signal efficiency and purity determined as a function of $r_{\text{pair}}$ . The S/B and significance are also shown with the cut value for maximum significance highlighted.	139
5.30	The invariant mass distribution of the correlated pairs from hadron contamination. The “signal” is extracted from the background only region $r_{\text{pair}} < 0.9$ . Notice, for instance, the lack of any sign of a $J/\psi$ peak and the clear $K_S^0 \rightarrow \pi^+\pi^-$ and $\phi \rightarrow K^+K^-$ peaks - suggesting that the excess $\mu^+\mu^-$ pairs are a result of correlated background.	143
5.31	The Systematic uncertainties for the $\mu^+\mu^-$ invariant mass spectra in $p + p$ collisions at $\sqrt{s} = 200$ GeV.	144
5.32	The relative uncertainty on the $\phi$ yield as a function of $p_T$ after all corrections due to several different sources of systematic uncertainty.	146
5.33	The hadronic cocktail for $p + p$ collisions at $\sqrt{s} = 200$ GeV. Each component is drawn separately. The sum is drawn in the black solid curve with the total uncertainty shown in the solid grey band.	148
6.1	Top: The $\mu^+\mu^-$ invariant mass spectra in $p + p$ collisions at $\sqrt{s} = 200$ GeV. The statistical uncertainties are shown in vertical bars and the systematic uncertainties are shown in shaded boxes. Bottom: The ratio of data to hadronic cocktail.	151
6.2	Top: The $\mu^+\mu^-$ invariant mass spectra in $p+\text{Au}$ collisions at $\sqrt{s_{NN}} = 200$ GeV. The statistical uncertainties are shown in vertical bars and the systematic uncertainties are shown in shaded boxes. Bottom: The ratio of data to cocktail.	152

6.3	Top: The branching ratio times the invariant yield of the $\phi$ meson measured through the $\phi \rightarrow \mu^+ + \mu^-$ decay channel. Bottom: The ratio of data to the Tsallis Blast-Wave scaled by the nominal $dN/dy$ value. The uncertainty on $dN/dy$ is shown in gray. . . . .	154
6.4	The mass of the $\phi$ measured in four $p_T$ bins. Error bars show only the uncertainties from the maximum likelihood fit. No energy loss correction is included. . . . .	155
6.5	The $J/\psi$ signal measured through the $\mu^+\mu^-$ channel in $p + p$ , $p+Au$ , and $Au+Au$ collisions at $\sqrt{s_{NN}} = 200$ GeV [164]. . . . .	156
6.6	The inclusive $J/\psi$ cross section measured in $0 < p_T < 14$ GeV/ $c$ . STAR dielectron and dimuon measurements are compared with model calculations. [165–167]. Figure taken from Ref. [168] . . . . .	157
6.7	The nuclear modification factor $R_{pAu}$ as a function of $p_T$ for $J/\psi$ measured through the dimuon channel [168]. . . . .	157
6.8	The nuclear modification factor $R_{AA}$ as a function of $p_T$ for $J/\psi$ measured through the dimuon channel [168]. LHC data points from Ref. [68, 173]. Model calculations from Ref. [67, 170, 171] . . . . .	158
6.9	The nuclear modification factor $R_{AA}$ as a function of $N_{part}$ for $J/\psi$ measured through the dimuon channel [168]. LHC data points from Ref. [68, 173]. Model calculations from Ref. [67, 170, 171] . . . . .	158
6.10	The single muon purity in 40 – 80% central $Au+Au$ collisions for various selection criteria. . . . .	159
6.11	The single muon purity in 20 – 40% central $Au+Au$ collisions for various selection criteria. . . . .	160
7.1	The cumulative number of events collected versus time through the run. . . . .	162
7.2	The expected hadronic cocktail components for ${}^{96}_{44}\text{Ru}+{}^{96}_{44}\text{Ru}$ or ${}^{96}_{40}\text{Zr}+{}^{96}_{40}\text{Zr}$ collisions at $\sqrt{s_{NN}} = 200$ GeV. The contributions from thermal production and the $\rho^0$ meson are not included. The $c\bar{c} \rightarrow e^+e^-$ contribution is from PYTHIA and scaled with a cross section of 800 $\mu\text{b}$ . . . . .	163

7.3	The hadronic cocktail component from $c\bar{c} \rightarrow e^+e^-$ Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. Two extreme scenarios for $c\bar{c}$ are shown: 1) default PYTHIA i.e. no modification from the medium, 2) fully decorrelated (random correlations). . . . .	163
7.4	The NA60 experiment combined a vertex offset fit (a) with an analysis of the pair $p_T$ to determine that the excess yield was not from Drell-Yan or $c\bar{c}$ but instead from a prompt thermal source. . . . .	165
7.5	The $c\bar{c}$ cross section versus collision energy ( $\sqrt{s}$ ) measured by several experiments and compared with the next-to-leading order (NLO) calculations (a). The measured excess $e^+e^-$ spectra near the $\rho$ meson mass in Au+Au collisions at $\sqrt{s_{NN}} = 19.6$ GeV from BES 1 compared with various theoretical curves (b). In addition, the expected reduction in uncertainty for BES II is shown for two scenarios, 1) for improved statistics with the current TPC hardware, and 2) for improved statistics with the inner TPC upgrade. . . . .	166
A.1	A simple perceptron with a binary output value. . . . .	167
A.2	A sigmoid perceptron is capable of outputting a continuous variable. . . . .	168
A.3	A simple ANN built from multiple perceptrons. Each perceptron has only one output. Multiple exiting lines signify that the output is given as input to many different neurons in the next layer. . . . .	169
B.1	The $J/\psi$ yield in 0 – 80% Au+Au collisions measured through the dimuon channel. The S/B ratio is $\sim 1/27$ . . . . .	173
B.2	The $e^+e^-$ invariant mass spectra in 0 – 80 Au+Au collisions. The lower panel shows the S/B ratio. In the $J/\psi$ region the S/B reaches $\sim 1/2 - 1$ [88]. . . . .	174
B.3	The purity fits are shown for $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . .	175

- B.4 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 175
- B.5 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 176
- B.6 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 176
- B.7 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 177
- B.8 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 177
- B.9 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 178
- B.10 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 178

B.11 The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively. . . . . 179



# Tables

3.1	A summary of the performance of the TOF system in STAR from the year 2012 to 2017. The total time resolutions marked with a “*” indicate that the VPD was not used for the start-time determination. Instead a bootstrapped start-time was calculated from the pions in the event using only the BTOF information. . . . .	52
3.2	Triggers included in the st_mtd data stream. The # of events, trigger ids, and sampled luminosity are given for each where available. . . . .	60
3.3	Cuts applied to select events from the data collected by the dimuon trigger condition. . . . .	60
4.1	Configuration options used for the MTD Simulation. The code can be found here: <a href="https://github.com/jdbrice/MtdSimGun">https://github.com/jdbrice/MtdSimGun</a> . . . . .	70
4.2	Cuts used for determining the signal and background $\Delta$ Time-of-Flight PDFs.	72
4.3	Cuts used to select $K_S^0 \rightarrow \pi^+ \pi^-$ decays. The daughter pions provide a $\pi$ -enhanced sample that can be compared to the $\pi$ Monte Carlo simulation.	73
4.4	Parameters used in the training phase for the shallow and deep neural networks.	77
5.1	Weak decay processes through which single muons may be created. . . . .	87
5.2	GEANT simulation of $\pi^\pm$ and $K^\pm$ in the STAR detector. These values represent the probability of each event occurring before exiting the STAR detector (specifically within the MTD, since it is the detector in STAR at largest radius). . . . .	88
5.3	Processes through which correlated dimuon backgrounds can be produced. .	103
5.4	Processes through which correlated dimuon backgrounds can be produced. .	105

5.5	The decay channels supported by the hadronic cocktail simulation code. For each decay channel, Table 5.5 shows the mass, $\Gamma_0$ , $dN/dy$ or $d\sigma/dy$ , branching ratio (BR), and $\Lambda^{-2}$ values are shown where applicable. Taken from Ref. [147, 151–158] . . . . .	122
5.6	The number of equivalent minimum bias events ( $N_{\text{MB}}^{\text{equivalent}}$ ) calculated for various different event selection cuts and efficiency factors. The result in the last row corresponds to the selection criteria used in these analyses. . . . .	132
5.7	Sources of systematic uncertainty on the measurement of the $\mu^+\mu^-$ invariant mass spectra in $p + p$ and $p+\text{Au}$ collisions at $\sqrt{s_{NN}} = 200$ GeV. . . . .	141
5.8	Sources of systematic uncertainty for the $\phi$ yield extraction. . . . .	145
5.9	The nominal value of $dN/dy$ along with its relative uncertainty is shown for each particle. [147] . . . . .	147
5.10	The nominal value of $\sigma$ along with its relative uncertainty is shown for $c\bar{c}$ , $b\bar{b}$ , and Drell-Yan. [147] . . . . .	147

## List of Listings

1	An example lineage tree for a primary $\pi$ that interacts with material and has a daughter $\pi$ hit the MTD. RC indicates that the MC track was reconstructed and MTD indicates that the MC track produced a hit in the MTD. The tuple after each particle is its $(p_T, \eta, \phi)$ . . . . .	70
2	An example of the lineage tree for a $\pi \rightarrow \mu + \nu$ decay . . . . .	71
3	The parameters used to instantiate PYTHIA for simulation of the correlated $c\bar{c} \rightarrow \mu^+ + \mu^-$ decays. The trigger $MSEL = 1$ was also used to check the line shape and magnitude of the $c\bar{c}$ contribution. . . . .	125
4	The parameters used to instantiate PYTHIA for simulation of the correlated $b\bar{b} \rightarrow \mu^+ + \mu^-$ decays. . . . .	125
5	The parameters used to instantiate PYTHIA for simulation of the Drell-Yan process, $q\bar{q} \rightarrow \gamma^*/Z^0 \rightarrow \mu^+ + \mu^-$ decays. . . . .	126
6	A python-like pseudo code implementation of the back propagation algorithm.	171

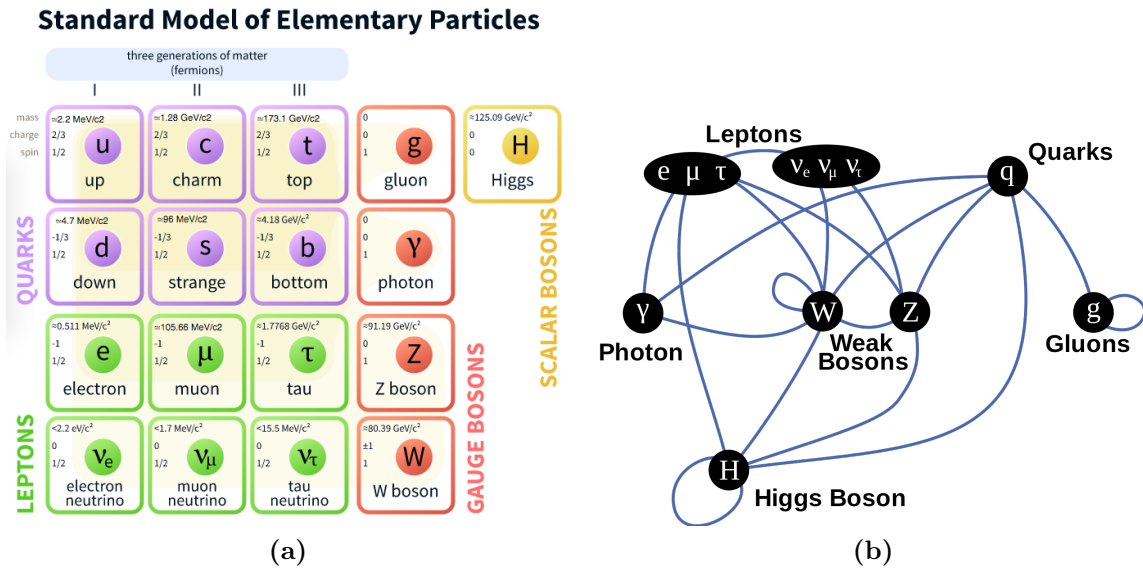
# Chapter 1

## Introduction

### 1.1 The Standard Model

Curious minds have long sought to understand the fundamental realities of nature. This task, once relegated to philosophers, has in recent times become the work of science. Physics, from the ancient Greek *φυσική* for “nature”, is the modern incarnation of this pursuit. Reductionism has been the favored tool of physics: perpetually explaining large complex systems by breaking them down into their constituent parts and interactions [4]. Over the course of hundreds of years, this discipline has slowly but surely shed light on the fundamental constituents of matter and the forces governing their interactions. The Standard Model (SM) of particle physics is the current state-of-the-art description capable of describing most matter and three of the four known fundamental forces (gravity is not yet included).

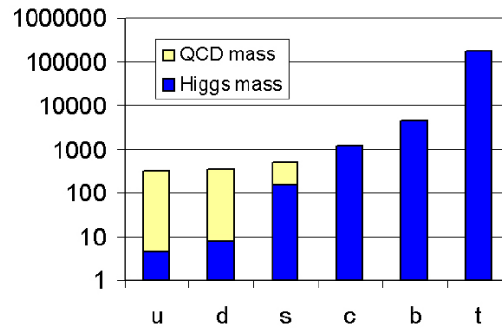
The Standard Model, shown in Fig. 1.1a, consists of matter particles (fermions, spin 1/2) and force carriers (bosons, spin 0 or 1). The matter particles interact through the fundamental forces via the force carrier particles (see Fig. 1.1b) and are sub-divided into quarks and leptons. James Maxwell first described classical electricity and magnetism in 1873 [5]. In the SM these forces, combined into the electromagnetic (EM) force and mediated by the photon ( $\gamma$ ), are described by Quantum Electrodynamics (QED). Shortly after, the theory of weak interactions ( $10^{-11}$  times weaker than the EM force) mediated by the massive  $W^\pm$  and  $Z^0$  particles was developed to explain Earnest Rutherford’s 1899 discovery of  $\beta$  decay [6, 7]. The SM’s current form was born in 1961 with Sheldon Glashow, Abdus Salam, and Steven Weinberg’s Nobel Prize winning accomplishment combining the descriptions of electromagnetism and the weak force [8–11]. Before that time the description of these forces stood somewhat alone as mostly disconnected islands of understanding. Later the third fundamental force in the SM, the strong force, was developed to explain how the dense nuclei of atoms remained bound despite the EM force pushing the positively charged protons apart. The strong force, mediated by the color charge carrying gluon, is described by Quantum



**Figure 1.1 :** (a) The particles included in the Standard Model of particle physics. The quarks and leptons make up the spin 1/2 matter particles called fermions. The spin 0 or 1 bosons make up the force carrier particles. (b) A schematic of the interactions between matter particles and force carriers in the Standard Model of particle physics.

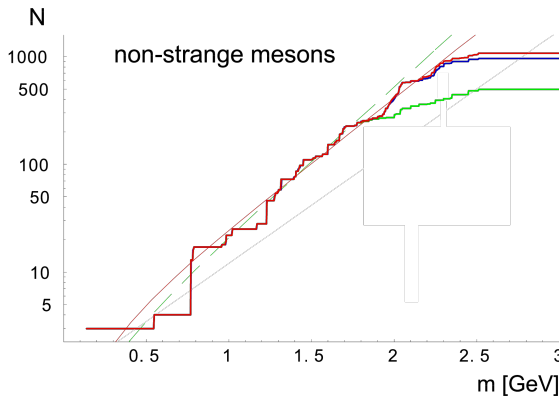
Chromodynamics (QCD) and is the main force of interest in this work. Despite providing some of the most accurate predictions of any scientific theory, the SM is not complete [12]. It does not include a description of the fundamental force of gravity or a viable candidate for dark matter, and until recently could not explain the mass of the fundamental matter particles. The Higgs mechanism, named after Peter Higgs but simultaneously proposed by several individuals in the 1960s, postulates the existence of the Higgs field responsible for giving finite masses to the  $W^\pm$  and  $Z^0$  force carriers as well as the mass of the quarks and leptons [13].

The most recent addition to the SM came on July 4th, 2012 when the Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS), experiments at the Large Hadron Collider (LHC), jointly announced that their search for the Higgs Boson led to the discovery of a new particle with a mass of  $\sim 125 \text{ GeV}/c^2$ . At the time it was not known for certain if the particle they had discovered was in fact the Higgs Boson, responsible for giving mass to the other fundamental particles. Since that time countless measurements have been combined to precisely measure the properties of this new particle, confirming that this is in fact the Standard Model Higgs [14, 15]. This discovery earned Peter Higgs and Francois Englert the 2013 Nobel Prize in Physics and marked a monumental achievement for the international scien-



**Figure 1.2 :** The masses of the 6 known flavors of quarks shown in  $\text{MeV}/c^2$ . The mass generated by the Higgs mechanism is shown in blue while the mass generated by the QCD vacuum is shown in taupe [16].

tific community. The discovery of the Higgs Boson represents decades of work by thousands of scientists around the world. The search for the Higgs Boson was motivated by a simple question: what gives fundamental particles like the  $W^\pm$  and the  $Z^0$  their mass? However, the mass of the proton and neutron, the constituents of almost all the matter we see around us, is not entirely explained by the Higgs mechanism. Unlike the fundamental particles, the mass of these objects is believed to be an emergent property of the fundamental characteristics of QCD. My own interest in high energy nuclear physics was largely born from this question of how the protons and neutrons gain their observed mass. Before returning to this question, an introduction to QCD and higher energy nuclear physics will be discussed.

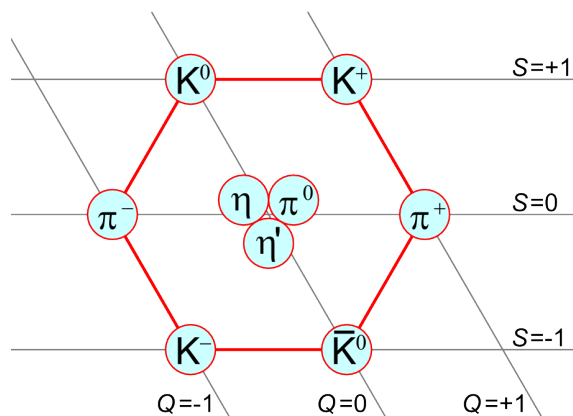


**Figure 1.3** : Step-like lines show the spectra of non-strange mesons as a function of mass. The lowest curve includes roughly the set of mesons that Hagedorn originally used. The higher two curves include measurements from later experiments. The solid and dashed lines are exponential fits of the form  $dN/dM \propto M^a e^{\frac{M}{T_H}}$  to the data. [17]

## 1.2 High Energy Nuclear Physics and QCD

### 1.2.1 Historical Origins

The origins of high energy nuclear physics and the development of QCD are intimately connected to advances in particle physics. The advent of higher and higher energy particle colliders led to an explosion in the number of hadronic states discovered in the 1960s and 1970s. So many states were found that Enrico Fermi is quoted, somewhat jokingly, saying: “If I could remember the names of all these particles, I would have been a botanist!” Rolf Hagedorn, a researcher of the Max Planck Institute at the time, began working on this problem to understand the exponential increase in observed hadronic states as a function of their mass. While few clues existed, he took note of the exponential drop in the transverse mass spectra of secondary particles, the lack of scaling with collision energy (expected from Boltzmann’s Law) of the momentum distribution and the exponential drop of elastic scattering at wide angles [17]. These clues led him to believe that the products of these reactions were emitted from a common thermal source or “fireball”. However, Hagedorn was not the first to propose a statistical mechanical description of these excited nuclear levels. Already in 1936 H. A. Bethe suggested that the density of states should rise exponentially ( $\rho(E) \propto e^{b\sqrt{AE}}$  with constant  $b$  and mass number  $A$ ) [18]. These early models failed to describe the spectra of the lightest hadrons though (specifically  $\pi^\pm$ ,  $m_{\pi^\pm} \approx 139.5 \text{ MeV}/c^2$ ). Hagedorn remedied these difficulties in the Statistical Bootstrap Model (SBM) of Hadronic States [17, 19, 20]. In



**Figure 1.4 :** The nonet of mesons with spin 0 predicted by the quark model.

his model (see Fig. 1.3), the fireball temperature, called the Hagedorn Temperature ( $T_H$ ), was extracted by fitting the hadronic mass distribution to an exponential. Using an exponential of the form  $dN/dM \propto M^\alpha e^{\frac{M}{T_H}}$ , the Hagedorn Temperature is found to be  $T_H \approx 160$  MeV ( $\approx 2 \times 10^{12}$  K) [20]. In Hagedorn's model, the heavy hadronic states, composed of lighter point-like constituents, form a gas that radiates like a thermal source at a constant temperature. The Hagedorn Temperature represents the maximum temperature of the gas at which further energy added to the system is added as latent heat needed to populate the higher excited states. As more and more excited states are populated the relevant degrees of freedom of the system grow.

The constituents of these heavy hadronic states were not known until the 1969 Nobel prize winning work of Murray Gell-Mann (and George Zweig among others) in which he introduced the quark model [21]. The quark model proposed fundamental particles called quarks (see Fig. 1.1) and organized the various hadron states according to their quark content. An example of the organization scheme of the quark model can be seen in Fig. 1.4. In the quark model, the lightest three flavors of quarks, up ( $u$ ), down ( $d$ ) and strange ( $s$ ) (with masses  $m_u \approx 2.3 < m_d \approx 4.8 \ll m_s \approx 95$  MeV/ $c^2$ ) are arranged in an  $SU(3)_{flavor}$  symmetry. The *flavor* subscript in  $SU(3)_{flavor}$  denotes that the symmetry is with respect to flavor e.g.  $u$ ,  $d$  and  $s$ . Since the strange quark is significantly more massive than the up or down, a more nearly exact  $SU(2)_{flavor}$  symmetry exists with only the two lightest flavors. This  $SU(2)_{flavor}$  symmetry, often called isospin, in analogy to quantum mechanical spin, is used to explain the symmetries between the proton and neutron. Since the  $u$  and  $d$  quark masses are not



actually identical, the  $SU(2)_{flavor}$  symmetry is not exact, thus the difference between the proton and neutron mass. The quark model allowed Gell-Mann to predict the existence of new hadronic states such as the spin  $\frac{3}{2}$   $\Omega^-$  baryon consisting of three strange quarks ( $sss$ ) with a mass of  $\sim 1.6$  GeV/ $c^2$ . The discovery of this state at Brookhaven National Lab in 1964 was the confirmation that the quark model needed. However, according to Pauli's exclusion principle, fermions should not be able to fill states with identical quantum numbers. Therefore, explaining the spin  $\frac{3}{2}$   $\Omega^-$  (and spin  $\frac{3}{2}$   $\Delta^{++}$  etc.) required the introduction of a new quantum number. This new quantum number, unlike the electromagnetic force which has a single charge (and its corresponding anti-charge), needed to have 3 "charges" to distinguish the 3 identical fermion states in e.g. the spin  $\frac{3}{2}$   $\Omega^-$ . The new quantum number was dubbed "color" making analogy to the 3 primary colors: red, green, and blue. Along with the quark model came the development of QCD. QCD is the theory describing the various quark flavors and the color force between them mediated by massless gluons. In QCD the color force is an exact  $SU(3)$  symmetry. Each quark carries a color charge while the gluons carry a combination color + anti-color charge resulting in 8 possible color combinations for the gluons.

### 1.2.2 Quantum Chromodynamics

At this point it is helpful to consider QCD in its modern formulation. The gauge invariant QCD part of the Standard Model Lagrangian can be written compactly as:

$$\mathcal{L}_{QCD} = \bar{\psi} (i\gamma_\mu D^\mu - m) \psi - \frac{1}{2} G_{\mu\nu}^a G_a^{\mu\nu} \quad (1.1)$$

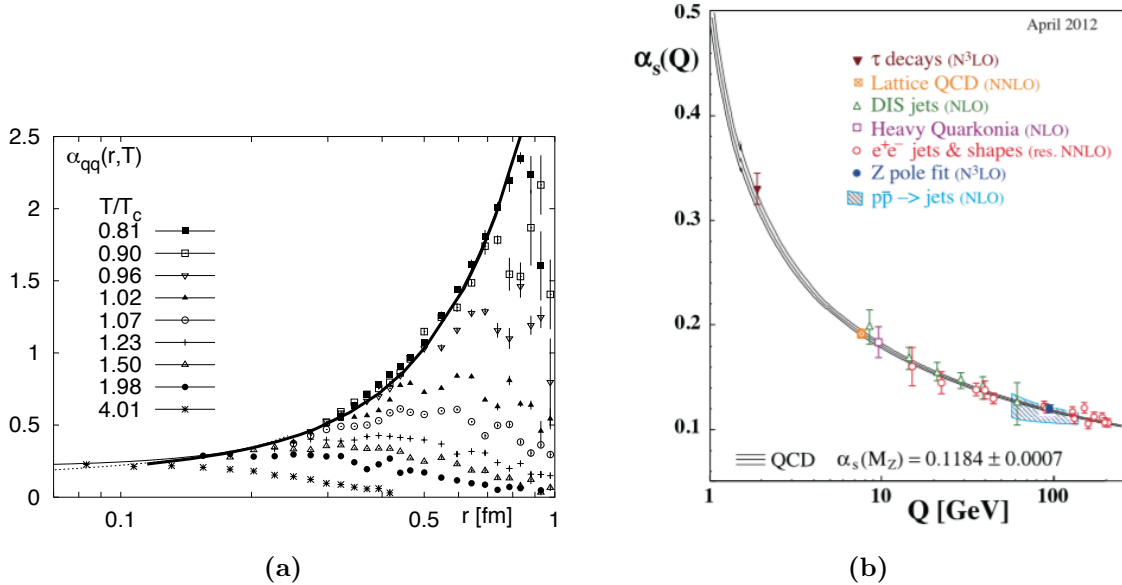
with

$$G_{\mu\nu}^a = \delta_\mu A_\nu^a(x) - \delta_\nu A_\mu^a(x) + gf_{abc} A_\mu^b(x) A_\nu^c(x) \quad (1.2)$$

$D_\mu$  is the gauge covariant derivative:

$$D_\mu = \delta_\mu - ig \frac{\lambda_a}{2} A_\mu^a(x) \quad (1.3)$$

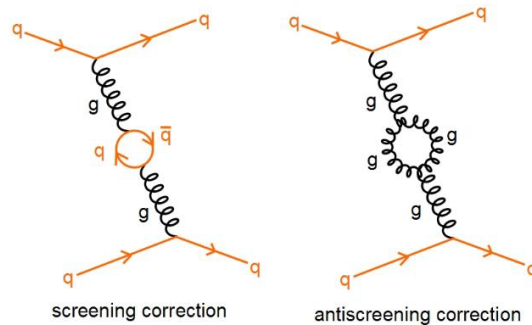
where  $g$  is the QCD coupling constant,  $f_{abc}$  are the  $SU(3)_{color}$  structure constants,  $\psi$  are the quark fields (implicit sums over quark flavor), and  $\gamma_\mu$  are the Dirac matrices.  $\lambda_a$  are the



**Figure 1.5 :** (a) The behavior of the strong coupling constant  $\alpha_s$  (shown here as  $\alpha_{qq}$ ) as a function of  $r$  for several different temperatures and at zero temperature (solid line). This is the result using the  $qq$  calculation scheme described in Ref. [22] (b) Experimentally measured values of  $\alpha_s$  for several different values of  $Q$  [23].

Gell-Mann matrices, a generalization of the Dirac matrices in  $SU(N)$ .  $G_{\mu\nu}^a$  are the gauge invariant field strength tensors and  $A_\mu^a$  are the gluon fields ( $a = 1, 2, \dots, 8$ ). We can gain further insight into the unique characteristics of QCD by considering its three distinctive attributes: 1) Confinement, 2) Asymptotic Freedom, and 3) Spontaneous Chiral Symmetry breaking.

**Confinement** refers to the fact that isolated quarks have never been observed. Instead hadronic states are always found as color neutral (or more precisely, color singlet) configurations of multiple quarks. Color singlet states can be formed as quark anti-quark pairs ( $q\bar{q}$ ) called mesons or as bound states of three quarks or anti-quarks ( $qqq$ ) called baryons. There was no direct evidence for color singlet bound states of more than 3 quarks until recently when the LHC Beauty (LHCb) experiment observed evidence for tetraquark ( $qqq\bar{q}$ ) and pentaquark ( $qqqq\bar{q}$ ) states through the decays of  $\bar{B}^0$  and  $\Lambda_B^0$  [24]. The fact that quarks are found only in color neutral bound states suggests that the interaction between quarks and gluons must be very strong at large distances. The natural length scale of the strong force is set by the size of the nucleus, dictating that it should be relatively short ranged, acting over distances of roughly 1 fm. According to its behavior at extremely short and long ranges, the static QCD potential can be expressed as:

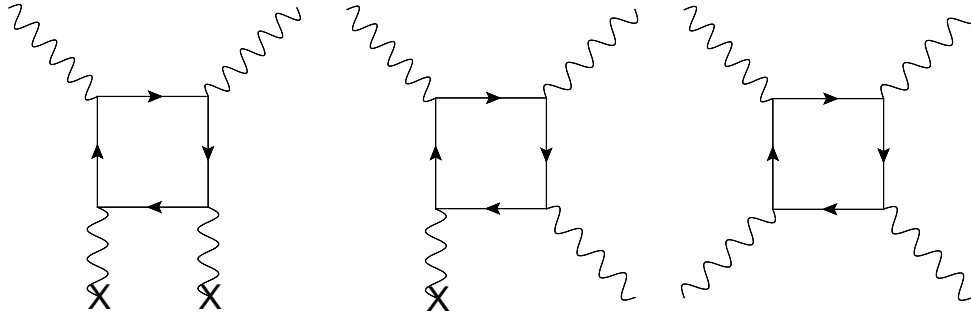


**Figure 1.6 :** Feynman diagrams for screening (left) and anti-screening (right) effects in QCD. The anti-screening interaction is only possible with gluon self-interactions.

$$V_s = -\frac{4}{3} \frac{\alpha_s}{r} + kr \quad (1.4)$$

where  $\alpha_s$  is the coupling constant of the strong force. The first term dominates at small distances and resembles the Coulomb potential between two charges in QED. At larger distances, the second term takes over causing the force between quarks and gluons to grow linearly with distance. As quarks are separated from one another, the color force between them concentrates into narrow tubes that grow stronger the more they are stretched (like a stretched rubber-band). At some distance it is more energetically favorable for a new  $q\bar{q}$  pair to be formed from the vacuum, severing the long distance tube-like color connections, than for the original quarks to continue to separate. This striking difference between QED and QCD arises from the behavior of the force carriers. In QED the force carrier is the photon which does not carry electric charge and therefore does not interact with itself (at least not directly, see Ref. [25]). However, in QCD the gluon does carry color charge and is therefore able to undergo self interactions. These self interactions give rise to an anti-screening effect which leads to the linearly growing potential as a function of distance. Figure 1.6 shows the diagrams that contribute to screening and anti-screening effects in QCD.

**Asymptotic Freedom** refers to the behavior of the QCD coupling constant  $\alpha_s$  at large momentum transfers  $Q$  and at small distances. An apparent contradiction to confinement was observed when deep inelastic scattering (DIS) experiments showed evidence that the quarks and gluons within hadrons behave as quasi-free point-like particles [26]. According to the static QCD potential in Eq. 1.4 the potential between two quarks should grow infinitely



**Figure 1.7 :** Diagrams for various  $\gamma\gamma \rightarrow \gamma\gamma$  interactions: Delbruck scattering (left), photon splitting (middle), and elastic light-by-light scattering (right). These types of interactions are forbidden in classical electromagnetism. The ATLAS detector at the LHC has recently found evidence for these types of reactions [25].

strong as  $r \rightarrow 0$  (Landau Pole). These apparent contradictions and the Landau Pole problem were resolved for QCD through the discovery of asymptotic freedom by David Gross, Frank Wilczek, and David Politzer in 1973, earning them a Nobel Prize in 2004 [27]. The effective coupling constant for the strong force can be written as:

$$\alpha_s(|q^2|) \equiv \frac{g_s^2(|q^2|)}{4\pi} \approx \frac{12\pi}{\beta_0 \ln(|q^2|/\Lambda_{QCD}^2)} \quad (1.5)$$

where  $\beta_0 = (11n_c - 2n_f)$  is a constant for a given number of colors ( $n_c$ ) and quark flavors ( $n_f$ ) active at the energy scale  $Q$ . The scale parameter,  $\Lambda_{QCD}$ , is a constant that must be determined experimentally and is found to be  $\Lambda_{QCD} \approx 250 \text{ MeV}/c$ . The behavior of  $\alpha_s$  as a function of  $Q$  can be seen in Fig. 1.5b. The value of  $\alpha_s$  drops from  $\alpha_s \sim 1$  for  $Q \sim 1 \text{ GeV}$  to  $\alpha_s \sim 0.118$  for  $Q \sim M_Z = 91.2 \text{ GeV}$  [23]. Asymptotic freedom can be succinctly summarized as an asymptotic weakening of the coupling constant as the energy scale increases or the distance scale decreases.

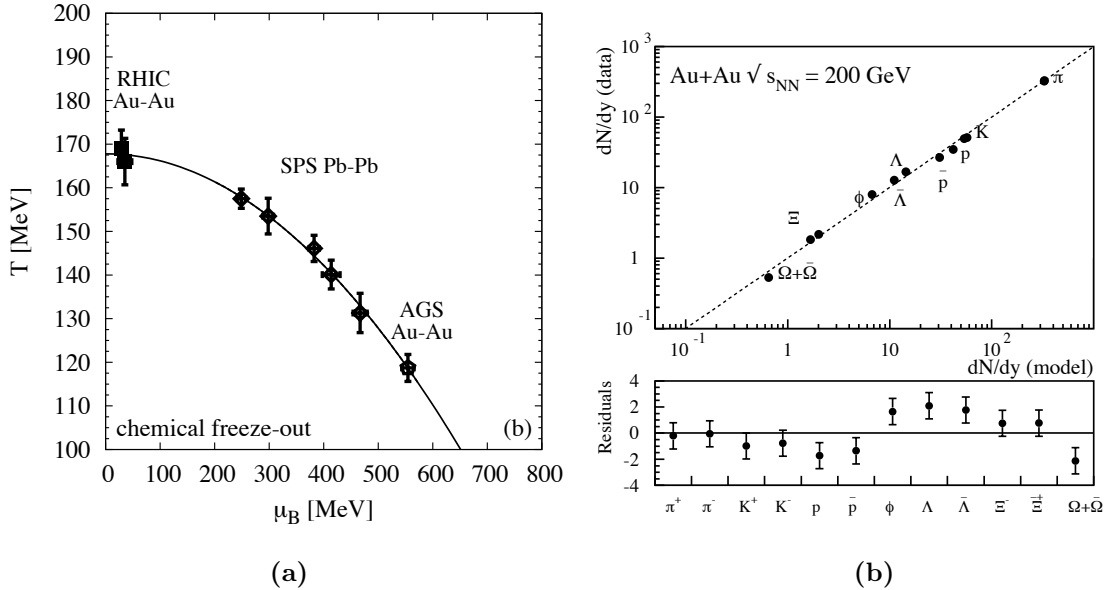
Asymptotic freedom has several important ramifications. Firstly, at high enough  $Q^2$  (or small enough distance scales) when  $\alpha_s \ll 1$  perturbative methods can be used to predict cross sections by expanding in powers of  $\alpha_s$ . This is in stark contrast to the low  $Q^2$  processes where  $\alpha_s \sim 1$  and higher order terms (in powers of  $\alpha_s$ ) will have larger and larger contributions resulting in a divergent approximation.

In the confinement regime, where  $\alpha_s \geq 1$ , perturbative QCD cannot be employed. Therefore other techniques must be found to solve QCD calculations. A large amount of success has been achieved performing QCD calculations in the low  $Q$  non-perturbative regime through a technique called Lattice QCD. Lattice QCD (LQCD) carries out calculations on a finite lattice in space and time (with lattice spacing parameter  $a$ ). In the limit of an infinite lattice with infinitely small spacing between grid-points ( $a \rightarrow 0$ ), LQCD can recover the continuum result [28]. Another technique for evaluating QCD calculations in the non-perturbative regime that has arisen out of string theory in recent years is the anti-de-Sitter/conformal field theory (AdS/CFT) correspondence [29–33]. This technique relates theories of quantum gravity to conformal quantum field theories and can be applied to some QCD-like theories.

### 1.2.3 The Quark Gluon Plasma

Another ramification of asymptotic freedom is the possibility that at extremely high temperature and/or densities quarks and gluons may become deconfined. In these extreme conditions a new state of matter would be formed. This state of matter, called the Quark Gluon Plasma (QGP) gets its name from analogy with an electric plasma. In an electric plasma, the force between bound electrons and ions becomes screened by the background cloud of other electrons and ions when the distance scales are comparable to the Debye screening length. This screening leads to a dissolution of bound states, allowing the electrons to move freely through the plasma leading to extremely high electrical conductivity. By analogy, the QGP is a plasma made of quarks and gluons in which the color charge between quarks and gluons within a hadron are screened sufficiently enough to liberate the quarks from the hadronic bonds. It is important to note that the conditions of the early universe (e.g. extremely high temperature) likely resembled those required to form a QGP.

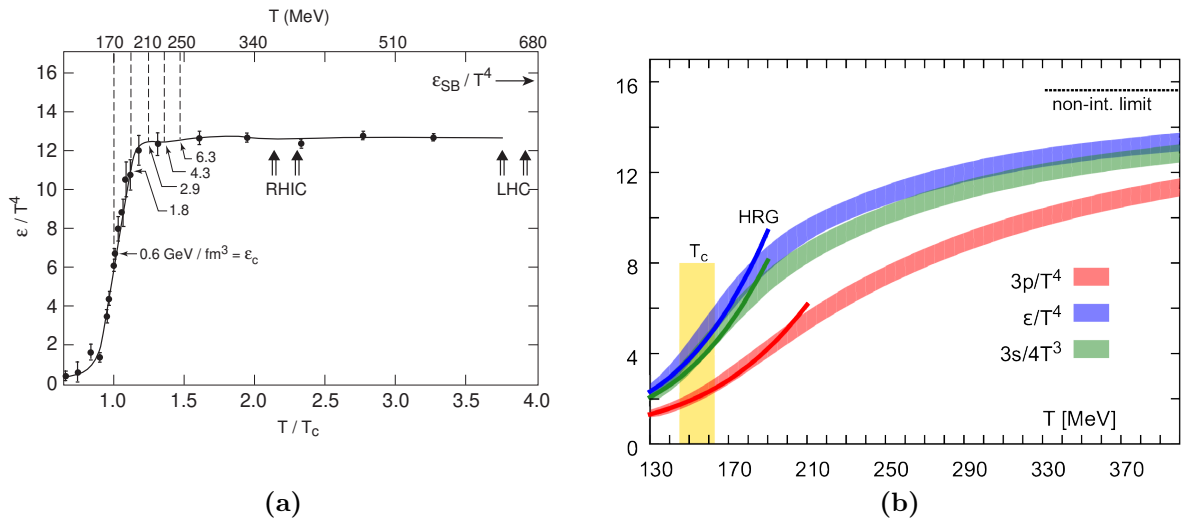
With LQCD and the QGP in mind we can now revisit the discussion of the Hagedorn Temperature, recognizing its continued importance to the study of the QGP today. In the Hadron Resonance Gas model,  $T_H$  represents the temperature of the system while energy is transferred into it to liberate excited hadronic states. Or from the QGP's point of view,  $T_H$  is the approximate freeze-out temperature at which point the fluid-like QGP evaporates into a hadron gas. Figure 1.8a shows the chemical freeze-out temperature measured in heavy ion collisions. At low baryon chemical potential ( $\mu_b$ ) the chemical freeze-out temperature is



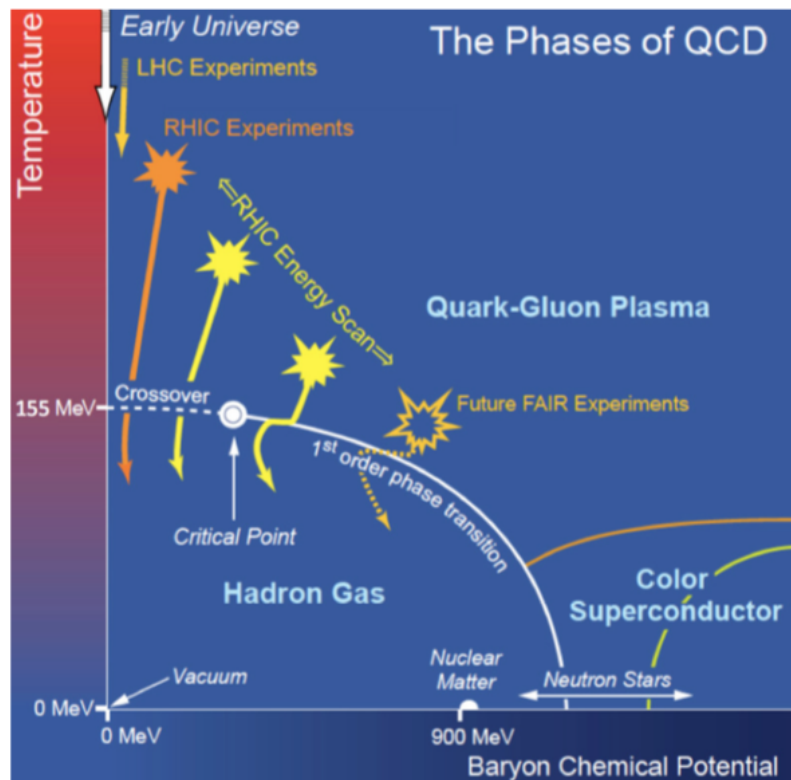
**Figure 1.8 :** The chemical freeze-out temperature versus  $\mu_b$  measured in heavy-ion collisions (a). The measured yield ( $dN/dy$ ) versus the yield from a grand canonical statistical model (b) [34]

similar to the Hagedorn Temperature. Figure 1.8b demonstrates that the yield of excited hadronic states can be well described by a statistical model [34]. Lattice QCD calculation shown in Fig. 1.9a, depict  $\varepsilon/T^4$  (proportional to the number of degrees of freedom of the system) as a function of  $T/T_c$  with values of the  $\varepsilon$  shown at various values of  $T$  [35]. The LQCD result is consistent with the HRG picture in which the number of degrees of freedom rise rapidly near some critical temperature,  $T_c \approx 150 - 180$  MeV, with the increasing number of hadronic states created. It is clear from the LQCD results that the temperature of the HRG is not actually constant as Hagedorn thought though. LQCD results indicate that the critical energy density is  $\varepsilon_c \sim 0.6 - 1$  GeV / fm<sup>3</sup> or about the energy density of a single proton inside a cube with 1 fm sides [36]. In this figure, the  $T^4$  scaling becomes apparent, requiring  $\sim 23$  times the energy density to reach  $T = 2T_c$ .

Above  $T_c$  the energy density (in units of  $T^4$ ) and the number of degrees of freedom begin to flatten as the system approaches degrees of freedom defined by deconfined quark and gluons. The LQCD calculation shown in Fig. 1.9a has almost no dependence on  $T$  above about  $T \sim 1.5T_c$ . A flattening of the energy density as a function of temperature is generally found in an ideal gas e.g in a non-interacting photon gas. For this reason, in the year 2000 when the



**Figure 1.9 :** (a) Lattice QCD calculations of the energy density in units of  $T^4$  (which scales with the number of degrees of freedom) vs. temperature [35]. The red squares show the result of the calculation with 4 temporal bins. The black circles show the result of 6 temporal bins (finer lattice spacing). (b) Normalized pressure, energy density, and entropy of the Hadron Resonance Gas (solid lines) compared to Lattice QCD (bands).



**Figure 1.10 :** An example QCD phase diagram in the temperature versus baryon chemical potential ( $T - \mu_b$ ) plane [37]. The points of reference are the limits of  $\mu_b$  i.e.  $\mu_b \rightarrow 0$ ,  $\mu_b \rightarrow \infty$  and the  $\mu_b$  of nuclear matter all at  $T = 0$ . Lattice QCD predicts a crossover transition from a hadron gas to a QGP at a temperature  $T_c \approx 155$  MeV at  $\mu_b = 0$ . The location and validity of the remaining features is largely speculative and is the source of much activity in the field.

Relativistic Heavy Ion Collider (RHIC) began operations, the community expected to find an ideal gas of nearly massless, quasi-free quarks and gluons. However, Fig. 1.9a and 1.9b both show that the  $\varepsilon/T^4$  curves do not reach the Stefan Boltzmann limit (non-interacting gas of massless quarks and gluons) indicated with arrows in the upper right of each plot. After years of study gaining evidence to the contrary at the RHIC and LHC, we now know that instead of an ideal gas, the QGP is more like an almost “perfect liquid” (defined by its shear viscosity to entropy ratio,  $\eta/s$  [38–43]). Looking back we can see that the deviation in  $\varepsilon T^4$  from LQCD calculations compared to the expectation in the Stefan Boltzmann limit is likely due to residual strong interactions between quarks and gluons.

Figure 1.10 shows an example of the QCD phase diagram in the temperature versus baryon chemical potential ( $T - \mu_b$ ,  $\mu_b \propto \rho^{2/3}$  in the Fermi gas model) plane [44]. Lattice QCD provides strong constraints along the  $T$  axis ( $\mu_b = 0$ ) suggesting that an analytic crossover transition from a hadronic gas to a QGP should occur at a critical temperature of about  $T_c \approx 155 - 180$  MeV [37, 45–47]. However, at finite values of  $\mu_b$  the LQCD technique becomes intractable due to the numerical sign problem (a rather generic difficulty when numerically integrating highly oscillatory functions of several variables) [48]. Though significant progress has been made to extend LQCD to finite  $\mu_b$ , LQCD is still not viable at intermediate baryon chemical potentials ( $\mu_b \approx 250$  MeV) in the region where a 1<sup>st</sup> order phase transition may be expected [45, 49]. If a 1<sup>st</sup> order phase transition does exist at finite  $\mu_b$ , Gibbs’ Phase Rule requires that there be a critical point connecting the smooth crossover predicted by LQCD with the 1<sup>st</sup> order phase transition. A system near a thermodynamical critical point would exhibit critical phenomena, namely the divergence of the correlation length, power law divergences of the susceptibilities, and universality to name a few. Currently there are no direct predictions from LQCD showing evidence of a 1<sup>st</sup> order phase transition and critical point. These indications are mostly motivated by other models which are arguably less robust and trustworthy than LQCD [50].

#### 1.2.4 Chiral Symmetry

In addition to the global  $SU(3)_{color}$  symmetry and the approximate  $SU(3)_{flavor}$  symmetries in the QCD Lagrangian, there exist other global and local symmetries as well. At momentum transfers of  $Q \sim 1$  GeV/ $c$  the lightest three quark flavors ( $u$ ,  $d$ , and  $s$ ) can be treated as



approximately massless. In the limit of massless quarks the Lagrangian in Eq. 1.1 (neglecting the gluon component) can be expressed as:

$$\mathcal{L} = i\bar{\psi}_f\gamma^\mu\delta_\mu\psi_f \quad (1.6)$$

where  $f = u, d, s$  indicates the quark flavor. The important characteristic of this Lagrangian is its global symmetry under *both* vector and axial-vector transformations:

$$\psi \rightarrow e^{i\alpha_V^a\frac{\lambda_a}{2}}\psi \quad (1.7)$$

$$\psi \rightarrow e^{i\gamma_5\alpha_A^a\frac{\lambda_a}{2}}\psi \quad (1.8)$$

The vector and axial-vector wave functions can be expressed in terms of their left and right chiral components as

$$\psi_{R,L} = \frac{1}{2}(1 \pm \gamma_5)\psi \quad (1.9)$$

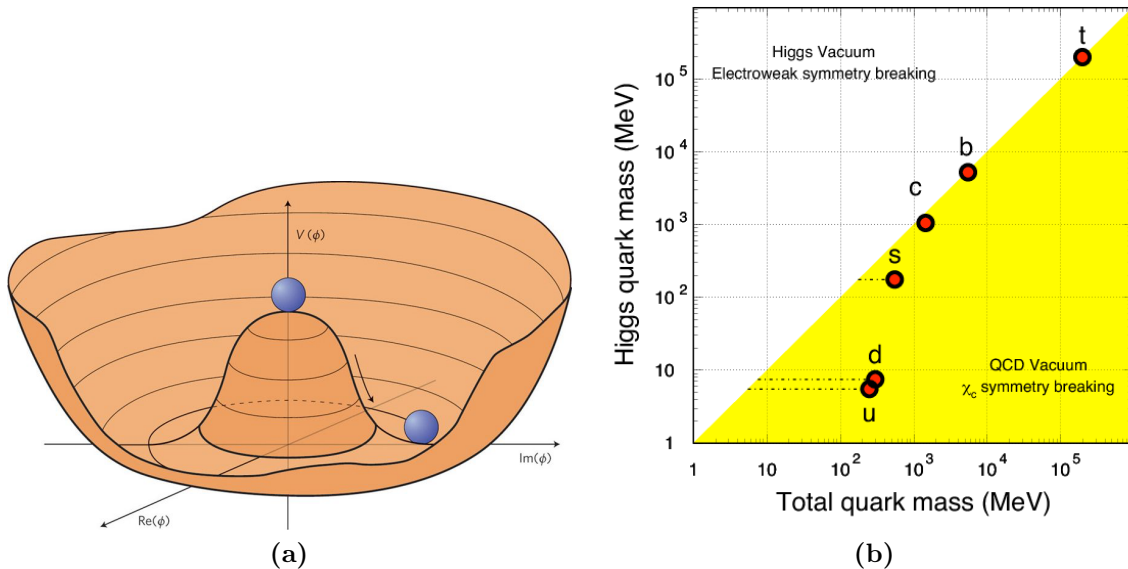
When considering the chiral forms  $\psi_R$  and  $\psi_L$  the transformations in Eq. 1.7 and 1.8 can be expressed as:

$$\psi_R \rightarrow e^{-ia_R^a\frac{\lambda_a}{2}}\psi_R, \psi_L \rightarrow \psi_L \quad (1.10)$$

$$\psi_L \rightarrow e^{-ia_L^a\frac{\lambda_a}{2}}\psi_L, \psi_R \rightarrow \psi_R \quad (1.11)$$

In this form the symmetry of the Lagrangian under vector and axial-vector transformations can clearly be seen to constitute a global symmetry with respect to the left and right-handed chiral components as well. This global  $SU(N_f)_R \times SU(N_f)_L$  symmetry in the massless quark limit is called *chiral symmetry*.

The non-zero quark masses in the QCD Lagrangian can be viewed as a perturbation of the Lagrangian ( $\delta\mathcal{L} = -m\bar{\psi}\psi$ ) in Eq. 1.6. The presence of non-zero quark masses breaks the symmetry under axial-vector transformation and therefore breaks the chiral symmetry. Symmetry breaking can arise as either explicit symmetry breaking or spontaneous symmetry breaking. In explicit symmetry breaking, the symmetry present in the Lagrangian is broken in the equations of motion. However, in spontaneous symmetry breaking, the equations of motion retain their invariance but the system's symmetry is broken because the lowest energy ground state of the system is not invariant.



**Figure 1.11** : (a) The Goldstone “Mexican Hat” potential  $V(\phi)$  [51]. (b) Comparison of the quark masses from the Higgs mechanism and from chiral symmetry breaking in QCD.

This situation can be visualized as a physical system like the one shown in Fig. 1.11a. In this picture, Goldstone’s “Mexican hat” is a symmetric potential but the state of the system (the ball in Fig. 1.11a) is very unstable [51]. Any small perturbation will cause the ball to fall into one of the infinite number of ground states in the lower energy trough. This instability results in spontaneously broken symmetry, in this case, of chiral symmetry. In the case with 3 massless quarks, we expect 8 degenerate, massless Goldstone bosons (with  $J^P = 0^-$ ). While there are in fact eight mesons with these quantum numbers ( $\pi^\pm$ ,  $\pi^0$ ,  $K^\pm$ ,  $K^0$ ,  $\bar{K}^0$ , and  $\eta$ ) the spontaneous breaking of chiral symmetry leads to non-zero masses. For instance, the lightest of these mesons, the  $\pi^0$  has mass  $m_{\pi^0} \approx 135 \text{ MeV}/c^2$  while the lightest mesons containing an  $s$  quark are the  $K^\pm$  with mass  $m_{K^\pm} \approx 494 \text{ MeV}/c^2$ . The broken chiral symmetry leads to a non-zero vacuum quark anti-quark condensate. This can be seen by considering the Gell-Mann-Oakes-Renner relation, which relates the pion mass ( $m_\pi$ ) to the quark condensate ( $\langle \bar{q}q \rangle$ ):

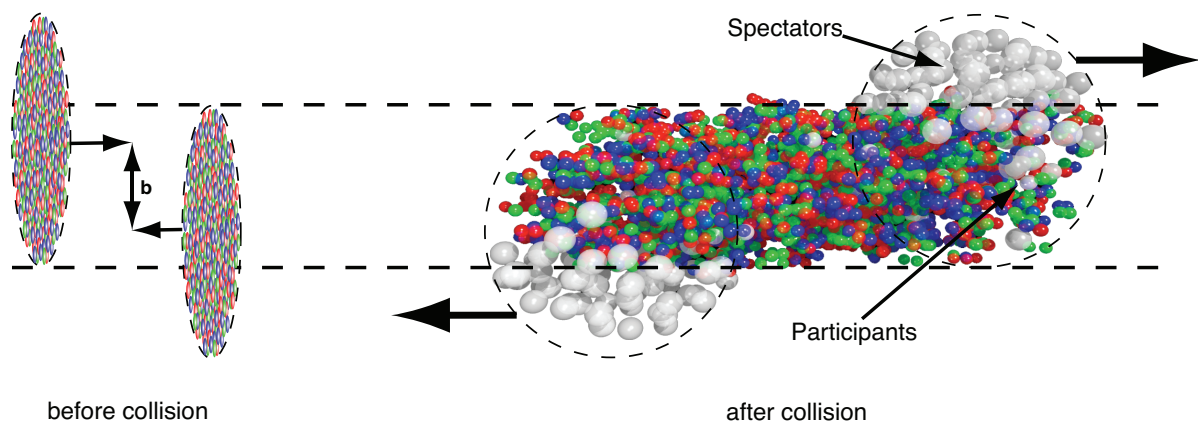
$$m_\pi^2 f_\pi^2 = -2\bar{m} \langle 0 | \bar{q}q | 0 \rangle \quad (1.12)$$

where  $\bar{m} \sim 6 \text{ MeV}/c^2$  is the mean mass of the  $u$  and  $d$  quarks. From this relation we find that the vacuum quark anti-quark condensate has a value  $\langle 0 | \bar{q}q | 0 \rangle \approx (-250 \text{ MeV})^3$ . A

non-zero quark anti-quark condensate is synonymous with chiral symmetry breaking. The effects of chiral symmetry breaking are also experimentally visible in the masses of chiral multiplets like the  $\rho^0(770)$  and the  $a_1(1260)$ . In the absence of chiral symmetry breaking these states would be degenerate. However, experimental measurements have shown that their masses are actually significantly different ( $m_{\rho^0} \approx 770 \text{ MeV}/c^2$ ,  $m_{a_1} \approx 1260 \text{ MeV}/c^2$ ). Such a large difference in the masses cannot be describe simply by the difference in the  $u$  and  $d$  quark masses but strongly suggests that chiral symmetry breaking is the source of the mass separation.

Modification of the quark anti-quark condensate is expected in a sufficiently hot ( $T > T_c^{\text{chiral}}$ ) and dense ( $n > n_c^{\text{chiral}}$ ) medium [52]. For this reason a phase transition from the hadronic phase (with spontaneous chiral symmetry breaking) to a chiral-restored phase is predicted in such environments [53]. It should be noted that the phase transition to chiral-restored QCD matter need not be coincident with the phase transition to a state of deconfined QCD matter like the QGP. In generally we may expect that the critical temperature follow  $T_c^{\text{chiral}} > T_c^{\text{QGP}}$  at a given baryon chemical potential. This implies that a QGP could exist even if chiral symmetry restoration is not achieved. On the other hand, this ordering states that chiral symmetry restoration requires deconfinement of quarks and gluons. Chiral symmetry restoration has the important experimental ramification of a re-established degeneracy in the masses of the  $\rho^0$  and  $a_1$  chiral partners. Measuring these two states, and the difference in their masses is therefore a direct experimental technique for observing chiral symmetry restoration.

At one extreme of the  $T - \mu_b$  phase diagram, such high energy density can be achieved by increasing the density of the system ( $\mu_b$ ) at  $T \approx 0$ . In this region of the QCD phase diagram, a degenerate Fermi gas of quarks with a condensate of Cooper pairs near the Fermi surface is expected [53]. Due to the presence of Cooper pairing, this phase of QCD is called a “color superconductor”. Rigorous perturbative QCD calculations in this regime are limited to the highest densities, the so-called color-flavor locked (CFL) phase. Compact stars, such neutron stars are expected to reach core densities sufficiently high to produce such a state of QCD matter. Astrophysical measurements of compact stars provide valuable constraints on the equation of state of super dense hadronic matter [54, 55].

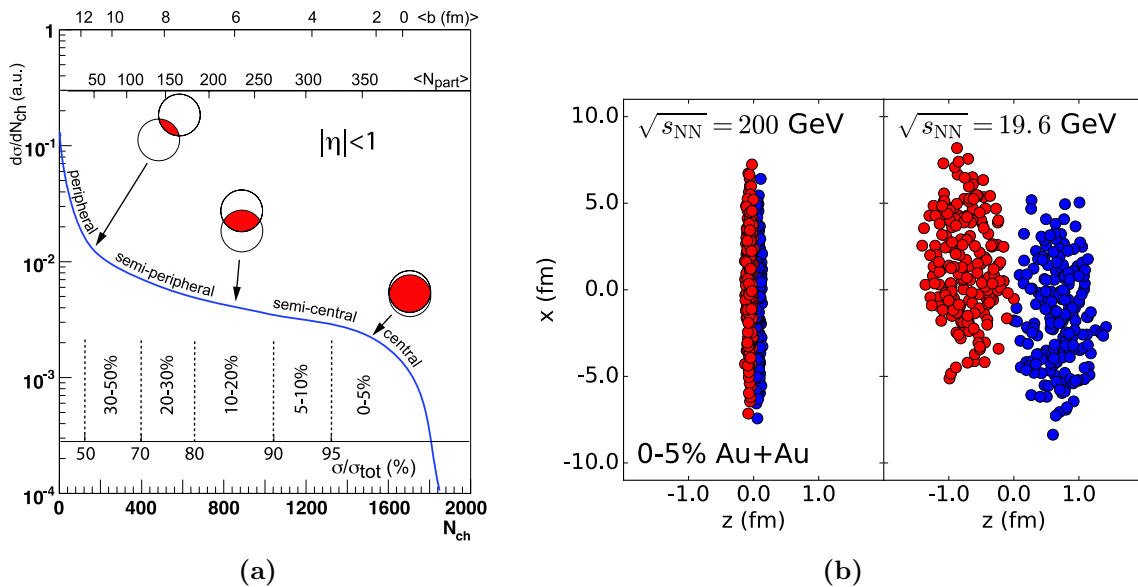


**Figure 1.12** : Left: Before colliding, length contracted ions approach with an impact parameter  $b$ . Right: During the collision, participant nucleons interact while the participant nucleons continue traveling almost undeflected in the forward and backwards directions. [56]

### 1.3 Heavy-Ion Collisions

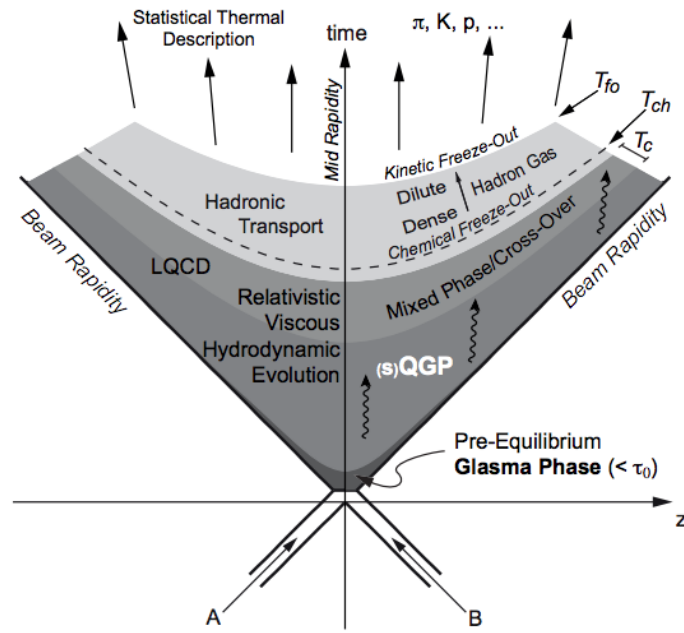
Unlike compact stars that are expected to exist at relatively low  $T$  but high  $\mu_b$ , heavy-ion collisions provide a technique for producing extremely hot QCD matter with relatively low  $\mu_b$ . Since each heavy-ion collision follows a specific trajectory through the  $T - \mu_b$  phase space, they can be used to explore the phase diagram of QCD matter. The approximate trajectory through phase space can be controlled by changing the collision species and by changing the collision energy per nucleon pair ( $\sqrt{s_{NN}}$ ). At RHIC top energies of  $\sqrt{s_{NN}} = 200$  GeV, gold ions ( $^{197}\text{Au}$ ) are accelerated to speeds greater than 99.99% the speed of light resulting in  $\gamma \approx 100$ .

At these relativistic speeds, the nearly spherical ions are length contracted into flattened discs. Each time the beams of flattened ions pass each other, there is a chance of collision with impact parameter  $b$  (see Fig. 1.13). Collisions with  $b \approx 0$  are called “central” collisions while those at the opposite extreme with  $b \approx 2R$  are called “peripheral” collisions. The nucleons that collide, called “participants” are often deflected at large angles while the non-interacting nucleons, called “spectator”, continue mostly unaffected. Since  $b$  cannot be measured directly, collisions at various impact parameters (centralities) must be selected based on the number of charged particles produced in the collision (see Fig. 1.13). Monte Carlo Glauber simulations are used to relate the number of charged particles to the initial impact parameter [57, 59]. By controlling collision energy and by selecting on centrality, the lifetime, system size, initial temperature, and  $\mu_b$  can be controlled.



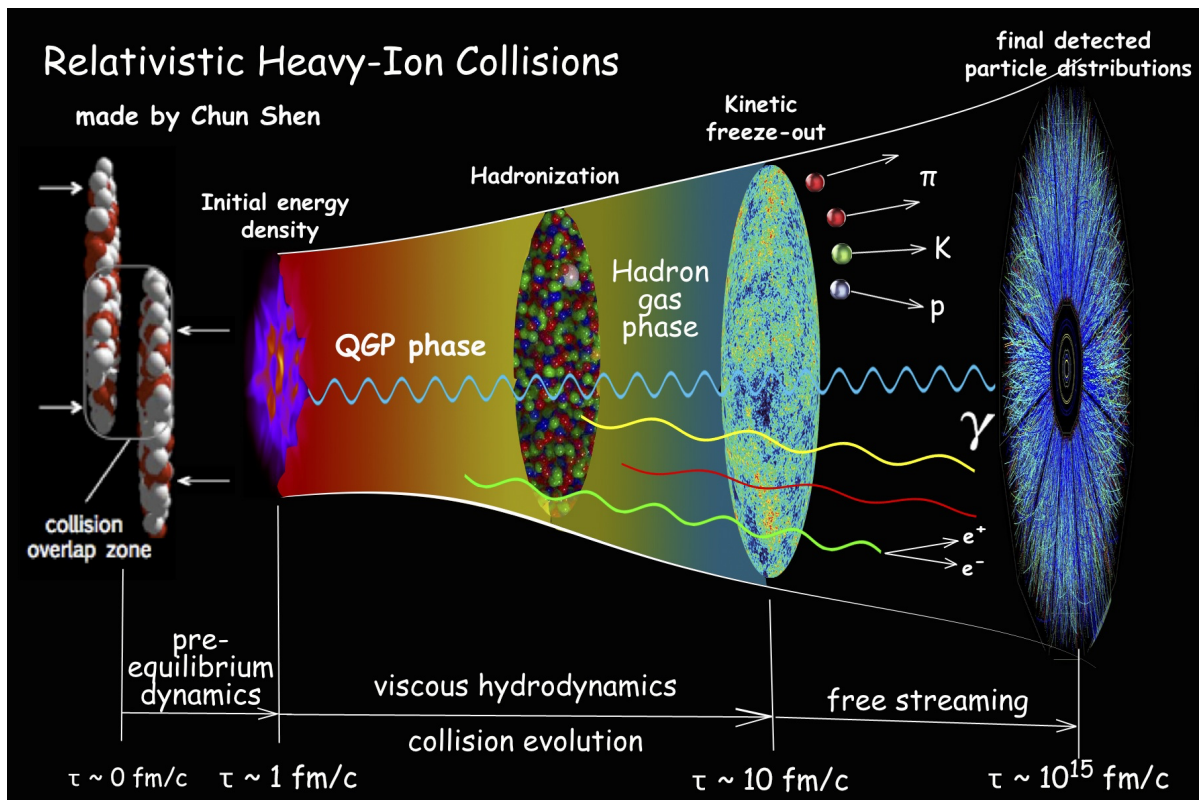
**Figure 1.13 :** (a) An example of the correlation between the geometric impact parameter  $b$  (not directly observable) and the final-state-observable charged-particle multiplicity ( $N_{ch}$ ). [57]. (b) Schematic view of the nucleon positions as a function of the longitudinal ( $z$ ) and one transverse direction ( $x$ ) shown for two different collision energies [58].

When heavy-ions collide, hot and dense matter is produced, beginning its journey through the QCD phase diagram. A space-time diagram showing the collisions various stages is shown in Fig. 1.14 while a more schematic diagram is shown in Fig 1.15. The initial collision proceeds through a combination of hard and soft scatterings between participant nucleons. Not much is known about the dynamics of the earliest moments ( $\tau < 1$  fm/ $c$ ) after the collision. These first moments, called the “pre-equilibrium” stage are described by various models such as the Color Glass Condensate (CGC) [60–62]. Within  $\approx 1$  fm/ $c$  approximate local thermalize is assumed [63] and a QGP is formed if the initial temperature and energy density are sufficiently high ( $T > T_c, \varepsilon > \varepsilon_c$ ) [63]. In recent years, increasingly sophisticated viscous hydrodynamical models have been employed to model the QGP phase of the system with great success [64]. As the system cools to below  $T_c$ , the QGP begins to freeze-out ( $\tau_{freezeout} \sim 15, 30$  fm/ $c$  at RHIC and LHC respectively) at which point a mixed phase may exist for a short time. As the system continues to cool and expand as a hadron gas, the inelastic scatterings capable of changing the particle species cease. This marks the chemical freeze-out of the system, at which point the relative yields of hadronic states are fixed (See Fig. 1.8b). After further expansion and cooling, the system reaches kinetic freeze-out when mean free path for scatters



**Figure 1.14 :** A space-time diagram of the evolution of a heavy-ion collision. [65]

is too long to allow any more interactions. After kinetic freeze-out the momenta spectra are fixed and the particles finish their journey by free-streaming to the experimental detectors. The entire evolution of a HIC must be reverse-engineered from the final-state-observables, namely the energy and momentum of long lived charged particles and the energy of neutral particles. Figure 1.15 nicely highlights the trajectories of different types of particles through the system's evolution. Photons and leptons, being inert to the strong force, are able to escape from the QCD matter with little disturbance. For this reason they make superb probes of the medium, as will be discussed in further detail in the next section.



**Figure 1.15** : A schematic of the evolution of a heavy-ion collision. [66]

## Chapter 2

### Dileptons

#### 2.1 Dilepton Production in Heavy Ion Collisions

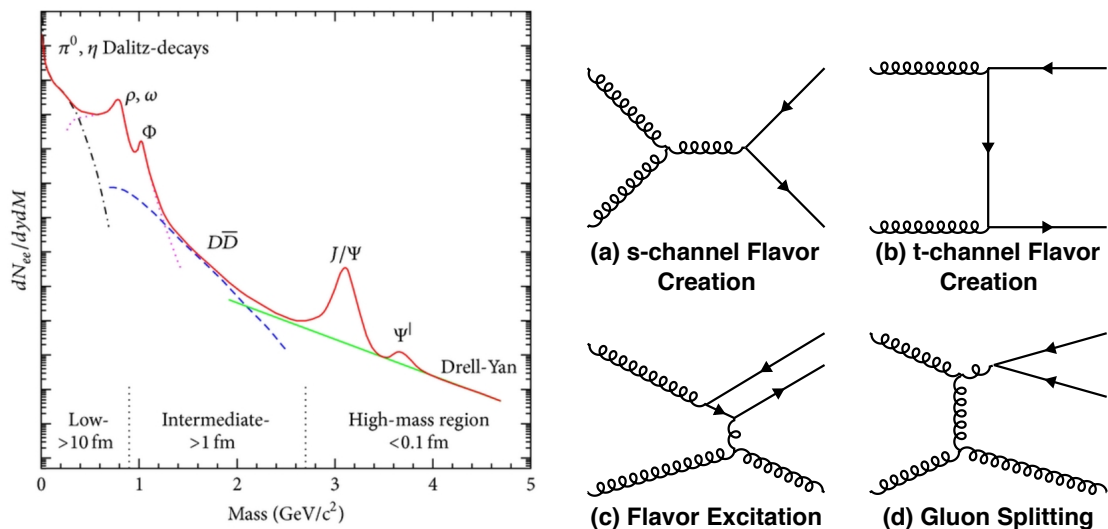
Since leptons are inert to the strong force, dileptons ( $l^+l^-$ ) make an ideal probe of the medium. Leptons interact through the EM force not the strong force, so dileptons carry pristine information about the system from even the earliest stages after the collision. The various production mechanisms responsible for dileptons can be distinguished through measurement of the invariant mass of the lepton pair. Dileptons are an especially essential tool for studying hot and dense QCD matter because the dilepton invariant mass spectra is the only experimental observation which can give direct access to measurements of an in-medium QCD spectral function [2].

The dilepton invariant mass spectra can be partitioned to isolate specific production mechanisms and time-periods of the collision. The majority of the dileptons in the high mass region (HMR,  $M_{ll} \geq M_{J/\psi}$ ,  $M_{J/\psi} = 3.096 \text{ GeV}/c^2$ ) result from the earliest stages of the collision when partons interact through hard scatters. The HMR is populated with dilepton pairs from Drell-Yan production ( $q\bar{q} \rightarrow \gamma^*/Z \rightarrow l^+l^-$ ), semi-leptonic heavy flavor ( $b\bar{b}$ ) decays, and the decay of heavy quarkonia states like the  $J/\psi$ ,  $\psi(2S)$ , and the  $\Upsilon$  states.

Quarkonium suppression due to the Debye screening of inter-quark potentials has been predicted as a direct signature of QGP formation for sometime [1, 67, 68]. The theoretical framework for the production and suppression of heavy quarkonia states is still under active investigation, but leading candidates include the Color Singlet Model (CSM) and the Color Evaporation Model (CEM) [69]. For this reason, precisely measuring the yield of heavy quarkonia states through their dilepton decays provides an important insight into the systems produced in HICs and the properties of the QGP.

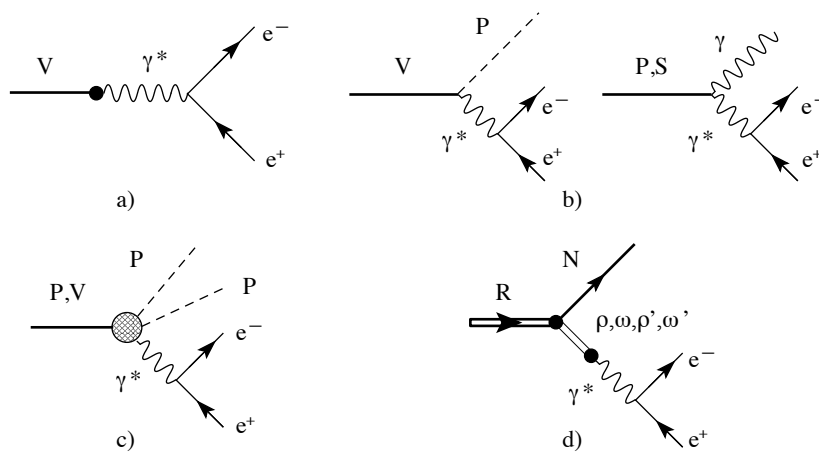
In HICs at RHIC and LHC energies, the intermediate mass range (IMR,  $M_\phi < M_{ll} < M_{J/\psi}$ ,  $M_\phi = 1.02 \text{ GeV}/c^2$ ) is dominated by production from correlated open heavy flavor decays ( $c\bar{c} \rightarrow l^+ + l^- + X$ ). Models also predict significant dilepton production from the





**Figure 2.1 :** The Feynman diagrams for flavor creation (a,b), flavor excitation (c), and gluon splitting (d) [70–72]

medium’s thermal radiation, resulting from  $q\bar{q}$  annihilation. However, separating the thermal radiation from the open heavy flavor decays is very challenging at RHIC and LHC energies. At lower energies like the SPS energy regime, the thermal sources and open heavy flavor contributions can be adequately separated. Other measurements in HICs at RHIC and LHC energies have shown evidence for substantial modification to the open heavy flavor production cross section [73]. This makes separating thermally produced dileptons in the IMR double challenging, since the medium induced modifications to the open heavy flavor contribution must be taken into account. One technique that has been proposed is to measure the  $c\bar{c} \rightarrow e + \mu$  spectra in order to constrain the  $c\bar{c}$  contribution in the  $l^+l^-$  decay channels. Another difficulty arises since the various possible  $c\bar{c}$  production mechanisms can lead to varying strength azimuthal correlations between the final state dileptons. Each of the open heavy flavor production Feynman diagrams shown in Fig. 2.2 contribute differently to the azimuthal correlations of the final state dileptons. Even in vacuum production (e.g in  $p + p$  collisions) the relative contributions from these processes is not well constrained [70]. Any change in the amount of correlation / de-correlation between the final state leptons directly affects the invariant mass density of the dileptons produced by open heavy flavor decays. Therefore, understanding the amount of correlation / de-correlation in-medium is important for extracting the dileptons from thermal radiation.



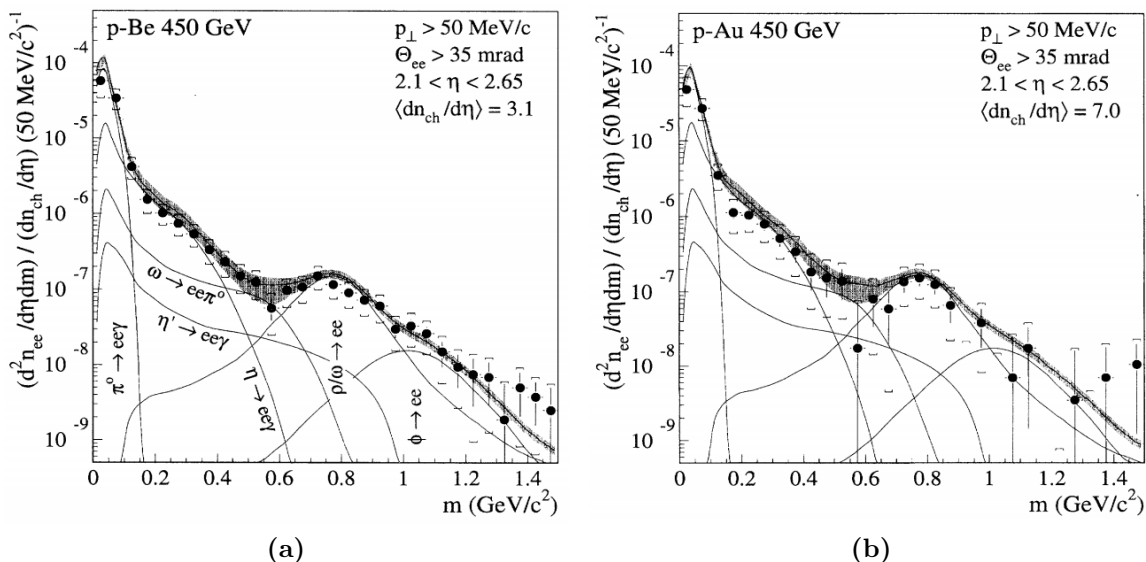
**Figure 2.2 :** The Feynman diagrams for various dilepton production mechanisms (shown for  $e^+e^-$ ): the direct decay of a vector meson ( $\rho, \omega, \phi$ ) through a virtual photon (a); the Dalitz decay of a vector (V), pseudo-scalar (P), or scalar (S) meson into a neutral particle and an  $e^+e^-$  pair (b); the four-body decay of a pseudo-scalar or vector meson into  $e^+e^-$  and two pseudo-scalar mesons ( $\pi$  or  $\eta$ ) through an intermediate state containing a virtual photon or vector meson (c); the decay of a nucleon or  $\Delta$  resonance into a nucleon and a vector meson which further decays into an  $e^+e^-$  pair. [74]

Measurement of the in-medium  $\rho^0$  meson offers an important link to chiral symmetry restoration in the QGP. Since the lifetime of the  $\rho^0$  ( $\tau_\rho \approx 1.3 \text{ fm}/c$ ) is so much shorter than the lifetime of the QGP ( $\tau_{QGP} \approx 10 \text{ fm}/c$  at top RHIC energies in central Au+Au collisions), the  $\rho^0$  spectral shape is significantly modified by interactions with the medium (mostly due to very strong coupling via  $\pi^+ + \pi^- \rightleftharpoons \rho$  channel). As discussed in Sec. 1.2.4, when chiral symmetry is restored, the masses of the  $\rho^0(770)$  and the  $a_1(1260)$  should become degenerate. The  $a_1$  is notoriously difficult to measure because, being an axial-vector meson, it decays preferentially to  $3\pi$  states. Isolating the dileptons from the  $\rho^0 \rightarrow l^+ + l^-$  decays is also challenging due to the significant background from the  $l^+l^-$  decays and Dalitz decays of long-lived hadronic states ( $\omega, \phi, \eta$  etc.) in the low mass region (LMR,  $M_{ll} \leq M_\phi$ ). The contribution from these hadron decays must be removed in order to isolate the in-medium  $\rho^0$ -meson invariant mass distribution. The invariant mass distribution of all known hadronic decays, often referred to as the “hadronic cocktail”, is generally simulated using the yields of each hadron measured through other decay channels. Finally, since the LMR region is populated with dileptons throughout the entire evolution of the system (both before and after  $T = T_c$ ), the excess yield over the hadronic cocktail is directly related to the lifetime of the emitting medium. The excess yield over the hadronic cocktail can therefore provide a

means for measuring the lifetime of the system [2].

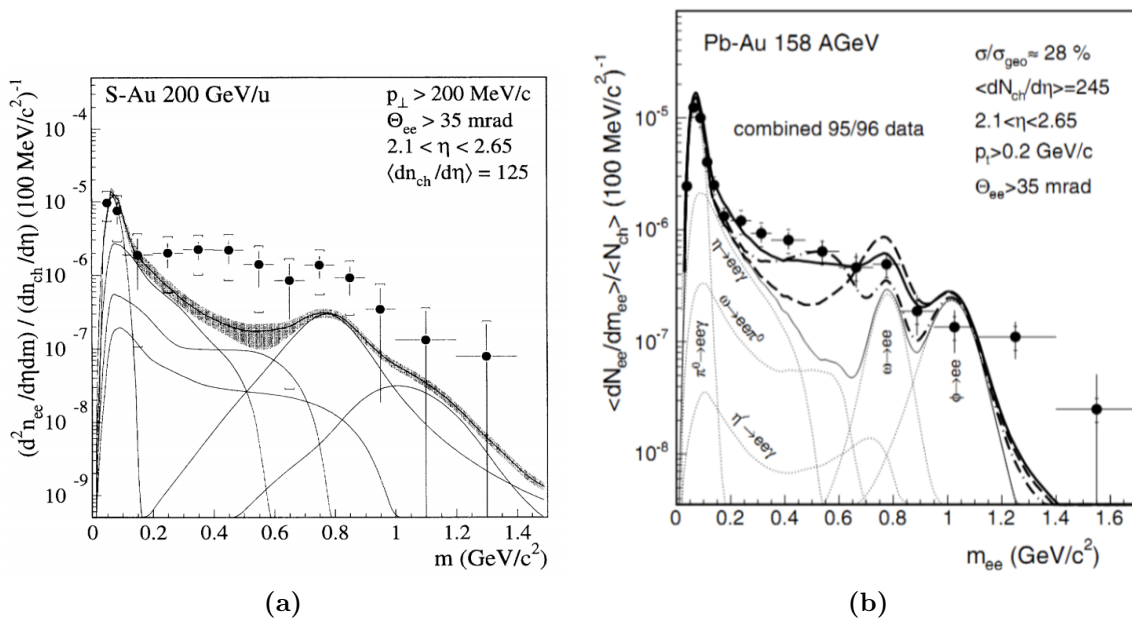
## 2.2 Survey of Past Dilepton Measurements

A wealth of dilepton measurements in various nuclear environments has been conducted over the years. For instance, the CERES/NA45 [75–77], HELIOS-3 [78], NA38/50 [79, 80] experiments at SPS; HADES at SIS18 [81–83]; PHENIX [84, 85] and STAR at the RHIC [86–89]; ALICE at the LHC [90, 91]. In this section, I will briefly survey the measurements (by experiment in roughly chronological order) that have marked milestones in our understanding and capabilities.



**Figure 2.3 :** Inclusive  $e^+e^-$  invariant mass spectra for  $p$ +Be (a) and  $p$ +Au (b) collisions at 450 GeV. The data points are shown in filled circles while the hadronic cocktail is shown in black curves [92]. The systematic uncertainty on the hadronic cocktail is shown as a shaded band.

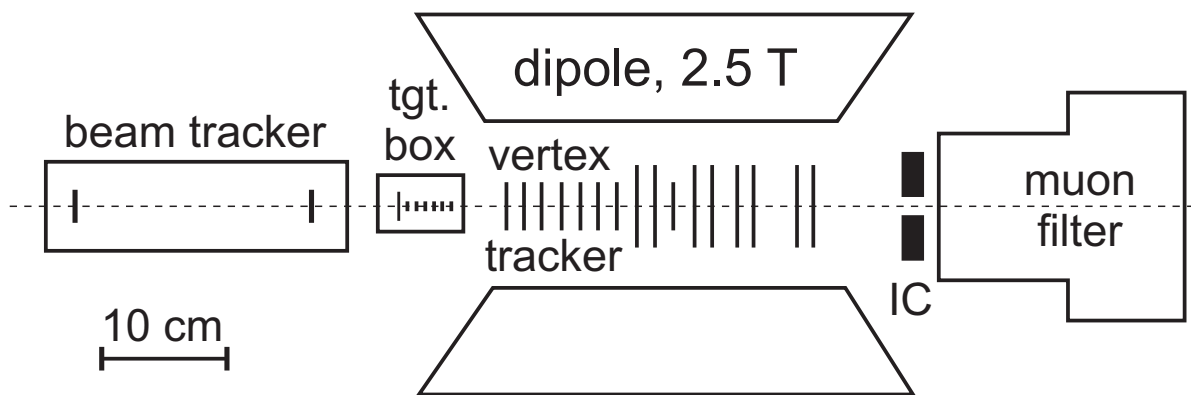
**Cherenkov Ring Electron Spectrometer (CERES)**, also known as NA45, was a fixed target experiment at the CERN Super Proton Synchrotron [75]. As the name suggests, the CERES experiment was an electron spectrometer dedicated to the study of low-mass  $e^+e^-$  pairs in nuclear collisions. Electron identification was carried out with two symmetric ring imaging Cherenkov detectors (RICH) [93]. Early measurements from CERES in  $p$ +A collisions demonstrated the capability of the detector for dielectron measurements. Figure 2.3 shows the CERES measurement of  $e^+e^-$  pairs in  $p$ +Be and  $p$ +Au collisions at proton beam energies of 450 GeV. Dilepton measurements are challenging in part due to the physical backgrounds and the immense number non-physical combinatorial pairs that must be removed.



**Figure 2.4 :** Inclusive  $e^+e^-$  invariant mass spectra in S+Au collisions at 200 A GeV (a) and in Pb+Au collisions at 158 A GeV. The dashed line shows the cocktail+vacuum  $\rho$  contribution. Models of the thermal radiation are shown for two competing scenarios: the dropping  $\rho$  mass (dot-dashed line), and the broadened  $\rho$  (solid line). While the vacuum  $\rho$  is disfavored by data, within the statistical uncertainties, the two competing thermal models cannot be distinguished [92].

The remaining real dilepton pairs result mostly from Dalitz decays of light neutral mesons. The measurement shown in Fig. 2.3 demonstrates that, after taking the various background sources into account, the yield of  $e^+e^-$  production can be well described by the expectation from the hadronic cocktail. These measurement in  $p$ +Au and  $p$ +Be collisions also demonstrate that initial state cold nuclear matter affects can be well described by the hadronic cocktail.

In addition to the  $p$ +Be and  $p$ +Au measurements, the CERES experiment also measured the  $e^+e^-$  invariant mass spectra in heavy-ion collisions. Figure 2.4a shows the  $e^+e^-$  invariant mass spectra in S+Au collisions at 200 A GeV. In this measurement a very clear excess is observed over the hadronic cocktail. The nature of this excess was further investigated by measuring the  $e^+e^-$  invariant mass spectra in Pb+Au collisions at 40 and 158 A GeV [77]. In this dataset the excess over the cocktail was found to scale strongly with centrality and to result mostly from low  $p_T^{ee}$  pairs. Several models attempted to explain this excess by incorporating the thermal radiation of the medium due to  $\pi^+ + \pi^- \rightleftharpoons \rho \rightarrow e^+ + e^-$ . However,

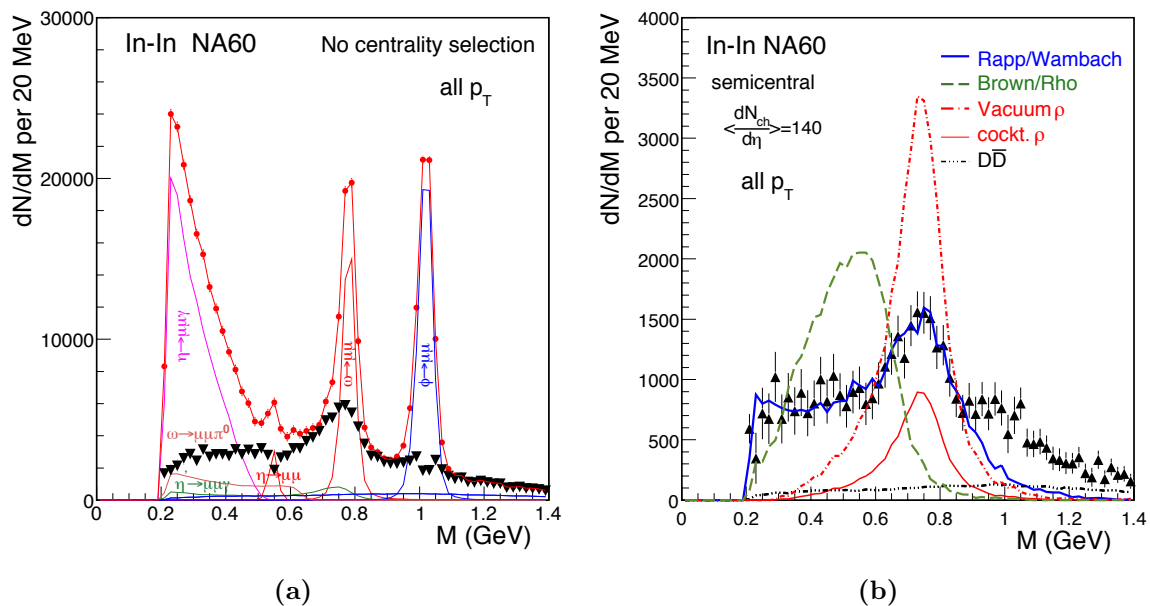


**Figure 2.5 :** The NA60 detector setup near the target. The addition of the high granularity silicon-based vertex tracker was crucial for the NA60  $\mu^+\mu^-$  measurements. [97]

models using the vacuum  $\rho$  spectral function were not able to adequately describe the excess yield. Instead, two different models for the in-medium  $\rho$  spectral functions were proposed that approximately equally described the data: a dropping  $\rho$  mass scenario [94], and a broadened  $\rho$  scenario [95]. The precision offered by the CERES  $e^+e^-$  invariant mass spectra in Pb+Au did not allow these two competing scenarios to be distinguished. The HELIOS-3 experiment measured a similar excess with respect to the hadronic cocktail in S+W collisions at 200 GeV through the  $\mu^+\mu^-$  channel [96]. Unfortunately, their precision was also too limited to distinguish between the dropping mass and broadening scenario.

**The NA60 experiment** was a fixed target experiment at SPS. NA60 was specifically designed to study the deconfinement phase transition in nuclear matter with a novel setup. In addition to its high quality muon spectrometer (inherited from NA38/50 [97]) NA60 also included a fast-readout, high granularity, silicon-based vertex tracker inside a 2.5 T dipole field. The setup of the vertex tracker can be seen in Fig. 2.5. Matching between the muon spectrometer hits and the vertex tracker hits allowed NA60 to achieve very good mass resolution on the order of  $20 \text{ MeV}/c^2$ . The precise tracking ( $\sigma_x < 10 \text{ }\mu\text{m}$  and  $\sigma_y < 15 \text{ }\mu\text{m}$ ) allowed each track's vertex offset to be measured. Precise vertex offset measurement is important for two reasons. First, it allows the substantial backgrounds from weak decays of  $\pi^\pm \rightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$  and  $K^\pm \rightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$  to be rejected. Second, precise vertex offset measurement is essential for separating the prompt (Drell-Yan and thermal radiation) and non-prompt (open heavy flavor decays) sources of dimuons. With this setup, NA60 was able to achieve a very high quality measurement of the low-mass  $\mu^+\mu^-$  invariant mass spectra

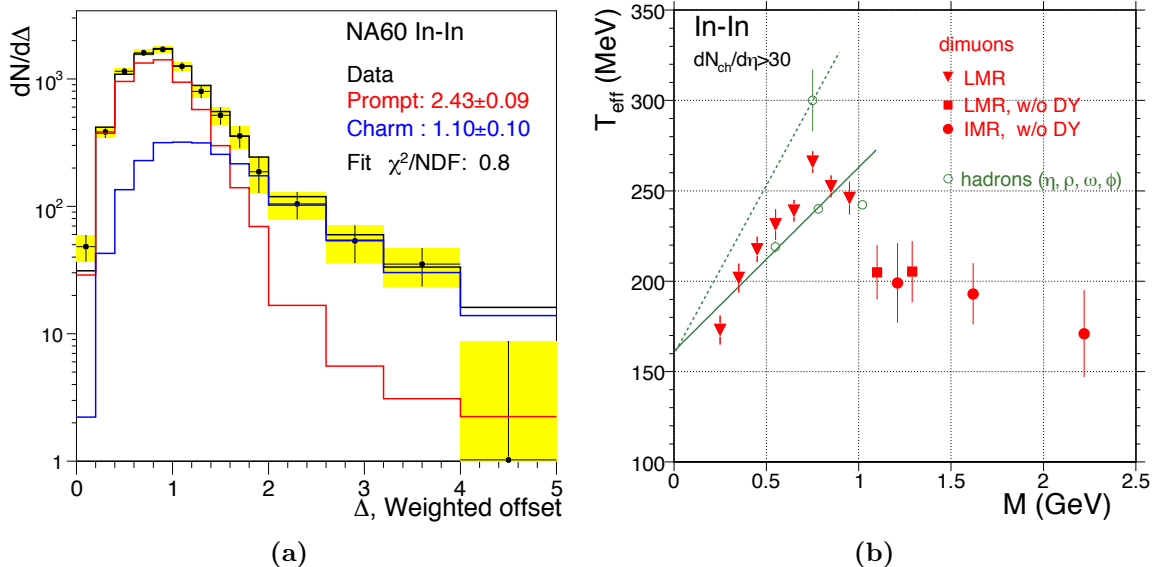
in In+In collisions at  $\sqrt{s_{NN}} = 17.3$  GeV. With this data set, NA60 had sufficient statistics to investigate the low-mass  $\mu^+\mu^-$  invariant mass spectra differentially in centrality. They found that the hadronic cocktail (including a vacuum  $\rho$  contribution) describe the spectra from peripheral collisions quite well [98]. However, an excess showing strong centrality and  $p_T^{\mu\mu}$  dependence was observed in semi-central collisions.



**Figure 2.6 :** (a) The  $\mu^+\mu^-$  invariant mass spectra in the acceptance of NA60 for In+In collisions without centrality selection before (red circles) and after (black triangles) subtracting the hadronic cocktail (without  $\rho$ ). (b) The excess  $\mu^+\mu^-$  invariant mass spectra in semi-central collisions for the acceptance of NA60. The predicted excess spectra is shown for several theoretical models.

Figure 2.6a shows the  $\mu^+\mu^-$  invariant mass spectra without any centrality selection after subtracting the background from combinatorial and physical sources. The  $\mu^+\mu^-$  invariant mass spectra is shown before (red circles) and after (black triangles) subtracting the hadronic cocktail of all relevant sources except the  $\rho^0$  meson. Figure 2.6b shows the  $\mu^+\mu^-$  invariant mass spectra in semi-central collisions after subtracting the cocktail contributions. Figure 2.6b also shows representative models for several scenarios: the cocktail  $\rho$ , the scaled vacuum  $\rho$ , the dropping  $\rho$  mass scenario, and the  $\rho$  broadening model. Unlike the CERES and HELIOS-3 measurements, the NA60  $\mu^+\mu^-$  measurement provides more than sufficient precision for separating between the dropping  $\rho$  mass scenario and the  $\rho$  broadening scenario.

This measurement clearly shows that the dropping  $\rho$  mass scenario is strongly disfavored by the data while the  $\rho$  broadening model describes the region from  $0.2 < M_{\mu\mu} < 0.8 \text{ GeV}/c^2$  reasonably well.



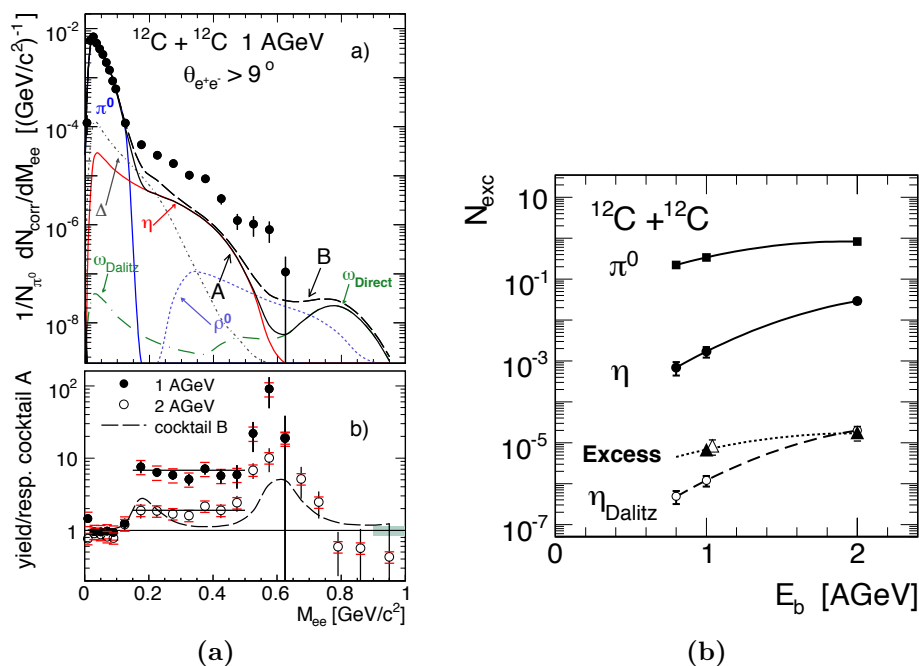
**Figure 2.7 :** (a) The  $\mu^+\mu^-$  invariant mass spectra in the acceptance of NA60 for In+In collisions without centrality selection before (red circles) and after (black triangles) subtracting the hadronic cocktail (without  $\rho$ ). (b) The excess  $\mu^+\mu^-$  invariant mass spectra in semi-central collisions for the acceptance of NA60. The predicted excess spectra is shown for several theoretical models.

An excess at higher masses ( $M_{\mu\mu} > 0.8 \text{ GeV}/c^2$ ) that was not consistent with any of the in-medium modified  $\rho$  scenarios was also observed. The open heavy flavor decay of  $c\bar{c} \rightarrow \mu^+ + \mu^-$  is the dominant contribution in this mass region. The charm cross section used for the hadronic cocktail was determined by appropriately scaling NA50's measurements of the charm cross section in  $p+\text{Al}$ ,  $p+\text{Ag}$ ,  $p+\text{Cu}$ , and  $p+\text{W}$  collisions [99]. The precise vertex offset measurement was essential for determining whether or not the observed excess resulted from an enhanced open heavy flavor contribution. By fitting the vertex offset distribution, shown in Fig. 2.7a, it was clear that the observed excess originated from prompt dimuons, not additional open heavy flavor decays. At this point, the next logical step was to investigate the possibility that the observed excess resulted from enhanced Drell-Yan production. By comparing the  $p_T^{\mu\mu}$  spectra from the excess with the  $p_T^{\mu\mu}$  spectra of Drell-Yan pairs, this possibility was ruled out. Instead, the  $p_T^{\mu\mu}$  spectra from the excess was found to be more



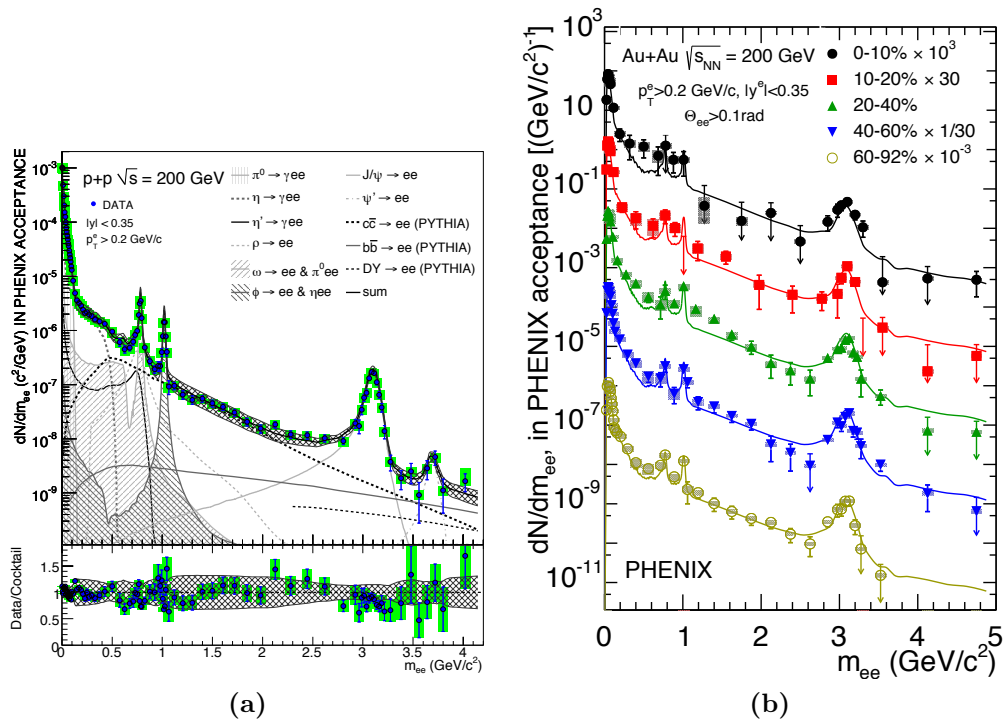
similar to the spectra from open heavy flavor decays than Drell-Yan.

The high statistics NA60 data also allowed a differential study of the  $m_T$ -spectra ( $m_T = (p_T^2 + M^2)^{1/2}$ ) for various invariant mass bins. Each mass bin was fit with the form  $1/m_T dN/dM \propto \exp(-m_T/T_{eff})$ , where the inverse slope parameter ( $T_{eff}$ ) is interpreted as the effective temperature of the medium. The value of  $T_{eff}$  (shown in Fig. 2.7b) was found to depend strongly on the invariant mass, rising at low invariant masses (consistent with the  $m_T$ -spectra of hadrons) before suddenly dropping for  $M_{\mu\mu} > 1.0 \text{ GeV}/c^2$ . The combination of the sudden drop in  $T_{eff}$  along with the other clues suggested that a hadron-like source could not explain the observed excess in the intermediate mass region. Instead, a more natural explanation is that of prompt thermal emission from a partonic source, i.e.  $q\bar{q} \rightarrow \mu^+\mu^-$ .



**Figure 2.8 :** (a) Top panel: The HADES measurement of the  $e^+e^-$  mass spectra in C+C collisions at 1 A GeV incident energy compared with expected sources. (a) Bottom panel: The ratio of the measured yield in 1 and 2 A GeV collisions to the respective cocktail A at each energy. (b) The inclusive multiplicity of the excess yield in the mass range  $0.15 < M_{ee} < 0.5 \text{ GeV}/c^2$ . The solid curves show the excitation functions for  $\pi^0$  and  $\eta$ . The dotted curve is the excitation function for the  $\pi^0$  scaled down by an arbitrary factor while the dashed curve shows the scaled contribution of the  $\eta$  excitation function [100].

**The High-Acceptance DiElectron Spectrometer (HADES)**, is a fixed target experiment operated at GSI with beams provided by the SIS18 synchrotron [101]. HADES has measured the  $e^+e^-$  invariant mass spectra in multiple collision systems, but most notably



**Figure 2.9** : First measurement of the  $e^+e^-$  invariant mass spectra. The hadronic cocktail describes the production over the full mass range. At RHIC energies the intermediate mass region is dominated by pairs from correlated open heavy flavor decays(a). The first PHENIX measurement of the  $e^+e^-$  invariant mass spectra in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV showed a very large excess in central collisions at low mass.

in  $^{12}\text{C}+^{12}\text{C}$  at 1 and 2 A GeV [81, 100]. The  $e^+e^-$  invariant mass spectra in the HADES acceptance is shown for 1 A GeV collisions in Fig. 2.8a. The comparison shown in 2.8a include a second hadronic cocktail with  $\Delta$  resonances and the  $\rho$  meson. Even with these additional components, the excess is not well described. Additionally, the enhancement over the cocktail is found to be stronger at 1 A GeV than at 2 A GeV. The scaling of the excess in  $0.15 < M_{ee} < 0.5$  GeV/ $c^2$  is shown in Fig. 2.8b as a function of the beam energy. For comparison, the excitation functions of the  $\pi^0$  and  $\eta$  are also shown. The evolution of the excess yield with  $E_b$  is found to be roughly consistent with the excitation for  $\pi^0$  but inconsistent with  $\eta$  Dalitz decay suggesting that the excess may be a result of light resonances like the  $\Delta$  contributing to the  $\pi^0$  yield. A more concrete conclusion requires further comparisons with models incorporating various affects.

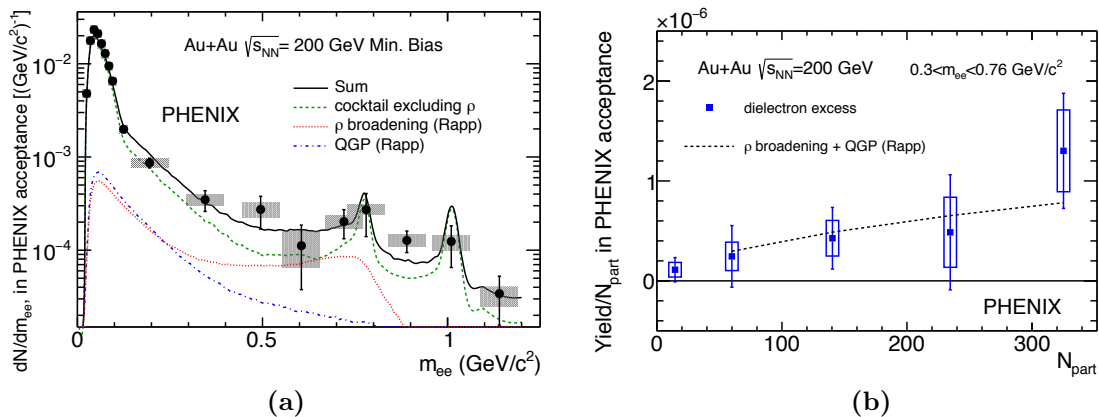
**The Pioneering High Energy Nuclear Interaction eXperiment (PHENIX)** is one of the two large collider experiments at RHIC. PHENIX was specifically designed with the

measurement of electrons, muons and photons in mind. A detailed description of the detector and its subsystems can be found in Ref. [102–104]. PHENIX has electron identification capabilities provided by a RICH detector and an electromagnetic calorimeter [105].

The dielectron spectra was first measured by PHENIX in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV. Figure 2.9a shows the measured inclusive  $e^+e^-$  invariant mass spectra in the PHENIX acceptance along with the cocktail of expected sources. The cocktail, which includes a very sizable  $c\bar{c}$  contribution at intermediate masses, is found to describe the spectra reasonably well across the entire mass range ( $M_{ee} < 4.0$  GeV/ $c^2$ ). The Pythia event generator (v6.42) was used to simulate the contributions from  $c\bar{c}$ ,  $b\bar{b}$ , and Drell-Yan.

PHENIX later measured the  $e^+e^-$  invariant mass spectra in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV [85,106]. A sizable excess over the hadronic cocktail was observed in the low mass region. The excess observed in the first PHENIX measurement of  $e^+e^-$  pairs in Au+Au collisions [85] was qualitatively consistent with the excess observed by NA60, but the yield was much larger than predicted. A later measurement [106] included updated techniques for rejecting hadron contamination and showed an excess consistent with STAR’s recent measurements in Au+Au collisions at the same energy [88]. Figure 2.10b shows that the excess observed by PHENIX scales strongly with centrality as NA60 observed. The dielectron yield in the low mass region for PHENIX acceptance is shown in Fig. 2.10a along with the expected contributions from the cocktail (excluding  $\rho$ ) and a model including  $\rho$  broadening and QGP thermal radiation. The uncertainties do not allow a strong statement to be made, but the excess is found to scale with  $N_{\text{part}}$  consistently with the broadened  $\rho$ +QGP model.

Recently, PHENIX has also published  $\mu^+\mu^-$  invariant mass spectra from  $p + p$  collisions at  $\sqrt{s} = 200$  GeV [70,107] seen in Fig. 2.11. In this measurement they sought to provide greater insight into the production mechanisms for dilepton pairs from open heavy flavor decays. Each of the diagrams shown in Fig. 2.2 result in different azimuthal correlations between the final state leptons. Changes in the correlations are important because they affect the density of pairs as a function of pair mass. In these measurements, PHENIX isolated  $c\bar{c}$  production via  $\mu^+\mu^-$  pairs in the intermediate mass region and  $b\bar{b}$  production via the high mass like-sign ( $\mu^\pm\mu^\pm$ ) pairs. With these measurements, PHENIX placed limits on the total charm and bottom cross sections. Figure 2.12 shows fits to the  $\Delta\phi$  performed to determine the relative contribution for the different production mechanisms which lead to



**Figure 2.10** : Comparison between the measured excess of  $e^+e^-$  pairs and the prediction from the broadened  $\rho + QGP$  model (a). The centrality dependence of the  $e^+e^-$  excess yield. [106]

different final state correlations.

**The Solenoidal Tracker at RHIC (STAR)** is one of the two large collider experiments at RHIC. Its dielectron program began in earnest with the installation of the Time-of-Flight (TOF) detector. By combining  $dE/dx$  information from the main tracking detector and  $\beta^{-1}$  measurements from the TOF detector, STAR is able to achieve  $94.6 \pm 2$  % electron purity in minimum bias Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV [88].

STAR first measured the  $e^+e^-$  invariant mass spectra from  $p + p$  collisions at  $\sqrt{s} = 200$  GeV collected in 2009. A later, significantly higher precision measurement shown in Fig. 2.13a was made in the same collision system and energy using the substantially larger data set collected in 2013 [109]. In these measurements the inclusive  $e^+e^-$  invariant mass spectra is found to be in good agreement with the sum of hadronic and open heavy flavor (from PYTHIA) sources. STAR has also measured the  $e^+e^-$  invariant mass spectra in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV using data collected in 2010 and 2011 (shown in Fig. 2.13b) [86, 110]. The excess yield observed by STAR within the  $\rho$ -like mass region ( $0.30 < M_{ee} < 0.76$   $GeV/c^2$ ) at  $\sqrt{s_{NN}} = 200$  GeV is  $1.76 \pm 0.06$  (stat)  $\pm 0.26$  (systematic)  $\pm 0.29$  (cocktail). The excess yield from the updated PHENIX results and the STAR measurement agree within uncertainties. The excess yield is compared with an effective many-body model and a dynamical microscopic transport model [95, 111–114].

STAR has systematically studied the  $e^+e^-$  production in Au+Au collisions as a function

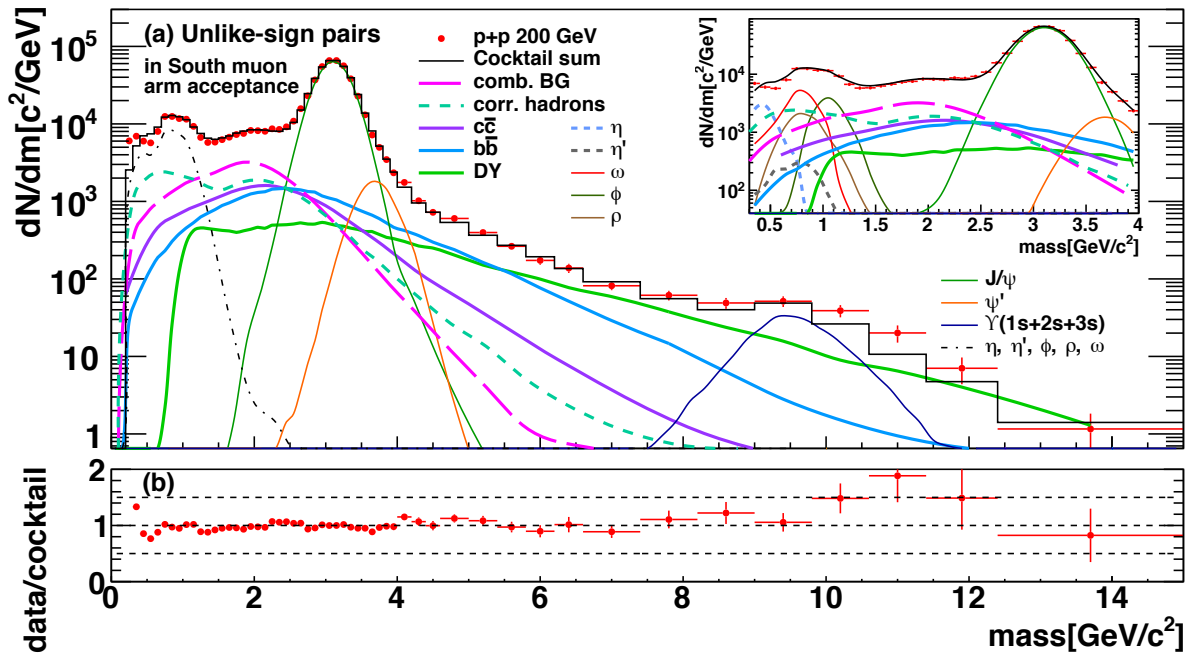


Figure 2.11 : The inclusive  $\mu^+\mu^-$  invariant mass distribution in  $p+p$  collisions at  $\sqrt{s}=200$  GeV within the PHENIX south arm acceptance. The lower panel shows the ratio of the measured yield to the hadronic cocktail [108].

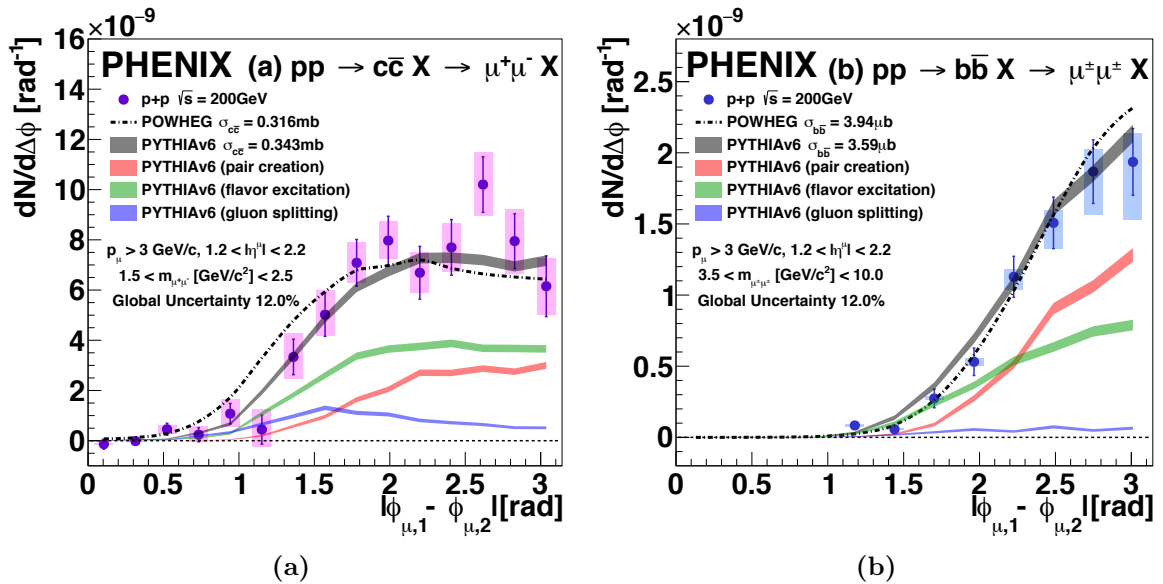
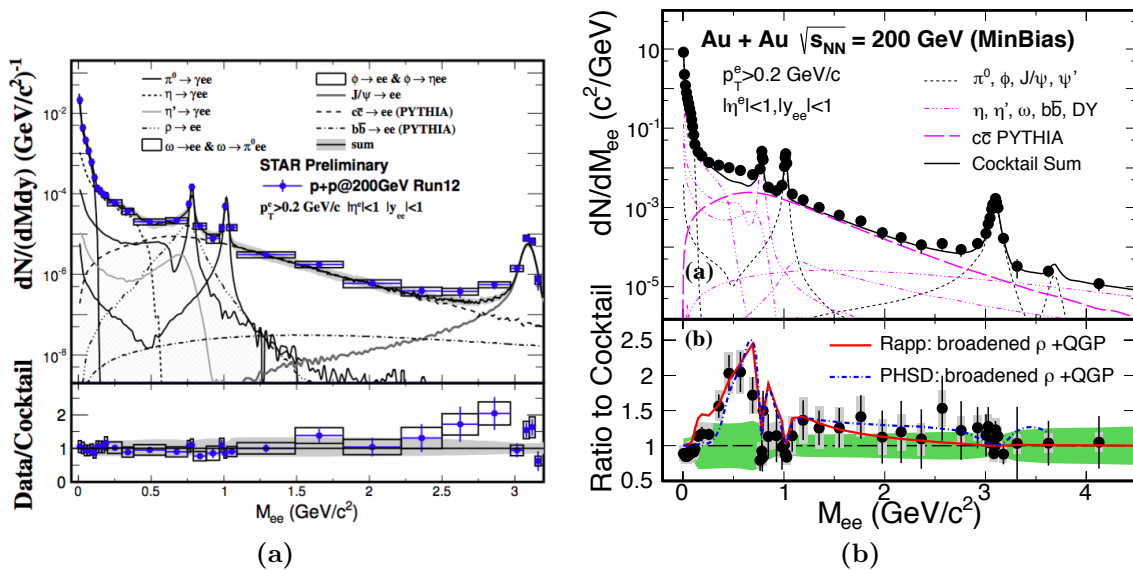
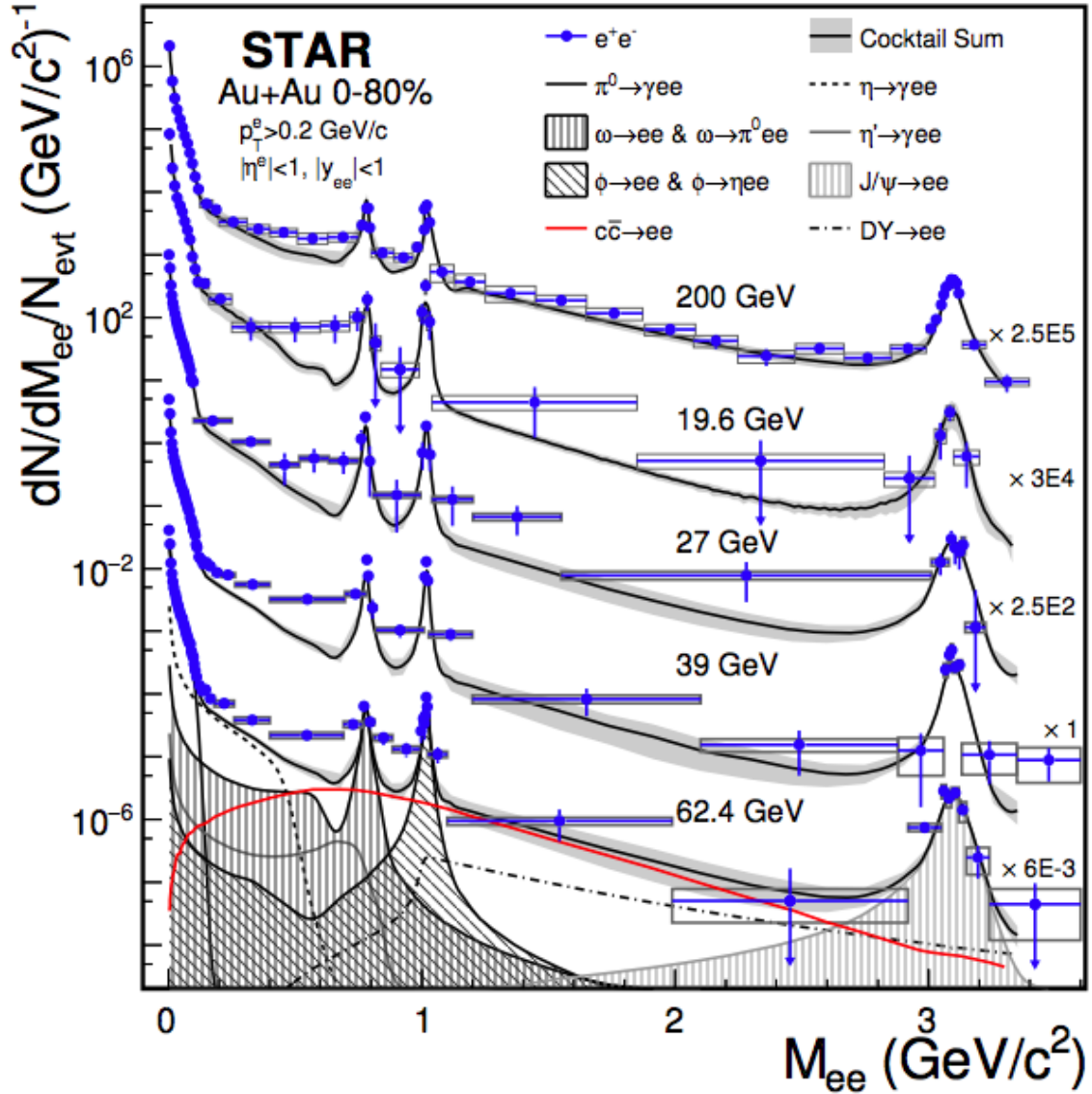


Figure 2.12 : Azimuthal correlations from  $c\bar{c}$  (a) and  $b\bar{b}$  measured through the dimuon channel. Measurements are compared with PYTHIA and POWHEG [70, 108]

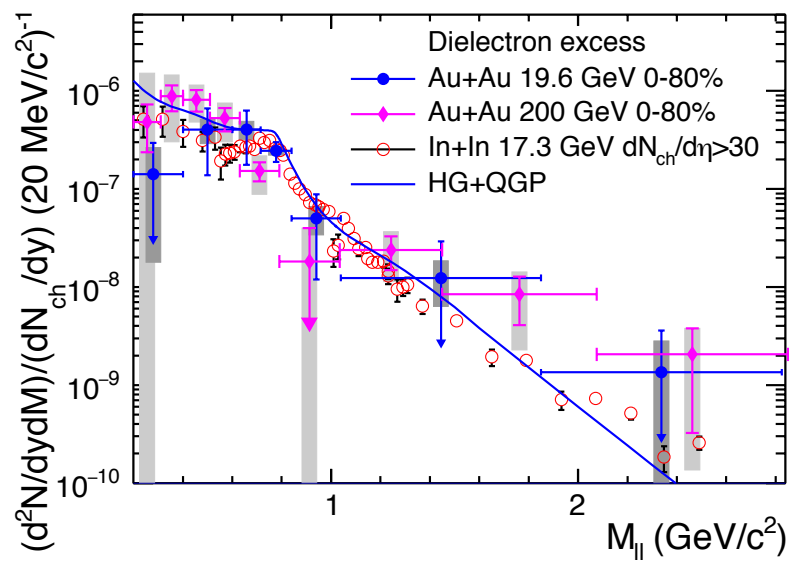


**Figure 2.13** : STAR measurement of the  $e^+e^-$  invariant mass spectra within STAR acceptance from  $p + p$  collisions at  $\sqrt{s} = 200$  GeV (a) [109]. Top panel of (b): The measurement of  $e^+e^-$  invariant mass spectra within STAR acceptance from 0 – 80% central Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. Lower panel (b): The ratio of the measured spectra to the hadronic cocktail (excluding  $\rho$ ) compared to an effective many-body model and a dynamical microscopic transport model [111, 113].

of center-of-mass energy for  $\sqrt{s_{NN}} = 19.6, 27, 39, 62.4, 200$  GeV. The invariant mass spectra measured at each energy in 0 – 80% central Au+Au collisions is shown in Fig. 2.14. At each energy, a significant excess over the hadronic cocktail is observed in the  $\rho$  mass region. Additionally, when volume effects are taken into account, the excess observed at  $\sqrt{s_{NN}} = 19.6$  GeV is consistent with the excess observed by NA60 in In+In collisions at  $\sqrt{s_{NN}} = 17.3$  GeV [110]. The volume normalized, full-phase space excess is shown in Fig. 2.15 for minimum bias (0–80%) Au+Au collisions at  $\sqrt{s_{NN}} = 19.6$  and 200 GeV from STAR, for In+In collisions at  $\sqrt{s_{NN}} = 17.3$  GeV from NA60, and for a theoretical model which includes an in-medium modified  $\rho$  spectral function and thermal emission from a QGP [111]. Agreement with the collection of STAR data over a wide range of energies demonstrates that the broadened  $\rho$  model is a robust description of the excess dilepton production mechanism.

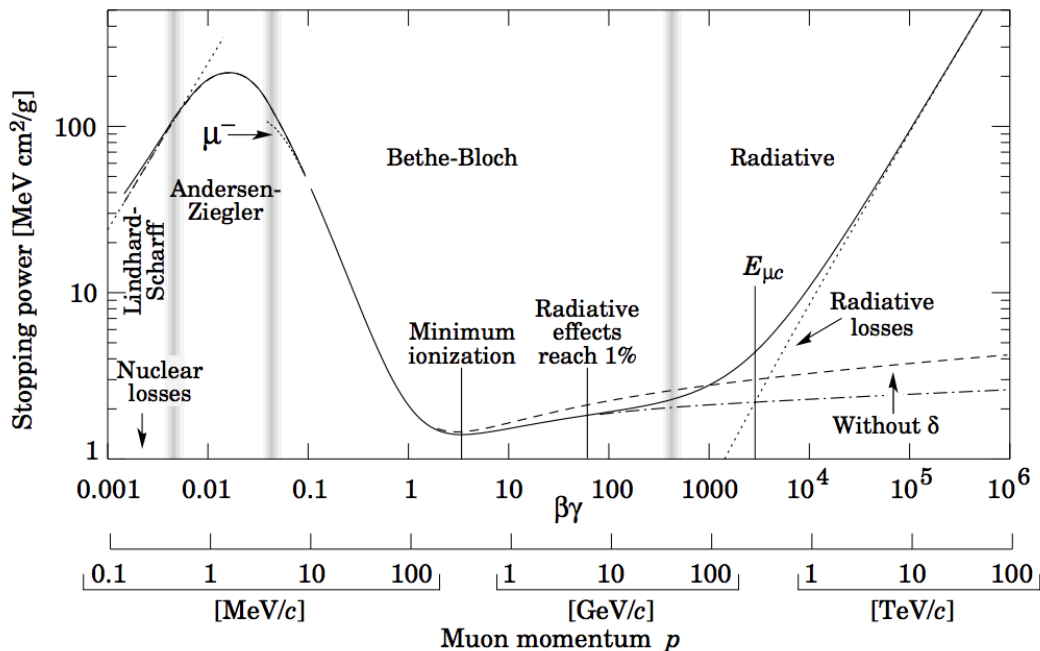


**Figure 2.14** : The  $e^+e^-$  invariant mass spectra within STAR acceptance from 0 – 80% central Au+Au collisions at  $\sqrt{s_{NN}} = 19.6, 27, 39, 62.4, 200$  GeV. Statistical (bars) and systematic (boxes) uncertainties are shown with the data points. The total hadronic cocktail is shown for each energy while the individual components are shown only for the  $\sqrt{s_{NN}} = 62.4$  GeV data [115].



**Figure 2.15** : The mid-rapidity charged particle multiplicity ( $dN_{ch}/dy$ ) normalized excess dielectron mass spectra in full phase-space for Au+Au collisions at  $\sqrt{s_{NN}} = 19.6$  and 200 GeV from STAR and for In+In collisions at  $\sqrt{s_{NN}} = 17.3$  GeV from NA60. The solid line shows a model calculation including a broadened  $\rho$  spectral function in a hadron gas (HG) and thermal radiation from a QGP [111].





**Figure 2.16** : The stopping power ( $-\langle dE/dx \rangle$ ) of  $\mu^+$  in copper as a function of  $\beta\gamma = p/Mc$  [116].

### 2.3 Motivations for $\mu^+\mu^-$ Measurements

There are several advantages to measuring the dilepton spectra through the dimuon channel in addition to the dielectron channel. The momentum smearing caused by bremsstrahlung energy loss of electrons directly affects the mass resolution achievable in the dielectron decay channel. The affect of bremsstrahlung on the mass resolution of the  $J/\psi$  can be clearly seen in STAR's dielectron results shown in Fig. 2.13a. The mass blurring caused by bremsstrahlung makes it more difficult to distinguish nearby states like the  $\Upsilon(1S)$ ,  $\Upsilon(2S)$ , and  $\Upsilon(3S)$ . At lower masses poor mass resolution can make it more difficult to distinguish between the many overlapping contributions to the dilepton spectra. The probability that a charged particle traveling through an electric field emits a photon as bremsstrahlung radiation is  $\propto 1/M^2$ . Since muons are about  $200\times$  heavier than electrons ( $m_\mu = 105 \text{ MeV}/c^2$  compared to  $m_e = 0.511 \text{ MeV}/c^2$ ) muons are far less likely to produce bremsstrahlung radiation (at equal momentum). Practically speaking, muon bremsstrahlung does not become a significant problem except for muon momenta at the TeV level as Fig. 2.16 shows.

At very low masses, the dielectrons invariant mass spectra can become swamped with  $e^+e^-$  pairs from  $\gamma \rightarrow e^+ + e^-$  conversion. For this reason, special purpose detectors like the

High Acceptance DiElectron Spectrometer (HADES) are specifically designed to minimize the amount of dead material to limit the photon conversion probability. Pair production of  $\mu^+\mu^-$  through  $\gamma$  conversion is not an issue due to the much higher mass of the muon compared to the electron. This is beneficial for the measurement of the lowest mass dimuon pairs (just above threshold).

Muons are also ideal for studying chiral symmetry restoration through  $\rho^0$ -meson decays. Again, the heavier mass of the muon compared to the electron limits the number of low mass mesons that can decay via the dimuon channel or through Dalitz decays with muons. With fewer sources in the LMR, the  $\rho^0$  meson is significantly easier to isolate. This benefit was powerfully demonstrated in NA60's measurement of the  $\rho^0$ -meson broadening in In+In collisions at  $\sqrt{s_{NN}} = 17.3$  GeV.

The advantage is not entirely on the side of the dimuon channel though. A serious disadvantage of the dimuon channel comes from the secondary muons resulting from weak decays. Since pions are the most common product in HICs and the  $\pi^+ \rightarrow \mu^+ + \nu_\mu$  branching ratio is  $\sim 99.9\%$ , secondary muons can be produced in copious amounts. The situation is slightly improved by the relatively long lifetime ( $2.6 \times 10^{-8}$  s) of pions since this decay proceeds via the weak force. This means that the majority of the secondary muons resulting from these decays will originate a finite distance from the primary interaction vertex. Detector complexes which combine muon detectors and precise (often silicon based) vertex tracking detectors can very effectively reject backgrounds from weak-decay secondary muons. This combination in NA60 is what allowed such a high quality result. By far the best option is to simply study the dilepton production in HICs through both the dielectron and dimuon channels.

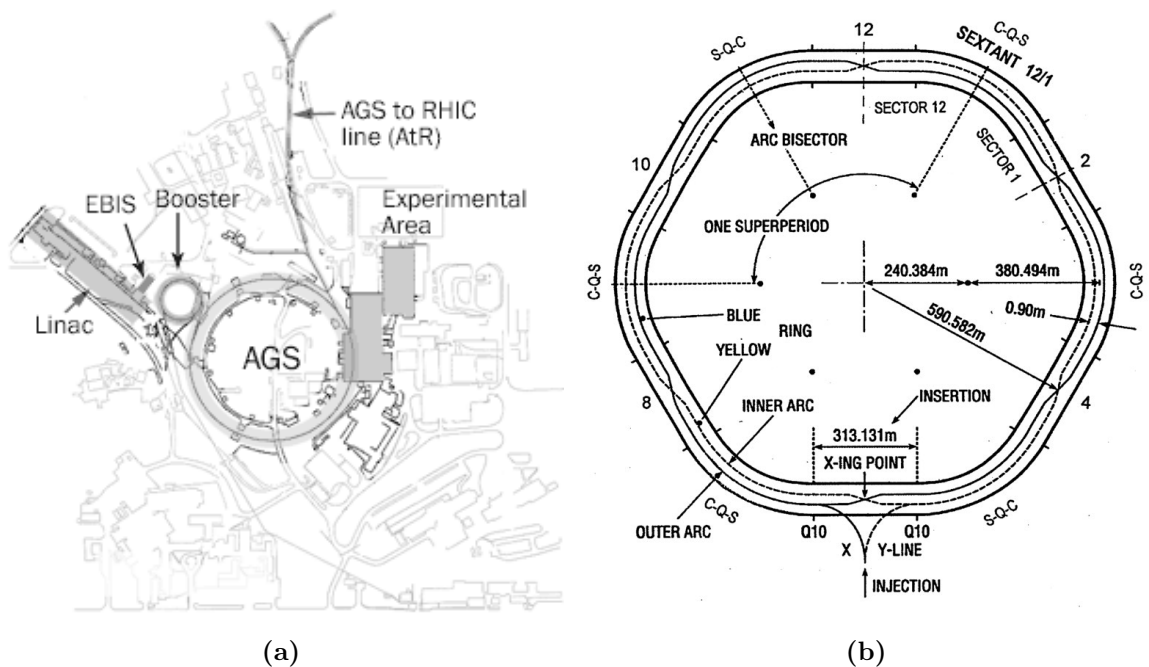
## Chapter 3

### Experimental Apparatus

#### 3.1 Relativistic Heavy Ion Collider (RHIC)

The Relativistic Heavy-Ion Collider located at Brookhaven National Laboratory consists two independent 3.8 km accelerator/storage rings. With the flexibility of two rings, RHIC is a versatile complex capable of colliding a diverse set of heavy ions and polarized protons. Since beginning operations in 2000, RHIC has collided systems of  $p + p$ ,  $p + \text{Au}$ ,  $p + \text{Al}$ ,  $d + \text{Au}$ ,  ${}^3\text{He} + \text{Au}$ ,  $\text{Cu} + \text{Cu}$ ,  $\text{Cu} + \text{Au}$ ,  $\text{Au} + \text{Au}$  and  $\text{U} + \text{U}$  at various energies. The top energy attainable for a  $p$  beam is  $\sim 500$  GeV, while the top energy for a Au ion is  $\sim 100/u$  (due to the  $Z/A$  ratio). The wide array of species and collision energies available at RHIC allow it unique avenues to study the properties of the QGP and the QCD phase diagram through ultra relativistic heavy ion collisions. In addition, as the world's only high energy polarized  $p + p$  collider, it provides unparalleled opportunities to study the fundamental spin structure of protons. A schematic of the RHIC accelerator complex is shown in Fig. 3.1b.

The RHIC accelerator is built off of the infrastructure that has existed for some time at BNL. The major components used to inject beams into RHIC are shown in Fig. 3.1a. Since the 1970s, the Tandem Van de Graaff at BNL has been providing ions for experimental research. As part of RHIC, the Tandem Van de Graaff provided the heavy-ions for many years until recently when an electron beam ion source (EBIS) was installed. The EBIS is now the primary source of ions for RHIC, though the Tandem Van de Graaff is still used in some special scenarios. In the EBIS the negatively charges gold ions are accelerated to 2 MeV/ $u$  (MeV per nucleon) and stripped to a charge of  $Q = +32$  before being delivered to the Booster Synchrotron. In the Booster Synchrotron, the ions are accelerated to  $\sim 100$  MeV/ $u$  and stripped to a charge of  $Q = +77$ . The Alternating Gradient Synchrotron (AGS) receives the beams from the Booster Synchrotron. Inside the AGS, the ions are fully stripped to a charge of  $Q = +79$  and accelerated to 8.89 GeV/ $u$  before being transferred through the AGS-to-RHIC (AtR) transfer line. Upon entering RHIC, the beams can be accelerated further



**Figure 3.1 :** The RHIC accelerator complex. The AGS area is shown in (a) along with the EBIS, the Linac, the Booster and the AGS-to-RHIC line. The RHIC rings are shown in (b) [119].

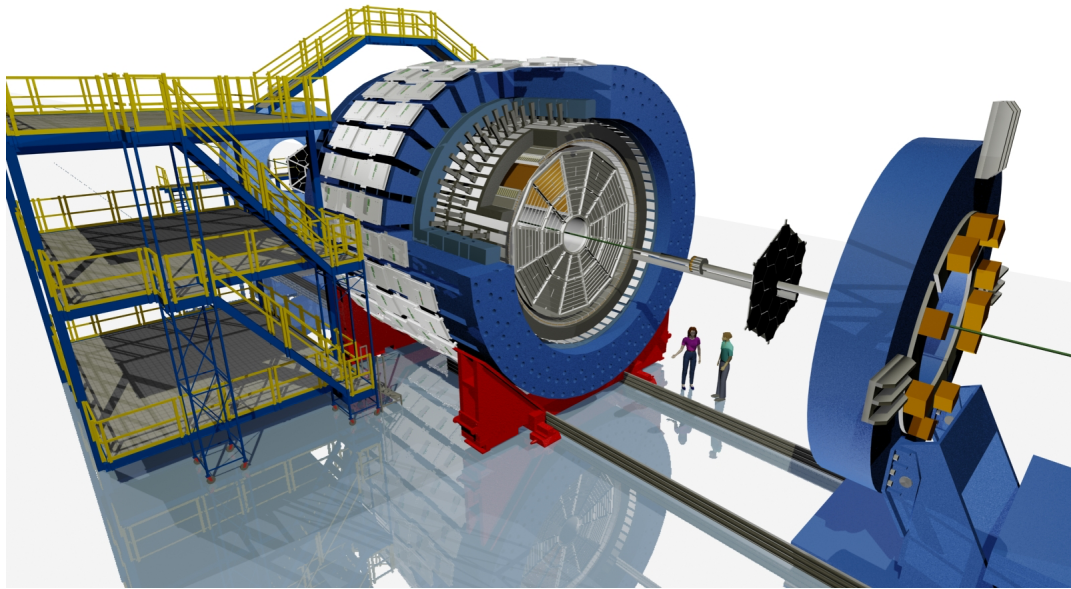
and stored for several hours.

RHIC has successfully delivered Au+Au collisions at energies ranging from  $\sqrt{s_{NN}}=7.7$  to 200 GeV. The targeted design luminosity for Au+Au collisions was  $2 \times 10^{26} \text{ cm}^{-2}\text{s}^{-1}$ , but in recent years the achieved average luminosity has been  $87 \times 10^{26} \text{ cm}^{-2}\text{s}^{-1}$ , 44 times the design goal. At a speed of  $\sim 99.995\%c$ , the beams take about  $12 \mu\text{s}$  to travel around the ring one time. When RHIC is fully injected, it has 111 bunches per beam filled with  $\sim 10^9$  ions/bunch. Each of RHIC's rings accelerate the bunches using powerful electric fields in resonant radio frequency cavities. The beams are steered with 1,740 superconducting magnets into six interaction points along the rings. At the start of its operations, RHIC had four main experiments: BRAHMS, PHOBOS, PHENIX and STAR. The PHOBOS experiment was decommissioned in 2005 and a year later the BRAHMS experiment was too. The PHENIX experiment was decommissioned and cleared from its interaction region in 2015 to make room for a new detector called sPHENIX [117,118]. STAR is the only one of the four experiments still in operation today.

## 3.2 Solenoidal Tracker at RHIC

The Solenoidal Tracker at RHIC (STAR) detector sits at roughly the 6 O'clock position in the RHIC ring. The origin of the STAR coordinate system is at the center of the solenoidal magnet. The x-axis points approximately south from the center of RHIC along its radius to where STAR is. The y-axis points upwards and the z-axis points in the direction of the clockwise circulating beam (west when viewed from above). The transverse momentum ( $p_T$ ) of charged particles are measured via the tracks' curvature in the x–y plane. The azimuthal angle  $\phi$  is measured in the x–y plane starting from the positive x-axis. The polar angle  $\theta$  is measured in the z–y plane from the positive z-axis. However, instead of using  $\theta$  the use of the pseudorapidity ( $\eta = -\ln(\tan(\theta/2))$ ) is far more common since it is additive under Lorentz boosts, similar to velocities under Galilean transformations. A render of the STAR detector from one corner can be seen in Fig. 3.2.

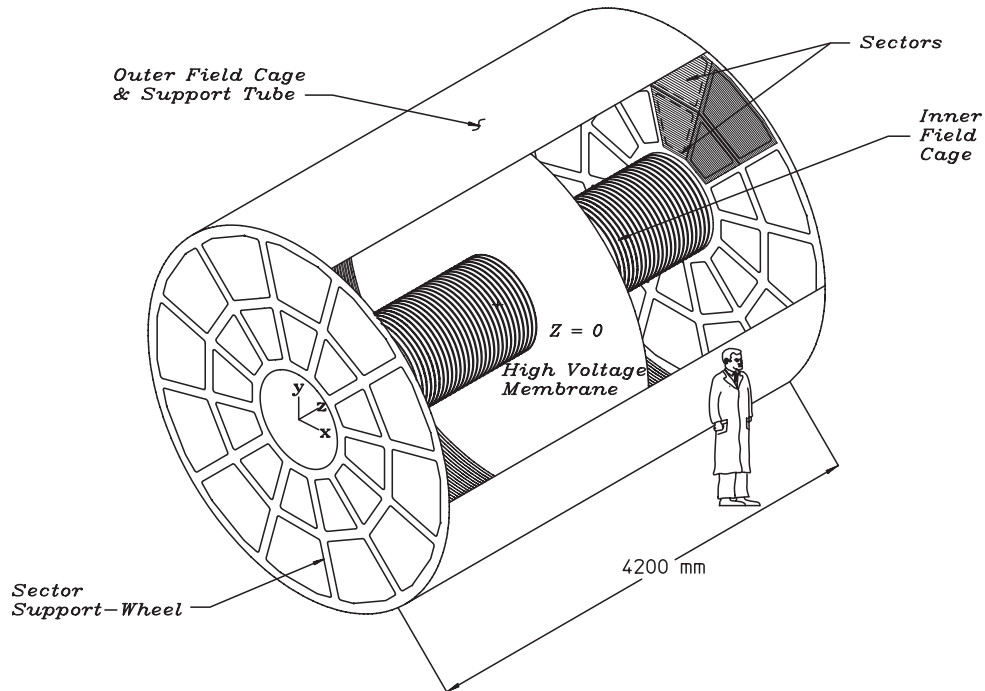
The STAR detector was built for the express purpose of studying high energy heavy-ion collisions [120]. To satisfy this purpose, STAR was built as a multi-purpose detector designed to have large, uniform acceptance in  $0 < \phi < 2\pi$  and  $|\eta| < 1$ . It was also important that STAR be able to provide charged particle tracking down to low momenta ( $\sim 100$  MeV/ $c$ ) and in high multiplicity environments. For this reason, the STAR detector is composed of a 0.5 T solenoidal magnet and a central Time Projection Chamber (TPC) for charged particle tracking. Outside the TPC are several other detector subsystems that provide advanced particle identification or energy measurement capabilities.



**Figure 3.2 :** The STAR detector system. [121]

The work reported in this thesis makes use of the Time Projection Chamber, the Vertex Position Detectors (VPD), the Barrel Time-of-Flight (BTOF) detector, and the Muon Telescope Detector (MTD). These detectors will be describe in more detail below. More information about STAR as a whole and about other STAR subsystems can be found in Ref. [120].

### 3.2.1 Time Projection Chamber



**Figure 3.3** : A schematic of the STAR Time Projection Chamber [122].

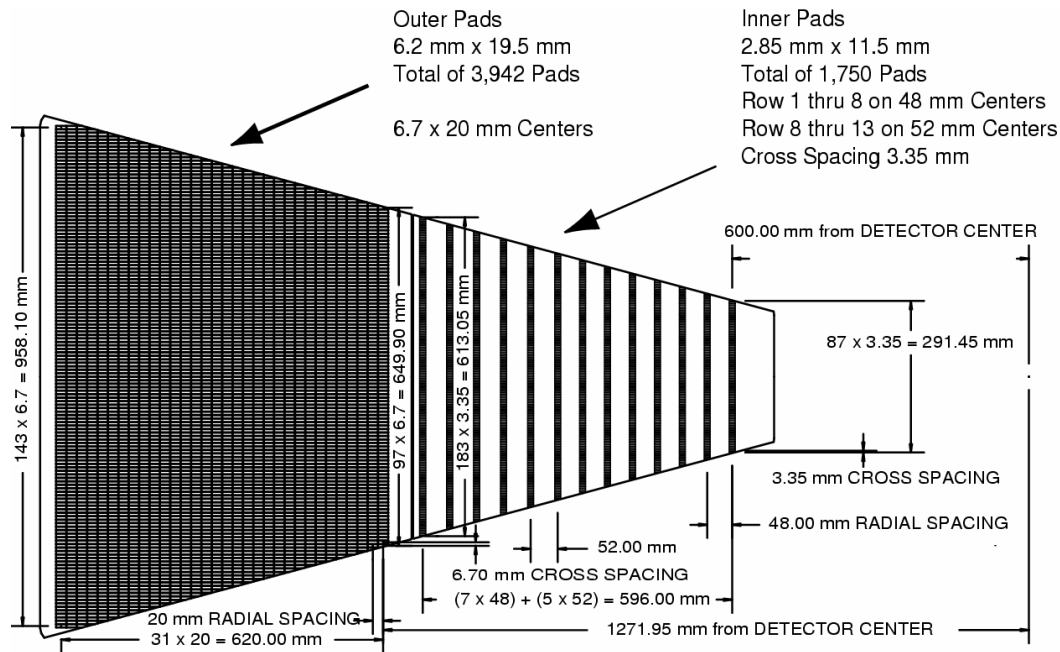
The Time Projection Chamber is at the heart of STAR with the rest of the detectors built around it. The TPC is itself a 4.2 m cylindrical detector built around the beam pipe. It has an inner radius of 50 cm and an outer radius of 200 cm. At the time of construction it was the largest TPC in the world, capable of tracking 3000 charge particles at a time [122]. One of the most desirable characteristics of the TPC is that it is mostly empty volume, making it cost-effective and providing a low material budget. The volume of the TPC is filled with P10 gas (90% Argon, 10% Methane) at +2 mbar above atmospheric pressure to assure that no water moisture enters the TPC exposing the internals to the dangers of oxidation. A -28kV central membrane and two grounded cathodes provide a well-defined constant electric field of  $\approx 135$  V/cm which permeates the TPC volume parallel to the z-axis. Inner and outer field cages use precision resistor chains to provide 182 equipotential regions within the TPC to keep the electric field constant. The  $\approx 135$  V/cm electric field results in an electron drift velocity through the P10 gas of 5.45 cm/ $\mu$ s.

The TPC uses multi-wire proportional chambers (MWPC) to readout the charges that drift toward either end. Twelve trapezoid-shaped MWPCs are used on either end-plane of the TPC to cover the full range in azimuth. The MWPC are segmented into an inner and an outer wire plane. The inner plane consists of 13 wires while the outer consists of 32 wires for a total of 45 readout locations in the radial direction. Each of these 45 pickup wires allow for a measurement of the position and ionization energy loss of a track traveling through the TPC.

The TPC accomplishes its goal of charged particle tracking via measurements of ionization in the P10 gas. A charged particle traversing the TPC volume will ionize Argon atoms freeing electrons along the way. These freed electrons drift away from the central membrane towards the read-out electronics near the cathodes. The spatial location in the  $x$ - $y$  plane of electron clusters are measured with a maximum of 45 spatial points by the MWPC array. Each cluster provides a position measurement with a resolution of  $\sim 500\mu\text{m}$ . At the same time the  $z$ -position is measured by time-binning ( $\sim 100$  ns time bins) the readouts and extrapolating based on the electron drift velocity.

Track trajectories are reconstructed in 3D from the 45  $x$ - $y$  hits plus the  $z$ -positions measured by the TPC. The conversion from discrete hits into helices is accomplished using the Kalman filtering process - a method of iteratively fitting helices to determine the most accurate track trajectories in a uniform magnetic field. Once the tracks have been reconstructed they constitute “global tracks” in the STAR jargon. Next, all of the global tracks are extrapolated to the  $z$ -axis and a  $\chi^2$  minimization technique is used to group tracks that point towards the same vertex. In this way one or more “primary” vertices are formed which represent the location of a heavy-ion collision. Sometimes one of the primary heavy-ion beams may cause a collision with the beam-pipe material which results in a real collision producing enough charged particles to reconstruct back to the collision vertex. This type of background event can generally be rejected (after the event is reconstructed) by requiring that the  $x$ - $y$  position of the primary vertex be within the empty vacuum region of the beam-pipe. Any particular event may have several primary vertices reconstructed.





**Figure 3.4**

In central heavy-ion collisions it is generally not too hard to pick out the “real” primary vertex since it may have more than 1000 charged particles associated with it. In peripheral heavy-ion collisions and  $p + p$  collisions this can be more challenging though and may require additional constraints. One such constraint is provided by the VPDs and will be discussed next. After determining which global tracks are associated with which primary vertices, the tracks are refit with the position of the primary vertex as an added point along their trajectories. The primary vertex generally has very precise position resolution, since it benefits from many track measurements. The addition of the high precision primary vertex position to the track helix fit significantly improves the momentum measurement and therefore improves the momentum resolution. With the addition of the primary vertex point along a particle’s track, the TPC obtains momentum resolution of  $\sim 1\text{--}2\%$  (depending on the particle species) at  $p_T \approx 1 \text{ GeV}/c$ . The direction of curvature of the track helix is also used to determine its charge, assumed to be  $\pm e$ .

In addition to charged particle tracking, the gas ionization measurements provided by the TPC are also useful for particle identification. In addition to simply measuring the location of electron clusters produced by ionization energy loss, the TPC is also able to measure the number of electrons in each cluster which is proportional to the amount of energy lost

per unit length traveled. The ionization energy loss per unit length traveled ( $dE/dx$ ) by a charged particle through a medium is described by the relativistic Bethe-Bloch formula:

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{c^2 m_e} \frac{n z^2}{\beta^2} \frac{e^4}{(4\pi\epsilon_0)^2} \left[ \ln \left( \frac{2c^2 \beta^2 m_e}{I \cdot (1 - \beta^2)} \right) - \beta^2 \right] \quad (3.1)$$

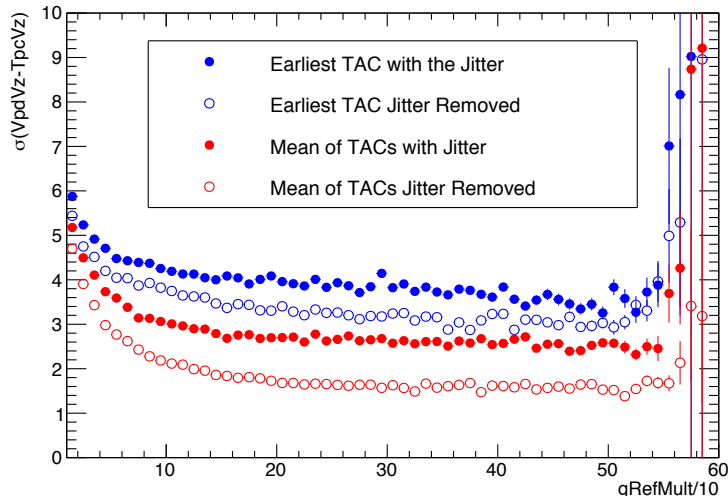
$$n = \frac{N_A \cdot Z \cdot \rho}{A \cdot M_u} \quad (3.2)$$

where  $c$  is the speed of light,  $\epsilon_0$  is the vacuum permittivity,  $\beta = v/c$ ,  $e$  is the electron charge,  $m_e$  is the electron mass,  $I$  is the mean excitation potential, and  $z$  is the particle's charge in units of  $e$ . The number density of the material is given by  $n$  where  $\rho$  is the density of the material,  $Z$  is its atomic number,  $A$  its atomic mass,  $N_A$  is Avogadro's number, and  $M_u$  is the molar mass constant. Since  $\beta\gamma = p/m$ , particles of different masses can be distinguished with a precise enough momentum measurement. The Bethe-Bloch formula was further improved upon by tailoring the equation specifically for the P10 gas in the TPC. The result of this tailoring was the Bichsel parameterization [123]. Using the Bichsel parameterization, the measured  $dE/dx$  values can be rescaled with respect to the expected value for a given particle. After being rescaled and divided by the resolution, the measured  $dE/dx$  value is called  $n\sigma$  (e.g.  $n\sigma_\pi$  for a  $\pi$ ). This formulation of the  $dE/dx$  measurement is very convenient since it allows particles to be roughly identified by selecting on ranges of  $n\sigma$ .

### 3.2.2 Vertex Position Detector

The Vertex Position Detectors are two identical photo-multiplier tube (PMT) based detectors covering the forward and backward regions ( $4.24 < |\eta| < 5.1$ ) approximately 5.7 m from the center of STAR [124]. Each VPD consists of nineteen identical detectors formed from a Pb converter, a fast plastic scintillator, and a PMT. The fast scintillators and PMT readouts along with high-quality signal digitization electronics allow the individual VPD tubes to measure particles passing through with a precision of  $\sigma_{single-tube} \approx 100$  ps. When all the hits are combined, the “start time” of an event can be measured with a precision of a few tens of ps. This level of timing precision can be recast as a longitudinal measurement of the collision vertex position with a precision of  $\sigma_z \approx 1$  cm (in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV) [124].

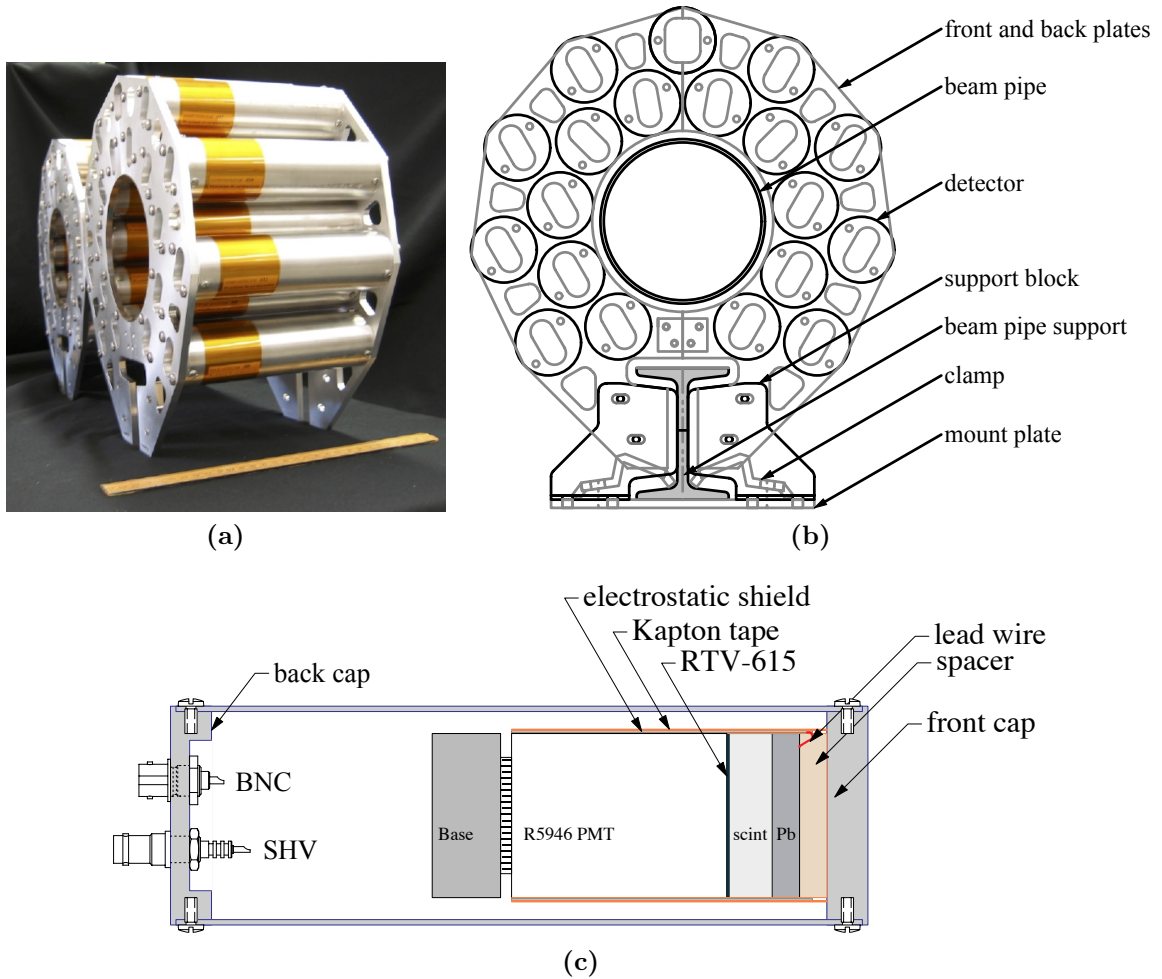
The precise timing and vertex position measurement provided by VPD are very important for “triggering” - i.e. almost instantly choosing which events to keep and which to



**Figure 3.5 :** The VPD vertex resolution using the trigger electronics as a function of  $g\text{RefMult}$  (the number of global tracks in  $\pm 0.5$  units of  $\eta$ ) in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. Four different configurations are shown. The default going into the 2016 run year provided  $\sigma_{vz} \approx 4$  cm. The improved version developed for the HFT provided  $\sigma_{vz} \approx 1.5$  cm at high multiplicities.

throw away. The VPD has two independent sets of electronics for digitizing the PMT signals. The electronics for the trigger logic use Time-to-Analog Converters (TACs) while the electronics used for offline data acquisition use CERN developed High Performance Time to Digital Converters (HPTDC). Since the VPD and TOF systems are largely a Rice University contribution to STAR, one of my main responsibilities as a Ph.D. student was to calibrate the electronic systems for the VPD. The response of the PMTs varies with collision species, so various VPD parameters must be tuned for each data taking period to ensure optimal performance.

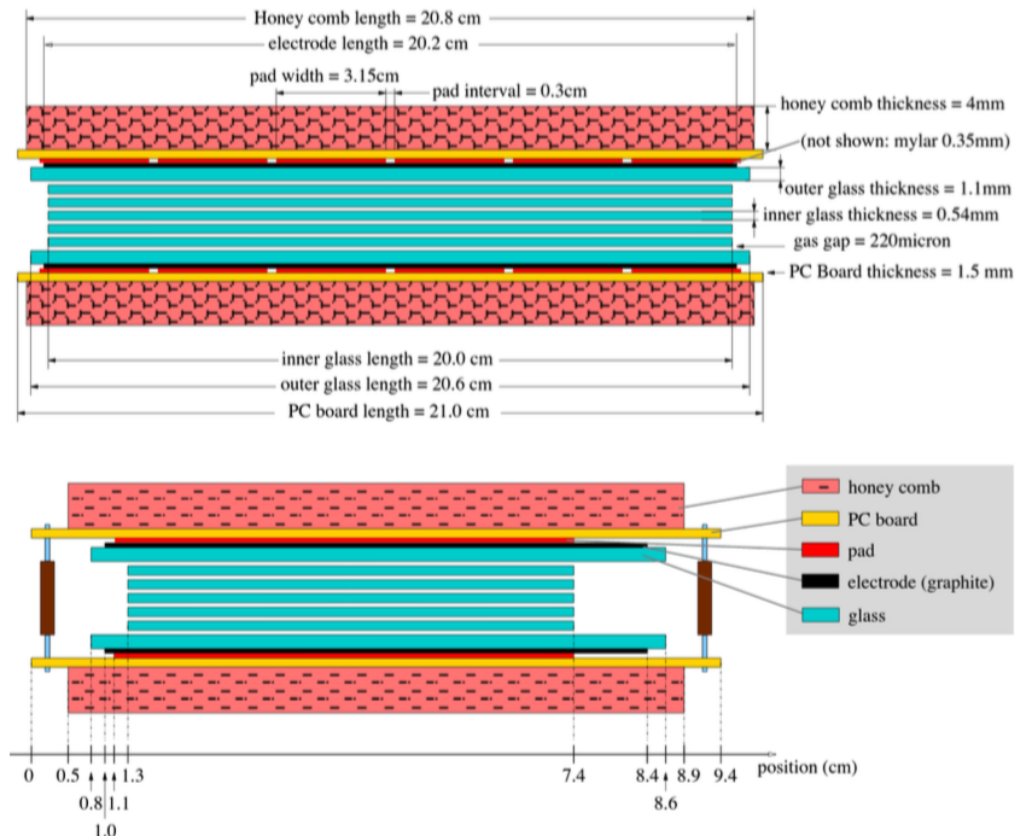
The calibration and definition of the VPD trigger was especially important during the 2016 collection of Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV data. For this dataset, a special silicon wafer-based Heavy Flavor Tracker (HFT) capable of providing additional track points with a spatial resolution of  $50\mu\text{m}$  was installed inside the TPC just around the beam-pipe. The addition of the HFT allows the reconstruction of secondary decay vertices from heavy flavor decays ( $D^0 \rightarrow K^-\pi^+$ ,  $c\tau \approx 120 \mu\text{m}$ ). However, the HFT only covered a longitudinal region of  $\pm 6$  cm. For this reason, the HFT could only provide tracking information for collisions within that region. The VPD was used to select in real-time only those collisions



**Figure 3.6 :** A photo of the Vertex Position Detectors before being installed around the beam-pipe (a). A schematic drawing the the VPD detector assembly, (b) and a schematic each single detector enclosure used by the VPD (c).

that occurred within  $-6 < v_z < 6$  cm. The improved calibration and trigger algorithm shown in Fig. 3.5 improved the HFT event purity from  $\sim 60\%$  to  $\sim 90\%$ .

The VPD also plays a crucial role for the MTD based triggers which will be discussed in detail below. The precise time measured by the VPD also acts as the “start-time” ( $t_0$ ) for the Barrel Time-of-Flight and MTD detectors. The Time-of-Flight detector measures the “stop-time” ( $t_1$ ) of tracks as they pass through the BTOF detector. Together, the VPD and BTOF detectors allow the precise measurement of a particle’s velocity. More details about the BTOF detector are discussed next.

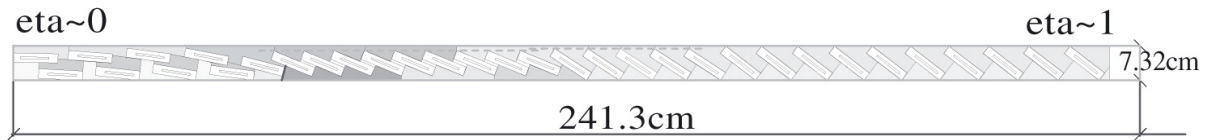


**Figure 3.7 :** A schematic of the MRPCs used in the STAR BTOF. [125]

### 3.2.3 Time-of-Flight Detector

The Barrel Time-of-Flight (BTOF) detector is a cylindrical detector installed just outside the TPC at a radius of 210 cm [125,126]. The BTOF detector is constructed from Multi-gap Resistive Plate Chambers (MRPC). An MRPC is essentially the evolution of the Resistive Plate Chamber (RPC) detector design. Each MRPC is a stack of thin layers (0.54 mm for inner layers, 1.1 mm for outer layers) of resistive glass ( $\rho \approx 10^{13} \Omega/\text{cm}$ ) separated by gas filled gaps of 0.22 mm. For an RPC the gas in the gaps is generally  $\sim 95\%$  Freon R-134a ( $\text{C}_2\text{H}_2\text{F}_4$ ) with  $\sim 5\%$  being isobutane and/or  $\text{SF}_6$ . Electrodes on either side of the MRPC are held at a potential difference of 10–15kV. Outside the electrodes, copper pads are used to pickup and readout the signals.

The RPC and MRPCs operate based on the production of Townsend avalanches within the gaseous gaps. When a charged particle passes through the gas filled gaps, electrons are freed through ionization. The strong electric field accelerates the free electrons, causing them to free more electrons via ionization of the gas' atoms. The MRPC has advantages over the



**Figure 3.8** : A schematic of a TOF tray showing the layout of the MRPCs within [126].

single gap RPC design. Making several narrow gaps instead of one large gap reduces the jitter caused by the primary ionization event. The existence of multiple gaps also leads to several independent avalanches, thereby allowing the final signal to be an average of multiple independent events thus improving the resolution. There is a balance that must be found though, since smaller gaps generally have a larger relative uncertainty due to the physical difficulty of defining very narrow separations uniformly. Also narrow gaps can increase the susceptibility to non-uniformities in the glass. A peak on the surface of a glass layer can cause thermionic emission induced by the breakdown of the electric current. This results in a noisy detector with fake signals that do not result from charged particles passing through. Over the past many years, RPC and MRPC-based detectors have been thoroughly studied [127–130] and their performance has been found to be very robust. A schematic diagram of the MRPCs used in the STAR Barrel Time-of-Flight can be seen in Fig. 3.7.

The BTOF detector is segmented into 120 essentially identical trays, with 60 on the east and 60 on the west side of STAR. Each tray measures  $241.3 \times 21.6 \times 8.9 \text{ cm}^3$ . The 120 trays cover  $2\pi$  in azimuth and approximately  $\pm 0.9$  units in pseudorapidity, similar to the coverage of the TPC. Each tray contains 32 individual MRPCs and each MRPC contains 6 channels for a total of  $120 \times 32 \times 6 = 23,040$  channels. The MRPCs within each tray are staggered in such a way as to point nearly perpendicular to the expected trajectory of a particle emanating from the origin of the STAR coordinate system. The side view of one tray can be seen in the schematic of Fig. 3.8. Like the VPD, the BTOF electronics need to be calibrated for each collision species and data taking period. One of my responsibilities as a student was to calibrate all 23,040 channels of the BTOF detector. A table summarizing the TOF performance for a selection of data sets can be seen in Table 3.1.

The purpose of the Time-of-Flight (TOF) system in STAR, composed of the BTOF and VPD detectors, is to provide precise timing information of particle trajectories for particle identification (PID). By measuring the time difference between the VPD start-time and

**Table 3.1 :** A summary of the performance of the TOF system in STAR from the year 2012 to 2017. The total time resolutions marked with a “\*” indicate that the VPD was not used for the start-time determination. Instead a bootstrapped start-time was calculated from the pions in the event using only the BTOF information.

Year	Collision System and $\sqrt{s_{NN}}$ (GeV)	VPD Single PMT Resolution (ps)	Stop-side resolution (ps)	Total Resolution (ps)
2012	$p + p$ 200	100	70	120
	$p + p$ 510	100	70	120
	U+U 193	70	70	75
	Cu+Au 200	92	75	78
2013	$p + p$ 510	133	80	150
2014	Au+Au 14.5	170	84	146*
	$^3\text{He} + \text{Au}$ 200	96	79	125*
	Au+Au 200	68	70	73
2015	$p + p$ 200	122	75	113
	$p + \text{Au}$ 200	122	90	98
2016	Au+Au 200	78	74	75
	$d + \text{Au}$ 200	114	85	88
	$d + \text{Au}$ 62.4	126	90	95
	$d + \text{Au}$ 39	152	83	121*
	$d + \text{Au}$ 19.6	160	83	128*
2017	$p + p$ 510	137	90	114
	$p + p$ 510 RHICf	125	82	101

the BTOF stop-time ( $\Delta t = t_1 - t_0$ ) the TOF system achieves a total time resolution of approximately  $\sigma_{TOF} \approx 100$  ps [124]. Particle velocity information alone has limited usefulness for separating particles of different masses. Unless the path length ( $\Delta s$ ) of the particle's trajectory and the momentum ( $p$ ) of the particle is also known, the TOF information cannot be used for PID. However, the TOF system works in concert with the TPC. The TPC tracking provides the momentum measurement and the trajectory which can be projected to the TOF volume to determine the total path length. With all three detectors working together (VPD for start-time, TPC for  $p$  and  $\Delta s$ , and TOF for stop-time) the particle velocity (expressed as  $\beta^{-1} = c/v$ ,  $c$  being the speed of light) can be measure with:

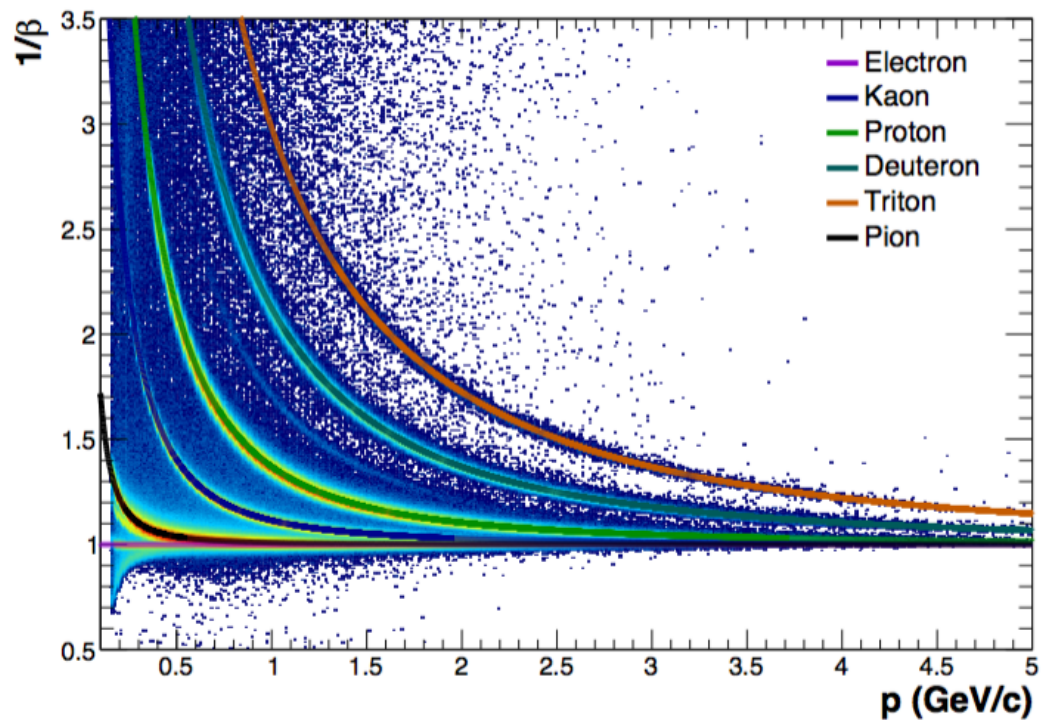
$$\beta^{-1} = c \frac{\Delta t}{\Delta s} \quad (3.3)$$

With the addition of the  $p$  measured by the TPC, the mass of the particle can be inferred with:

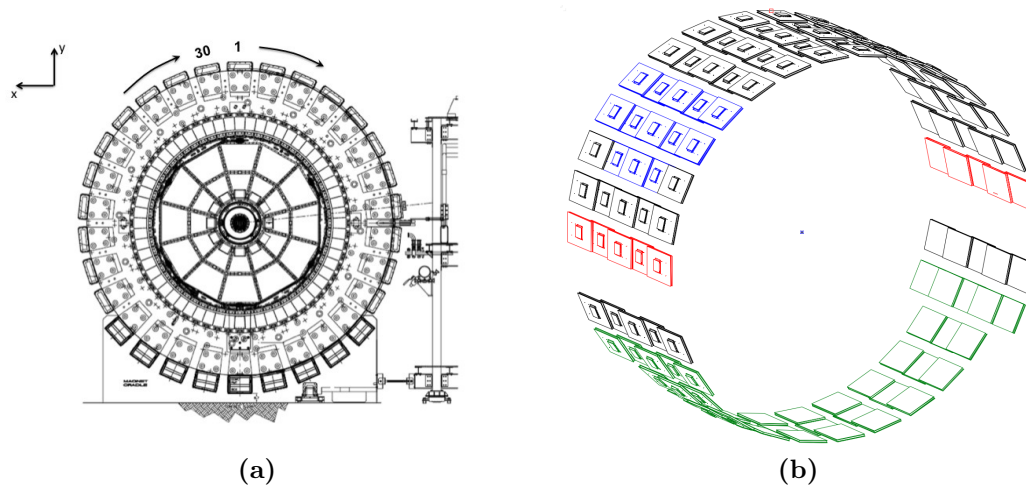
$$m = \frac{p}{\gamma\beta c} = \frac{p\sqrt{1-\beta^2}}{\beta c} = \frac{p}{c}\sqrt{\beta^{-2}-1} \quad (3.4)$$

making the separation and identification of  $\pi^\pm$ ,  $K^\pm$ , and  $p/\bar{p}$  possible out to  $p \approx 1.5$  GeV/ $c$ . An example of the TOF's particle identification capabilities can be seen in Fig. 3.9 for Au+Au collisions at  $\sqrt{s_{NN}}=7.7$  GeV. Information from the TOF detector is not used directly for muon identification in these studies. Instead the use of  $1/\beta$  measurements from the TOF detector are used in Sec. 4.3 to separate various sources of hadron punch through that leave signals in the MTD.





**Figure 3.9 :** The  $\beta^{-1}$  vs.  $p$  measured by the TOF system in STAR for Au+Au collisions at  $\sqrt{s_{NN}} = 7.7$  GeV. The expected  $\beta^{-1}$  is shown for each particle species. Taken from Ref. [131]

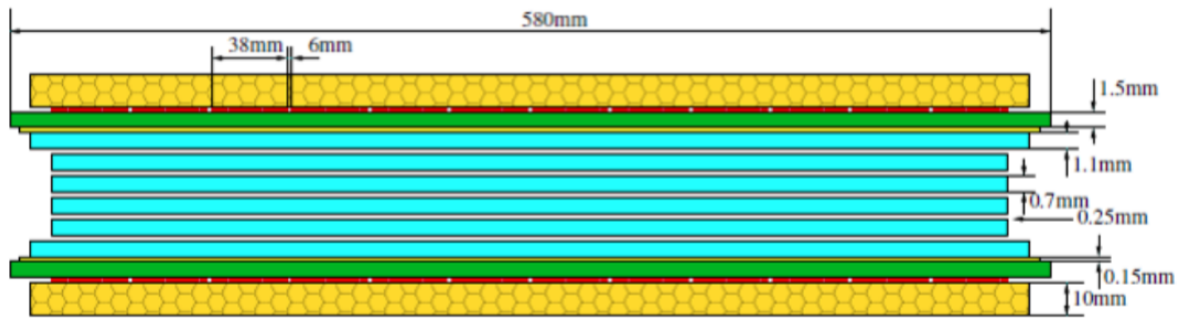


**Figure 3.10** : Schematic views of the MTD geometry. The magnet return steel is shown in (a) with numbers indicating the order of the MTD backlegs. An isometric view of the MTD modules (b) showing that some backlegs contain 5 modules and others contain only 3 modules due to space constraints.

### 3.2.4 Muon Telescope Detector

The Muon Telescope Detector (MTD) is a Multi-gap Resistive Plate Chamber (MRPC) based detector installed outside the magnet's return yoke steel at a radius of 410 cm [132]. The 60 cm thick magnet steel acts as the hadron absorber, offering  $\sim 5$  interaction lengths of material at its thickest. The MRPC-based design allows the MTD to be a large coverage, cost effective detector [133]. The view of the magnet and MTD from the east side of STAR is shown in Fig 3.10. The MTD divides the azimuthal angle into 30 backlegs with each backleg covering 12 degrees. The active area covered by the MTD in each backleg is about 8 degrees with 2 degrees of gap on either side of the active region. In the longitudinal direction (beam direction) 5 modules are installed in backlegs 1-8, 10, 21-22, and 24-30. Backlegs 13-20 have only the central 3 modules installed due to physical constraints. Backlegs 9 and 23 do not have modules installed.

The backlegs with 5 modules installed cover approximately  $-0.6 < \eta < 0.6$  and the backlegs with only 3 modules cover approximately  $-0.3 < \eta < 0.3$  for collisions at the center of the STAR magnet. On average the MTD system covers  $\sim 45\%$  of the azimuthal direction within  $|\eta| < 0.5$ . Each MTD module is an MRPC with 12 double-ended readout strips. Each strip is 87 cm long and 3.6 cm wide. There is a gap of about 0.6 cm between each strip. The

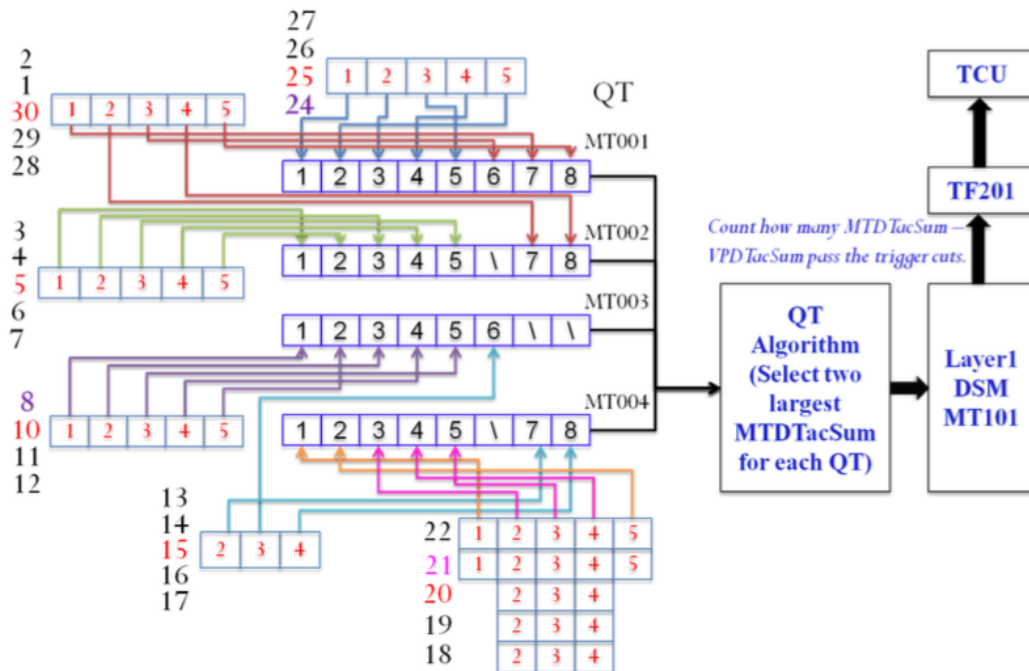


**Figure 3.11** : A schematic of the side-view of an MTD LMRPC module. [133]

MTD was installed progressively over the course of several years. In 2012 the first 12 modules (10%) were installed. The next year an additional 63 modules were installed (61%) before completing the installation of all 122 modules in 2014.

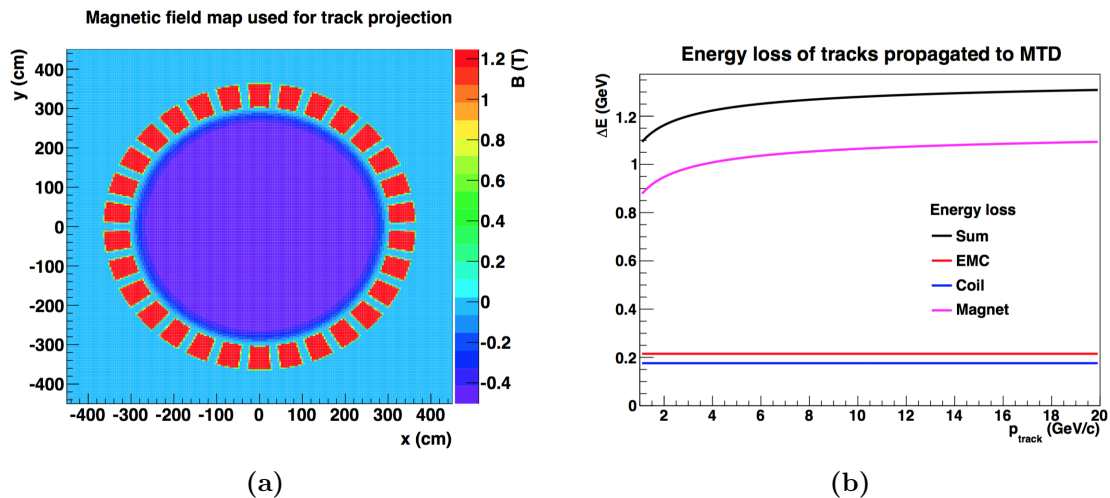
The MRPCs used in the MTD are similar to the ones used in the BTOF but with a few important differences. While the BTOF used small MRPCs with single ended readouts, each of the modules in the MTD are Long-Strip MRPC (LMRPC) with readout electronics on both ends [133]. In tests conducted at Fermilab, this design provided better than 70 ps time resolution and spatial resolution of approximately 1 cm along the readout strips [134, 135]. The LMRPC design also showed better than 95% detection efficiency. A schematic of the LMRPC design used for the MTD is shown in Fig. 3.11. The LMRPCs have four inner panes of glass (0.7 mm thick) and two outer panes of glass (1.1 mm thick) for a total of five  $250\mu\text{m}$  gas gaps. Special highly resistive glass was used with a surface resistivity of  $\sim 10^{13}\Omega\cdot\text{cm}$ . High voltage electrodes are attached on either side of the outer panes of glass. The total active area of the LMRPC, defined by the area of the inner glass panes, is  $875 \times 543 \text{ mm}^2$ .

Each of the twelve double-ended readout strips in the LMRPCs measure 870 mm long by 38 mm wide with a gap of 6mm between neighboring strips. Like the VPD and BTOF detectors, the MTD is a fast detector used for triggering. For this reason the MTD has two parallel but separate sets of electronic readout equipment. One set of electronics is used to readout the MTD modules as part of the trigger logic while the other is used to readout information into the offline data acquisition system. The offline electronics use the same HPTDC chips as the BTOF detector electronics. The trigger electronics divide the 122 modules into 28 “trigger patches”. Most trigger patches are comprised of 5 MTD modules,



**Figure 3.12 :** A schematic drawing of the MTD trigger logic. Most trigger patches consists of 5 modules while some consist of only 3. The modules in a trigger patch are always grouped from position of equal  $\eta$ . The red and magenta trigger patch numbers mark the location of various electronics boxes. The violet numbers mark trigger patches with some channels disabled where there is no material guarding the module [136].

while some contain only 3 modules. Modules from the same  $\eta$  position but different azimuthal angles are grouped together in each trigger patch. The grouping in  $\eta$  is meant to help ensure that all signals will arrive at the MTD trigger logic electronics at roughly the same time, since tracks with equal  $\eta$  should have roughly equal track lengths. The first layer of trigger logic is implemented in QT boards. The signal from all of the trigger patches are received in the QT boards MT001-MT004. The QT boards at this stage determine the fastest hits (largest TAC value) from each trigger patch. The fastest signals from MT001-MT004 are sent to MT101 where the times are compared with the trigger electronics' start-time provided by the VPD. The final stages of the trigger logic (TF201) tests to determine if two or more of the time differences (from two or more different trigger patches) fall into the pre-defined timing window. If acceptable times are found then the event is sent to the Trigger Control Unit (TCU) to be kept. A schematic diagram of the MTD trigger logic can be found in Fig. 3.12.

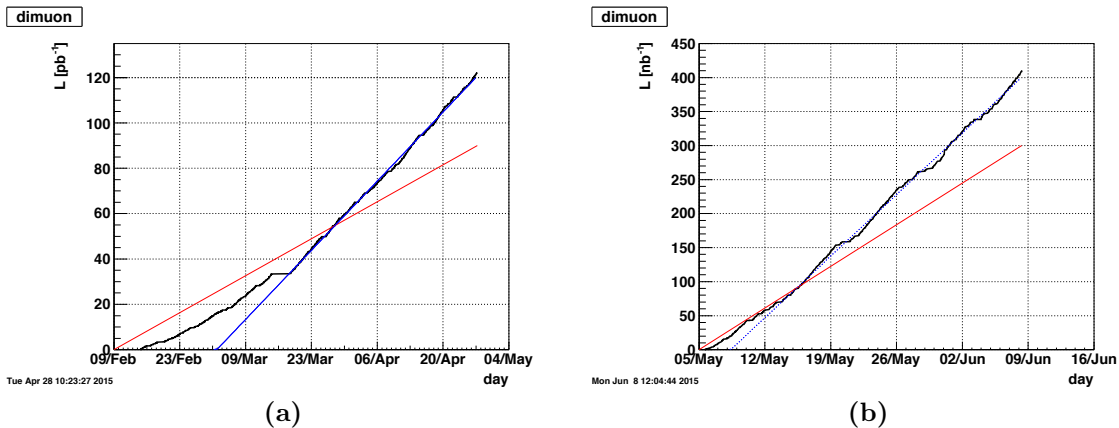


**Figure 3.13** : A map of the magnetic field of the STAR detector in the transverse plane. (a). The energy loss parameterization used in the STAR MTD track matching algorithm (b). The energy loss is parameterized as a function of track momentum by averaging the energy loss from GEANT over many events for several sub-systems [136].

### MTD Track Matching

The most important information that the MTD can provide for muon identification is whether or not a given track produced a hit in the MTD. The matching of tracks reconstructed from hits in the TPC to MTD hits is accomplished by projecting the TPC tracks out to the MTD radius. The MTD modules' double-ended readout strips allow the local  $Z$  position of hits to be measured via the difference in time between the signals on either ends of the strip. The local  $Y$  position is determined simply based on which strip registered the signal. The  $\Delta Z$  and  $\Delta Y$  variables are formed from the residuals between the local  $Y$  and  $Z$  measured position and the track's projected position. If MTD hits are found in the vicinity of the track projection's local position at the MTD then the closest MTD hit is assigned as the match (smallest  $\Delta Z$  and  $\Delta Y$ ). If multiple tracks are able to match to the same MTD hit, then the closest track is chosen.

The track projection procedure takes small steps along an ideal helix to determine a tracks position at the MTD radius. After each small step taken along the track's helix, the energy loss is calculated and used to modify the trajectory. Since the magnetic field direction changes between the inside and outside of the magnet, the magnetic field direction is checked at each step. The magnetic field map in the transverse ( $x$ - $y$ ) plane is shown in Fig. 3.13a.



**Figure 3.14 :** The luminosity of collected dimuon triggered events as a function of time through the run. The red line shows the initial luminosity goal, the black shows what was actually collected, and the blue shows a projection based on the first several days of data collection. The p+p dataset is shown in (a) and the p+Au dataset is shown in (b).

Within the inner magnet, the magnetic field is a uniform  $-0.5$  T in the  $z$  direction. Inside the magnet steel the field strength is  $+1.6$  T and 0 outside the magnet. Instead of computing the energy loss on a track-by-track basis, the average energy loss based on a GEANT [137] simulation was parameterized as a function of the track momentum. This parameterization, produced by averaging over the energy loss contribution from multiple detectors is shown in Fig. 3.13b.

### 3.3 Data Sets and Event Selection

This study uses data from p+p collisions at  $\sqrt{s} = 200$  GeV, p+Au and Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV collected during the years 2014 and 2015. Data from the MTD related triggers is stored in the `st_mtd` data stream. See Table 3.2 for the triggers included in the `st_mtd` data stream along with the number of events collected, their trigger IDs, and the sampled luminosity. All of the studies in this work use data from only the dimuon trigger condition.

Events were selected from those triggering the dimuon trigger condition by applying a cut on the primary vertex position in the longitudinal direction, by requiring that the difference between the primary vertex  $z$  position and the  $z$  position measured by the Vertex Position Detector was less than 6 cm, and by requiring a primary vertex ranking greater than or equal

**Table 3.2** : Triggers included in the st\_mtd data stream. The # of events, trigger ids, and sampled luminosity are given for each where available.

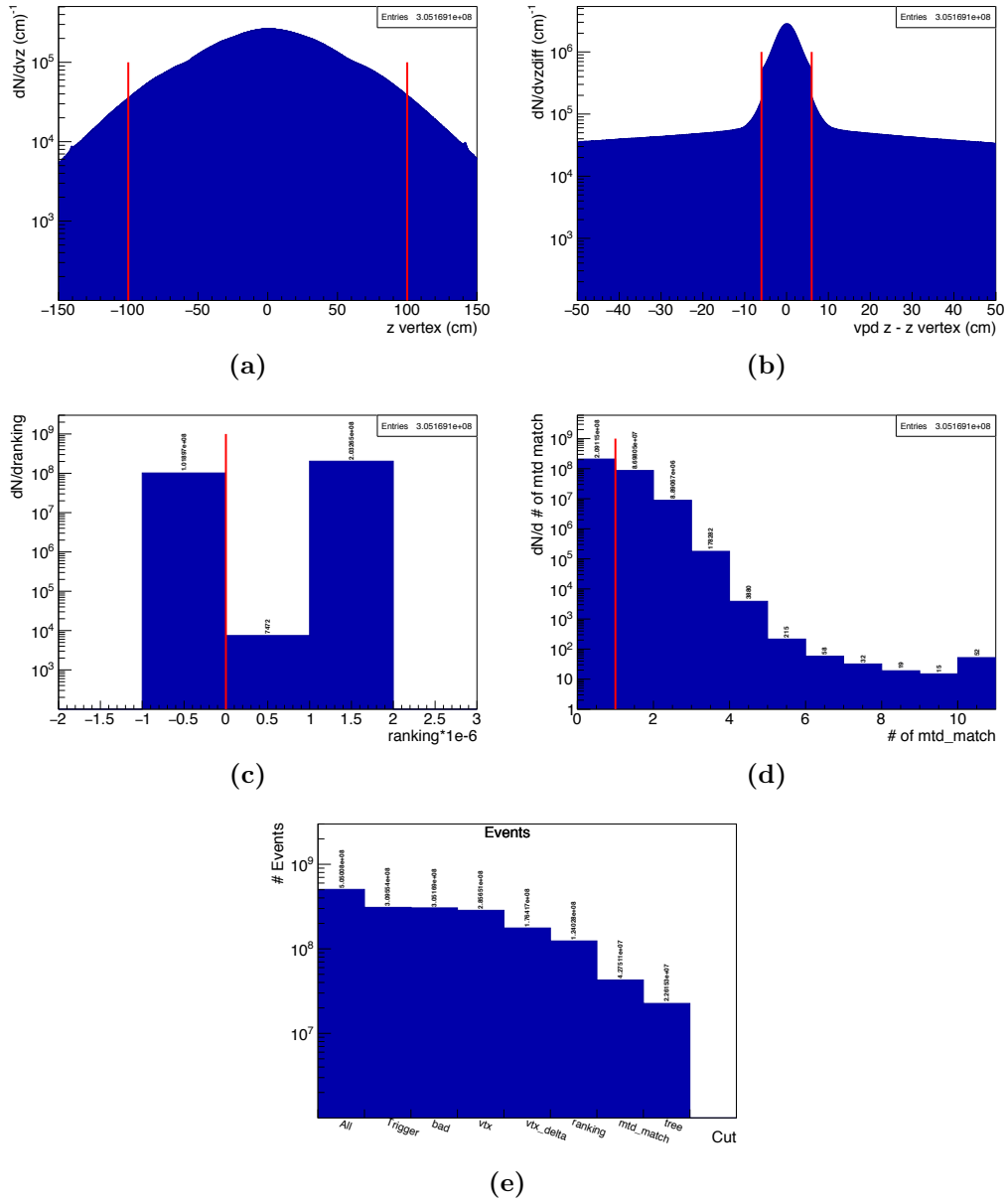
System	Trigger	# of Events (M)	Trigger IDs	Sampled Luminosity
$p + p$	single-muon	74.1	470600, 480600, 490600	$0.475 \text{ pb}^{-1}$
$p + p$	dimuon	320.4	470602, 480602, 490602	$122.13 \text{ pb}^{-1}$
$p + p$	e-muon	102.6	470601, 480601, 490601	$45.92 \text{ pb}^{-1}$
$p + \text{Au}$	single-muon	14.8	500600	$0.74 \text{ nb}^{-1}$
$p + \text{Au}$	di-muon	266.2	500602	$409.97 \text{ nb}^{-1}$
$p + \text{Au}$	e-muon	53.5	500601	$60.62 \text{ nb}^{-1}$

**Table 3.3** : Cuts applied to select events from the data collected by the dimuon trigger condition.

Quantity	Cut Criteria
$ z_{\text{TPC}} $	$< 100 \text{ (cm)}$
$ z_{\text{VPD}} - z_{\text{TPC}} $	$< 6 \text{ (cm)}$
ranking	$\geq 0$
# of MTD matched tracks	$\geq 1$

to zero. The detailed event selection cuts are shown in Table 3.3.

Each event level distribution is shown in Fig. 3.15 before any of the cuts are applied. The cuts applied to each distribution are shown with solid lines. In Fig. 3.15a and Fig. 3.15b the values between the lines are accepted. In Fig. 3.15c and Fig. 3.15d the events with values to the right of the lines are accepted. Finally, a summary of the number of events passing the application of each cut is shown in Fig. 3.15e.



**Figure 3.15 :** The Event selection cuts shown for the data collected from p+p collisions at  $\sqrt{s} = 200$  GeV. The primary vertex position ( $|z_{\text{TPC}}|$ ) before cuts is shown in (a). The difference between the TPC and VPD vertices ( $|z_{\text{VPD}} - z_{\text{TPC}}|$ ) is shown in (b). The primary vertex ranking is shown in (c). The multiplicity of MTD-matched tracks is shown in (d). In (a) and (b), the red lines show the cut applied to those quantities with the region between the red lines accepted. In (c) and (d) the red lines show the cut applied, with all events to the right of the red line being accepted. Summary of events surviving each cut (e). The bin labels correspond to : All events, dimuon triggered (after day 50), pass bad run rejection, pass  $|z_{\text{TPC}}| < 100$  cm, pass  $|z_{\text{VPD}} - z_{\text{TPC}}| < 6$  cm, pass ranking  $\geq 0$ , has at least one MTD matched track, and has at least one MTD matched primary track. The events surviving all cuts in the "tree" bin are saved to a tree for further analysis.



## Chapter 4

### Muon Identification

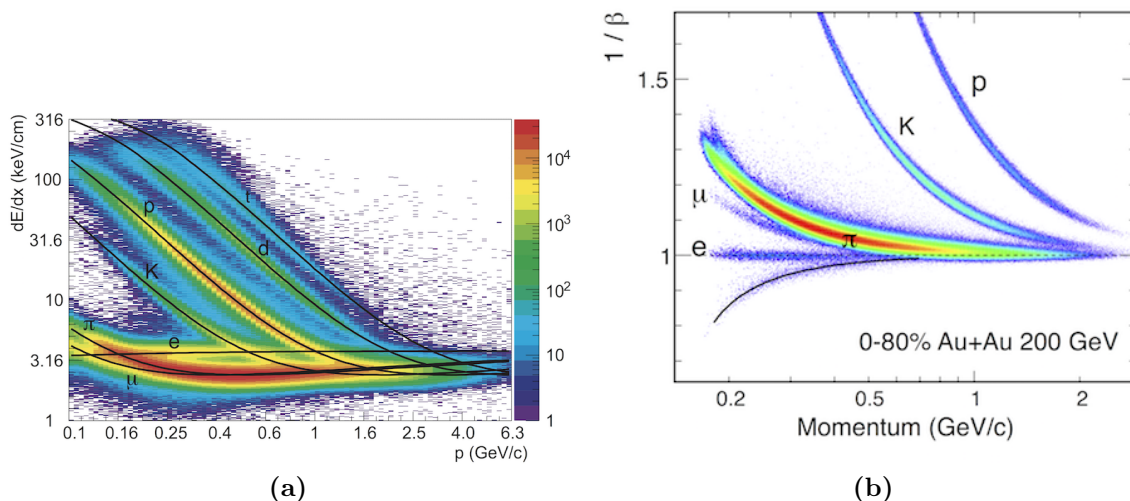
Charged particle identification in STAR has traditionally been carried out using a combination of  $dE/dx$  and Time-of-Flight. However, these techniques are not viable for distinguishing muons from charged pions since the charged pion mass ( $m_{\pi^{+/-}} = 139 \text{ MeV}/c^2$ ) and the muon mass ( $m_{\mu} = 105 \text{ MeV}/c^2$ ) are so similar. Since their masses are so similar the muon and charged pion  $dE/dx$  bands almost completely overlap (See Fig. 4.1a). Similarly, for  $p > 0.35 \text{ GeV}/c$  the  $\beta^{-1}$  bands for charged pions and muons merge and become indistinguishable (See Fig. 4.1b). Before any measurement can be done with muon pairs, muon identification must take place. The identification of low momentum ( $p < 0.35 \text{ GeV}/c$ ) muons will be discussed first. Next, the identification of muons at higher momenta using the MTD will be discussed.

#### 4.1 Identifying Low Momentum Muons with the TPC and TOF

The time of flight for a particle as a function of mass and momentum is given by :

$$\beta^{-1} = \sqrt{1 + \left(\frac{m}{p}\right)^2} \quad (4.1)$$

where  $m$  is the particle's mass and  $p$  its momentum. Due to this relation between  $\beta^{-1}$  and a particle's mass, a precise measurement of  $\beta^{-1}$  can be used to distinguish particles of different masses. It can be seen however, that for any mass  $m$ , when  $p \gg m$  the  $\beta^{-1} \approx 1.0$ . Since the charged pions and muons have very similar masses they become practically indistinguishable for  $p > 0.35 \text{ GeV}/c$ . Below this momentum, it is possible to distinguish muons and charged pions with a reasonable purity. With  $\beta^{-1}$  alone, electron contamination would make it impossible to separate muons from electrons. However, a cut on  $dE/dx$  can be used to efficiently reject electrons which would otherwise significantly overlap with the muon  $\beta^{-1}$  band. In principle, the mean of the  $\beta^{-1}$  distribution for  $\pi$  and  $\mu$  could be obtained by simply evaluating Eq. 4.1 with the respective masses. However, due to imperfect calibrations the



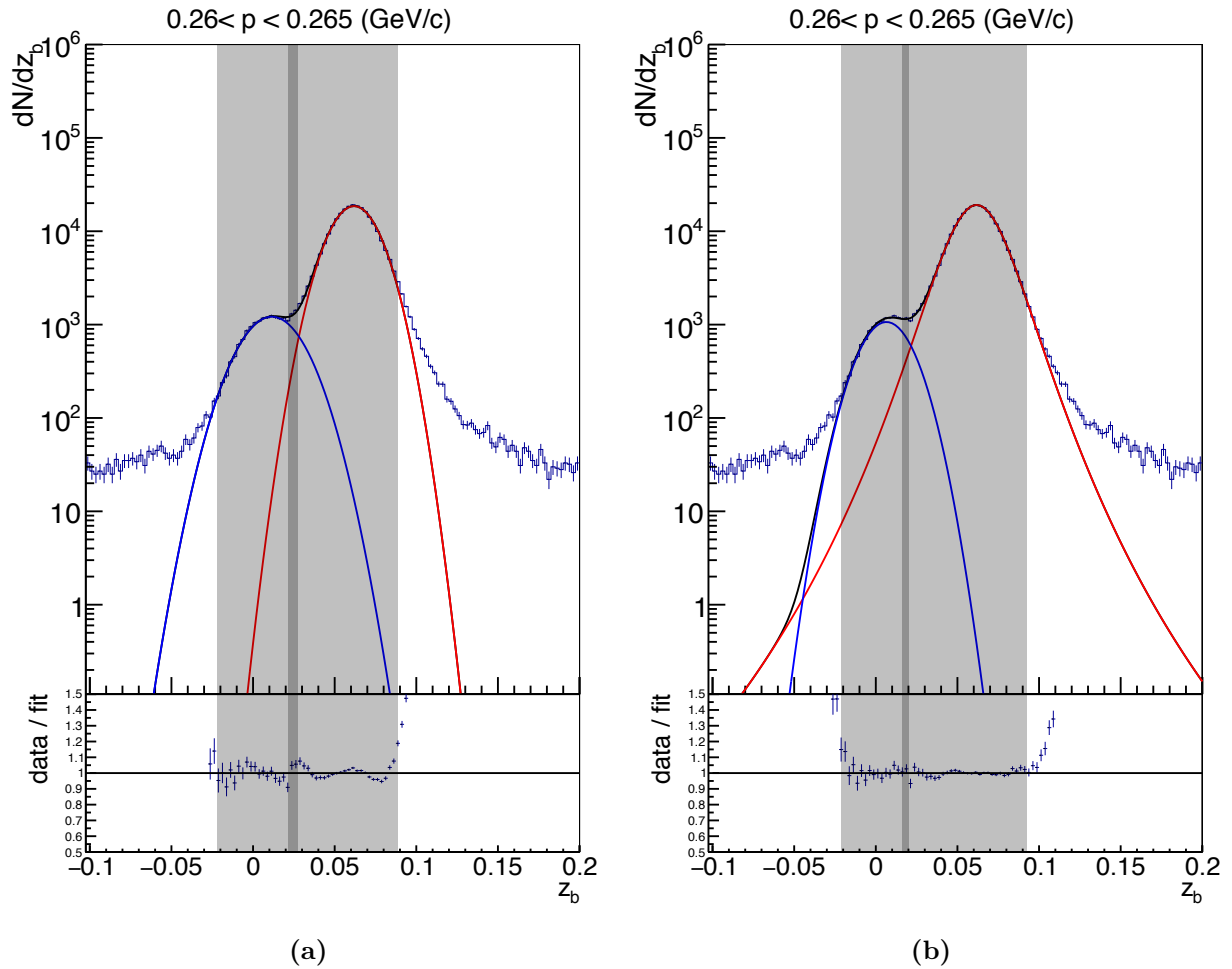
**Figure 4.1 :** The  $dE/dx$  (a) and  $\beta^{-1}$  (b) for charged particles. Since both of these quantities are proportional to the particle's mass, they can be used to identify charged particles.

$\beta^{-1}$  distributions may have significant momentum dependent shifts and broadening. In order to select  $\mu$  using the  $\beta^{-1}$  distributions we need to determine the precise location of the  $\pi$  and  $\mu$  distributions along with their widths ( $\sigma$ ). To ease in fitting and visualization the  $\beta^{-1}$  distribution is centered around the expected  $\beta^{-1}$  for  $\pi$  according to:

$$z_b = \sum_{s=\pi,\mu} [w_s \cdot (\beta_{\text{measured}}^{-1} + \mu^s(\langle p \rangle) - \mu^s(p))] - \mu^\pi(\langle p \rangle) \quad (4.2)$$

Where  $\beta_{\text{measured}}^{-1}$  is the value measured,  $\mu^s(\langle p \rangle)$  is the expected mean of the  $\beta^{-1}$  distribution for species  $s$  found by evaluating Eq. 4.1 with the mean momentum in the given momentum bin,  $\mu^s(p)$  is the expected mean of the distribution for species  $s$  found by evaluated Eq. 4.1 with the track's measured momentum, and  $w_s$  is a factor which weights the contributions for each species  $s$  based on proximity to that species expected mean. Finally, the factor  $\mu^\pi(\langle p \rangle)$  is the expected mean value for the  $\pi$  evaluated at the mean momenta in the bin. This term is subtracted off the re-weighted distribution to center it around the expected mean of the  $\pi$  peak.

After applying the re-weighting and re-centering procedure, the  $z_b$  distribution can be fit to determine the precise mean( $\mu$ ) and sigma ( $\sigma$ ) of the  $\pi$  and  $\mu$  contributions. Ideally the contribution to the  $z_b$  distribution for each species would be Gaussian. However, in practice this is not the case. In p+p collisions at  $\sqrt{s} = 200$  GeV the start time for the  $\beta^{-1}$  measurement



**Figure 4.2 :** Comparison of Gaussian (a) versus a Pearson VII (b) distribution for the pion and muon  $\beta^{-1}$  shapes. The Gaussian significantly underestimates the tails of the distributions. Since the yield of the  $\pi$  is much larger than the  $\mu$ , underestimating the  $\pi$  tails could significantly effect the fit to the  $\mu$  contribution. The Pearson VII distribution, with a free parameter related to the distribution's kurtosis, provides a much better description of the tails in the  $\beta^{-1}$  distributions.

comes from the Vertex Position Detector (VPD). Events with varying event activities will lead to varying number of hits in the VPD. The time resolution of the VPD's start-time improves with additional hits in its photo multiplier tubes (PMTs). For this reason, events with more hits in the VPD will have a better  $\beta^{-1}$  resolution than events with fewer hits in the VPD (See Fig. 3.5). In addition to the start time contribution, the non-gaussian tails of the  $\beta^{-1}$  distribution could also result from combining measurements for tracks of various lengths, variations in momentum resolution, and from an imperfect calibration of the 23,040 cells in the barrel Time of Flight detector.

Since the particle peaks in the  $z_b$  distribution have non-gaussian tails, we investigate alternatives to the Gaussian distribution for fitting the  $\pi$  and  $\mu$  peaks. The Pearson VII distribution was chosen as a possible improvement over the Gaussian distribution because it has a free parameter which can account for the kurtosis ( $\kappa$ ) of the distribution. The Pearson VII distribution is a special case of the family of Pearson distribution, in which the skewness ( $\nu$ ) is fixed at 0 to enforce a symmetric distribution. The density of the Pearson VII distribution is given by:

$$p(x|\lambda, \sigma, m) = \frac{1}{\alpha B(m - 1/2, 1/2)} \left[ 1 + \left( \frac{x - \lambda}{\alpha} \right)^2 \right]^{-m} \quad (4.3)$$

Where B is the Beta function and  $\alpha = \sigma\sqrt{2m - 3}$ . In this parameterization, the free parameters  $\lambda$ ,  $\sigma$ , and  $m$  control the mean, variance ( $\sigma^2$ ), and the kurtosis ( $\kappa$ ) of the distribution. The distribution is only well defined for  $m > 3/2$ . It should be noted that if the  $\lambda$  and  $\sigma$  parameters are held constant while the parameter  $m$  approaches infinity, the standard normal distribution arises. Additionally, the Pearson VII distribution is equivalent to the Student's t-distribution under the transformations :

$$\begin{aligned} \lambda &= \mu \\ \alpha &= \sqrt{\nu\sigma^2} \\ m &= \frac{\nu + 1}{2} \end{aligned} \quad (4.4)$$

with the requirement that  $\nu > 0$ . Figure 4.2 shows a comparison of the fits to the  $z_b$  distribution in  $0.26 < p < 0.265$  GeV/c for the Gaussian and Pearson VII distributions.

It can be seen that the Gaussian distribution significantly underestimates the tail of the  $\pi$  contribution while the Pearson VII distribution provides significantly better agreement as is visible in the ratio plots in the lower panels. In these figures the fit regions are shown with a shaded grey box. The  $\pi$  fit contribution is shown in red and the  $\mu$  fit contribution is shown in blue. The total fit for  $\pi$  and  $\mu$  is shown in black. In each figure the lower panel shows the ratio of the data to the fit result.

## 4.2 Muon Identification Information from the MTD

At momenta above  $\sim 0.35$  GeV/c muons can no longer be separated from charged pions and electrons with using  $dE/dx$  and  $\beta^{-1}$  alone. For this reason, muon identification at higher momenta employs the MTD. The MTD provides several variables that aide in the identification of muons and the rejection of pions, kaons, and other hadrons. The information provided by the MTD consists of precise timing ( $\sim 100$  ps resolution) and position ( $\sim 1$ -2 cm resolution). A total of 11 variables (6 MTD variables and 5 track variables) are used for MTD based muon identification. The full list of quantities measured by the TPC, and MTD detectors that are used in this study for muon identification are:

- $\Delta TOF$  - Difference between the calculated time-of-flight using a muon hypothesis versus the time-of-flight measured by the MTD.
- $\Delta Z$  - Difference between the local Z position calculated using a muon hypothesis versus the position measured by the MTD.
- $\Delta Y$  - Difference between the local Y position calculated using a muon hypothesis versus the position measured by the MTD from the center of the matched strip.
- cell - the geometric strip index ranging from 0 - 11 with 0 and 11 at the outside edges of each module. The average amount of steel between the interaction point and the MTD module is lowest at the edges.
- module - the geometric module index ranging from 0 - 4 along the z direction of the STAR coordinate system.
- backleg - the geometric backleg index ranging from 0 - 29. The amount of material

between the interaction point and the MTD backlegs varies as a function of backleg since the detector is not fully symmetric in the  $\phi$  direction.

- $n\sigma_\pi$  - the  $dE/dx$  information measured by the TPC. The value normalized by the expected value for the  $\pi$  and corrected for detector resolution is used for simplicity. The value of  $n\sigma_\pi$  for muons is on average  $\sim+0.5$ .
- DCA - Distance of closest approach of the track to the primary collision vertex.
- $p_T$  - Transverse momentum of the track. The  $\Delta TOF$ ,  $\Delta Y$ , and  $\Delta Z$  resolutions depend strongly on  $p_T$ .
- $q$  - the track charge measured from its curvature.

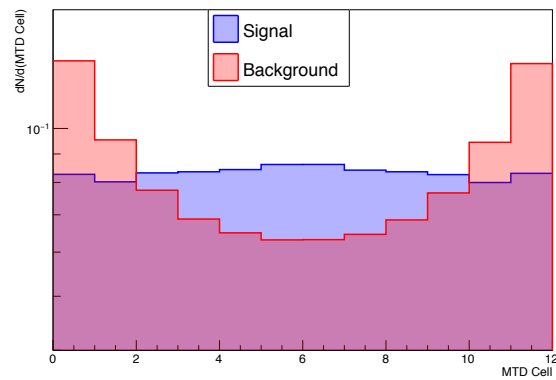
These "features" will be used as the inputs when training neural network classifiers in Sec. 4.4.

## 4.3 Preparation of Training Datasets

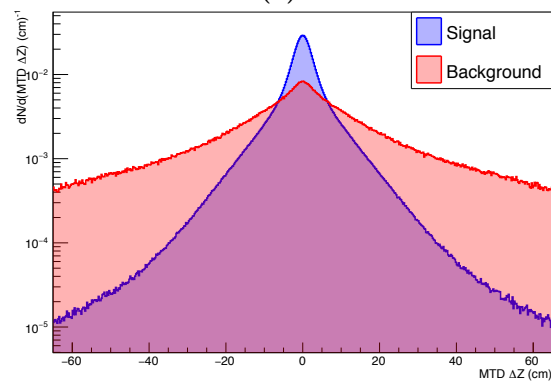
### 4.3.1 Simulation of Muon Telescope Detector

In Section 4.4 the training and use of ANNs to perform a two-class classification problem to distinguish signal muons from various types of backgrounds is discussed. This type of ANN based classification is an example of supervised learning and therefore requires labeled datasets for the training phase. A Monte Carlo (MC) simulation procedure is used to generate the labeled signal and background datasets needed to train the supervised learning algorithms discussed in Section 4.4. We define our signal class to be muon tracks reconstructed within the tracker, associated with the primary event vertex, and that are matched to a hit in the MTD. In contrast, the background class includes all other sources of tracks that may match to a signal in the MTD and result in a reconstructed track in the tracker. The main sources of background are a result of:

- Punch through hadrons: e.g.  $\pi^\pm$ ,  $K^\pm$ , and  $p/\bar{p}$
- Charged pion decays:  $\pi \rightarrow \mu + \nu$
- Charged Kaon decays:  $K \rightarrow \mu + \nu$



(a)



(b)

**Figure 4.3 :** Simulated MTD cell (a) and  $\Delta Z$  distributions for signal (primary  $\mu s$ ) and background sources. The signal and background distributions are normalized for comparison. The effect of varying amounts of steel in the  $\phi$  direction can be clearly seen in the cell distribution. Hadrons are significantly more likely to punch through the steel guarding the edge cells (0 & 11) than the central cells.

It should be noted that, since particle decay is a stochastic process, some fraction of  $\pi$  and K will decay at or near the primary vertex producing a daughter  $\mu$  that travels in almost the same trajectory as the parent and is reconstructed as part of the same track in the tracker. With only the PID variables considered here, these types of tracks will be indistinguishable from muons created at the primary interaction point. Secondary decays can be better distinguished from primary produced muons with the help of a tracker with finer DCA resolution such as STAR's Heavy Flavor Tracker (HFT) [138]. However, due to compounding efficiency and acceptance, the number of tracks reconstructed with both the HFT and the MTD information is far too limited for a physics analysis.

The procedure used to forward model the signal and backgrounds consists of three main steps: kinematic event generator, a simulation of the STAR detector, and full event reconstruction. First, events are generated with  $\sim 20$  particles/event to mimic the multiplicity of primary tracks in a  $p+p$  collision at  $\sqrt{s} = 200$  GeV. Each track in the event is randomly chosen to be a  $\mu$ ,  $\pi$ , K, or  $p$ . The kinematics of each particle are sampled from flat distributions in  $0 < p_T < 10.0$  GeV/c,  $|\eta| < 0.8$ , and  $-\pi < \phi < \pi$ . The particle species and kinematics are then fed into a GEANT3 [137] based simulation of the full STAR geometry. The GEANT3 simulation performs decays of unstable particles, models energy loss of particles traversing media, and interactions with detector materials. Finally, full event reconstruction is performed on the result of the GEANT3 based simulation. This step performs charged particle reconstruction using the simulated hits in the TPC, determines the event's primary interaction vertex, and computes the dE/dx of reconstructed tracks. After tracking is complete the tracks are matched to the simulated MTD hits. The result of this simulation is a set of the PID variables for each of the signal and background processes. An example of the MTD  $\Delta Z$  and cell variables are shown for signal and background in Figure 4.3.

The simulation code discussed in this section was used to investigate the types of processes that contribute to the background candidates. The MonteCarlo simulation provides information about all generated tracks and interactions. With this information we can look at the lineage of any  $\pi$  matched to the MTD. Using this technique we find that the  $\pi$  background can be separated into three rough categories including 1) clean punch-through, 2) material interaction, 3) secondary decays.

The clean punch-through tracks are pions that do not interact hadronically with any



Option	Value
STAR Library Version	SL16j
STAR Geometry Version	y2015c
Magnetic Field	RFF
Chain Options	DbV20160418 y2015c ReverseField fzin mtdSim tpcX AgML tpcDB TpcHitMover Idst ITTF UseXgeom BAna VFMinuit l3onl emcDY2 fpd trgd ZDCvtx btof mtd mtdCalib BEmcChkStat CorrX OSpaceZ2 OGridLeak3D tpcrs TpxRaw VFMCCE TpxClu big MakeEvent IdTruth McEvent McAss

**Table 4.1** : Configuration options used for the MTD Simulation. The code can be found here: <https://github.com/jdbrice/MtdSimGun>.

material including the magnet steel before crossing the MTD’s active regions. These tracks will not loose energy due to hadronic interactions and therefore will not deviate significantly from the expected trajectory of a muon track making them difficult to reject with the MTD  $\Delta Y$ ,  $\Delta Z$  or  $\Delta TOF$  variables.

The second class of  $\pi$  tracks interact with the beam pipe, magnet, or other material. In these cases the primary  $\pi$  track will cause a shower of secondary particles that may match to the MTD. These tracks are most likely to have a large MTD  $\Delta Y$ ,  $\Delta Z$  or  $\Delta TOF$  variable and are therefore most easily distinguished.

```

pi+ (RC) (4,-0.42,1.7)
|--pi+ (MTD) (0.38,0.55,2.1)
|  |--PROTON(0.18,0.54,2.4)
|  |--pi+(0.11,-0.54,-0.71)
|     |--mu+(0.023,-0.76,-1.5)
|        |--e+(0.027,-0.76,-0.92)
|           |--NEUTRINO(0.032,0.19,-2.8)
|           |--NEUTRINO(0.034,0.47,1.1)
|           |--NEUTRINO(0.023,0.76,1.7)
|           |--PROTON(0.095,-1.4,-2.5)

```

**Listing 1:** An example lineage tree for a primary  $\pi$  that interacts with material and has a daughter  $\pi$  hit the MTD. RC indicates that the MC track was reconstructed and MTD indicates that the MC track produced a hit in the MTD. The tuple after each particle is its  $(p_T, \eta, \phi)$

The third class of  $\pi$  background is the case in which the  $\pi \rightarrow \mu + \nu$  decay occurs outside the tracking detector such that the  $\pi$  track is reconstructed as primary (by construction).

Secondary muons that result from decays so early that the muon track is fit to the primary vertex are not distinguished by this stage of the muon identification. A very tight DCA cut is needed to reject muons resulting from decays very near the primary vertex. The DCA resolution of the TPC does not allow rejection of such secondary muons.

```

pi+ (RC) (3.8,-0.51,1.2)
|--mu+ (MTD) (2.5,-0.52,1.1)
|--NEUTRINO(1.2,-0.52,1.2)

```

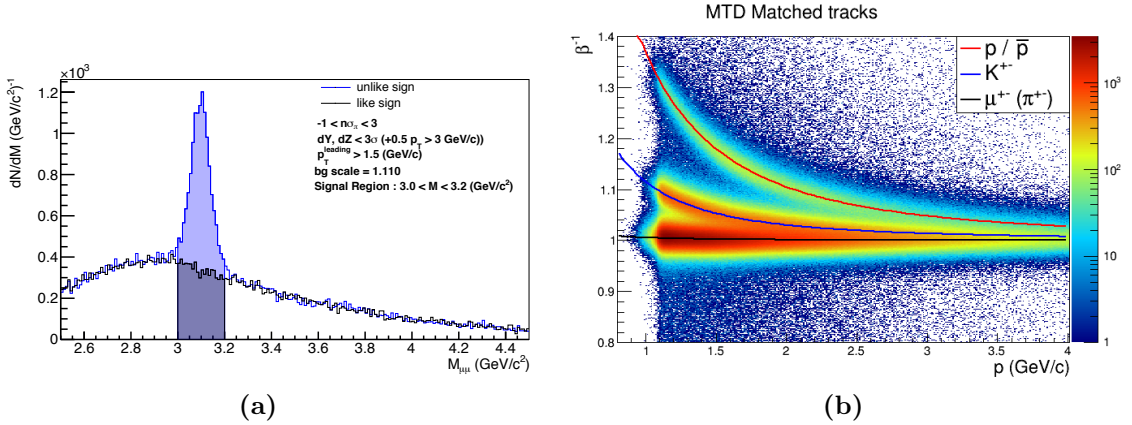
**Listing 2:** An example of the lineage tree for a  $\pi \rightarrow \mu + \nu$  decay

### 4.3.2 Extracting $\Delta TOF$ distributions from Data

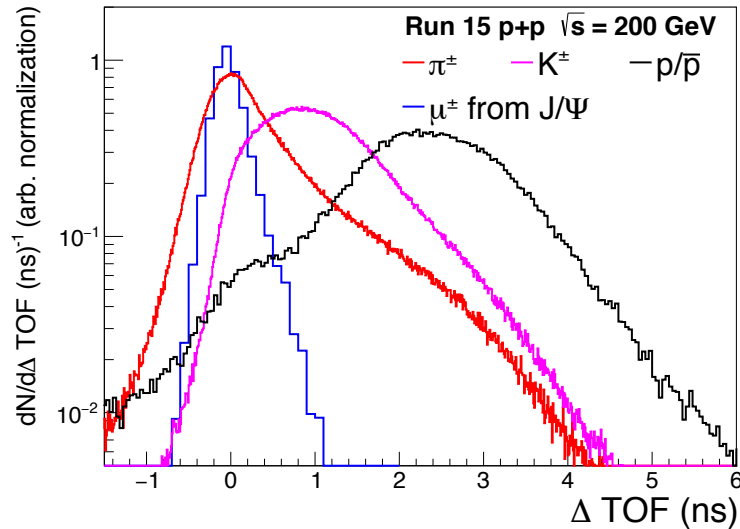
Due to complications simulating the timing electronics, the MTD simulation does not provide time-of-flight information. Since the  $\Delta TOF$  distributions are not provided by the MTD simulation a data-driven approach is employed to determine these distributions separately for the signal and background classes.

For this procedure 1D cuts are applied to all PID variables except  $\Delta TOF$ . See Table 4.2 for the cut values used for signal and background classes. With these cuts we can obtain a relatively pure  $J/\psi$  sample from which to extract the  $\Delta TOF$  probability distribution functions (PDFs) for signal. Specifically the signal PDF is extracted from the  $J/\psi$  mass peak ( $3.0 < M < 3.2 \text{ GeV}/c^2$ ) with the background under the peak estimated using the like-sign pairs in the same mass region. The  $\Delta TOF$  from the like-sign background is properly scaled and subtracted from the peak region to remove background contributions. The signal  $\Delta TOF$  PDF is shown in Figure 4.5. The background  $\Delta TOF$  PDF is extracted from tracks passing an inverted set of cuts meant to exclude all signal muons. These cuts are shown in Table 4.2.

The background  $\Delta TOF$  distribution is further separated into the contributions for  $\pi$ , K, and  $p$ . The sub-sample of tracks which match to both the MTD and BTOF are used to extract the MTD  $\Delta TOF$  distribution for  $\pi$ , K, and  $p$  separately. The  $\beta^{-1}$  distribution for all tracks passing basic track cuts matched to both the MTD and BTOF are shown in Fig. 4.4b. In this figure, there are clear  $\beta^{-1}$  bands corresponding to  $\pi + \mu$ , kaons and  $p$ . The MTD  $\Delta TOF$  distribution for these three distributions were extracted by cutting around a



**Figure 4.4 :** (a) The invariant mass distribution for unlike-sign and like-sign pairs near the  $J/\psi$  mass. The N-1 cut technique is used to maximize the  $J/\psi$  significance by cutting on all MTD PID variables except the  $\Delta TOF$  distribution. A  $p_T^{\text{leading}} > 1.5$  (GeV/c) cut is applied to further improve the purity in the  $J/\psi$  mass region. The  $\beta^{-1}$  vs. momentum distribution for all tracks passing basic QA cuts that are matched to hits in the MTD and the BTOF detectors. (b) The  $\beta^{-1}$  calculated from the BTOF information shows clear contributions from  $\pi$ , K, and  $p/\bar{p}$ .



**Figure 4.5 :** The  $\Delta TOF$  distributions for  $\mu^\pm$  from  $J/\psi$ ,  $\pi^\pm$ ,  $K^\pm$ , and  $p/\bar{p}$ .

$J/\psi$ Selection Cuts
$3.0 < M_{\mu\mu} < 3.2$ ( $\text{GeV}/c^2$ )
$\text{DCA} < 1.0$
$-1 < n\sigma_\pi < 3$
$ \Delta Y  < 3\sigma$ ( $+0.5, p_T > 3.0(\text{GeV}/c)$ )
$ \Delta Z  < 3\sigma$ ( $+0.5, p_T > 3.0(\text{GeV}/c)$ )
$p_T^{\text{leading}} > 1.5$ ( $\text{GeV}/c$ )

**Table 4.2 :** Cuts used for determining the signal and background  $\Delta$ Time-of-Flight PDFs.

$K_S^0$ Selection Cuts
$0.472 < M_{\pi\pi} < 0.522$
decay length $> 2.7$ cm
daughter mutual DCA $< 1.5$ cm
pointingAngle $< 0.162 + 0.1123p_T + 0.025p_T^2$
$ n\sigma_\pi  < 3$

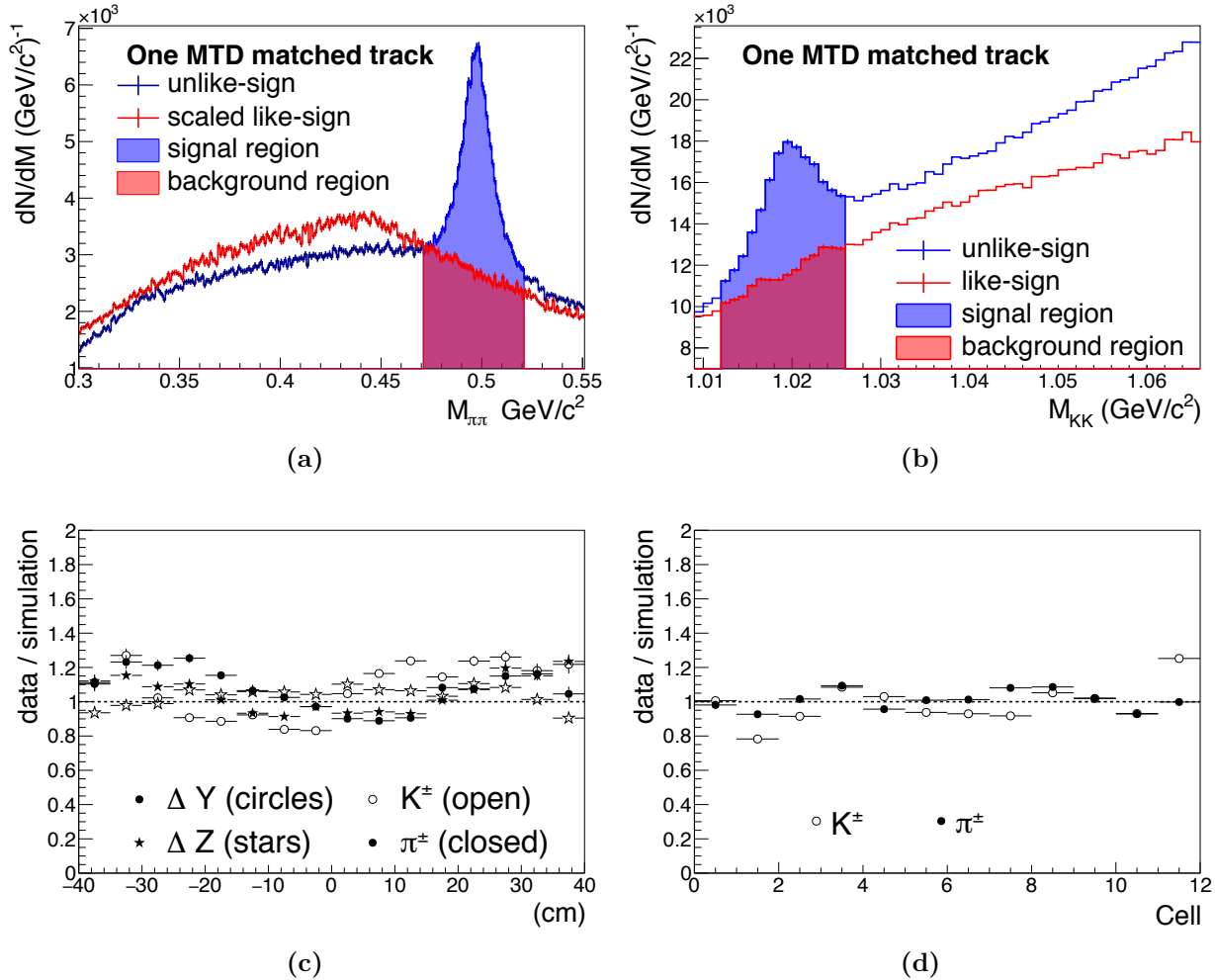
**Table 4.3 :** Cuts used to select  $K_S^0 \rightarrow \pi^+\pi^-$  decays. The daughter pions provide a  $\pi$ -enhanced sample that can be compared to the  $\pi$  Monte Carlo simulation.

given species  $\beta^{-1}$  distribution from BTOF.

### 4.3.3 Background MC Closure Test Using Identified $K_S^0 \rightarrow \pi^+\pi^-$ and $\phi \rightarrow K^+K^-$ Decays

Selecting  $K_S^0 \rightarrow \pi^+\pi^-$  decays in data provides a  $\pi^\pm$  enhanced sample that can be used to test the validity of the MC simulation procedure for the  $\pi^\pm$  background sources. The selection of  $K_S^0$  candidates is carried out by applying the topological selection cuts listed in Table 4.3. In order to increase the available statistics for comparison, only one of the  $K_S^0$  daughters is required to have a matching hit in the MTD. Figure 4.6a shows the  $\pi^+\pi^-$  invariant mass distribution near the  $K_S^0$  mass used to select  $\pi^\pm$  daughter tracks. The  $\pi^\pm$   $\Delta Y$ ,  $\Delta Z$ , and Cell distributions are computed using the unlike-sign distribution minus the scaled like-sign distribution for each variable in the  $K_S^0$  mass region ( $497 \pm 25$  MeV/ $c^2$ ).

Distributions with an enhanced kaon yield can be selected from daughters of  $\phi \rightarrow K^+K^-$  decays. The  $K^+K^-$  invariant mass distribution around  $M_\phi$  is shown in Fig. 4.6b for the case in which one track is matched to an MTD hit. The  $K^\pm$   $\Delta Y$ ,  $\Delta Z$ , and Cell distributions are computed using the unlike-sign distribution minus the scaled like-sign distribution for each variable in the  $\phi$  mass region ( $1.019 \pm 0.007$  MeV/ $c^2$ ). The comparison between the  $\Delta Y$ ,  $\Delta Z$  and MTD Cell distributions from MC and data for  $\pi^\pm$  and  $K^\pm$  tracks are shown in Fig. 4.6c and 4.6d. The data / simulation ratios show that the  $\Delta Y$ ,  $\Delta Z$  and MTD Cell distributions agree within  $\sim \pm 20\%$ .



**Figure 4.6 :** The  $M_{\pi^+\pi^-}$  distribution near the  $K_S^0$  mass shown for the cases in which only one track is matched to an MTD hit (a) and the  $M_{K^+K^-}$  distribution near the  $\phi$  mass shown for the cases in which only one track is matched to an MTD hit (b). The  $\Delta Y$  and  $\Delta Z$  data / simulation ratio for both  $\pi^\pm$  and  $K^\pm$  (c). The MTD Cell data / simulation ratio for both  $\pi^\pm$  and  $K^\pm$  (d).

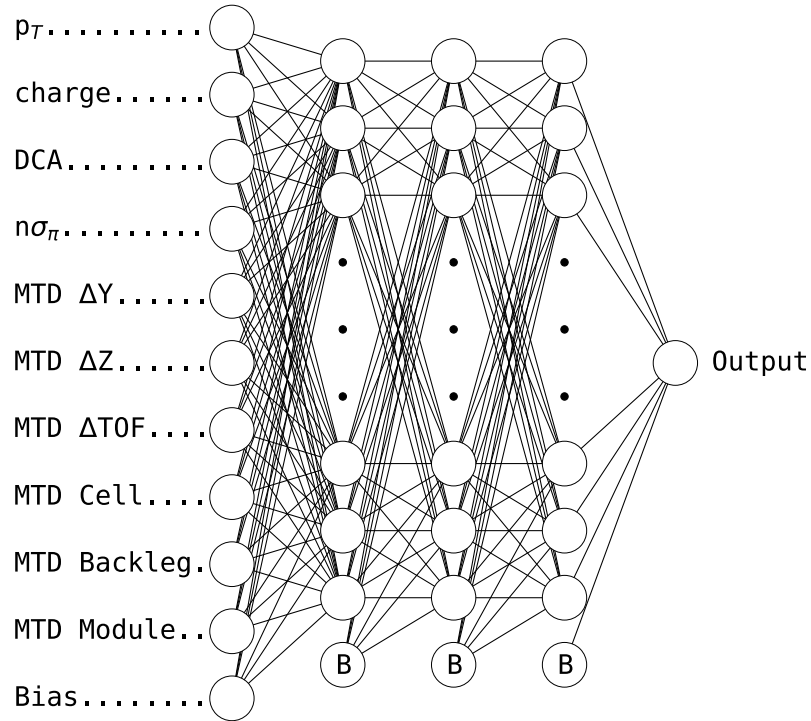
## 4.4 Training and Evaluation of Artificial Neural Networks

### 4.4.1 Introduction to Artificial Neural Networks and Machine Learning

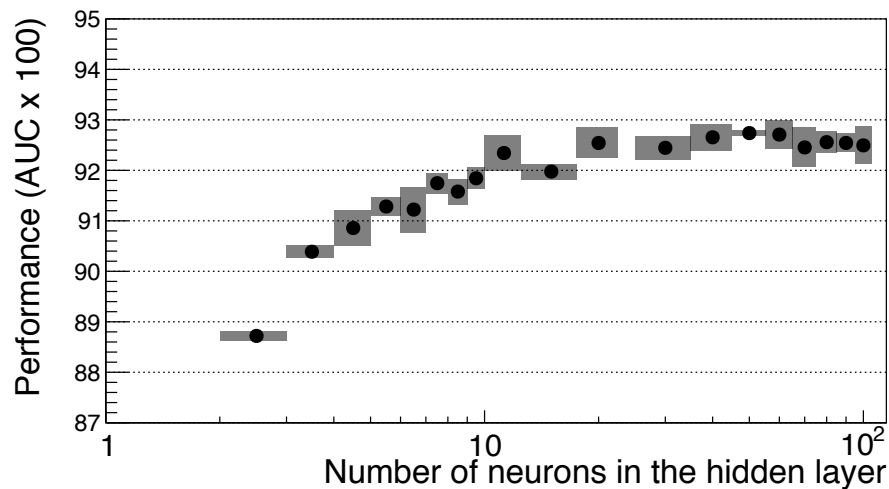
In this section dense Multilayer Perceptrons (MLP), a type of feedforward ANN, are trained as continuous classifiers for the purpose of muon identification. Back-propagation is used to train all of the neural networks in these studies. Since back-propagation is the key to training an artificial neural network, a description of the technique is included in Appendix A. First, shallow artificial neural networks (SNN) will be discussed. A shallow artificial neural network is defined by the presence of a single hidden layer of neurons between the input and output layers. The universal approximation theorem [139,140] states that a feedforward ANN with certain activation functions and at least one hidden layer containing a finite number of neurons can approximate any continuous function on compact subsets of  $\mathbb{R}^n$ . However, the universal approximation theorem makes no claim about the size of the hidden layer required to approximate a given function. In practice the number of neurons in the hidden ( $N_H$ ) layer may need to be intractably large to approximate the desired function with acceptable error. In addition, with increasing number of neurons the risk of over training can increase resulting in a model capable of representing the input data with small error but with very poor generalization performance.

### 4.4.2 Shallow Neural Networks

In this section an exploration of the performance of a large set of SNNs as a function of the number of neurons in their hidden layer is presented. The models are trained using the Toolkit for Multivariate Data Analysis with ROOT (TMVA) [141]. Table 4.4 lists the parameters used in the training phase for all models. Each model is trained on a random subset of 100K signal events and 100K background events. A disjoint testing sample is drawn from 250K signal and 250K background events to test the model's response and to evaluate the over training score. The use of a Monte Carlo generator for producing the labeled training samples allows an essentially unlimited number of labeled data sets and allows independent samples for training and testing. When an unlimited labeled data set is not available, bootstrapping techniques can be used to evaluate model performance [142,143]. The SNN models include a bias neuron in the each of the input and hidden layers to account for trivial offsets in the



**Figure 4.7 :** An example of a dense multilayer perceptron neural network architecture. The shallow neural networks have only a single hidden layer of neurons between the input and output layers. The deep neural networks have two or more. Bias neurons in the hidden layers are marked with a “B”.



**Figure 4.8 :** The signal vs. background rejection power as a function of the number of neurons ( $NN_{HL}$ ) in the hidden layer of a shallow neural network. The performance of the SNN’s are quantified using the AUC - the area under the background rejection vs. signal efficiency curve (See 4.4.4 in text). The points are the mean value of 10 models trained with different random samples. The uncertainties show the  $1\sigma$  variance of the models assuming a Gaussian variance.

**Table 4.4 :** Parameters used in the training phase for the shallow and deep neural networks.

Parameter	Value
Neuron Activation Function	tanh
Estimator Type	Mean Square
Neuron Input Function	sum
Training Method	Back-Propagation
Learning Rate	0.02
Decay Rate	0.01
Learning Mode	Sequential
Max # Training Cycles	500
Testing Rate	100

mean value of the data. The bias neuron is always "on" - i.e. it provides an input of 1 so that weights between it and other neurons are constant factors. The use of a bias node in the input and each hidden layer has become standard practice in neural network architecture design. An example of one SNN from the set of SNN with 15 neurons in its hidden layer can be seen in Fig. 4.7 along with its classifier response.

Shallow neural networks were trained with 1 to 500 neurons in the hidden layer. For each value of  $N_H$ , 10 models were trained with different randomized training and testing samples. The result of the SNN scan are summarized in Fig. 4.8 where the background rejection power quantified through the AUC, the area under the background rejection vs. signal efficiency curve (higher is better), is shown as a function of  $N_H$ . Each point shows the mean response of 10 models with uncertainties that show the  $1\sigma$  variation between the response of the 10 models assuming a Gaussian variance. The background rejection power of the SNN shows clear improvement as  $N_H$  is increased until  $N_H \approx 30$ . Above  $N_H \approx 30$ , adding more and more neurons provides relatively smaller and smaller improvement in the background rejection power.

#### 4.4.3 Deep Neural Networks and Hyper-parameter Optimization

Deep neural networks, in contrast to SNN which contain only a single hidden layer, contain two or more hidden layers. The additional hidden layers can allow a network to learn complex relationships between input features with far fewer neurons and connections than a shallow network would need. Depending on the application it is also common for DNNs to combine various types of layers, such as convolutional layers, to promote the learning of specific types

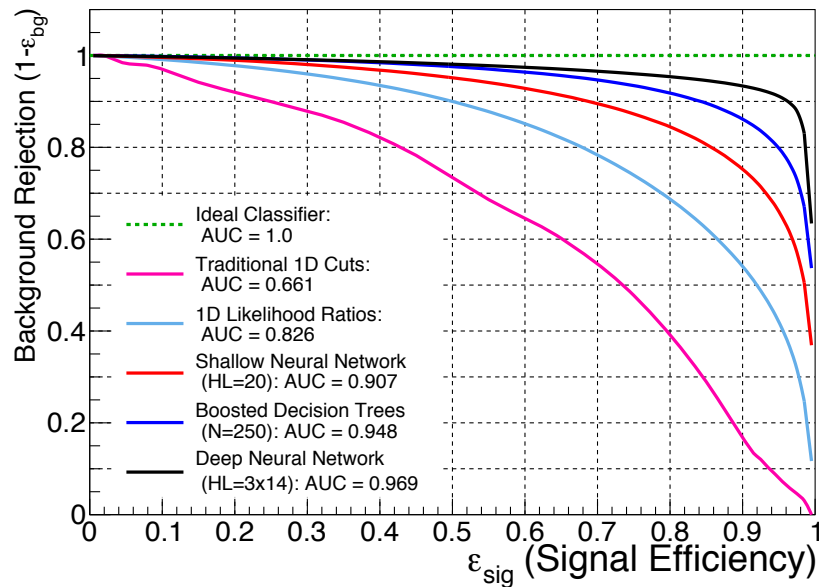


of relationships. The term, “Deep Learning” is often used when a DNN contains several hidden layers with varying types of relationships. In this case, only the simplest type of DNNs is explored, specifically dense multi-layer perceptrons.

For the case of muon identification, the goal is to determine if DNNs can provide a better classification performance than SNNs with a reasonable number of neurons. Answering this question is not trivial though, since the performance and response of a deep MLP can depend strongly on the number of hidden layers and the number of neurons in each layer. The process of determining the optimal DNN architecture is often referred to as hyperparameter optimization. A grid-search strategy is used in this case to search the optimal DNN architecture on a grid of the # of hidden layers and the number of neurons in each layer. The order of the hidden layers was also encoded so that a network with hidden layers HL=5,6,7 (i.e. 5 neurons in the first hidden layer, 6 in the second, and 7 in the third) would be a distinct grid-point compared to one with HL=7,6,5 despite having the same # of hidden layers and number of neurons. For each grid-point a DNN was trained and evaluated based on the following criteria:

- Signal vs. background rejection power
- Prefer simplest NN architecture (fewer # of neurons is better and fewer # of hidden layers is better)
- Prefer monotonically increasing signal-to-background ratio as a function of NN response

These three criteria were considered in order to determine the optimal set of DNN hyperparameters. Each DNN was trained using the parameters listed in Table 4.4 with only the architecture related parameters varying. Training DNNs can require significantly more time and larger labeled samples compared to SNNs to reach convergence. The DNNs were trained with 1M signal and 1M background events and took between 10 and 100 times longer to train than the set of SNNs depending on the specific architecture. However, the time-cost required to train DNNs can be greatly reduced by employing modern libraries like TensorFlow that have been heavily optimized for parallelized network training using GPUs [144].

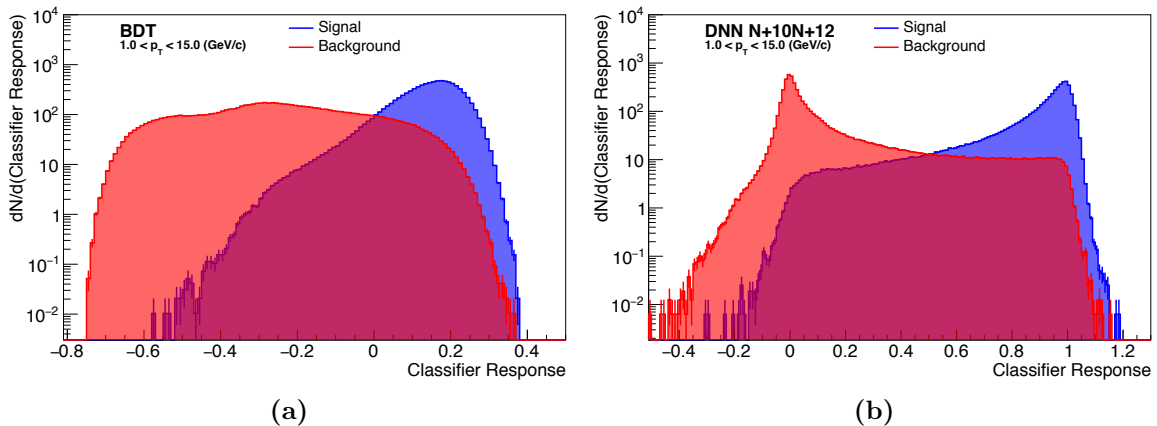


**Figure 4.9 :** A Neural Network with two hidden layers containing  $N=30,20$  neurons (a) and its output for simulated signal and background tracks (b).

#### 4.4.4 Comparison of Multivariate Classifiers

In the previous section neural networks were trained as classifiers for the purpose of separating signal muons from various background sources. The performance of the neural network based classifiers are compared using modified receiver operating characteristic (ROC) curves in Fig. 4.9 by plotting the background rejection power ( $1 - \epsilon_{bg}$ ) vs. the signal efficiency ( $\epsilon_{sig}$ ). The performance of a classifier can be succinctly summarized with the area under the curve (AUC) of the background rejection vs. signal efficiency curve. An ideal classifier is able to reject 100% of the background while providing 100% signal efficiency and has an AUC of 1. On the other hand, a random guess classifier should have a 50/50 chance of correctly guessing the class and has an AUC of 0.5.

The neural network classifiers shown in Fig 4.9 are also compared with classifiers employing optimized 1D cuts, 1D likelihood ratios, and boosted decision trees (BDTs). The cuts used in the 1D cut classifier were optimized on the  $J/\psi$  peak in p+p collisions at  $\sqrt{s} = 200$  GeV. Both the 1D likelihood ratio classifier and the BDTs were trained using the TMVA package. The 1D likelihood ratio classifier was trained with default parameters using spline interpolation when building the feature PDFs. The track  $p_T$  and charge ( $q$ ) variables were removed from the 1D likelihood classifiers since they should not be used directly for muon



**Figure 4.10** : The classifier response for the BDT (a) and DNN (b). These are the responses for all simulated tracks, so the distribution is mostly dominated by the high  $p_T$  tracks.

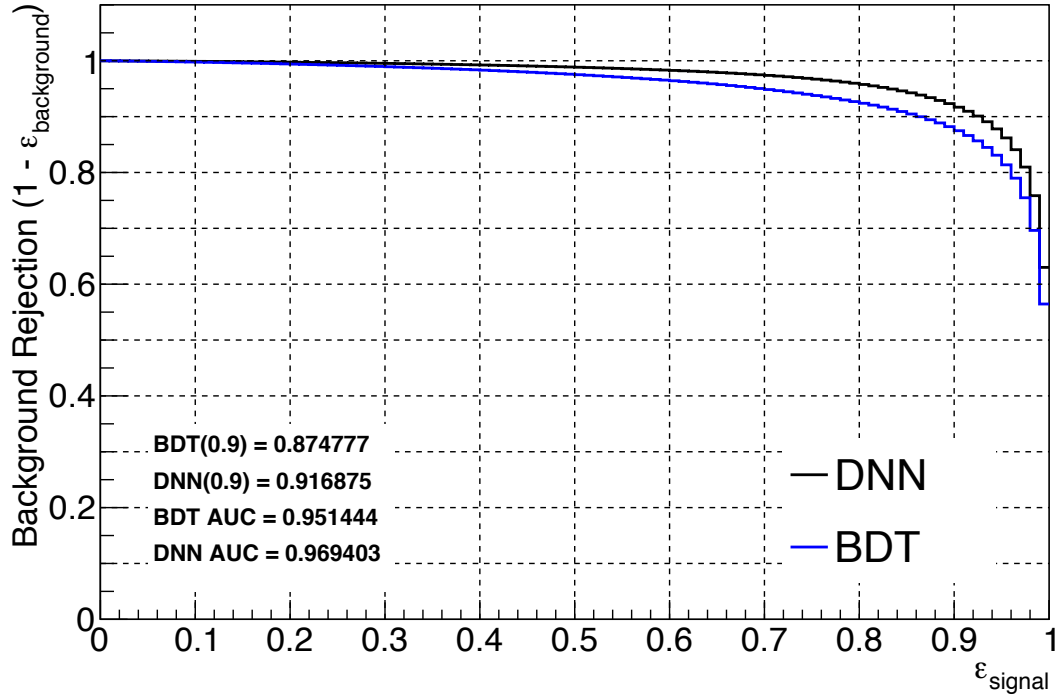
identification. Additionally, since 1D likelihoods cannot properly incorporate the  $p_T$  dependence of the  $\Delta\text{TOF}$ ,  $\Delta Y$ , and  $\Delta Z$  features, the 1D likelihood classifier was evaluated only for tracks in a narrow  $p_T$  range ( $1.4 < p_T < 1.6$ ). A more thorough look at using likelihood ratios for muon identification with the MTD can be found in [145]. The BDT classifier was trained with  $N\text{Trees}=250$  and  $MaxDepth=5$  with all other parameters set to the defaults.

#### 4.4.5 Identification of Muon Pairs

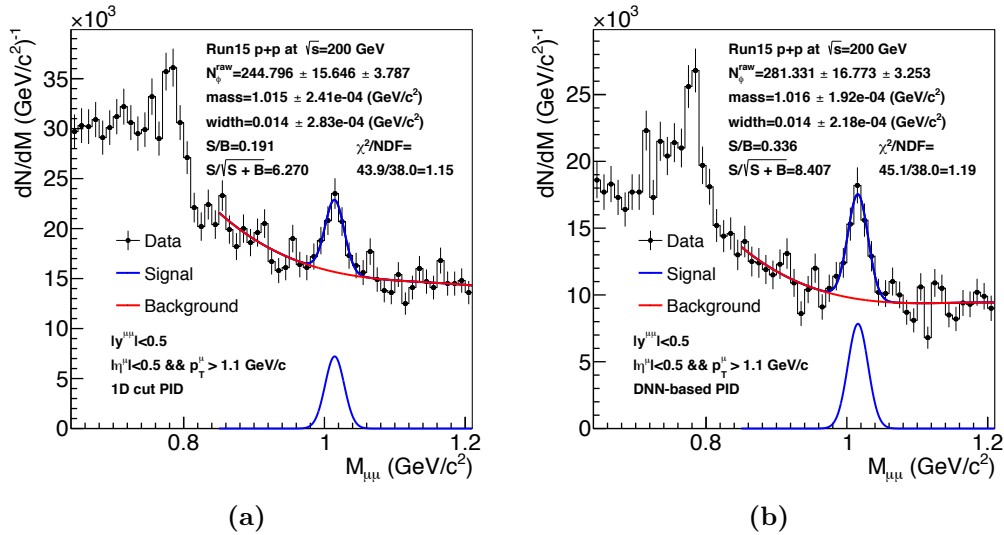
The DNN classifier out-performed the other multivariate classifiers investigated in Sec. 4.4 based on an analysis of the background rejection power vs. signal efficiency evaluated on a testing sample of simulated events. We can further test the performance of the DNN classifier by applying it to the dimuon data collected from p+p collisions at  $\sqrt{s} = 200$  GeV. The decay of resonances to muons, like the  $\phi \rightarrow \mu^+\mu^-$  decay, provides a self-analyzing set of data for testing muon identification techniques. Muon pairs are selected in the data by first evaluating the DNN response for all muon candidates in an event. Pairs are then formed from oppositely charged muons. Signal pairs are selected based on the pair DNN response  $r_{\text{pair}}$ :

$$r_{\text{pair}} = \sqrt{r_a^2 + r_b^2} \quad (4.5)$$

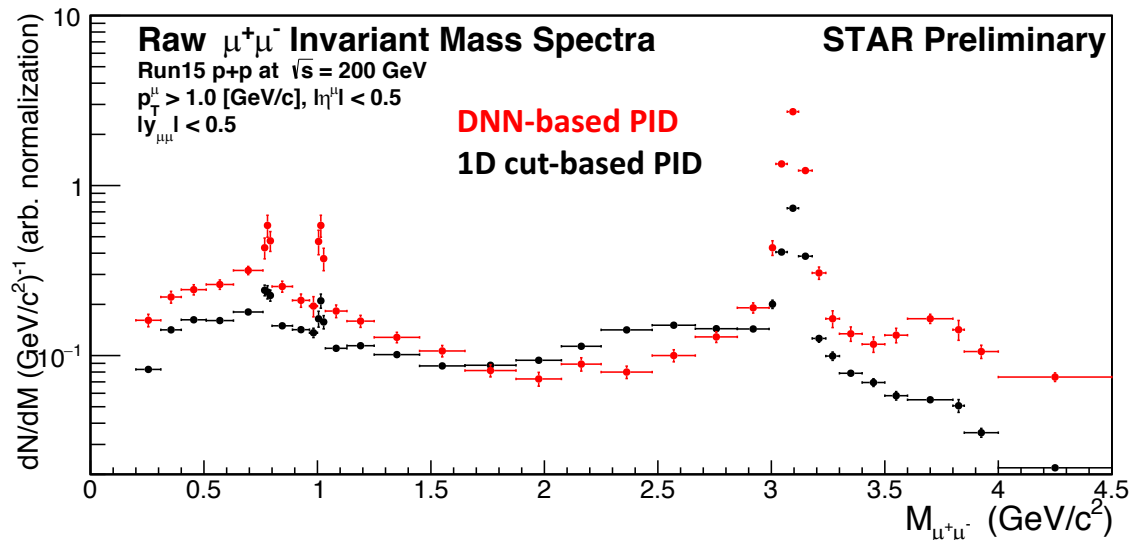
where  $r_a$  and  $r_b$  are the DNN responses for paired muons  $a$  and  $b$ , respectively. The DNN was specifically optimized to promote a response of  $r \approx 1$  for signal muons and a response of  $r \approx 0$



**Figure 4.11** : Comparison of the performance of the DNN and BDT (nTrees=800) for all simulated tracks ( $1.0 \leq p_T \leq 15.0$  GeV/c). The DNN out-performs the BDT at mid to high signal efficiencies.



**Figure 4.12** : Raw yield extraction of the  $\phi$  meson using optimized traditional 1D PID techniques (a) compared to the DNN based PID (b).



**Figure 4.13** : The raw  $\mu^+\mu^-$  invariant mass spectra from optimized traditional PID and from deep neural network based PID. The distributions are scaled in the intermediate mass region to make comparison easier.

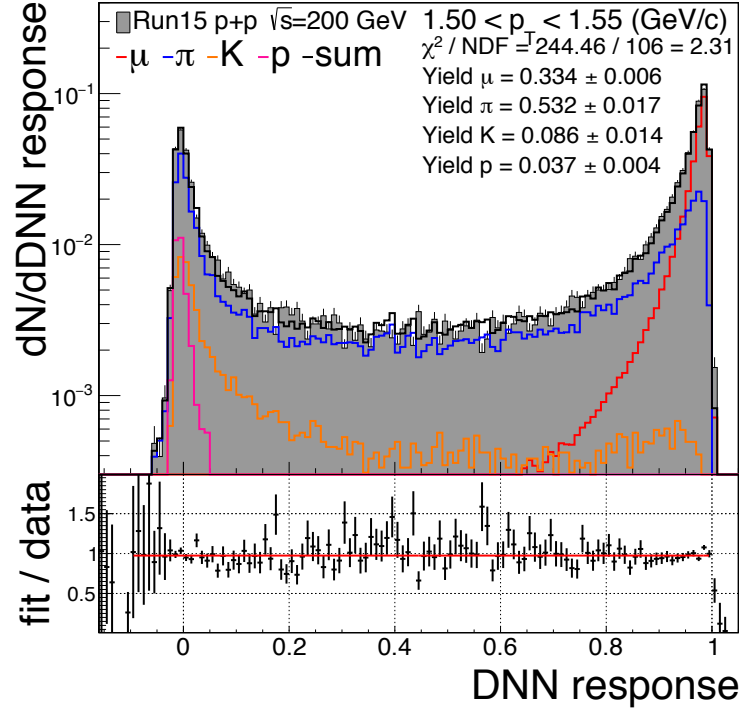
for background sources. For this reasons the pair response for a  $\mu^+\mu^-$  pair will be  $r_{\text{pair}} \approx \sqrt{2}$ . The optimal  $r_{\text{pair}}$  cut for selecting  $\phi \rightarrow \mu^+\mu^-$  decays was determined by maximizing the  $\phi$  significance ( $S/\sqrt{S+B}$ ) in steps of  $r_{\text{pair}} = 0.01$ . The signal and background contributions were extracted by fitting the raw  $\mu^+\mu^-$  invariant mass spectra in  $0.85 < M_{\mu\mu} < 1.5 \text{ GeV}/c^2$ . A 4th-order polynomial was used to model the background and a Gaussian was used for the  $\phi$  peak. The optimal cut was found to be  $r_{\text{pair}} > 1.36$  which provides a  $\phi$  significance of  $\sim 8.3$  and a  $S/B$  ratio of 0.33. Figure 4.12 shows the raw  $\phi$  yield extraction fits using traditional 1D cuts optimized on the  $J/\psi$  for muon identification and using the DNN-based muon identification. The DNN-based muon identification simultaneously provides higher S/B ratio, significance, and signal efficiency compared to the optimized 1D muon identification. In Fig. 4.13, the raw  $\mu^+\mu^-$  invariant mass spectra in the range  $0 < M_{\mu\mu} < 4.5 \text{ GeV}/c^2$  is shown for optimized 1D cut-based muon identification and compared with the DNN-based muon identification. In addition to improving S/B and significance of the  $\omega$  and  $\phi$  mesons, the DNN-based muon identification allows the  $\psi(2S)$  at  $M_{\mu\mu} \approx 3.7 \text{ GeV}/c^2$  to be visible.

#### 4.4.6 Data-Driven Muon Purity Measurements

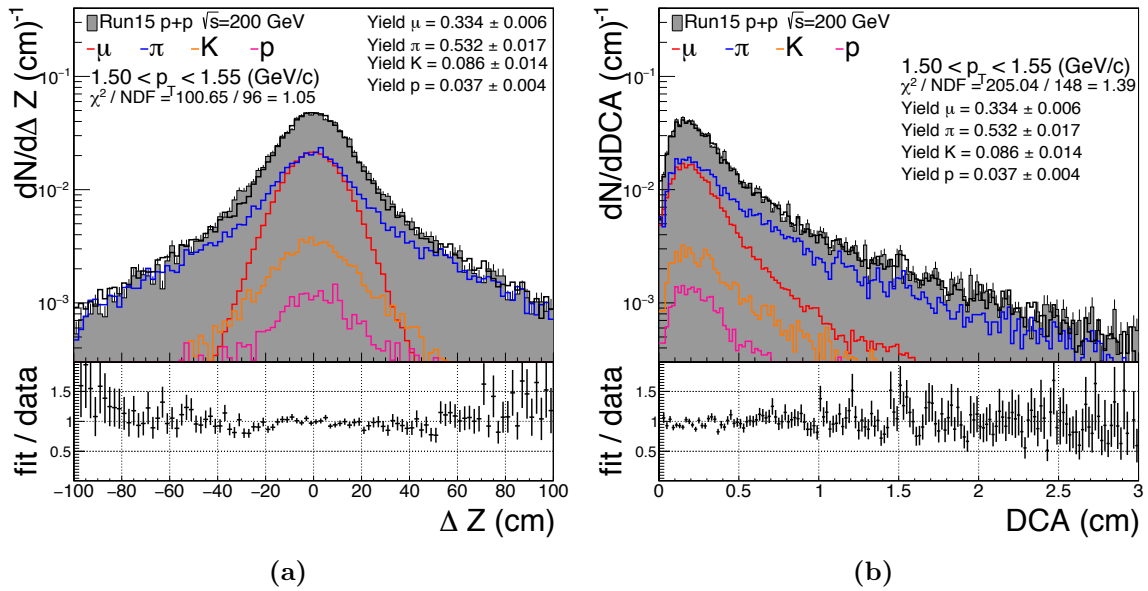
Since no individual feature among the set of PID features clearly separates signal from background contributions, it is not possible to fit any one of the features in order to extract the muon purity of tracks in data. Given the signal and background PDFs for each of the 8 PID features (neglecting  $p_T$  and  $q$ ), one could in principle conduct a simultaneous fit to all 8 distributions in order to extract the yield of signal and background contributions. Since each distribution would need to be fit to a  $\mu$ ,  $\pi$ , K, and  $p$  contribution it would require simultaneously fitting 8 distributions with 32 templates constrained by 4 free yield parameters. While possible, in practice a simultaneous fit with so many distributions and templates is technically challenging and often proves unstable.

Instead, the complexity of the problem can be greatly reduced by simply fitting the DNN response for muon candidates with the template shapes for signal and background components. In this setup the DNN combines all PID features, so only a single distribution needs to be fit with the 4 template shapes for signal and background each with a free yield parameter. Figure 4.14 shows the result of this procedure applied to muon candidate tracks in the range  $1.5 < p_T < 1.55$  GeV/c. The template for each component is computed by evaluating the DNN on simulated tracks in the same kinematic regions as those in the data. The data/fit ratio shown in the lower panel of Fig. 4.14 shows that the fit is capable of describing the DNN response for muon candidates to within  $\sim 20\%$  over the entire range of DNN responses.

After determining the yield of each signal and background contribution, the DNN response can be projected back onto all of the 8 PID features to verify that the DNN is properly combining the information from all variables. Ensuring that the projection onto each PID feature results in a good description of the data is a strong demonstration that the DNN is not over-training on artifacts in the training samples. Projections onto the  $\Delta Z$  and DCA features are shown in Fig. 4.15a and 4.15b. This technique allows the increased signal vs. background separation power provided by the DNN-based muon identification to be leveraged for data-driven muon purity measurements. At the same time, the ability to project the muon purity fit results back onto the PID features provides a data-driven strategy to test for over-training and poor model generalization.



**Figure 4.14 :** The top panel shows the DNN response for muon candidates in the range  $1.5 < p_T < 1.55$  GeV/c. A template fit is conducted to extract the contributions from  $\mu$  (red),  $\pi$  (blue), K (orange), and  $p$  (magenta). The lower panel shows the ratio of the data over the sum of the contributions.



**Figure 4.15 :** The result of the DNN response fit for  $\mu$ ,  $\pi$ , K, and  $p$  contributions projected back onto the  $\Delta Z$  (a) and DCA (b) distributions. The ratio of fit / data is shown in the lower panels of each figure.

## Chapter 5

### Dimuon Analysis Details

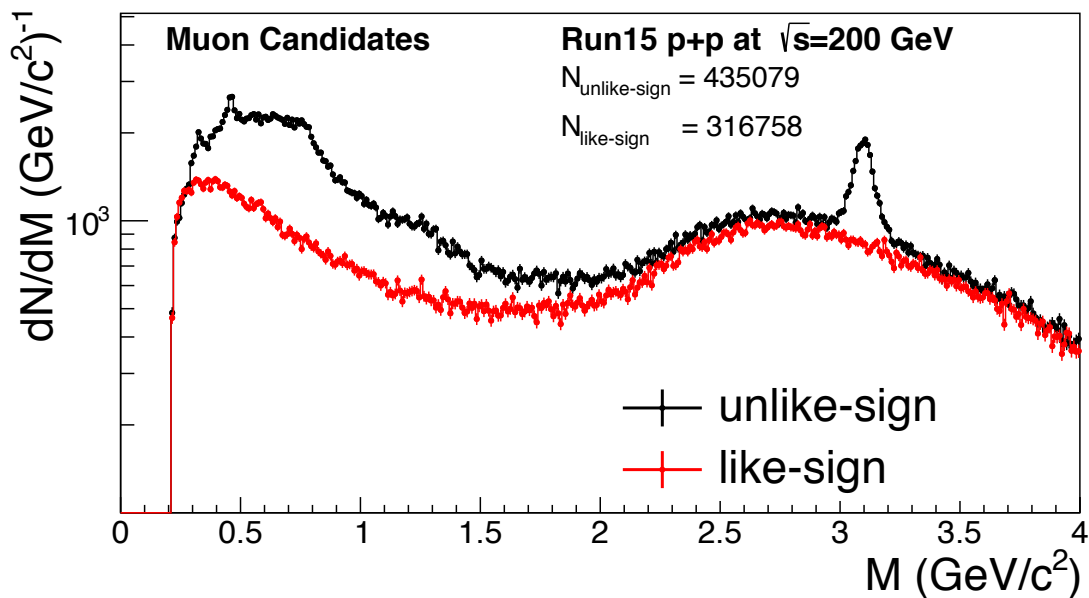
This chapter describes the steps involved in the analysis of the dimuon invariant mass spectra in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV and  $p + \text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV. Specifics of the inclusive  $\phi \rightarrow \mu^+ \mu^-$  yield measurement for  $|y| < 0.5$  are given as well. The organization of the chapter is as follows: 1) the cuts used for muon candidate selection and track quality assurance are discussed, 2) the muon purity in data is measured as a function of transverse momentum for all muon candidates. Upper bounds on the muon purity are investigated through the application of increasingly tighter muon identification cuts. 3) techniques for extending the single track DNN-based PID technique for pair identification are introduced and compared, 4) backgrounds from combinatorial and physical sources are discussed. A statistical analysis is presented to motivate the techniques that will be used for background estimation in later sections. A toy MonteCarlo event generator is used to further demonstrate the nuances that must be taken into account in the background estimation strategies. 5) The data-driven background measurements are presented and the signal dimuons are extracted. 6) The techniques used to simulate the hadronic cocktail resulting from light hadron decays, open heavy flavor, and Drell-Yan are presented. 7) The determination of the efficiency and acceptance corrections using a folding procedure is developed. 8) The estimation of systematic uncertainties are discussed.

#### 5.1 Track Selection and Muon Identification

Muon candidates are formed from all tracks matched to a hit in the MTD that pass basic track quality cuts:

- $mtdMatchFlag > 0$  : Require that the track be matched to an MTD hit.
- $mtdTriggerFlag > 0$  : Require that the MTD hit fire the trigger (dimuon trigger).
- $nHitsFit > 15$  : The number of hit points used to reconstruct the track must be



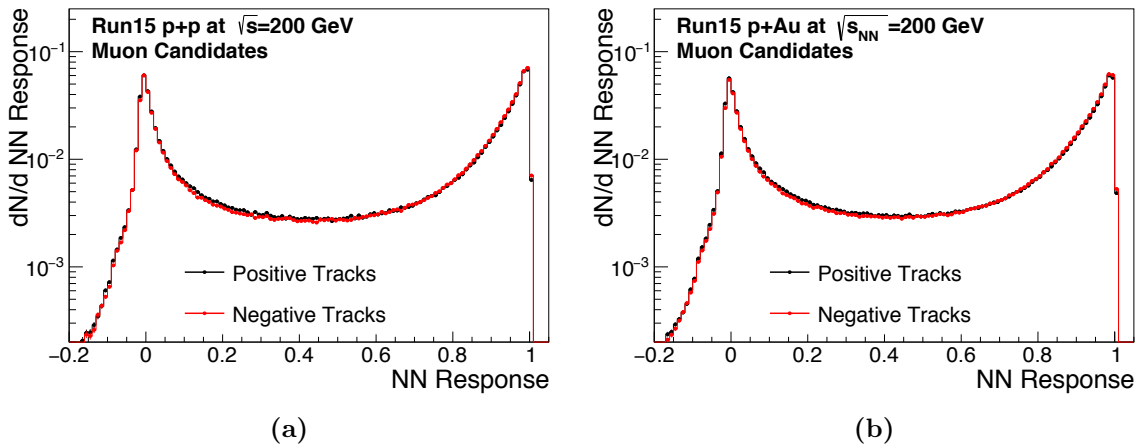


**Figure 5.1** : The invariant mass of muon candidate tracks shown for unlike-sign pairs and like-sign pairs.

greater than 15. This cut rejects low quality tracks and helps ensure that the tracks have consistent momentum resolution and DCA resolution.

- $nHitsDedx > 15$  : The number of hits used to determine the  $dE/dx$  information must be greater than 15. This cut improves the  $dE/dx$  resolution, aiding in particle identification.
- $nHitsFit/nHitsMax > 0.52$  : The ratio of number of hits used to reconstruct track compared to the possible number of hits must be greater than 0.52. This cut ensures that a single tracks is not split into two track segments during reconstruct.
- $DCA < 3$  cm : The distance to closest approach of the global track to the primary vertex must be less than 3 cm. In the STAR reconstruction software, tracks associated with the primary collision vertex must be within  $\sim 3$  cm.
- $p_T > 1.1$  GeV/c : The track's transverse momentum must be greater than 1.1 GeV/c. This cut is used to remove tracks with very low MTD matching and response efficiency in the range  $0.9 < p_T < 1.1$  GeV/c.

The invariant mass of all muon candidate tracks is shown in Fig. 5.1. The  $J/\psi$  peak is



**Figure 5.2 :** The neural network response for positive and negative tracks in  $p+p$  collisions at  $\sqrt{s}=200$  GeV (a) and  $p+Au$  collisions at  $\sqrt{s_{NN}}=200$  GeV (b).

**Table 5.1 :** Weak decay processes through which single muons may be created.

Process	Branching Ratio
$\pi^+ \rightarrow \mu^+ + \nu_\mu$	99.9 %
$K^+ \rightarrow \mu^+ + \nu_\mu$	63 %
$K^+ \rightarrow \pi^0 + \pi^+ \rightarrow \pi^0 + \mu^+ + \nu_\mu$	3.8 %

clearly visible but the low mass region is highly obscured by background sources. In Chapter 4, techniques for training deep neural networks (DNN) for muon identification were discussed. The trained DNN can be applied to the tracks from data and the output can be used to select muons and reject background.

## 5.2 Hadron Contamination and Muon Purity

Since the MTD has only a single layer of steel providing  $\sim 5$  interaction lengths at its thickest, hadron contamination from punch-through pions, kaons, and protons is a major source of background in this analysis. Hadron punch-through is not the only significant source of contamination though. For the purpose of measuring  $\mu^+\mu^-$  pairs, a significant background arises from secondary muons that are produced as singles - i.e. not in pairs. The dominant source of secondary muons come from the decays:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad (5.1)$$

**Table 5.2 :** GEANT simulation of  $\pi^\pm$  and  $K^\pm$  in the STAR detector. These values represent the probability of each event occurring before exiting the STAR detector (specifically within the MTD, since it is the detector in STAR at largest radius).

	No Decay	Muonic Decay	Hadronic Decay	Elastic	Inelastic	Absorption
$\pi^+$	86.6%	3.7%	—	0.03%	9.7%	—
$\pi^-$	86.8%	4.2%	—	0.05%	8.9%	0.01%
$K^+$	79.6%	9.5%	4.7%	0.01%	6.2%	—
$K^-$	80.6	6.8%	3.5%	0.01%	8.3%	0.87%

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \quad (5.2)$$

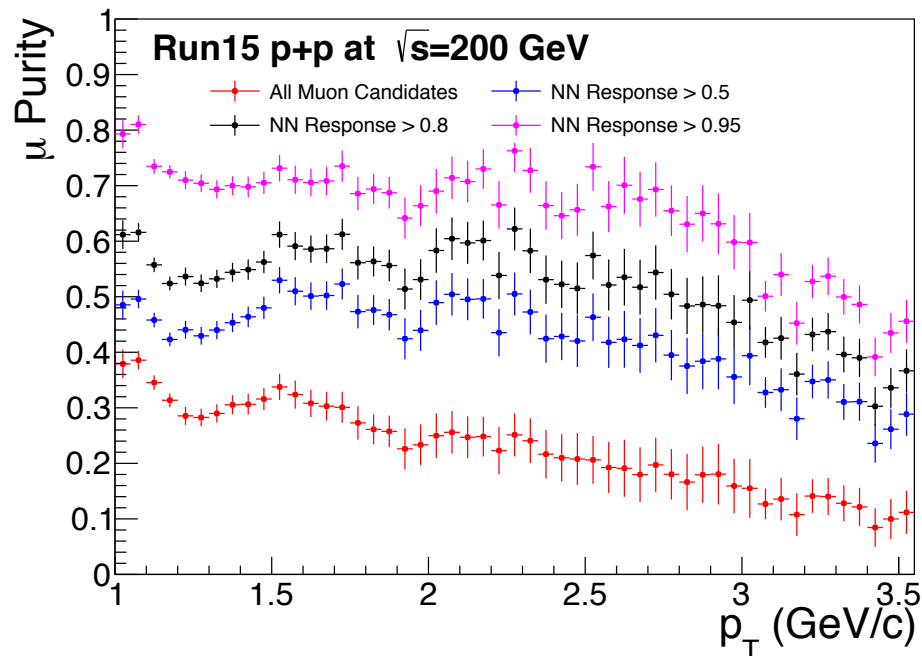
$$K^+ \rightarrow \mu^+ + \nu_\mu \quad (5.3)$$

$$K^- \rightarrow \mu^- + \bar{\nu}_\mu \quad (5.4)$$

Since the charged pion and kaons decay into muons via the weak interaction they are long lived with  $\tau(\pi^\pm) \approx 26$  ns and  $\tau(K^\pm) \approx 12$  ns. The decay length is given by:

$$L = \gamma\tau c \quad (5.5)$$

Therefore, at 1 GeV ( $\gamma \approx 10$ ) the charged pion decay length is  $\approx 76$  m and the charged kaon decay length is  $\approx 35$  m. Since these decay lengths are much longer than the radius of the TPC, most pions and kaons will escape the TPC before they decay. However, since particle decay is a stochastic process, a significant number may still decay either before they exit the TPC or after exiting the TPC but before entering the MTD. Table 5.2 shows the results of a GEANT simulation in which  $\pi^\pm$  and  $K^\pm$  were embedded into real events. Flat  $1 < p_T < 10$  GeV/ $c$ ,  $-1 < \eta < 1$ , and  $0 < \phi < 2\pi$  kinematic distributions were used for the simulated  $\pi^\pm$  and  $K^\pm$  tracks. Each track was simulated from the primary collision vertex until it exited the volume enclosed by the MTD. Interactions were identified using GEANT's codes for decay, elastic collisions, inelastic collisions, and absorption. The decays were further identified using the decay products where applicable. Even though the charged kaons have a lower muonic decay branching ratio than charged pions, the shorter lifetime contributes to a higher overall percentage of muonic decays within the MTD volume. However, since the  $dN/dy$  of pions is  $\approx 50$  times larger than that of kaons in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV, the absolute yield of secondary muons from charged pion decays is larger.



**Figure 5.3 :** The muon purity as a function of track  $p_T$  for all muon candidate tracks and for various DNN-base pid cuts.

In Chapter 4 a technique was presented for measuring the muon purity in data by leveraging the separation power of the DNN-base muon identification. This technique can be applied systematically to determine the muon purity as a function of track  $p_T$ .

For this procedure, the DNN response template for signal and each background source is determined using the simulated data samples. Since the shape of the DNN response varies as a function of  $p_T$ ,  $\eta$  and  $\phi$ , the templates are extracted from the simulated events using the same kinematic cuts as the given data sample. The purity is determined in slices of  $p_T$  for all  $\eta$  and  $\phi$ . In each  $p_T$  slice, the fits are conducted as a maximum likelihood fit with the TMINUIT minimization library. The only free parameters are the yields of each template. For  $p_T < 2.0$  GeV/c the  $\pi$ , K, and  $p$  sources are used as background sources. The proton contribution template is no longer used above  $p_T > 2.0$  GeV/c and the K contribution is no longer used above  $p_T > 3.0$  GeV/c. Dropping the proton and kaon contributions at higher  $p_T$  values where they are no longer significant sources helps improve the quality and stability of the fit.

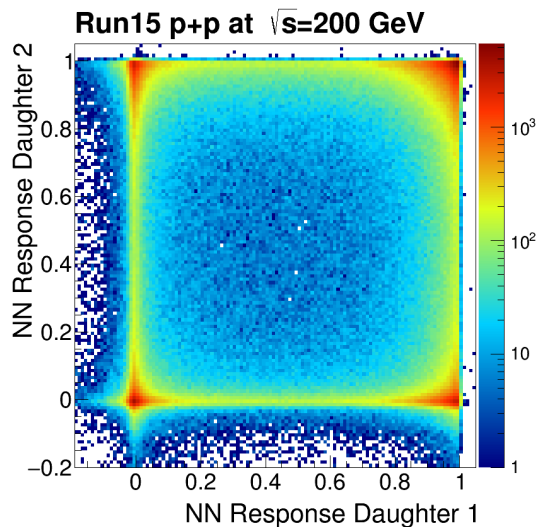
While the hadron punch through can be reduced using the DNN-based pid, there remains a significant amount of impurity. We can see from Fig. 5.3 that the average muon purity, even

with a tight DNN cut, is  $\sim 70\%$ . For comparison, the STAR dielectron analysis in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV achieved approximately 95% electron purity average for 0–80% central collisions. The relatively high electron purity allowed the dielectron analyses to consider hadron contamination only as a source of systematic uncertainty. In this analysis, the low muon purity necessitates data driven techniques to estimate the contamination from the significant hadron contamination. These techniques will be discussed in the remainder of this chapter. To make matters worse, the purity quoted in in Fig. 5.3 includes a significant portion of secondary muons resulting from the weak decay of pions and kaons (via process listed in Table 5.1). These decays can be very efficiently removed with a dedicated high-precision vertex tracker for reconstructing secondary decay vertices. In the previous discussion of NA60’s dimuon results, it was noted that the novel setup including a silicon-based vertex tracker near the target was essential for that measurement.

The lack of tracking information from a high precision tracker near the collision vertex in these datasets makes it impossible to separate muons produced in the initial collision from those that result from a decay at or near the collision vertex. Furthermore, the very small mass of the neutrino results in a decay muon that is very collimated with the parent particles trajectory. For this reason, even decays that happen a significant distance from the collision vertex, or even outside the TPC, may still produce a hit in the MTD close to the projected track position. The presence of hadron-punch through contamination and secondary muons introduces a significant amount of background into the dimuon invariant mass. These background sources and the techniques for removing them will be discussed further in Sec. 5.5.1

### 5.3 Identification of Muon Pairs

In this section the identification of muon pairs is discussed. Even if the DNN-based PID provides an ideal discriminator for selecting single muons, there are still some variations in the way that muon pairs can be selected. Figure 5.4 shows the DNN response of daughter 1 ( $r_1^{NN}$ ) versus the DNN response of daughter 2 ( $r_2^{NN}$ ). One way to select pairs would be to simply require that  $r_1^{NN} > x$  and  $r_2^{NN} > x$  where  $x$  is a given cut value. This simplified approach may not be optimal though. Since the single muon selection still leaves a large amount of contamination from hadrons and secondary muons, choosing an optimal technique

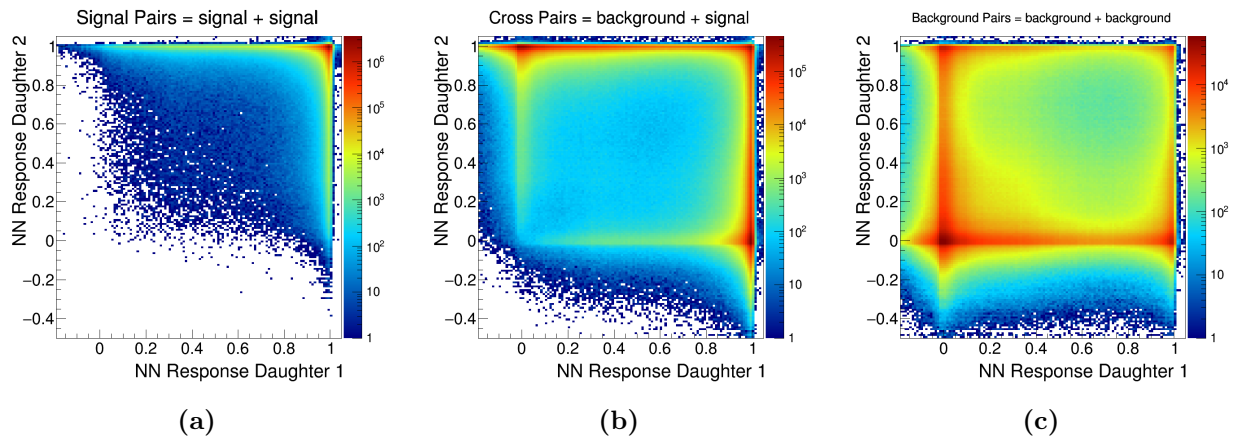


**Figure 5.4 :** The DNN response of daughter 1 ( $r_1^{NN}$ ) versus the DNN response of daughter 2 ( $r_2^{NN}$ ) for  $p + p$  collisions at  $\sqrt{s} = 200$  GeV.

for pair selection is crucial. Additionally, since the response of the DNN has some variation as a function of the  $p_T$  of each track cutting on a single  $r^{NN}$  value for all  $p_T$  may not be optimal.

The templates used for muon purity fits can also be used to aid in the determination of the optimal pair selection criteria. If we consider that the tracks can fall into one of two categories - signal or background (combining  $\pi$ , K, and  $p$  together), then we have a total of three possible pair types:

- **Background Pair:** The result of pairing a background track with another background track. These pairs are the easiest to reject since they are generally well discriminated by the DNN.
- **Cross Pair:** The result of pairing one background track with one signal track. These pairs can be very similar to real signal pairs, since the DNN template for background has a large tail near  $r^{NN} \approx 1$ . Some amount of these pairs are indistinguishable from real signal pairs. Techniques for removing them will be discussed further.
- **Signal Pairs:** The result of pairing two signal tracks together. These pairs are most likely the result of a real dimuon pair, though combinatorial backgrounds and secondary muons may still be included.

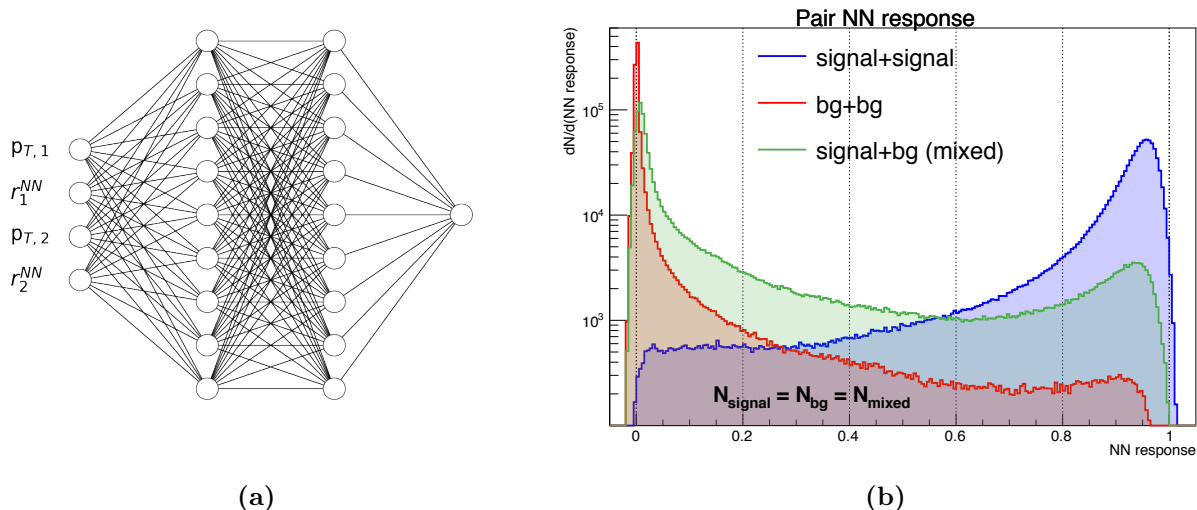


**Figure 5.5 :** The three categories of pairs: Signal pairs resulting from signal + signal (a), cross pair resulting from background + signal (b), and background pair resulting from background + background (c).

These three types of pairs can be seen in Figures 5.5a, 5.5b, and 5.5c for signal pairs, cross pairs, and background pairs, respectively. A deep neural network, similar to those used in Chapter 4 for single track identification, can be trained to output the optimal pair discrimination value while at the same time normalizing differences in single track DNN response as a function of  $p_T$  if needed. The purpose of this procedure is not necessarily to use a neural network for evaluating the pair PID score, but to test various naive techniques for pair selection against the pair neural network. If the naive selection techniques perform just as well as the pair neural network, the naive selection can be safely used without fear that it is sub-optimal.

Figure 5.6a shows the architecture chosen for this task. The input features for the pair DNN are the single track DNN response  $r^{NN}$  and  $p_T$  for each daughter track. While a formal grid-search minimization was not performed, several different neural network architectures of decreasing complexity were tested (by using 2 hidden layers and decreasing the number of neurons in each layer). The templates shapes for signal pairs, cross pairs, and background pairs shown in Fig. 5.5 were used in the training phase. The output of the trained pair DNN is shown in Fig. 5.6b for each of the three categories.

In order to determine what the pair DNN learned, various linear and nonlinear combinations of the single track DNN response can be compared with the pair DNN output. If the single track DNN had no variation in  $p_T$  then the sum in quadrature would be nearly optimal



**Figure 5.6 :** The neural network architecture used for the selection of signal pairs (a). The inputs are the  $p_T$  and neural network response  $r^{NN}$  of each daughter track in the pair. The output of the pair DNN for the three different categories of pairs (b). Note: In this figure mixed pairs=cross pairs.

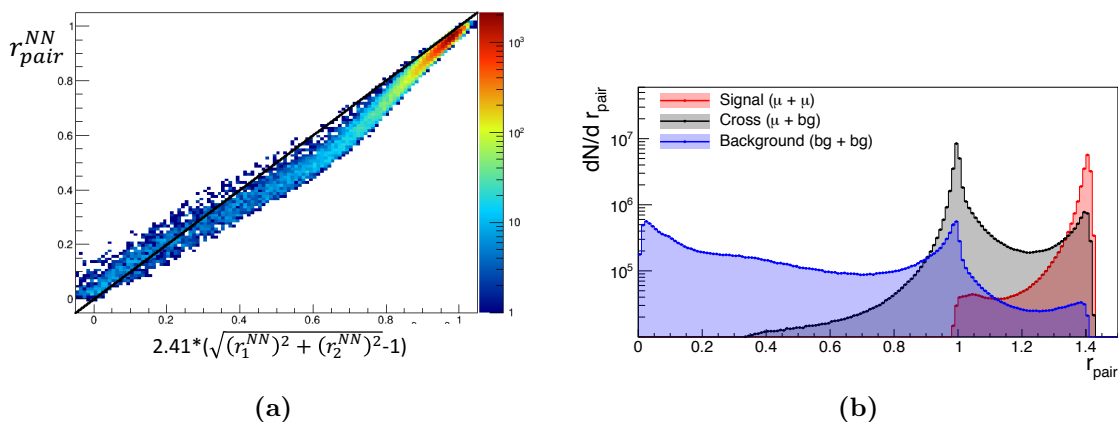
for separating signal and background pairs. The pair DNN response can be compared to the sum in quadrature to determine if the pair DNN has learned any additional relationship between the single track DNN responses. This comparison is shown in Fig. 5.7a with the sum in quadrature (x-axis) scaled to be in the range of 0 to 1. The comparison shows that the quadrature sum performs nearly as well as the pair DNN. Figure 5.7b shows the three pair categories using the simpler sum in quadrature instead of the pair DNN response.

## 5.4 Sources of Background

The foreground dimuon spectra made by pairing all muon candidates is shown in Fig. 5.1. This distribution includes both the signal pairs (S) and various background contributions. Isolating the yield of signal dimuons pairs requires removing the pairs from background processes. Background pairs can be organized into two main categories:

- **Uncorrelated combinatorial background pairs (CB):** this background arises from the random pairing of muons from different parents. The existence of this uncorrelated combinatorial background is by construction in the sense that it results from the intentional pairing of all positive muon candidates with all negative muon candidates in





**Figure 5.7 :** The correlation between  $r_{pair}^{NN}$  and  $2.41 \times (\sqrt{(r_1^{NN})^2 + (r_2^{NN})^2} - 1)$  (a). The three types of pairs shown using the  $r_{pair}$  projection (b).

each event. In high multiplicity environments, like Au+Au collisions, the combinatorial background can easily account for upwards of 90% of the total background pairs. At lower multiplicities, such as  $p + p$  and  $p$ +Au collisions at  $\sqrt{s_{NN}} = 200$  GeV there are few events with more than one positive and one negative muon candidate, reducing the relative contribution of uncorrelated combinatorial background to the total background.

- Correlated background pairs (CP):
  - **Jet-like pairs (JP)** : correlated background pairs may result from jet or jet-like pairs in various ways. Jets may produce back-to-back hadrons that punch-through to the MTD. Another possibility is to produce a spray of particle which travels through one of the regularly spaced gap in the steel producing multiple hits in the MTD detector. These hits are most likely in the edge-cells of nearby backlegs. If tracks from the initial jet or other event activity are matched to the MTD hits then a correlated pair is produced. Further details are discussed below.
  - **Hadron decays (HD)**: The large amount of contamination from punch-through hadrons results in a significant susceptibility to correlated pairs from hadron decays. For instance, the  $\rho \rightarrow \pi^+ \pi^-$  decay (with a branching ratio of  $\sim 100\%$ ). More examples of these background sources will be discussed in greater detail below.
  - **Secondary hadron decays (SHD)**: In addition to the correlated pairs resulting from hadron decay (mentioned above), the large branching ratio for  $\pi^\pm$  and  $K^\pm$

decay to muons causes correlated pairs from secondary decays of the hadron decay products. This can be illustrated using the previous example of the  $\rho \rightarrow \pi^+\pi^-$  decay. Since the branching ratio for  $\pi^\pm$  decay to a muon is  $\sim 99\%$ , a large number of correlated decays will be formed in which one or both of the daughter pions further decays. The resulting pair, either a  $\mu^\pm\pi^\mp$  pair or a  $\mu^+\mu^-$  pair, will be less strongly correlated than the HD pairs, but still present in the foreground dimuon spectra. On the one hand, these types of pairs will be enhanced with respect to HD due to the significantly larger efficiency for muons to match to the MTD than for hadrons. On the other hand, the lifetime of the pion and kaon significantly reduce the probability that the secondary decays will occur within the volume of the MTD.

Based on these sources, the foreground dimuon spectra ( $FG_{+-}$ ) can be expressed as (notation partially borrowed from Ref. [106]):

$$FG_{+-} = S + CB_{+-} + JP_{+-} + HD_{+-} + SHD_{+-} \quad (5.6)$$

while the like-sign spectra ( $\mu^+\mu^+$  and  $\mu^-\mu^-$  separately) can be written as:

$$FG_{++} = CB_{++} + JP_{++} \quad (5.7)$$

$$FG_{--} = CB_{--} + JP_{--} \quad (5.8)$$

$$FG_{\pm\pm} = FG_{++} + FG_{--} \quad (5.9)$$

#### 5.4.1 Uncorrelated Combinatorial Background

From the formulation it is clear that the foreground like-sign spectra will not contain contributions from hadron decays and secondary hadron decays. In past STAR dielectron analyses, the foreground like-sign spectra are subtracted from the foreground unlike-sign spectra to isolate the signal pairs. This technique is suitable when the detector acceptance is identical for like-sign and unlike-sign pairs (e.g. for  $2\pi$  acceptance in azimuth) and when the combinatorial background is charge-symmetric. In these cases the shape of the foreground like-sign spectra can be expected to closely reproduce the shape of the combinatorial background in

the foreground unlike-sign spectra. This approach can be summarized as follows. When averaging over a large number of events, the mean number of unlike-sign pairs from combinatorial pairing in a given invariant mass bin can be expressed as:

$$\langle \text{CB}_{+-} \rangle = \langle n_+ \rangle \langle n_- \rangle \quad (5.10)$$

and the number of like-sign pairs in the same invariant mass bin is:

$$\begin{aligned} 2\langle \text{CB}_{++} \rangle &= \langle n_+(n_+ - 1) \rangle \\ &= \langle n_+^2 \rangle - \langle n_+ \rangle \end{aligned} \quad (5.11)$$

$$\begin{aligned} 2\langle \text{CB}_{--} \rangle &= \langle n_-(n_- - 1) \rangle \\ &= \langle n_-^2 \rangle - \langle n_- \rangle \end{aligned} \quad (5.12)$$

We can notice that in the special case that positive and negative particles are produced in pairs and are distributed according to a Poisson distribution, then the following equality holds:

$$\langle n^2 \rangle = \langle n \rangle^2 + \langle n \rangle \quad (5.13)$$

Therefore, the above simplify to:

$$2\langle \text{CB}_{++} \rangle = \langle n_+ \rangle \langle n_+ \rangle \quad (5.14)$$

and

$$2\langle \text{CB}_{--} \rangle = \langle n_- \rangle \langle n_- \rangle \quad (5.15)$$

Finally it can be seen that the combinatorial like-sign pairs can be used to estimate the combinatorial unlike-sign pairs as:

$$\langle \text{CB}_{+-} \rangle = 2\sqrt{\langle \text{CB}_{++} \rangle \langle \text{CB}_{--} \rangle} \quad (5.16)$$

This result shows that in ideal cases, the geometric mean of the foreground like-sign pairs can be used to estimate the combinatorial background in the foreground unlike-sign pair. However, in practice this procedure can not be applied directly due to detector acceptance

affects which cause asymmetries in the invariant mass spectra of like-sign versus unlike-sign combinatorial pairs. The important take away from this exercise is that two assumptions are needed to ensure the applicability of this approach. They are:

- The positive and negative particles are produced in pairs
- The underlying production distribution is Poisson

The first assumption is clearly broken when considering  $\mu^+\mu^-$  pairs, due to the large contribution from weak decays of pions and kaons. This implies that the foreground like-sign spectra - even corrected for acceptance affects - is not capable of reproducing the shape and magnitude of the combinatorial background present in the foreground unlike-sign distribution.

We now consider in full rigor the derivation of the proper background scaling for the combinatorial pairs with the previous assumptions neglected. Namely, we allow some positive and some negative particles to be produced independently of one another. We further generalize by assuming a binomial distribution (instead of Poisson) for the number of reconstructed pairs in order to incorporate finite efficiency effects. The derivation in this section is based heavily on the appendix A in Ref. [85]. To begin, we assume that the true number of pairs produced in a given event is found by sampling a Poisson distribution:

$$N_{\text{pairs}} = P(\langle N_{\text{pairs}} \rangle) \quad (5.17)$$

where  $P$  is the Poisson distribution whose PDF is:

$$P(k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (5.18)$$

with  $k$  being the number of events sampled and  $\lambda$  being the mean of the distribution. In order to determine the number of reconstructed pairs we must take into account the efficiency of reconstructing the individual tracks in the pair. Therefore the number of reconstructed pairs is given by sampling a binomial distribution with the pair efficiency:

$$p(n_{\text{pairs}}^{\text{reco}}) = B(n_{\text{pairs}}^{\text{reco}}; N_{\text{pairs}}, \varepsilon_{\text{pair}}) \quad (5.19)$$

where  $\varepsilon_{\text{pair}} = \varepsilon_+ \cdot \varepsilon_-$  with  $\varepsilon_+$  and  $\varepsilon_-$  being the single track efficiencies for positive and negative

tracks, respectively.  $B$  is the binomial distribution whose PDF is given by:

$$B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k} \quad (5.20)$$

with  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$  and  $k$  is the number of successes given  $n$  total trials each with a probability of success given by  $p$ .

For convenience we let:

$$n_{\text{pairs}}^{\text{lost}} = N_{\text{pairs}} - n_{\text{pairs}}^{\text{reco}} \quad (5.21)$$

Where  $n_{\text{pairs}}^{\text{lost}}$  represents the number of pairs that are not fully reconstructed, either because one track was lost or both. Given the number of pairs reconstructed in a given event the probability for the remaining observable outcomes is given by sampling a Multinomial Distribution:

$$p(n_+, n_-) = M(n_+, n_-; n_{\text{pairs}}^{\text{lost}}, \varepsilon_+ \cdot (1 - \varepsilon_-), \varepsilon_- \cdot (1 - \varepsilon_+)) \quad (5.22)$$

$$p(n_-) = \sum_{n_+=1}^{n_{\text{pairs}}^{\text{lost}}} M(n_+, n_-; n_{\text{pairs}}^{\text{lost}}, \varepsilon_+ \cdot (1 - \varepsilon_-), \varepsilon_- \cdot (1 - \varepsilon_+)) \quad (5.23)$$

$$p(n_+) = \sum_{n_-=1}^{n_{\text{pairs}}^{\text{lost}}} M(n_+, n_-; n_{\text{pairs}}^{\text{lost}}, \varepsilon_+ \cdot (1 - \varepsilon_-), \varepsilon_- \cdot (1 - \varepsilon_+)) \quad (5.24)$$

As noted before,  $n_{\text{pairs}}^{\text{lost}}$  are the number of pairs that are not fully reconstructed and may have a single track reconstructed, either positive or negative, or be lost entirely. The Multinomial distribution's Probability Mass Function (PMF) is given by:

$$M(x_1, \dots, x_k; p_1, \dots, p_k) = \frac{\Gamma\left(\sum_i x_i + 1\right)}{\prod_i \Gamma(x_i + 1)} \prod_{i=1}^k p_i^{x_i} \quad (5.25)$$

The Multinomial distribution can be thought of as modeling the outcomes of rolling a  $k$  sided dice  $n$  times where each side of the dice has a probability of landing,  $p_1, \dots, p_k$  with all  $p_i > 0$ . Each of the  $k$ -sides of the dice are mutually exclusive outcomes. Therefore, applied to this case, sampling a multinomial distribution ensures that:

$$p(n_+) + p(n_-) + p(n_{\text{lost}}) = 1 \quad (5.26)$$

That is, that the probability of the three possible outcomes for pairs that are not fully reconstructed sums to unity. The total number of unlike sign pairs is then computed as:

$$\begin{aligned} n_{+-} &= n_{\text{pairs}}^{\text{reco}} \cdot n_{\text{pairs}}^{\text{reco}} + n_{\text{pairs}}^{\text{reco}} \cdot n_{+} + n_{\text{pairs}}^{\text{reco}} \cdot n_{-} \\ &= S + BG_{+-} \end{aligned} \quad (5.27)$$

the total number of positive like-sign pairs is given by:

$$\begin{aligned} n_{++} &= (n_{\text{pairs}}^{\text{reco}} + n_{+}) \cdot (n_{\text{pairs}}^{\text{reco}} + n_{+} - 1)/2 \\ &= BG_{++} \end{aligned} \quad (5.28)$$

and the total number of negative like-sign pairs is given as:

$$\begin{aligned} n_{--} &= (n_{\text{pairs}}^{\text{reco}} + n_{-}) \cdot (n_{\text{pairs}}^{\text{reco}} + n_{-} - 1)/2 \\ &= BG_{--} \end{aligned} \quad (5.29)$$

At this point we can note that when averaging over a large number of events we can recover the earlier results, i.e. that the geometric mean of the foreground like-sign pairs is a good approximation for the combinatorial background in the foreground unlike-sign pairs. That is in this notation:

$$\langle BG_{+-} \rangle = 2\sqrt{\langle BG_{++} \rangle \cdot \langle BG_{--} \rangle} \quad (5.30)$$

We can make the simple addition of adding the possibility of single particle production independent of the pair production and reconstruction. That is, a certain number of single positive  $n_{+}^{\text{single}}$  and single negative  $n_{-}^{\text{single}}$  are produced according to:

$$p(n_{+}^{\text{single}}) = B(n_{+}^{\text{single}}; N_{+}^{\text{single}}, \varepsilon_{+}) \quad (5.31)$$

and

$$p(n_{-}^{\text{single}}) = B(n_{-}^{\text{single}}; N_{-}^{\text{single}}, \varepsilon_{-}) \quad (5.32)$$

with the true number positive and negative tracks ( $N_{\pm}^{\text{single}}$ ) produced according to a Poisson distribution:

$$N_{\pm}^{\text{single}} = P(\langle N_{\pm}^{\text{pairs}} \rangle) \quad (5.33)$$

With the addition of independently produced positive and negative particles the foreground unlike-sign pairs becomes:

$$n_{+-} = n_{\text{pairs}}^{\text{reco}} \cdot n_{\text{pairs}}^{\text{reco}} + n_{\text{pairs}}^{\text{reco}} \cdot n_+ + n_{\text{pairs}}^{\text{reco}} \cdot n_- + n_{\text{pairs}} \cdot n_+^{\text{single}} + n_- \cdot n_+^{\text{single}} + n_{\text{pairs}} \cdot n_-^{\text{single}} + n_+ \cdot n_-^{\text{single}} \quad (5.34)$$

and the foreground like-sign pairs become:

$$n_{++} = (n_{\text{pairs}}^{\text{reco}} + n_+ + n_+^{\text{single}}) \cdot (n_{\text{pairs}}^{\text{reco}} + n_+ + n_+^{\text{single}} - 1)/2 \quad (5.35)$$

$$n_{--} = (n_{\text{pairs}}^{\text{reco}} + n_- + n_-^{\text{single}}) \cdot (n_{\text{pairs}}^{\text{reco}} + n_- + n_-^{\text{single}} - 1)/2 \quad (5.36)$$

This leads to the conclusion that when single particle production is present:

$$\langle BG_{+-} \rangle > 2\sqrt{\langle BG_{++} \rangle \cdot \langle BG_{--} \rangle} \quad (5.37)$$

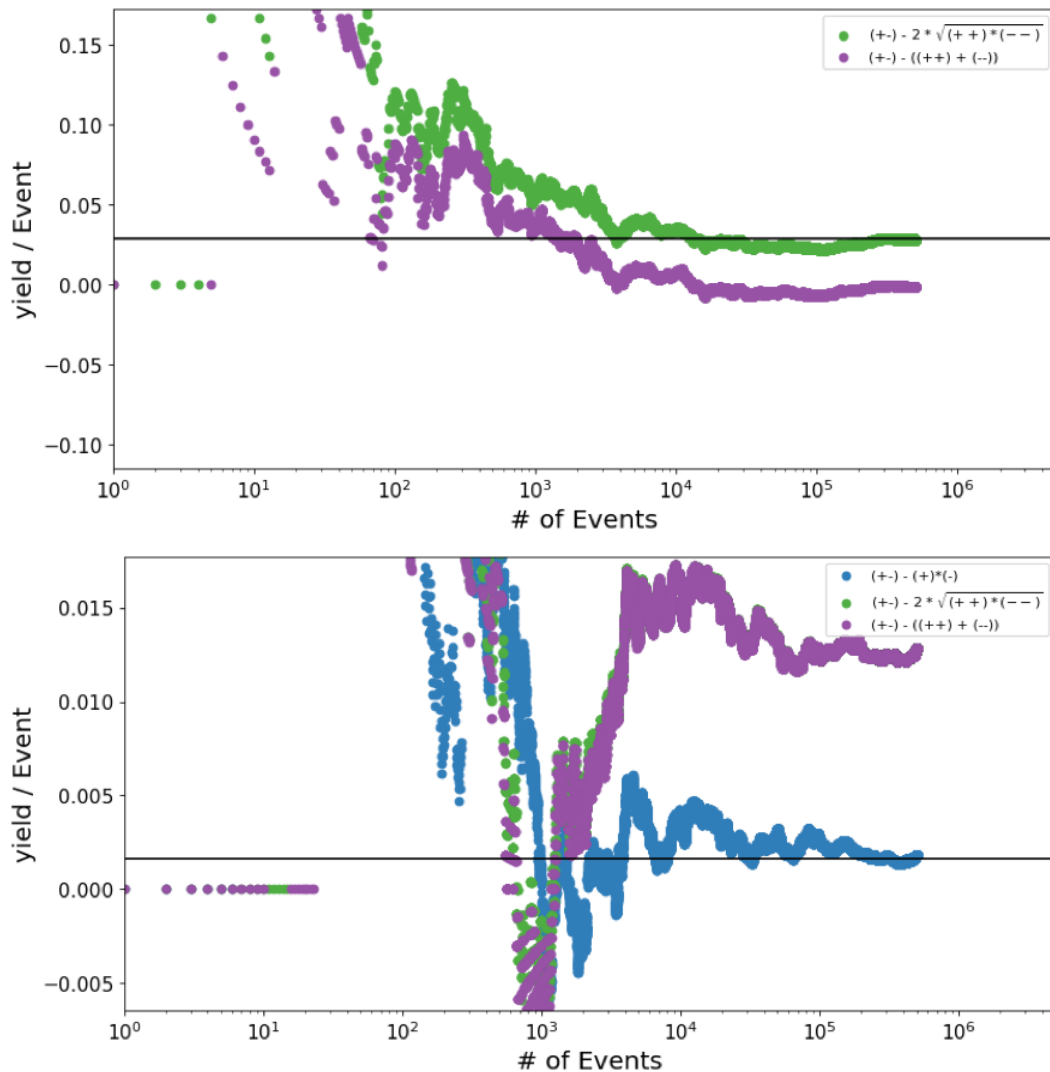
and instead we may only estimate that, in the limit where  $N_{\text{pairs}} \ll N_+^{\text{single}} \cdot N_-^{\text{single}}$  the combinatorial background, averaged over many events can be estimated by the product of the mean multiplicities:

$$\langle BG_{+-} \rangle \approx \langle n_{\text{pairs}}^{\text{reco}} + n_+ + n_+^{\text{single}} \rangle \cdot \langle n_{\text{pairs}}^{\text{reco}} + n_- + n_-^{\text{single}} \rangle \quad (5.38)$$

The results of this mathematical derivation were implemented in a toy Monte Carlo simulation in order to visualize the various background estimation techniques. Figure 5.8 shows the yield per event versus the cumulative number of events for two scenarios. In the top panel, the  $e^+e^-$  case is shown, in which all particles are produced in pairs. A charge asymmetry is produced by letting  $\varepsilon_+ \neq \varepsilon_-$  to demonstrate that the direct mean (lower violet curve) does not estimate the background well and therefore under predicts the signal yield. However, the geometric mean does converge to the signal yield (black line) even with a charge asymmetry. The lower panel of Fig. 5.8 shows the case in which single particle production is included. In this case, both the direct and geometric mean under predict the background and therefore over predict the signal. When using the background estimated by the mean multiplicity (blue) the measured yield converges to the signal yield. The under estimation of

the background by the geometric mean is a general result - it can be expected that in data the mean like-sign distribution will underestimate the background, not over estimate it.





**Figure 5.8 :** Plots of the yield / event from a toy Monte Carlo simulations with various types of combinatorial backgrounds present. The solid black line shows the true amount of signal / event. Top: Combinatorial background resulting from an  $e^+e^-$ -like case, i.e. approximately Poisson pair production. In this case the geometric mean of the like-sign reproduced the background correctly. A charge asymmetry is introduced to demonstrate that in this case the direct mean does not reproduce the background correctly. Bottom: Combinatorial background in the case with independently produced singles. The direct mean and geometric mean both underestimate the background, leading to an overestimation of the signal. The combinatorial background estimated via the mean multiplicities gives a more accurate estimate of the unlike-sign background.

**Table 5.3** : Processes through which correlated dimuon backgrounds can be produced.

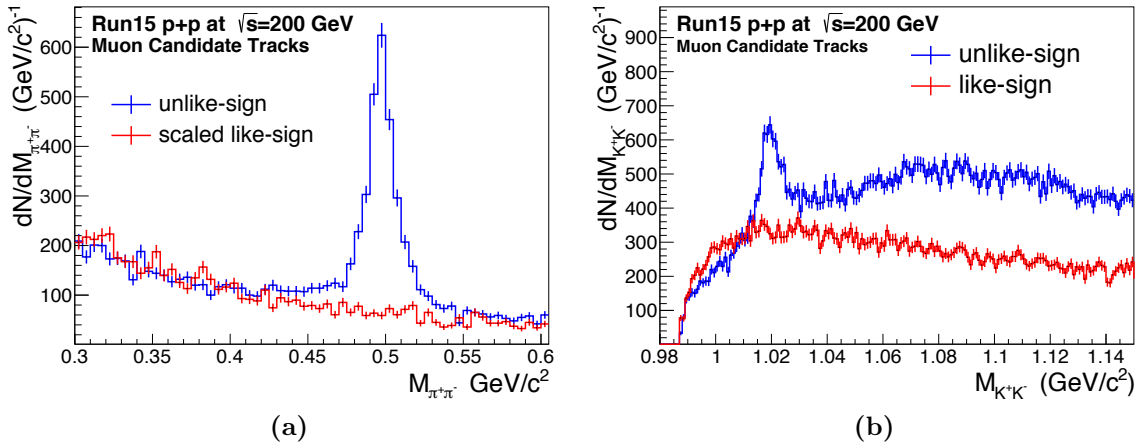
Hadron Decay Processes	Branching Ratio
$\rho^0 \rightarrow \pi^+\pi^-$	$\sim 100\%$
$K_S^0 \rightarrow \pi^+\pi^-$	69%
$\Lambda \rightarrow p + \pi^-$	63%
$\phi \rightarrow K^+K^-$	49%

### 5.4.2 Physical Background Sources

In addition to the uncorrelated combinatorial backgrounds, there exist several sources of physical backgrounds as mentioned earlier. The background pairs originating from physical sources will result in some degree of correlation between the final state tracks. Each category of background will be addressed individually below.

**Jet-like pairs** result from jets or high- $p_T$  hadrons that produce hits in the MTD. Back-to-back jets may lead to correlated pairs of punch-through hadrons (or decay muons) with a large opening angle. Due to the regular gaps in the spacing of the magnet steel, the spray of particles from a single jet or from a high- $p_T$  hadron interaction with material may cause multiple hits in the MTD. This type of background is most likely to produce hits in the edge-cells of neighboring backlegs. These hits can in turn be matched to tracks in the event. While the tracks themselves may be random products of other event activity, the MTD hits have a small opening angle. For this reason, the resulting pairs formed will peak at very specific invariant mass values. Also, the significant acceptance differences between like-sign pairs and unlike-sign pairs causes the specific invariant masses to be different. While the most background is observed in the edge cells, in principle there is a slightly different background mass distribution for every possible combination of cells, i.e. edge + edge, edge + center, etc. These pairs will contribute to the correlated background at low masses ( $M_{\mu\mu} < 1.5 \text{ GeV}/c^2$ ).

**Hadron decays** produce correlated hadron pairs. Due to the significant level of hadron contamination present in the muon candidates, correlated hadron pairs are likely to exist in the foreground unlike-sign dimuon distribution. The dominant channels contributing to the hadron decay background are listed in Table 5.3. For instance, the  $\rho^0 \rightarrow \pi^+\pi^-$  decay (with a branching ratio of  $\sim 100\%$ ) will produce a relatively large amount of correlated  $\pi^+\pi^-$  pairs. Since the efficiency for a  $\pi^\pm$  to have a match in the MTD is about 1%, these pairs will be



**Figure 5.9 :** The  $M_{\pi^+\pi^-}$  distribution near the  $K_S^0$  mass (a) and the  $M_{K^+K^-}$  distribution near the  $\phi$  mass (b). Only MTD matched tracks passing all quality cuts are used.

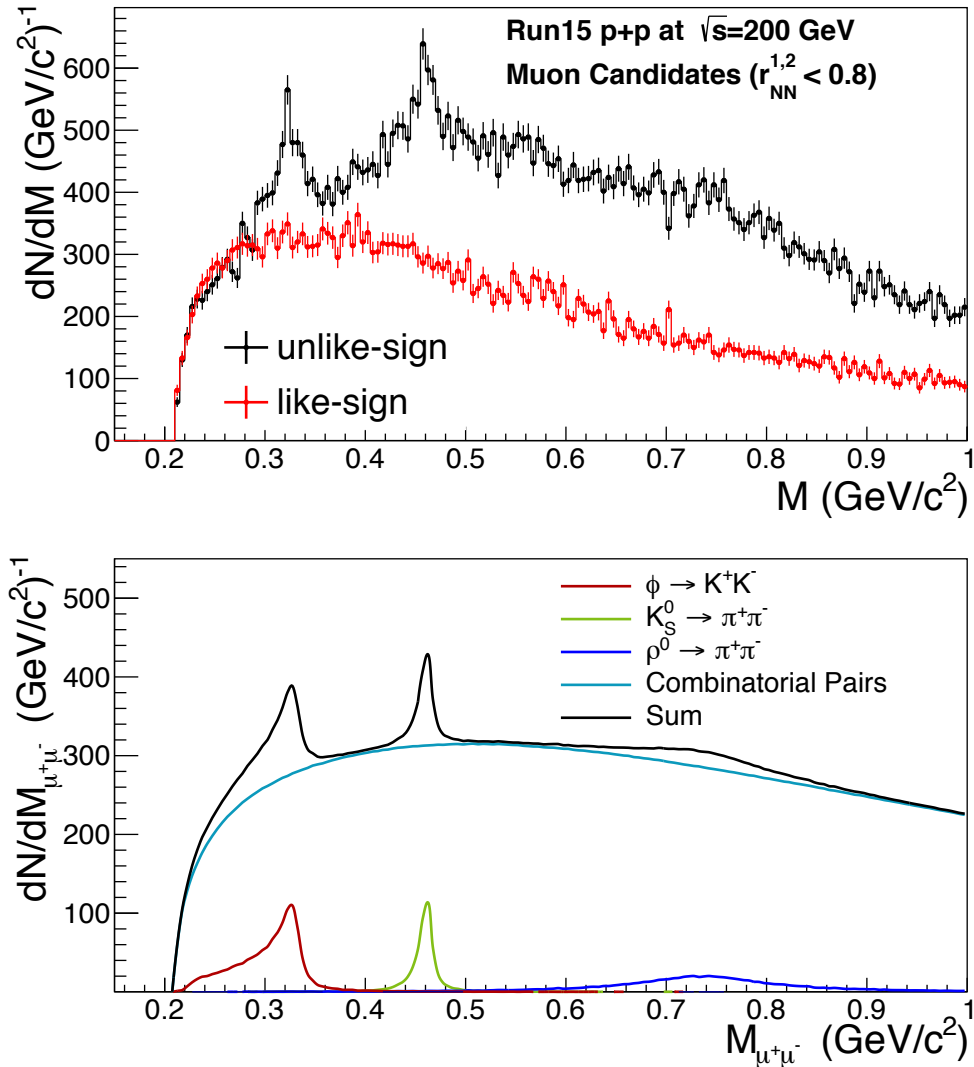
suppressed but still present. Despite the low reconstruction efficiency, these decays are still significantly more common than the signal decay  $\rho^0 \rightarrow \mu^+\mu^-$  which occurs with a branching ratio of only  $\sim 4.55 \times 10^{-5}$ . When this type of correlated hadron pair is reconstructed as part of the dimuon invariant mass spectra, the pair mass will be shifted due to the assumption that the daughters have the mass of muons. Since the charged pion and muon masses are so similar, the pair mass shift is small. On the other hand, the mass shift for a  $K^+K^-$  pair is much larger since the charged kaon mass is several hundred  $\text{MeV}/c^2$  larger than the muon mass. Very obvious peaks are visible from  $K_S^0 \rightarrow \pi^+\pi^-$  decays (Fig. 5.9a) and from  $\phi \rightarrow K^+K^-$  decays (Fig. 5.9b) despite the requirement that both tracks are matched to the MTD and pass quality cuts. As described above, the peaks from the background sources are shifted as part of the dimuon invariant mass spectra with the  $K_S^0$  peak being most shifted. The low mass dimuon invariant mass distribution from  $p + p$  collisions at  $\sqrt{s} = 200$  GeV is shown in the top panel of Fig. 5.10. In the lower panel of Fig. 5.10 simulated background contributions from combinatorial pairs,  $\phi \rightarrow K^+K^-$  decays,  $K_S^0 \rightarrow \pi^+\pi^-$  decays, and  $\rho^0 \rightarrow \pi^+\pi^-$  decays are shown. The scaling is arbitrary and meant only to show that the vast majority of the pairs from muon candidate can be explained with background-only sources.

**Secondary hadron decays** produce correlated pairs, though with a weaker degree of correlation between final state particles compared to a single decay. Examples of the decay processes contributions to this type of background are listed in Table 5.4. Generally speaking, the weaker the correlation between final state particles, the more smeared-out the

**Table 5.4 :** Processes through which correlated dimuon backgrounds can be produced.

Secondary Hadron Decay Processes	Total Branching Ratio
$\rho^0 \rightarrow \pi^+\pi^- \rightarrow \pi^\pm\mu^\mp + X$	$(\sim 100\% \times 99\%) = 99\%$
$K_S^0 \rightarrow \pi^+\pi^- \rightarrow \pi^\pm\mu^\mp + X$	$(69\% \times 99\%) = 68\%$
$\Lambda \rightarrow p + \mu^- + X$	$(63\% \times 99\%) = 62\%$
$\phi \rightarrow K^+K^- \rightarrow K^\pm\mu^\mp + X$	$(49\% \times 63\%) = 30\%$
$\rho^0 \rightarrow \pi^+\pi^- \rightarrow \mu^+\mu^- + X$	$(\sim 100\% \times (99\%)^2) = 98\%$
$K_S^0 \rightarrow \pi^+\pi^- \rightarrow \mu^+\mu^- + X$	$(69\% \times (99\%)^2) = 67\%$
$\phi \rightarrow K^+K^- \rightarrow \mu^+\mu^- + X$	$(49\% \times (63\%)^2) = 19\%$

invariant mass distribution will be. For this reason, the hadron decays will produce the most prominent peaks in the dimuon distribution. The secondary hadron decays will contribute much broader background distributions.



**Figure 5.10** : Top: The invariant mass of muon candidates in the low mass region ( $M < 1.0 \text{ GeV/c}^2$ ) with a cut on the NN response of both daughter 1 and 2 ( $r_{NN}^{1,2} < 0.8$ ) to select predominantly background-like pairs. Bottom: The simulated dimuon invariant mass distribution for hadron decay sources and combinatorial backgrounds. Scaling is arbitrary, meant to show that the distribution in data can be qualitatively described by background-only sources.

## 5.5 Signal Extraction

### 5.5.1 Like-Sign Ratio Background Estimation Technique

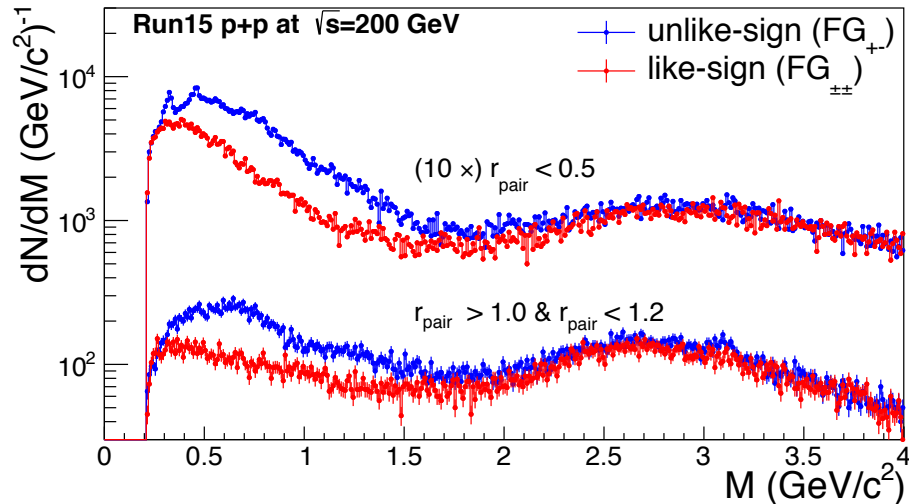
In the previous section, the contributions from uncorrelated combinatorial backgrounds and physical correlated backgrounds were discussed. In order to isolate the dimuon signal from these background sources, a technique is needed to accurately estimate the absolute yield of background sources as a function of invariant mass. While Monte Carlo simulations could provide approximate measures of these contributions, in order to reach the level of precision required, we instead explore data-driven techniques.

In this section the *like-sign ratio* (LSR) background estimation method is developed. Throughout the course of this research several techniques were investigated but none were found to perform as well as this method. The basic idea of this strategy is to parameterize the ratio of background in the foreground unlike-sign distribution to the foreground like-sign distribution. By using the LSR instead of the mean (direct or geometric) of the like-sign distribution, multiple effects can be taken into account simultaneously. For instance, this method naturally accounts for the significant acceptance differences between like-sign pairs and unlike-sign pairs. It is also capable of simultaneously accounting for uncorrelated background sources and physical background sources.

The LSR method consists of a few steps:

1. Compute the ratio of unlike-sign background to foreground like-sign,  $R(M; background)$ , in a region where the unlike-sign background can be isolated.
2. Compute the transfer function for  $R(M; background) \rightarrow R(M; signal)$ , where  $R(M; signal)$  is the ratio of unlike-sign background to foreground like-sign for signal pair selection criteria.
3. Correct the foreground like-sign distribution with signal selection criteria as  $FG_{++/--}(M; signal) \otimes R(M; signal)$

The first step in this procedure, computing  $R(M; background)$ , is accomplished by selecting background-only pairs using the pair response  $r_{\text{pair}}$ . Figure 5.11 clearly demonstrates that the background-only, foreground unlike-sign distribution ( $FG_{+-}$ ) contains significant sources that are absent in the like-sign distribution ( $FG_{\pm\pm}$ ). The excess is observed mostly at low

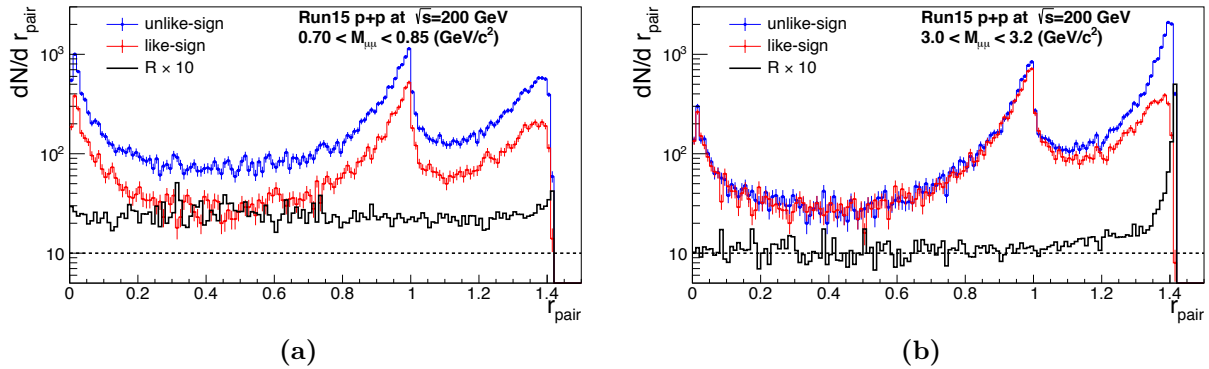


**Figure 5.11** : The  $FG_{+-}$  and  $F_{\pm\pm}$  distributions for two different background-only  $r_{\text{pair}}$  selection criteria. In both cases the  $FG_{\pm\pm}$  distribution fails to describe the  $FG_{+-}$  distribution.

to intermediate masses, which is consistent with our expectations from physical background sources. This figure also shows that at high masses, where combinatorial background dominates, the foreground like-sign distribution underestimates the unlike-sign distribution as expected. The lower set of curves in Fig. 5.11 show that  $R(M; \text{background})$  clearly changes as a function of the  $r_{\text{pair}}$  cuts applied, even in background-only regions.

The second step in this method depends heavily on the continuous PID response provided by the DNN-based PID. The pair response,  $r_{\text{pair}}$ , computed from the DNN-pid can be used to approach the signal PID region in a smooth, continuous manner. In terms of the  $r_{\text{pair}}$  value,  $\sim 0$  is a very high probability of being a background pair,  $\sim 1$  is a high probability of being a cross pair, and a value of  $\sim \sqrt{2}$  is a high probability of being a signal pair. It is also very important to point out that while the physical background, like  $K_S^0 \rightarrow \pi^+\pi^-$ , will produce correlated background pairs, the PID response of each daughter in the pair will be independently sampled from the underlying PID distribution. This means that the yield of cross pairs, which contribute the vast majority of the background for  $r_{\text{pair}} > 1.0$ , can be constrained with the yield of background-only pairs. Additionally, it means that the PID template shape for cross pairs (even for correlated pairs) can be accurately produced by independently sampling the single-track background and signal PID templates.

In order to develop the transfer function for computing  $R(M; \text{signal})$  a more differential approach can be used. Instead of expressing the ratio as a function of mass (i.e.  $R = R(M)$ )



**Figure 5.12** : An example of determining the like-sign ratio as a function of  $r_{\text{pair}}$ , i.e.  $R = R(r_{\text{pair}})$  in the  $\omega$  mass region,  $0.7 < M_{\mu\mu} < 0.85 \text{ GeV}/c^2$  (a) and the  $J/\psi$  mass region,  $3.0 < M_{\mu\mu} < 3.2 \text{ GeV}/c^2$ . The blue points show the  $FG_{+-}$  distribution, the red points show the  $FG_{\pm\pm}$  distribution, and the black distribution shows the ratio  $R \times 10$  to make it visible on the same axis. A dotted black line is added to show where  $R = 1$  would be.

for discrete ranges of  $r_{\text{pair}}$ , the ratio can be expressed as a function of  $r_{\text{pair}}$  for a given mass range, e.g.  $R = R(r_{\text{pair}}; 0.7 < M_{\mu\mu} < 0.85 \text{ GeV}/c^2)$  for part of the  $\omega$  mass range. Figure 5.12 shows the  $FG_{+-}$  (blue) and  $FG_{\pm\pm}$  (red) distributions as a function of  $r_{\text{pair}}$  for the mass range  $0.7 < M_{\mu\mu} < 0.85 \text{ GeV}/c^2$ . In black  $10 \times R(r_{\text{pair}}) = FG_{+-}/FG_{\pm\pm}$  is shown for the same mass range. A dotted line is drawn to show where  $R = 1$  would appear. From this plot and the value of  $R$ , it is clear that across the entire range of  $r_{\text{pair}}$ ,  $FG_{+-}$  has  $\sim 2\times$  excess over the  $FG_{\pm\pm}$  distribution. A small (large) uptick in the ratio is visible for the highest values of  $r_{\text{pair}}$  for the  $\omega$  ( $J/\psi$ ) mass region due to the presence of signal pairs.

Expressing the like-sign ratio as a function of  $r_{\text{pair}}$  shows that the value of  $R$  in the background only region cannot be used in the signal region directly. However, the value of  $R$  cannot be measured directly in the signal region, since the presence of signal will increase  $R$  by construction. For this reason, some technique is needed to transfer the value of  $R(r_{\text{pair}}; \text{background})$  into the signal region. While a simple zeroth-order or first-order polynomial appears to be a good choice in some mass bins, these are clearly not a good choice for other mass bins. For instance, a linear extrapolation function used in the  $\phi$  mass region can give values of  $R(r_{\text{pair}}; \text{signal})$  that vary by over 60% depending on the specific fit range used. Instead of guessing at some fit function, it would be better to understand the evolution of the three categories of pair types (background, cross, and signal pairs) as a function of  $r_{\text{pair}}$  so that the exact contributions can be computed in the signal region.



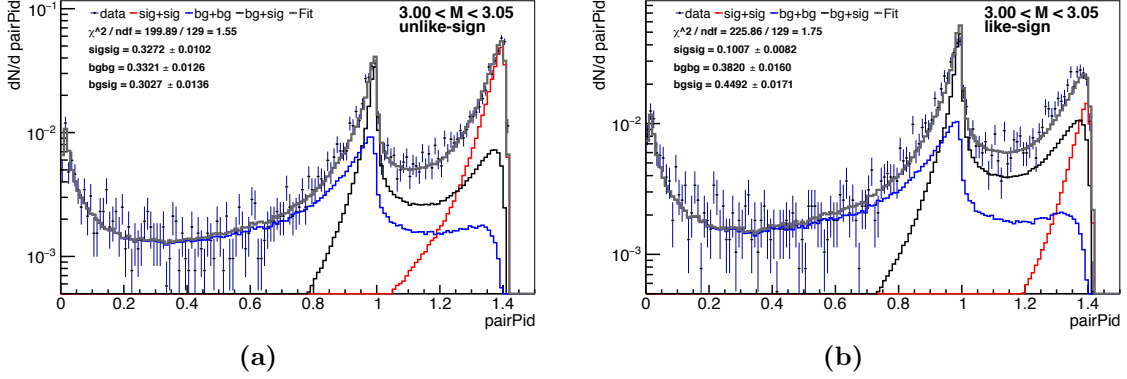
A template fitting strategy similar to the one used to extract the muon purity can be used to determine the yield of background, cross, and signal pairs from a  $FG_{+-}(r_{\text{pair}})$  distribution. In order to fit the  $r_{\text{pair}}$  distribution, the corresponding template shapes must be built from the simulated signal and background sources. It was crucial to match the exact kinematic ranges used in data when generating the single track templates, otherwise the shapes would not agree well and the fit quality would suffer. In the case of pairs the single track kinematics as well as the correlations in  $p_T$ -space must be taken into account in order to match the shapes found from pairs in data. Since the single track background shape changes depending on how much  $\pi$ , K, and  $p$  fractions are present, this also must be taken into account. Therefore, for each fit of  $FG_{+-}(r_{\text{pair}})$  or  $FG_{\pm\pm}(r_{\text{pair}})$  pair templates are generated according to:

$$\begin{aligned}
T_{\text{background}}^{\text{pair}}(p_{T,1}, \eta_1, \phi_1, p_{T,2}, \eta_2, \phi_2) = & \\
& \left( T_{\text{background}}^{\text{single-track}}(p_{T,1}, \eta_1, \phi_1, Y_\pi(p_{T,1}), Y_K(p_{T,1}), Y_p(p_{T,1})) \right. \\
& \left. \otimes T_{\text{background}}^{\text{single-track}}(p_{T,2}, \eta_2, \phi_2, Y_\pi(p_{T,2}), Y_K(p_{T,2}), Y_p(p_{T,2})) \right) \cdot w(p_{T,1}, p_{T,2})
\end{aligned} \tag{5.39}$$

$$\begin{aligned}
T_{\text{cross}}^{\text{pair}}(p_{T,1}, \eta_1, \phi_1, p_{T,2}, \eta_2, \phi_2) = & \\
& \left( T_{\text{background}}^{\text{single-track}}(p_{T,1}, \eta_1, \phi_1, Y_\pi(p_{T,1}), Y_K(p_{T,1}), Y_p(p_{T,1})) \right. \\
& \left. \otimes T_{\text{signal}}^{\text{single-track}}(p_{T,2}, \eta_2, \phi_2) \right) \cdot w(p_{T,1}, p_{T,2})
\end{aligned} \tag{5.40}$$

$$\begin{aligned}
T_{\text{signal}}^{\text{pair}}(p_{T,1}, \eta_1, \phi_1, p_{T,2}, \eta_2, \phi_2) = & \\
& \left( T_{\text{signal}}^{\text{single-track}}(p_{T,1}, \eta_1, \phi_1) \right. \\
& \left. \otimes T_{\text{signal}}^{\text{single-track}}(p_{T,2}, \eta_2, \phi_2) \right) \cdot w(p_{T,1}, p_{T,2})
\end{aligned} \tag{5.41}$$

where  $T_{\text{signal}}^{\text{single-track}}(\dots)$  and  $T_{\text{background}}^{\text{single-track}}(\dots)$  are the single track template shapes for signal and background, respectively. Each single-track template is a function of the daughter track's kinematic variables  $(p_T, \eta, \phi)$ . In order to take into account the variations in background template shape due to different hadron contamination levels, the single-track background template is also a function of the relative yields of  $\pi$ , K, and  $p$  ( $Y_\pi$ ,  $Y_K$ , and  $Y_p$  respectively)



**Figure 5.13** : Fits to the  $r_{\text{pair}}$  (pairPid) distributions  $FG_{+-}(r_{\text{pair}})$  (a) and  $FG_{\pm\pm}(r_{\text{pair}})$  (b).

for the given  $p_T$  value. The  $\otimes$  operator represents an outer product between the template shapes. Additionally, the transverse momentum space correlations between the daughter particles is taken into account. These rather complex forms were developed iteratively by adding features to improve the correspondence between the pair template shapes and the data. An example fit using the pair templates can be seen in Fig. 5.13 for the mass range  $3.0 < M_{\mu\mu} < 3.05 \text{ GeV}/c^2$ . Each distribution is fit to:

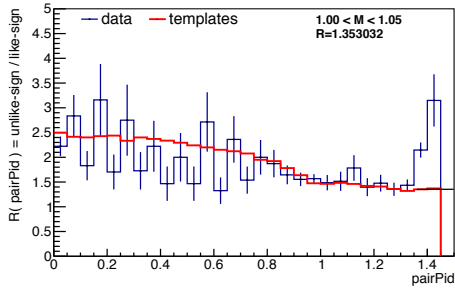
$$f(Y_{\text{signal}}, Y_{\text{cross}}, Y_{\text{background}}) = Y_{\text{signal}} \cdot T_{\text{signal}}^{\text{pair}} + Y_{\text{cross}} \cdot T_{\text{cross}}^{\text{pair}} + Y_{\text{background}} \cdot T_{\text{background}}^{\text{pair}} \quad (5.42)$$

Where  $Y_{\text{signal}}$ ,  $Y_{\text{cross}}$ , and  $Y_{\text{background}}$  are the only free parameters. The  $\chi^2/ndf$  values are computed as a metric of the goodness-of-fit using only the statistical uncertainty on the data points. Figure 5.13 shows that the fits work equally well for the  $FG_{+-}$  distributions as the  $FG_{\pm\pm}$  distributions.

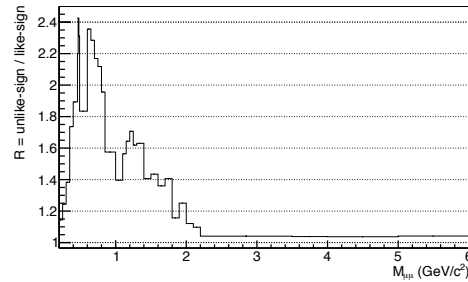
With template fits to the  $FG_{+-}(r_{\text{pair}})$  and  $FG_{\pm\pm}(r_{\text{pair}})$  distributions,  $R(r_{\text{pair}})$  can be computed without needing to guess at the proper extrapolation functions. In order to calculate  $R(r_{\text{pair}})$  without the contribution from real signal pairs,  $R$  is computed as follows:

$$R(r_{\text{pair}}) = \frac{Y_{\text{background}}^{+-} \cdot T_{\text{background}}^{+-} + Y_{\text{cross}}^{+-} \cdot T_{\text{cross}}^{+-} + Y_{\text{signal}}^{\pm\pm} \cdot T_{\text{signal}}^{\pm\pm}}{Y_{\text{background}}^{\pm\pm} \cdot T_{\text{background}}^{\pm\pm} + Y_{\text{cross}}^{\pm\pm} \cdot T_{\text{cross}}^{\pm\pm} + Y_{\text{signal}}^{\pm\pm} \cdot T_{\text{signal}}^{\pm\pm}} \quad (5.43)$$

This is equivalent to taking the ratio of unlike-sign to like-sign template fits, but with the substitution of the like-sign signal yield into the unlike-sign result. This ensures that



**Figure 5.14** :  $R_{\text{pair}}$  computed from the data (blue), which includes contribution from signal pairs and  $R_{\text{pair}}$  computed from templates (red).

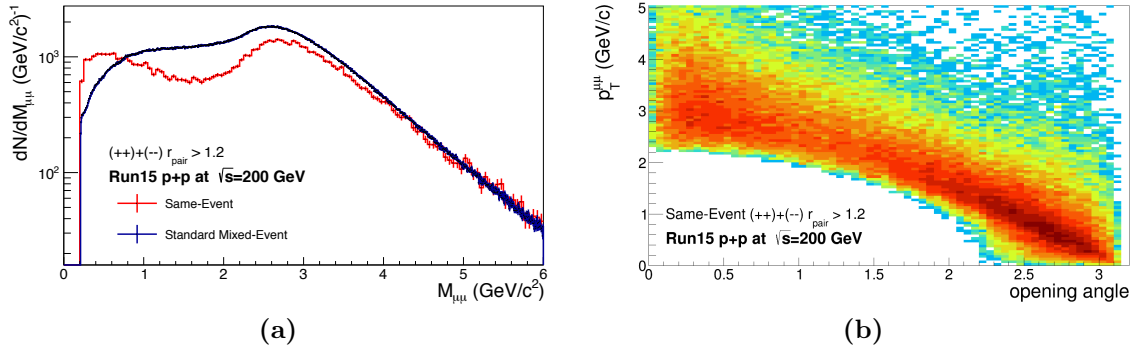


**Figure 5.15** :  $R_{\text{pair}}$  computed from the templates and evaluated in the signal selection region shown as a function of  $M_{\mu\mu}$ . The effect of the  $K_S^0$  can be clearly seen around  $M_{\mu\mu} = 0.45 \text{ GeV}/c^2$ . The value at high mass stabilizes to a small enhancement due only to the statistical effect from single muon production.

the real signal in the  $FG_{+-}$  is not used when computing  $R$ , which should only measure the background sources. An example of the result of this procedure is shown in Fig. 5.14. In this figure, the ratio of  $FG_{+-}$  to  $FG_{\pm\pm}$  is also shown for comparison since they should be approximately equal in the background only region. The computed value of  $R(M; \text{signal})$  using this technique is shown in Fig. 5.15. At this point, the full background for the  $FG_{+-}$  signal distribution could be computed using  $R(M; \text{signal})$  and the  $FG_{\pm\pm}$  distribution. However, the incredibly limited statistics available in the  $FG_{\pm\pm}$  after applying the signal selection pid cuts result in large statistical fluctuations in the  $FG_{\pm\pm}$  distribution, especially in small mass bins like those used under the resonances.

### 5.5.2 Correlation Weighted Event Mixing Technique

The standard event mixing technique is to make pairs from tracks in different events, thus breaking any correlation between the two tracks. Using this method a nearly infinite number of uncorrelated combinatorial pairs can be produced with the same acceptance constraints as the same-event pairs. In past STAR dielectron analyses, the event-mixing technique has been used to compute the acceptance differences between like-sign and unlike-sign pairs to arbitrary precision. In some cases the uncorrelated combinatorial mixed-event background can describe the background in the foreground unlike-sign distribution sufficiently well enough

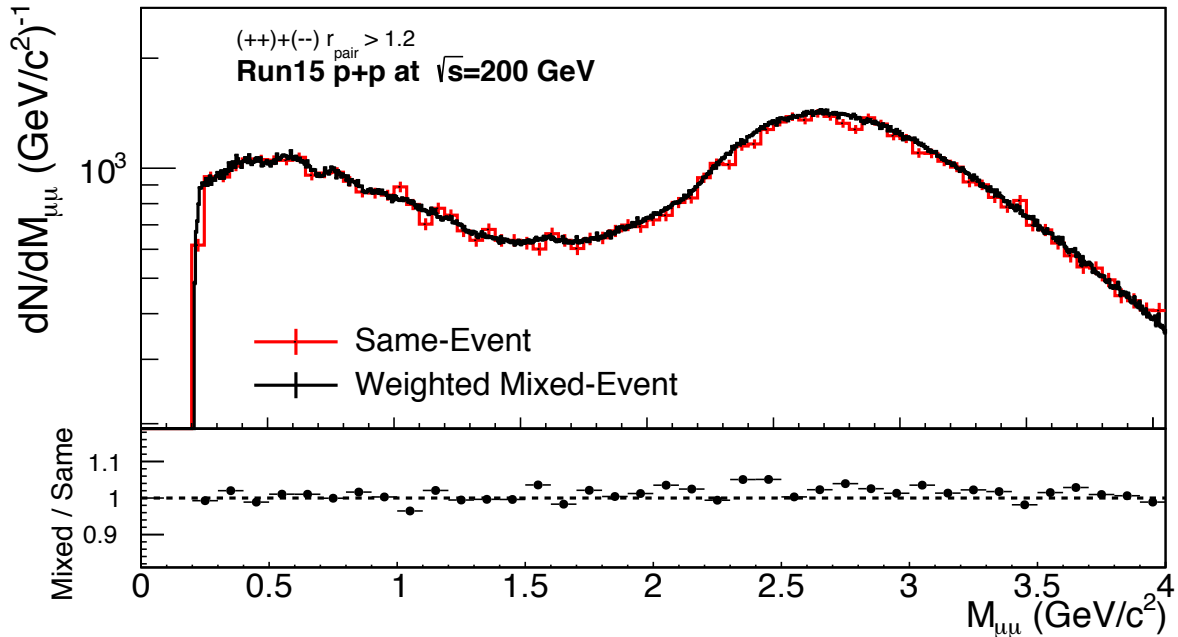


**Figure 5.16 :** The standard mixing-event technique compared with the same-event like-sign distribution (a). The mixed-event distribution is able to match the same-event shape only at very high  $M_{\mu\mu}$ . The opening angle versus pair  $p_T$  used to weight the mixed-event distribution in order to produce a mixed-event that reproduces the shape of the same-event distribution (b).

to be used as the background estimate.

In this analysis, the goal is to produce a replacement for the same-event like-sign distribution. This replacement, hopefully free of statistical fluctuations, would then be used as the baseline background in the like-sign ratio background method discussed in the previous section. However, producing a replacement distribution for the  $FG_{\pm\pm}$  using standard event-mixing is not possible since there exist correlated pairs even in the  $FG_{\pm\pm}$ , though not as many as in  $FG_{+-}$ . This can be clearly seen in Fig. 5.16a, where the same-event  $FG_{\pm\pm}$  distribution is plotted along with the distribution resulting from a standard event-mixing procedure ( $ME_{\pm\pm}$ ). The  $ME_{\pm\pm}$  distribution is somewhat able to reproduce the shape at large invariant masses, but completely fails to at low and intermediate masses. Since the goal is to replace the  $FG_{\pm\pm}$  distribution especially at low mass under the  $\omega$  and  $\phi$  resonances, the standard event-mixing is not sufficient.

Instead of using the standard mixing event technique, the  $ME_{\pm\pm}$  distribution can be intentionally weighted to achieve the same correlations as found in the  $FG_{\pm\pm}$  distribution. One way this can be accomplished by measuring the same-event two-particle correlations  $\Delta\eta$  and  $\Delta\phi$  and then re-weighting the contribution from each mixed event pair as the  $ME_{\pm\pm}$  distribution is built. This approach was found to be only partially successful though. The  $\Delta\eta$  vs.  $\Delta\phi$  map needs a sufficiently high granularity to reproduce the shape of the same-event correlations. However, with too many bins the statistical uncertainty originally embedded in



**Figure 5.17 :** The mixed-event distribution weighted by the pair opening angle as a function of pair  $p_T$ .

the  $FG_{\pm\pm}$  invariant mass distribution, becomes embedded in the correlation map and simply transferred to  $ME_{\pm\pm}$  when the weighting is applied. The end result is a distribution with the correct correlations but with the same statistical fluctuations as the  $FG_{\pm\pm}$ .

Instead it was found that weighting the distribution according to the pair opening angle as a function of  $p_T^{\mu\mu}$  was more effective. Using the opening angle ( $\alpha$ ) to measure the correlation between tracks reduces the degree of the problem to 1D. Additional statistical stability is gained due to the symmetry inherent in measuring the opening angle ( $\alpha$  and  $-\alpha$  are interchangeable). The same-event  $\alpha$  vs.  $p_T^{\mu\mu}$  distribution is shown in Fig. 5.16b. This distribution can be used to weight the standard  $ME_{\pm\pm}$  distribution, resulting in a mixed-event distribution with the same correlations as in same-event, but with significantly better statistical stability. This result can be seen in Fig. 5.17 for like-sign pairs with  $r_{\text{pair}} > 1.2$ , though the technique works equally well for all PID ranges and also for background only unlike-sign pairs.

### 5.5.3 Raw Signal Extraction

The raw signal extraction procedure will be discussed first for the  $\phi$  meson and then for the invariant mass distribution. The  $\phi \rightarrow \mu^+\mu^-$  spectra is extracted in four momentum bins:  $p_T = 2.2-3.0, 3.0-3.5, 3.5-4.0,$  and  $4.0-6.0$  GeV/ $c$ . In each momentum bin, the raw number of  $\phi$ ,  $N_\phi$  is extracted through a maximum likelihood fit in which a Gaussian is used for the  $\phi$  resonance while a fourth-order polynomial is used for the background shape. The TMINUIT library is used for the minimization procedure. The MINOS algorithm was used to compute the uncertainties of the fit [146]. The raw data points along with the maximum likelihood fits is shown for each  $p_T$  bin in Fig. 5.18.

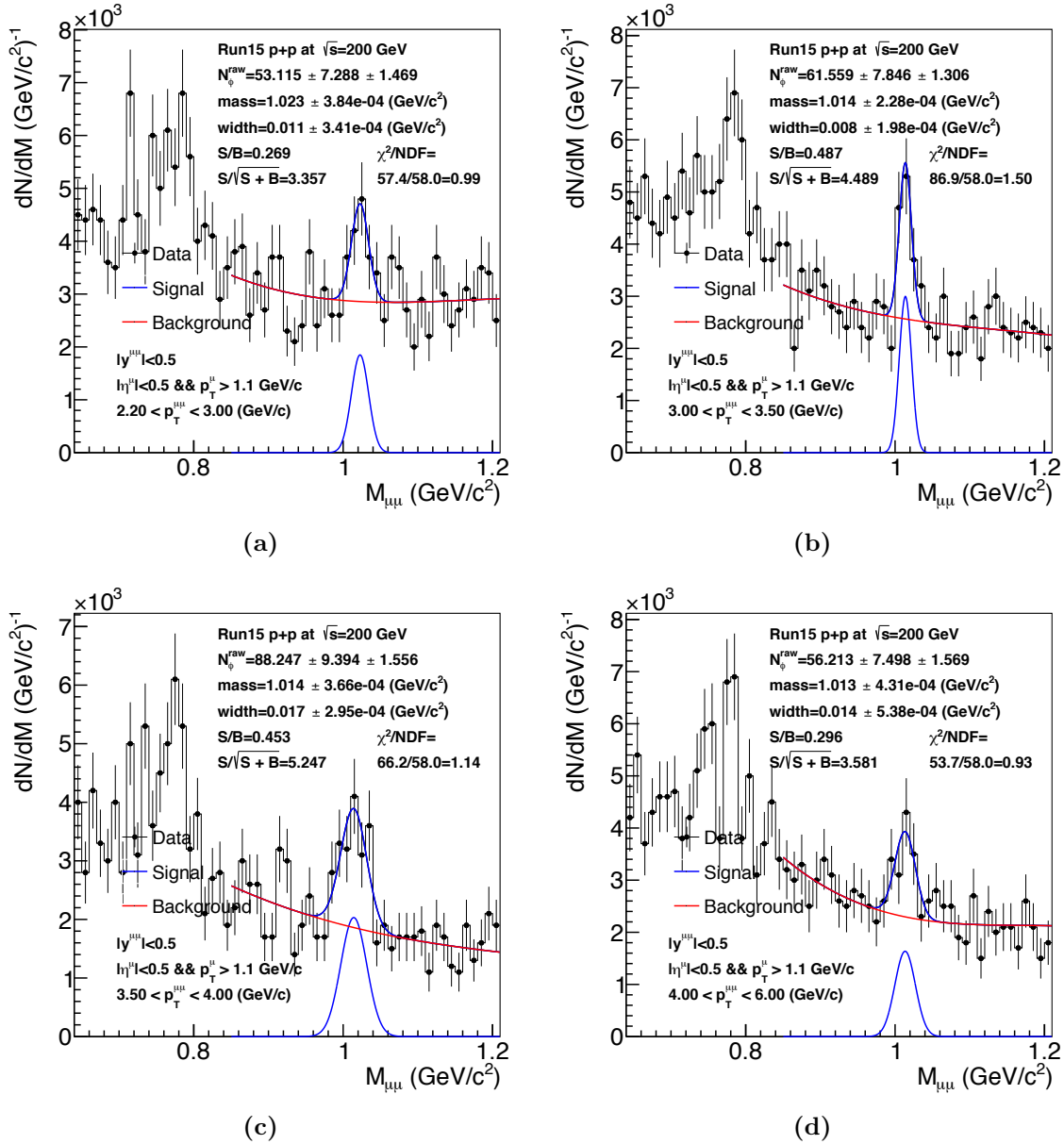
The signal dimuons can be extracted from the foreground  $\mu^+\mu^-$  distribution by subtracting off all background sources. The like-sign ratio method, discussed in Sec. 5.5.1 is used to correct the foreground like-sign distribution. The signal is given by:

$$S = FG_{+-} - FG_{\pm\pm}(M) \circ R(M; signal) \quad (5.44)$$

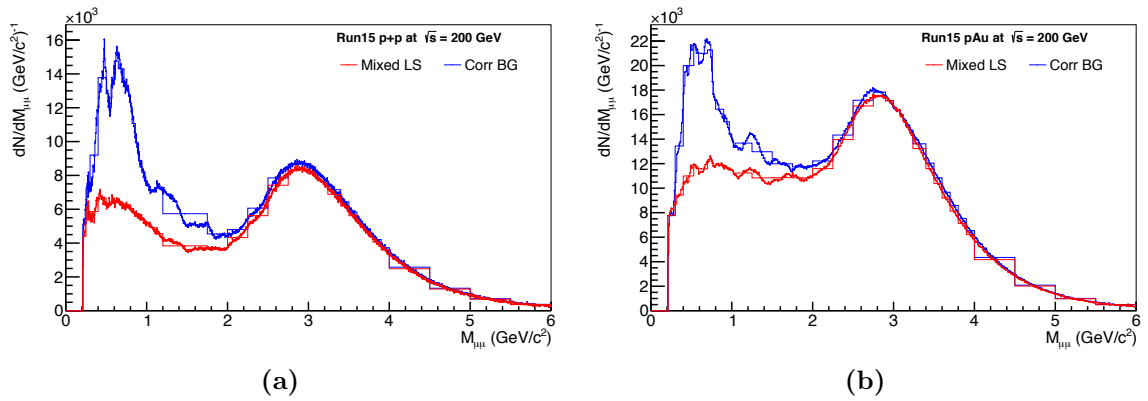
where  $R(M; signal)$  is the correction factor needed to incorporate the missing correlated backgrounds that are not present in the  $FG_{\pm\pm}$  distribution. Where the  $\circ$  symbol is used for the Hadamard product, i.e. that the correction is applied on a bin-by-bin bases. The low muon purity necessitates a very tight cut on  $r_{\text{pair}}$  which reduces the already limited statistics of the data samples. The statistical fluctuations in  $FG_{\pm\pm}$  are prohibitively large, especially in the small mass bins used underneath the resonances. For this reason, the correlation weighted mixed-event technique is employed to produce a statistically stable mixed-event distribution ( $ME_{\pm\pm}$ ) that reproduces the major features of the  $FG_{\pm\pm}$  distribution. Therefore, instead of using the same-event foreground like-sign distribution to extract the signal we can use the mixed-event distribution as follows:

$$S = FG_{+-} - ME_{\pm\pm}(M) \circ R(M; signal) \quad (5.45)$$

The  $ME_{\pm\pm}$  distributions is shown in Fig. 5.19 before and after applying the correction factor  $R(M; signal)$  needed to reproduce the correlated backgrounds. The  $FG_{+-}$  distribution is shown in Fig. 5.20 along with the  $ME_{\pm\pm}$  distribution and the correlated background estimate. The resulting raw signal dimuon distribution is also shown in that figure. Figure



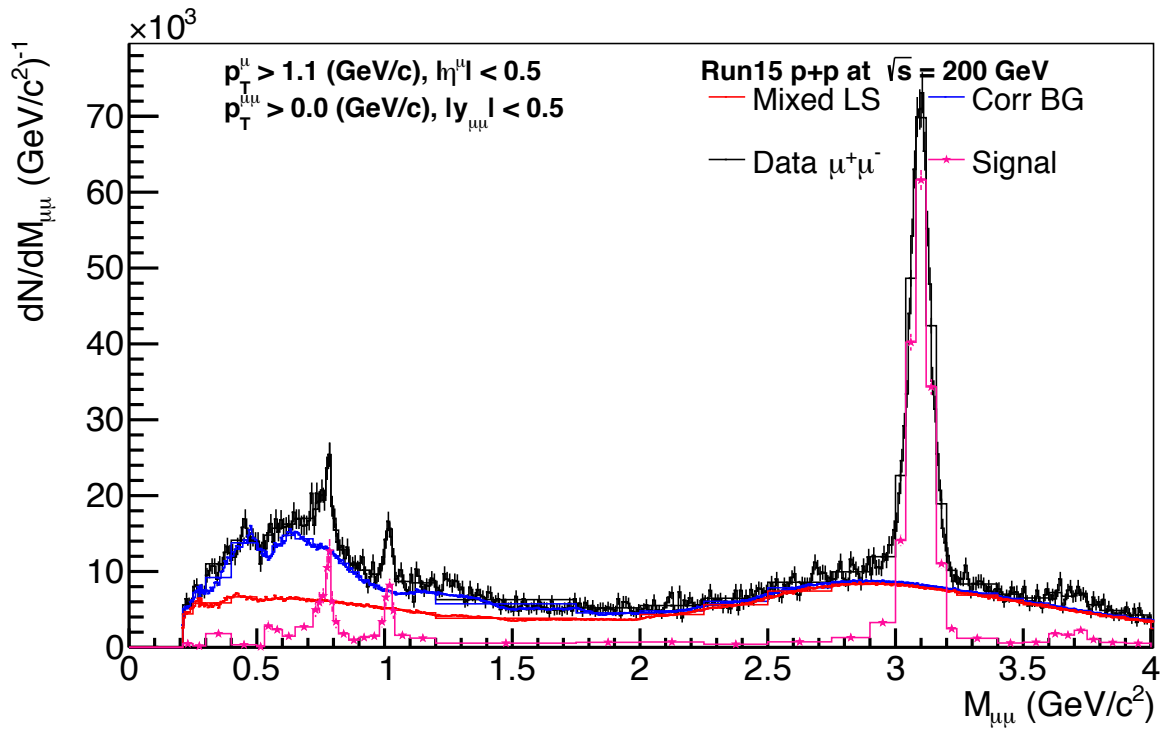
**Figure 5.18** : The maximum likelihood fits used to extract the raw  $\phi \rightarrow \mu^+ + \mu^-$  yields in four momentum bins. Each fit used a Gaussian for the  $\phi$  peak and a fourth-order polynomial for the background shape. The nominal fit range was  $0.85 < M_{\mu\mu} < 1.2 \text{ GeV}/c^2$ .



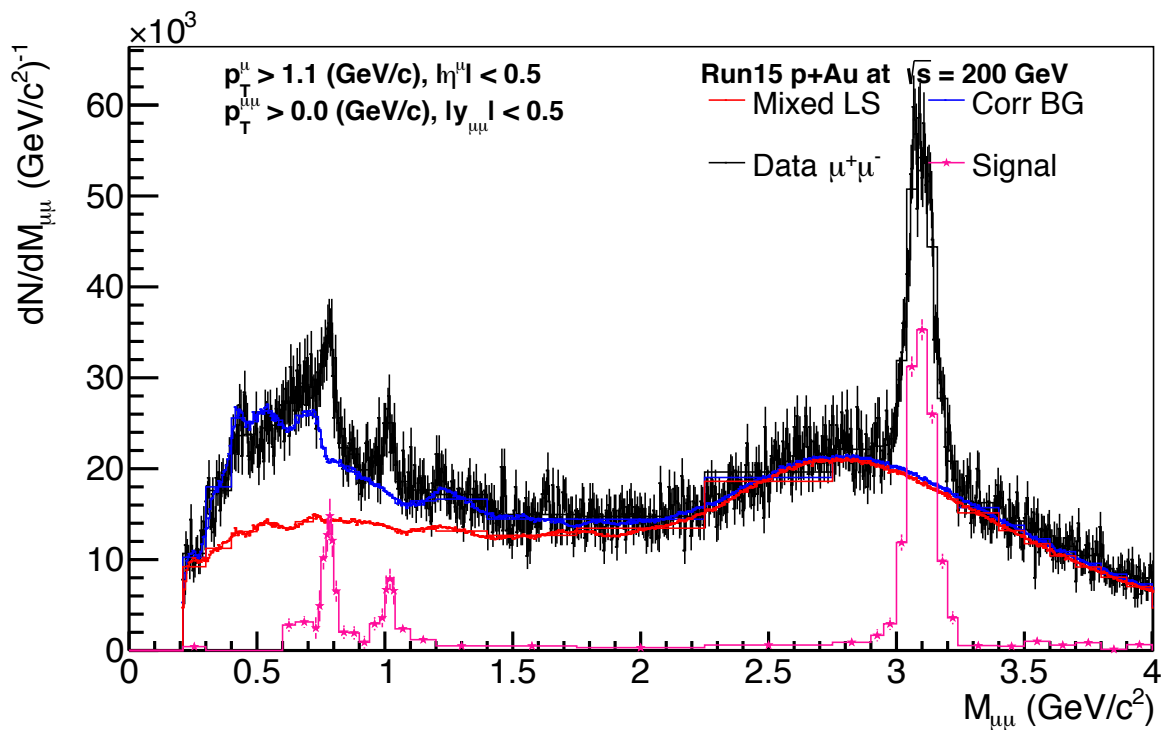
**Figure 5.19** : The raw  $ME_{\pm\pm}$  (red) and the correlated background (blue) for  $p + p$  (a) and  $p + \text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV.

5.21 shows the resulting signal-to-background ratio for the raw signal with respect to the correlated background estimate in  $p + p$  and  $p + \text{Au}$  collisions.



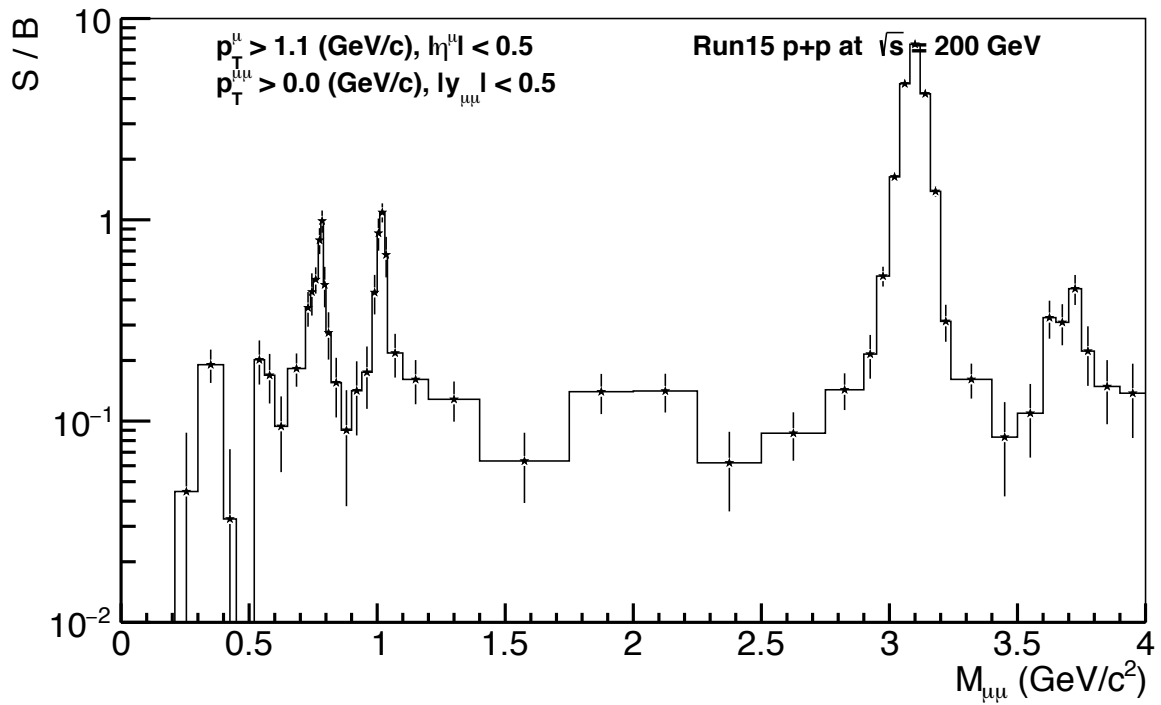


(a)

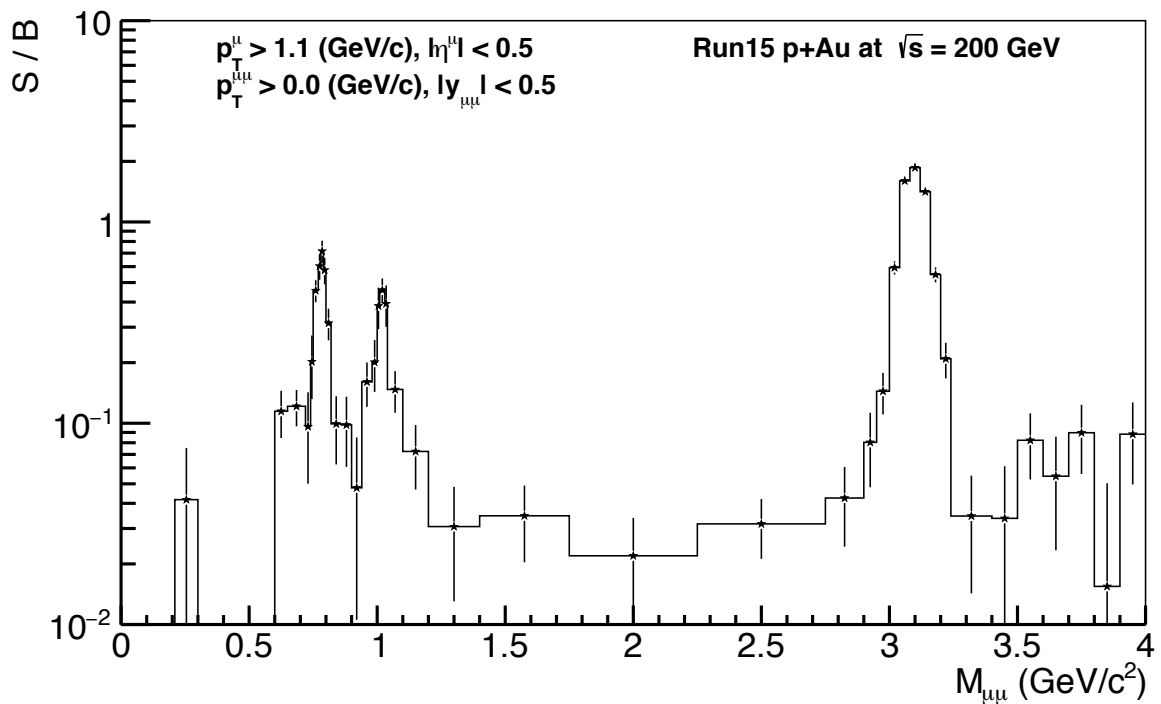


(b)

**Figure 5.20** : The  $FG_{+-}$  distribution (black) along with the  $ME_{\pm\pm}$  distribution (red) and the correlated background estimation (blue). The signal dimuons are shown in magenta.



(a)



(b)

**Figure 5.21** : The signal over background ratio as a function of invariant mass for  $p + p$  (top) and  $p+\text{Au}$  (bottom).

## 5.6 Contributions from Hadronic Decays

### 5.6.1 Decay Channels, Branching Ratios, and Kinematics

Light and heavy hadron decays account for a significant amount of the  $\mu^+\mu^-$  pairs produced in  $p+p$  and  $p+\text{Au}$  collisions at  $\sqrt{s_{NN}}=200$  GeV. The significant sources for the  $\mu^+\mu^-$  channel are listed below:

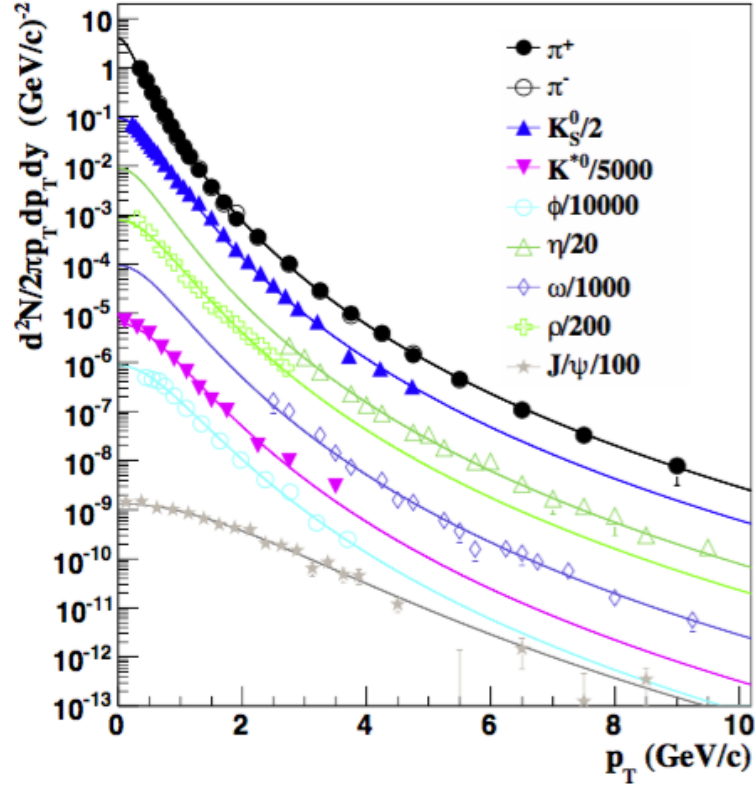
- Di-muon decays:  $\eta \rightarrow \mu^+\mu^-$ ,  $\omega \rightarrow \mu^+\mu^-$ ,  $\rho \rightarrow \mu^+\mu^-$ ,  $\phi \rightarrow \mu^+\mu^-$ ,  $J\psi \rightarrow \mu^+\mu^-$ ,  $\psi(2S) \rightarrow \mu^+\mu^-$
- Dalitz decays:  $\eta \rightarrow \gamma\mu^+\mu^-$ ,  $\omega \rightarrow \pi^0\mu^+\mu^-$ ,  $\eta' \rightarrow \gamma\mu^+\mu^-$
- Heavy flavor decays:  $c\bar{c} \rightarrow \mu^+\mu^-$ ,  $b\bar{b} \rightarrow \mu^+\mu^-$
- The Drell Yan process.

In this section and the next the details of the hadronic cocktail simulation for the light hadron contributions,  $J/\psi$ , and  $\psi(2S)$  will be discussed. In the following section the contributions from the correlated heavy flavor decays and the Drell-Yan process will be discussed. A dedicated Monte Carlo code was written to simulate the two-body and three-body hadron decays. The PYTHIA event generator is used to simulate the heavy flavor decays and Drell-Yan processes.

The kinematic distributions for each parent particle is needed by the hadronic cocktail simulation code to simulate each decay into final products. The simulation code for the hadronic cocktail expresses the kinematics in terms of the  $p_T$ ,  $\eta$  and  $\phi$ . Measurement of each particles' transverse momentum spectra are fit simultaneously to a Tsallis Blast-Wave parameterization [147–149]. The Tsallis Blast-Wave parameterization is given by:

$$\frac{dN}{m_T dm_T} \propto m_T \int_{-Y}^{+Y} \cosh(y) dy \int_{-\pi}^{+\pi} d\phi \int_0^R r dr \left( 1 + \frac{q-1}{T} (m_T \cosh(y) \cosh(\rho) - p_T \sinh(\rho) \cos(\phi)) \right)^{-1/(q-1)} \quad (5.46)$$

where  $\rho = \tanh^{-1}(\beta_s(r/R)^n)$  is the flow profile which grows with the  $n$ -th power from the center of the collision to  $\beta_s$  at the hard-sphere radius of  $R$  along the transverse radial



**Figure 5.22 :** The measured transverse momentum spectra of a large number of mesons in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV. The solid lines are Tsallis Blast-Wave fits to each particle species [147].

direction  $r$ . The measured values and their Tsallis Blast-Wave fit can be seen in Fig. 5.22. The simultaneously fit Tsallis Blast-Wave does a good job of describing the spectra of light hadrons and can even be used to predict the spectra of unmeasured particle species based on their mass.

The rapidity distribution of the parent particles is also needed in order to simulate the decay into final state leptons. While a flat distribution could be used, a more accurate description is given by the CERES parameterization [150]:

$$dN/dy = \cosh^{-2} \left( \frac{3y}{4\sigma_{Landau} \left(1 - \frac{y^2}{2\sqrt{s}/m}\right)} \right) \quad (5.47)$$

$$\sigma_{Landau} = \sqrt{\log(\sqrt{s}/(2m_N))} \quad (5.48)$$

**Table 5.5 :** The decay channels supported by the hadronic cocktail simulation code. For each decay channel, Table 5.5 shows the mass,  $\Gamma_0$ ,  $dN/dy$  or  $d\sigma/dy$ , branching ratio (BR), and  $\Lambda^{-2}$  values are shown where applicable. Taken from Ref. [147, 151–158]

Meson	M (GeV/c <sup>2</sup> )	$\Gamma_0$ (GeV/c <sup>2</sup> )	dN/dy	Channel	BR	$\Lambda^{-2}$ (GeV <sup>-2</sup> )
$\eta$	0.547	$1.31 \times 10^{-6}$	$1.70 \times 10^{-1}$	$\gamma\mu^+\mu^-$	$3.1 \times 10^{-4}$	1.95
$\eta$				$\mu^+\mu^-$	$5.8 \times 10^{-6}$	—
$\omega$	0.782	$8.49 \times 10^{-3}$	$1.33 \times 10^{-1}$	$\pi^0\mu^+\mu^-$	$1.3 \times 10^{-4}$	2.24
$\omega$				$\mu^+\mu^-$	$9.0 \times 10^{-5}$	—
$\eta'$	0.957		$4.07 \times 10^{-2}$	$\gamma\mu^+\mu^-$	$1.08 \times 10^{-4}$	1.8396
$\rho^0$	0.775	$1.491 \times 10^{-1}$	$2.22 \times 10^{-1}$	$\mu^+\mu^-$	$4.55 \times 10^{-5}$	—
$\phi$	1.019	$4.26e \times 10^{-3}$	$1.73 \times 10^{-2}$	$\gamma\mu^+\mu^-$	$1.4 \times 10^{-5}$	3.8
$\phi$				$\mu^+\mu^-$	$2.87 \times 10^{-4}$	—
$J/\Psi$	3.096	$9.29 \times 10^{-5}$	$2.44 \times 10^{-5}$	$\mu^+\mu^-$	$5.96 \times 10^{-2}$	—
$\Psi(2S)$	3.68	$3.04 \times 10^{-4}$	$3.38 \times 10^{-6}$	$\mu^+\mu^-$	$8.0 \times 10^{-3}$	—

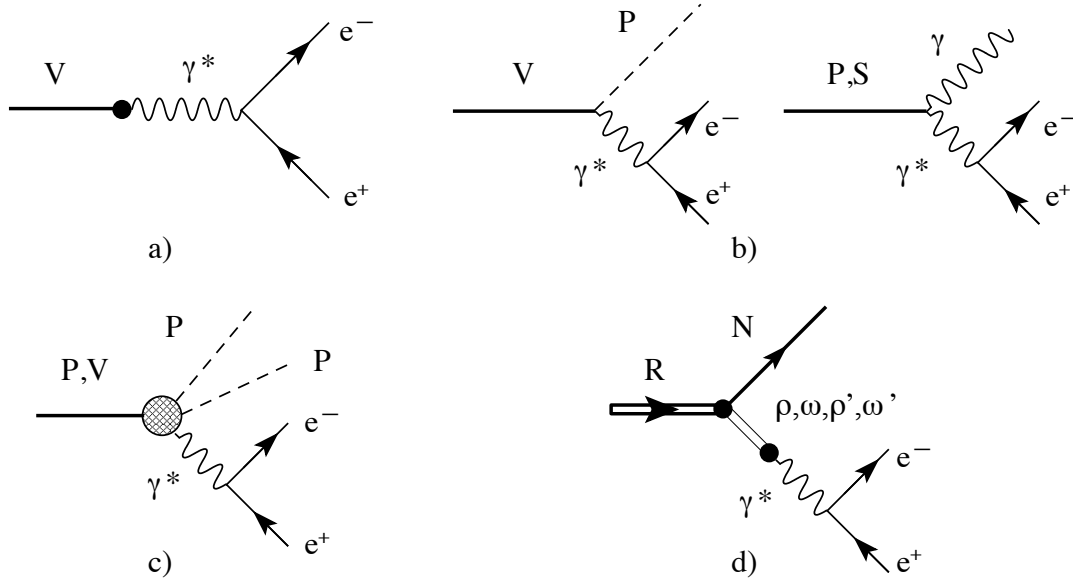
where  $\sqrt{s}$  is the center of mass energy per nucleon,  $m$  is the particle species mass, and  $m_N$  is the nucleon mass. The azimuthal distribution of each particle is assumed to be flat in  $0 < \phi < 2\pi$ . Table 5.5 shows all of the decay channels that are simulated by the hadronic decay simulation code. For each decay channel, Table 5.5 shows the mass,  $\Gamma_0$ ,  $dN/dy$  or  $d\sigma/dy$ , branching ratio, and  $\Lambda^{-2}$  values which are used to simulate the decays.

### 5.6.2 Decays of Vector Mesons and Pseudo Scalar Mesons

For the decay of narrow resonances decaying through a two-body decay channel, the non-relativistic Breit-Wigner line shape can be used. The non-relativistic S-wave Breit-Wigner has the form:

$$\frac{dN}{dM_{\mu\mu}} = \frac{2\Gamma_0}{(M_{\mu\mu} - M_0)^2 + \Gamma_0^2/4} \quad (5.49)$$

where  $M_0$  is the mass of the resonance and  $\Gamma_0$  is its width. This line shape is normalized and used by the hadronic cocktail code as the probability density function for the mass of the parent particle upon decay. All of the two-body decay channels use the non-relativistic Breit-Wigner line shape except the  $\rho$ -meson. Since the  $\rho$ -meson width is comparable to its mass ( $M_\rho = 0.775$  GeV/c<sup>2</sup>,  $\Gamma_0 = 0.149$  GeV/c<sup>2</sup>) the non-relativistic Breit-Wigner is not a good description of its line shape. Instead the relativistic S-wave Breit-Wigner can be used



**Figure 5.23 :** The Feynman diagrams for various dilepton production mechanisms (shown for  $e^+e^-$ ): the direct decay of a vector meson ( $\rho, \omega, \phi$ ) through a virtual photon (a); the Dalitz decay of a vector (V), pseudo-scalar (P), or scalar (S) meson into a neutral particle and an  $e^+e^-$  pair (b); the four-body decay of a pseudo-scalar or vector meson into  $e^+e^-$  and two pseudo-scalar mesons ( $\pi$  or  $\eta$ ) through an intermediate state containing a virtual photon or vector meson (c); the decay of a nucleon or  $\Delta$  resonance into a nucleon and a vector meson which further decays into an  $e^+e^-$  pair. [74]

as the line shape for decays. The relativistic S-wave Breit-Wigner has the form:

$$\frac{dN}{dM_{\mu\mu}dp_T} \propto \frac{M_{\mu\mu}M_\rho\Gamma_{\mu\mu}}{(M_\rho^2 - M_{\mu\mu}^2)^2 + M_\rho^2(\Gamma_{\pi\pi} + \Gamma_{\mu\mu}\Gamma_2)^2} \times PS(p_T) \quad (5.50)$$

$$\Gamma_{\pi\pi} = \Gamma_0 \frac{M_\rho}{M_{\mu\mu}} \left( \frac{M_{\mu\mu}^2 - 4M_\pi^2}{M_\rho^2 - 4M_\pi^2} \right)^{1/2} \quad (5.51)$$

$$\Gamma_{\mu\mu} = \Gamma_0 \frac{M_\rho}{M_{\mu\mu}} \left( \frac{M_{\mu\mu}^2 - 4M_\mu^2}{M_\rho^2 - 4M_\mu^2} \right)^{1/2} \quad (5.52)$$

$$PS(p_T) = \frac{M_{\mu\mu}}{\sqrt{M_{\mu\mu}^2 + p_T^2}} e^{-\frac{\sqrt{M_{\mu\mu}^2 + p_T^2}}{T}} \quad (5.53)$$

where  $M_\rho$  is  $0.775 \text{ GeV}/c^2$ ,  $M_\pi$  is the mass of the  $\pi^\pm$ ,  $\Gamma_0$  is  $0.149 \text{ GeV}/c^2$ ,  $\Gamma_2$  is the branching ratio of  $\rho^0 \rightarrow \mu^+ + \mu^-$ ,  $PS(p_T)$  is the Boltzmann phase space factor, and  $T$  is 160 MeV. The relativistic Breit-Wigner form depends on the particle's momentum through the

phase space factor (PS). The hadronic cocktail code evaluates the phase space factor at the mean- $p_T$  of the  $\rho$  instead of computing the phase-space for all values of  $p_T$ . Since the  $p_T$  range of the  $\rho$  in this analysis is limited by the MTD acceptance, the affect of this approximation is negligible.

The Kroll-Wada line shape is used for the Dalitz decays of the  $\eta$ ,  $\omega$ , and the  $\eta'$  [159,160]. The Kroll-Wada form consists of three distinct components: the QED factor, the phase space factor, and the transition form factor. The QED factor is given by:

$$QED = \sqrt{1 - \frac{4M_\mu^2}{M_{\mu\mu}^2}} \left(1 + \frac{2M_\mu^2}{M_{\mu\mu}^2}\right) \frac{1}{M_{\mu\mu}} \quad (5.54)$$

which is only a function of the mass of the lepton in question, the  $\mu$  in this case. The Phase Space term for a Dalitz decay with a massive neutral daughter is:

$$PS = \left[ \left(1 + \frac{M_{\mu\mu}^2}{M_h^2 - M_n^2}\right)^2 - \frac{4M_h^2 M_{\mu\mu}^2}{(M_h^2 - M_n^2)} \right]^{3/2} \quad (5.55)$$

where  $M_h$  is the mass of the parent hadron undergoing decay, and  $M_n$  is the mass of the neutral decay product e.g.  $M_n = M_{\pi^0}$  in the case of the  $\omega \rightarrow \pi^0 + \mu^+ + \mu^-$  Dalitz decay. If the neutral daughter particle is massless, e.g. in the case of  $\eta \rightarrow \gamma + \mu^+ + \mu^-$ , the phase space term above simplifies to:

$$PS = \left(1 - \frac{M_{\mu\mu}^2}{M_h^2}\right)^3 \quad (5.56)$$

The transition form factor is :

$$|F(M_{\mu\mu}^2)|^2 = \frac{1}{(1 - M_{\mu\mu}^2 \Lambda^{-2})^2 + \Gamma_0^2 \Lambda^{-2}} \quad (5.57)$$

where  $\Gamma_0$  and  $\Lambda^{-2}$  are found in Table 5.4. For more information on the measurements of the transition form factors please see Ref. [160–163].

The full Kroll-Wada form is given by the combination of these three components:

$$dN/dM_{ll} \propto \sqrt{1 - \frac{4M_l^2}{M_{ll}^2}} \left(1 + \frac{2M_l^2}{M_{ll}^2}\right) \frac{1}{M_{ll}} \left[ \left(1 + \frac{M_{ll}^2}{M_0^2 - M_{\pi^0}^2}\right)^2 - \frac{4M_0^2 M_{ll}^2}{(M_0^2 - M_{\pi^0}^2)} \right]^{3/2} \frac{1}{(1 - M_{ll}^2 \Lambda^{-2})^2 + \Gamma_0^2 \Lambda^{-2}} \quad (5.58)$$

### 5.6.3 Contributions from $c\bar{c}$ , $b\bar{b}$ , and Drell-Yan

```
pythia->SetMSEL(4);
pythia->SetPARP(91,1.0); //<kT>
pythia->SetPARP(67,1.0); //mstp*4
```

**Listing 3:** The parameters used to instantiate PYTHIA for simulation of the correlated  $c\bar{c} \rightarrow \mu^+ + \mu^-$  decays. The trigger  $MSEL = 1$  was also used to check the line shape and magnitude of the  $c\bar{c}$  contribution.

The intermediate mass region between the  $\phi$  and  $J/\psi$  mesons is dominated by the semileptonic decay of open heavy flavor. The semileptonic decay of  $c\bar{c}$  and  $b\bar{b}$  pairs, correlated through flavor conservation, result in correlated final state leptons that appear as signal in the dilepton invariant mass spectra. Additional contributions of correlated lepton pairs arise from the Drell-Yan process. The PYTHIA version 6.416 event generator is used to compute the contributions from correlated open heavy flavor decays and Drell-Yan at leading order (LO). The default Parton Distribution Functions (PDFs) used in the version 6.416 of PYTHIA are the updated CTEQ5L PDFs.

```
pythia->SetMSEL(5);
pythia->SetPARP(91,1.0); //<kT>
pythia->SetPARP(67,1.0); //mstp*4
```

**Listing 4:** The parameters used to instantiate PYTHIA for simulation of the correlated  $b\bar{b} \rightarrow \mu^+ + \mu^-$  decays.

The  $c\bar{c}$ ,  $b\bar{b}$ , and Drell-Yan processes are each simulated separately using different PYTHIA settings. The settings and parameter tunes used for the simulation of  $c\bar{c}$ ,  $b\bar{b}$ , and Drell-Yan are found in Listings 3, 4, and 5 respectively. For  $c\bar{c}$  and  $b\bar{b}$  events were selected that had exactly 2 strings in the event resulting from the desired heavy flavor quark. Events which decayed semileptonically into  $\mu^+\mu^-$  pairs were counted with weights according to the decay chain responsible. Drell-Yan events were accepted from those in which a  $q\bar{q} \rightarrow \gamma^*/Z^0 \rightarrow \mu^+ + \mu^-$  decay was present. The tune parameters for the Drell-Yan processes are chosen to match



measurements as described in Ref. [84]. The cross-sections used to scale the Drell-Yan,  $b\bar{b}$ , and  $c\bar{c}$  are shown in Table 5.10.

The contribution to the  $\mu^+\mu^-$  invariant mass spectra from  $c\bar{c}$ ,  $b\bar{b}$ , and Drell-Yan below  $M_{\mu\mu} \sim 2.2 \text{ GeV}/c^2$  is limited to only pairs with large  $p_T^{\mu\mu}$  due to the minimum transverse momentum cut-off of the MTD. This makes the contributions in the lower and intermediate mass regions very sensitive to the  $p_T$  shape used to describe the inclusive cross-section as a function of  $p_T$ . To put it in context, the total  $\sigma_{c\bar{c}}$  is already given a 20% uncertainty, so the uncertainty at higher  $p_T^{\mu\mu}$  must grow from there. The same is true of the  $b\bar{b}$  and Drell-Yan contributions, though some constraints used to guide the assignment of uncertainties are taken from Ref. [70, 84, 106, 107].

```

pythia->SetMSEL(11);
pythia->SetMSTP(43,1); // qqbar --> gamma* --> l+l-
pythia->SetCKIN(1,1.); // M parameter
pythia->SetMSTP(33,1); // common kt factor
pythia->SetMSTP(32,4); // Q^2 scale
pythia->SetMSTP(51,7); // switch to the CTEQ5L PDFs
pythia->SetPARP(31,1.8); // D->1.5
pythia->SetPARP(91,1.5); // <kt>
//switch off all decays
for(Int_t i=162; i<=189; i++) pythia->SetMDME(i,1,0);
// turn back on mu decays
pythia->SetMDME(171,1,1);
pythia->SetMDME(184,1,1);

```

**Listing 5:** The parameters used to instantiate PYTHIA for simulation of the Drell-Yan process,  $q\bar{q} \rightarrow \gamma^*/Z^0 \rightarrow \mu^+ + \mu^-$  decays.

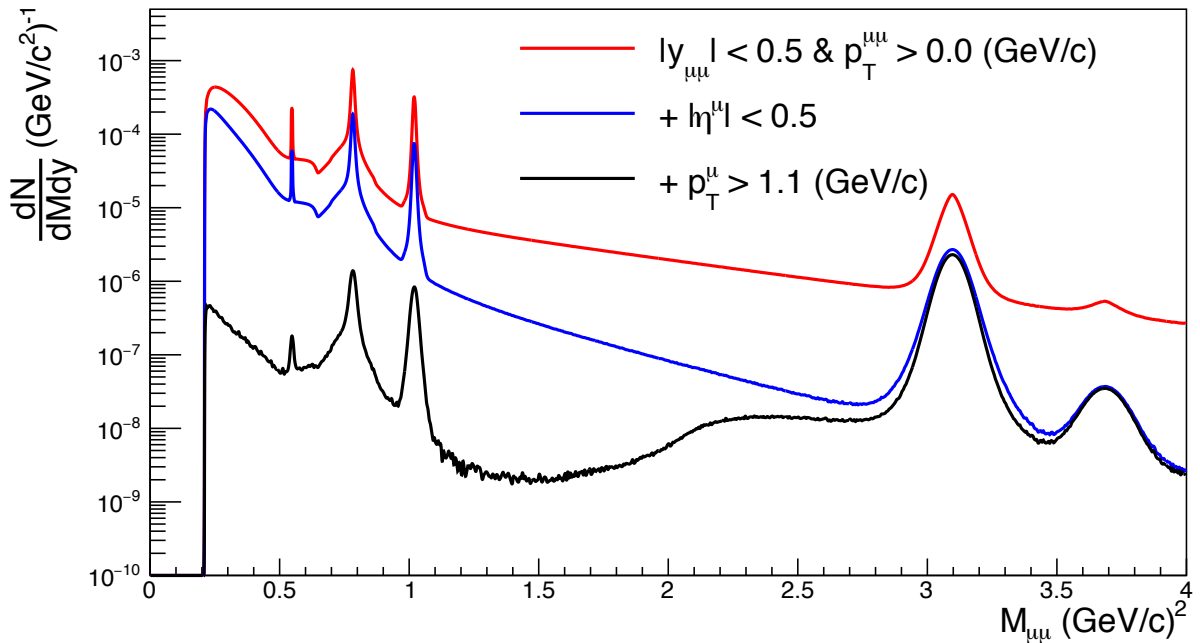
#### 5.6.4 Hadronic Cocktail for the MTD Acceptance

The kinematic acceptance of the MTD is taken into account when simulating the contributions for the hadronic cocktail. The kinematic cuts imposed by the MTD are:

- Pair rapidity:  $|y^{\mu\mu}| < 0.5$
- Daughter pseudo-rapidity:  $|\eta^\mu| < 0.5$
- Daughter  $p_T$ :  $p_T^\mu > 1.1 \text{ GeV}/c$

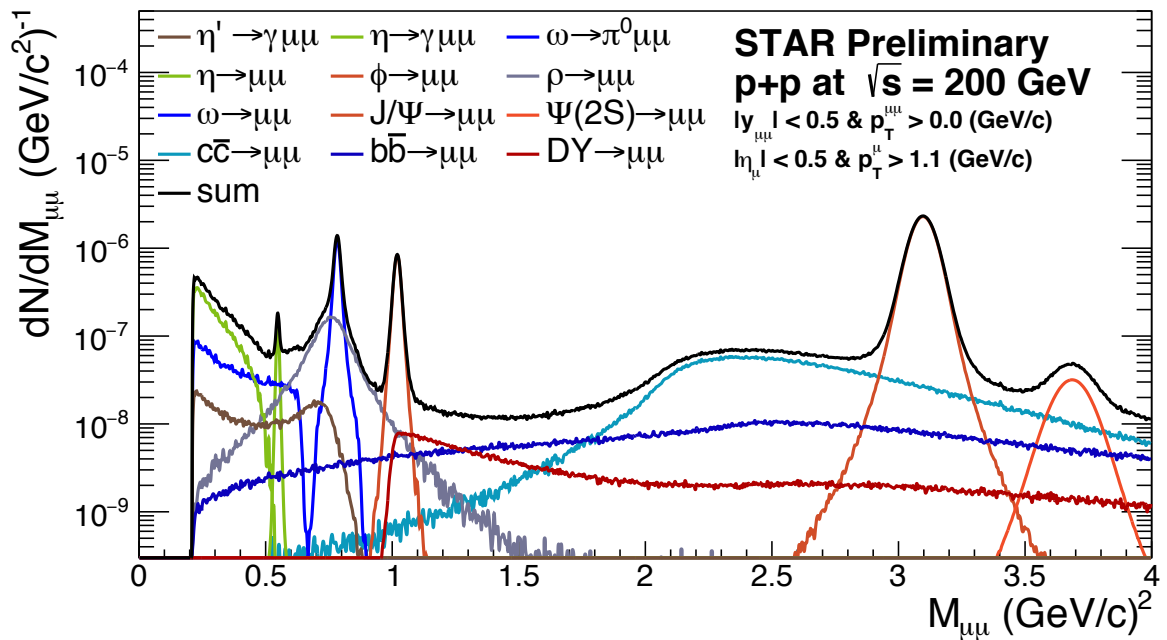
The affect of these three acceptance cuts can be seen on the sum of all hadronic cocktail components in Fig. 5.24. This figure clearly shows the harsh impact that the kinematic cuts imposed by the MTD have on the low mass region. The  $p_T^\mu > 1.1 \text{ GeV}/c$  requirement effectively rejects low mass hadrons with a  $p_T$  less than  $\sim 2.2 \text{ GeV}/c$ . The higher mass mesons like the  $J/\psi$  are not as affected by this requirement, since their masses are sufficient to provide the momentum kick to the daughters needed to reach  $p_T^\mu > 1.1 \text{ GeV}/c$ . Taking these kinematic requirements into account when generating the cocktail means that comparison with the data will depend mostly on efficiency corrections and not on gross acceptance corrections. The next section will discuss the method of determining the efficiency corrections to be applied to data. The hadronic cocktail enters this calculation, since the efficiency factors as a function of  $p_T^{\mu\mu}$  and  $M_{\mu\mu}$  depend on the kinematics of the underlying daughter particles.

In order to allow a one-to-one comparison with the measured  $\mu^+\mu^-$  invariant mass spectra, the cocktail needs to be smeared with a mass resolution comparable to that provided by STAR. The mass smearing is taken into account via momentum smearing of the daughter particles. The momentum resolution for tracks reconstructed in the TPC is relatively good ( $\sim 1\%$ ) for  $\mu$  tracks with a momentum of  $\sim 1 \text{ GeV}/c$ . The momentum resolution worsens almost linearly as a function of  $p_T$  reaching values of  $\sim 10\%$  at values of  $p_T > 7 \text{ GeV}/c$ . Since muons at these energies do not experience significant energy loss through Bremsstrahlung radiation, a Gaussian shape is used to smear the momentum. The quality of the momentum smearing can be judged through the correspondence between the  $J/\psi$  in the cocktail and data. Since the  $J/\psi$  width is so narrow ( $\Gamma_0 \approx 9.29 \times 10^{-5} \text{ GeV}/c^2$ ) its effective width is entirely dominated by the mass resolution and therefore momentum smearing of affect of the track reconstruction.



**Figure 5.24 :** The total hadronic cocktail for  $p + p$  collisions at  $\sqrt{s} = 200$  GeV for a progression of kinematic cuts. Requiring only that  $|y^{\mu\mu}| < 0.5$  results in the solid red curve. The addition of the  $|\eta^{\mu\mu}| < 0.5$  cut on each daughter  $\mu$  is shown in the solid blue curve. Finally, the addition of the  $p_T^{\mu\mu} > 1.1$  GeV/ $c$  cut is shown in the solid black curve.

After simulating each decay the full hadronic cocktail results from the properly scaled combination of all contributions. The  $dN/dy$  or  $\sigma$  values are used to scale each contribution and the final result is reported in terms of the yield / event. The hadronic cocktail for  $p + p$  collisions at  $\sqrt{s} = 200$  GeV Fig. 5.25. The cocktail for  $p + \text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV is identical to the  $p + p$  cocktail except that the yield of each contribution is also scaled by the average number of binary collisions per event  $\langle N_{col}^{p+Au} \rangle$ . The full hadronic cocktail for  $p + \text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV is obtained by scaling the hadronic cocktail from  $p + p$  collisions by  $\langle N_{coll} \rangle = 4.7$ . The full cocktail for  $p + \text{Au}$  collisions is shown in Fig. 6.2 in Chapter 6.



**Figure 5.25** : The hadronic cocktail for  $p + p$  collisions at  $\sqrt{s} = 200$  GeV. Each component is drawn separately. The sum of all contributions is drawn in the black solid curve. The hadronic cocktail for  $p + \text{Au}$  collisions is identical, except that the entire cocktail is scaled by the number of binary collisions,  $N_{col}$ .

## 5.7 Efficiency and Acceptance Corrections

### 5.7.1 Trigger Correction

The dimuon trigger records only those events which satisfy the trigger conditions. In order to compare with the hadronic cocktail, and express cross-sections in absolute terms, a conversion factor is needed to determine the total number of collisions (total luminosity) sampled by the trigger. This can be accomplished by relating the number of dimuon triggered events to the equivalent number of minimum bias (MB) events. Accomplishing this requires an understanding of the data recording process, the live-time of the detectors and the difference in efficiencies between the dimuon trigger and the MB triggers. Taking all of these factors into account results in an expression for the number of equivalent minimum bias events given by:

$$N_{\text{MB}}^{\text{equivalent}} = \sum_{\text{run}} N_{\text{MB,run}} \cdot \frac{\text{PS}_{\text{MB,run}}}{\text{PS}_{\text{dimuon,run}}} \cdot \frac{\text{LiveTime}_{\text{dimuon,run}}}{\text{LiveTime}_{\text{MB,run}}} \cdot \frac{N_{\text{dimuon,run}}^{\text{analyzed}}}{N_{\text{dimuon,run}}^{\text{recorded}}} \cdot \varepsilon_{\text{good}} \quad (5.59)$$

where:

- The sum is over all runs (data collection periods) collected and analyzed.
- $N_{\text{MB,run}}$  is the number of MB events in each run after applying the vertex and ranking cuts. It is calculated as the number of recorded MB events multiplied by the vertex cut efficiency measured in data from the same run.
- $\text{PS}_{\text{MB,run}}$  is the pre-scale of the minimum bias trigger for the given run. The pre-scale factor is used to alleviate bandwidth issues. Since the data acquisition system cannot read-out every event, a certain number of events that pass the trigger condition must be dropped in order to reduce the bandwidth load. In recent years, the STAR data acquisition system has been able to run at a rate of  $\sim 2.3$  kHz.
- $\text{PS}_{\text{dimuon,run}}$  is the pre-scale of the dimuon trigger for the given run.
- $\text{LiveTime}_{\text{dimuon,run}}$  is the live time for the dimuon trigger for a given run. The live-time refers to the fraction of time that the detectors were able to read out data and form events. Due to bandwidth limitations, the data acquisition system is not able to read

out every event. For this reason, if two events passing the dimuon trigger condition come too close together in time, the later event may be dropped due to the dead-time induced by reading out the first event.

- $\text{LiveTime}_{\text{MB,run}}$  is the live time for the minimum bias trigger for a given run.
- $N_{\text{dimuon,run}}^{\text{analyzed}}$  is the number of dimuon triggered events that are analyzed (pass event cuts) for a given run.
- $N_{\text{dimuon,run}}^{\text{recorded}}$  is the number of dimuon triggered events that are recorded for a given run.
- $\varepsilon_{\text{good}}$  is the fraction of good events measured by fitting the  $|z_{\text{TPC}} - z_{\text{VPD}}|$  distribution (more details below).

The  $N_{\text{MB}}^{\text{equivalent}}$  based on this relation are listed in Table 5.6. The result in the last row corresponds to the selection criteria used in these analyses. This number of minimum bias events is used in the comparison between the measured  $\mu^+\mu^-$  invariant mass spectra and the hadronic cocktail. It is also used to calculate the  $\text{BR} \times dN/dy$  of the  $\phi \rightarrow \mu^+ + \mu^-$ . Most of the variables listed in Eq. 5.59 are obtained from the meta-data recorded during the data collection process. For instance, the pre-scale and live-time of detectors is recorded during the data acquisition process. The number of recorded and analyzed events are trivial to calculate from the data.

The determination of  $\varepsilon_{\text{good}}$  is a little more involved however. The collider system is not perfect so various types of background collisions can occur due to beam imperfections that cause collisions with the beam-pipe or other material. If the trigger systems were perfect then background collisions would always be rejected. However, in practice each trigger condition has a different susceptibility to background. Specifically, the minimum bias trigger condition is found to be more susceptible to background events than the dimuon trigger condition. This difference needs to be taken into account when determining the correspondence between the number of dimuon triggered events and minimum bias events. This factor was estimated using  $z_{\text{TPC}} - z_{\text{VPD}}$ , i.e. the difference between the vertex as measured by the TPC (from reconstructed tracks) and the vertex position detectors (based on timing of hits). Background events may produce a large number of tracks but are more likely to be out-of-time resulting in a shifted vertex as reported by the vertex position detectors. On average, the dimuon trigger

**Table 5.6 :** The number of equivalent minimum bias events ( $N_{\text{MB}}^{\text{equivalent}}$ ) calculated for various different event selection cuts and efficiency factors. The result in the last row corresponds to the selection criteria used in these analyses.

Cuts	$N_{\text{Minimum Bias}}^{\text{equivalent}}$
$p + p$	
$z_{\text{TPC}} - z_{\text{VPD}} < 6 \text{ cm}$ and $z_{\text{TPC}} < 100 \text{ cm}$	4.375e11
$z_{\text{TPC}} - z_{\text{VPD}} < 6 \text{ cm}$ and $z_{\text{TPC}} < 100 \text{ cm}$ and ranking $\geq 0$	3.56e11
$z_{\text{TPC}} - z_{\text{VPD}} < 6 \text{ cm}$ and $z_{\text{TPC}} < 100 \text{ cm}$ and ranking $\geq 0$ and $\varepsilon_{\text{good}}$	3.382e11
$p+\text{Au}$	
$z_{\text{TPC}} - z_{\text{VPD}} < 6 \text{ cm}$ and $z_{\text{TPC}} < 100 \text{ cm}$ and ranking $\geq 0$ and $\varepsilon_{\text{good}}$	1.401e11

was found to provide  $\sim 95\%$  good event selection based on this method. As a cross-check the  $J/\psi$  yield per event was compared in MB events and dimuons triggered events to gauge the efficiency for selecting good events.

### 5.7.2 Tracking and MTD Efficiency

In order to compare the  $\mu^+\mu^-$  invariant mass measurement with the simulated hadronic cocktail the measurement must be corrected for inefficiency. The general strategy is to compute the single  $\mu^\pm$  efficiencies by embedding simulated tracks into real data events. Each detectors' response is simulated for the embedded tracks and added to the information read-out for real data. After combining the real data and simulated data, the full event is reconstructed using the normal track reconstruction and event building software. The efficiencies can then be computed as a function of kinematic variables by comparing the number of correctly reconstructed tracks to the number originally embedded. After the single particle efficiencies have been measured, the pair efficiencies can be determined by a folding procedure using the single particle efficiencies. First, the individual components that make up the total efficiency will be discussed. Then the folding procedure will be presented along with the final efficiency corrections.

The efficiency for reconstructing tracks from a signal pair with the TPC and having them matched to MTD hits can be separated into multiple contributions:

$$\varepsilon^{\text{total}} = Acc^{\text{MTD}} \cdot \varepsilon^{\text{TPC}} \cdot \varepsilon^{\text{MTD Matching}} \cdot \varepsilon^{\text{MTD Electronics}} \cdot \varepsilon^{\text{Trig}} \cdot \varepsilon^{\text{Trigger Unit}} \cdot \varepsilon^{\text{PID}} \quad (5.60)$$

where  $Acc^{\text{MTD}}$  is the kinematic acceptance factor for the MTD,  $\varepsilon^{\text{TPC}}$  is the TPC track reconstruction efficiency,  $\varepsilon^{\text{MTD Matching}}$  is the MTD matching efficiency,  $\varepsilon^{\text{MTD Electronics}}$  is the MTD electronics' response efficiency,  $\varepsilon^{\text{Trig}}$  is the efficiency of the trigger timing-window cut,  $\varepsilon^{\text{Trigger Unit}}$  is the efficiency of the MTD's trigger unit system, and  $\varepsilon^{\text{PID}}$  is the efficiency of the muon identification. All of these efficiencies (unless otherwise noted) must be computed as a function of the track's  $p_T$ ,  $\eta$ , and  $\phi$  in order to incorporate the complex geometry of the MTD. Each of these efficiencies are discussed in detail below.

As mentioned in Sec. 5.6, the gross acceptance of the TPC and MTD are already factored out when generating the hadronic cocktail. Therefore, they do not need to be corrected for a comparison between the  $\mu^+\mu^-$  invariant mass spectra and the hadronic cocktail. The acceptance corrections are needed for the measurement of the  $\text{BR} \times dN/dy$  of the  $\phi$  meson however. Since the TPC covers  $\sim 2\pi$  in azimuth,  $|\eta| < 1$ , and  $p_T > 0.2 \text{ GeV}/c$  the acceptance is determined by the kinematic acceptance of the MTD. The acceptance for the  $\phi$  with  $|y_\phi| < 0.5$  is measured using the simulation code developed for the hadronic cocktail. First,  $\phi$  mesons are generated according to a flat azimuthal distribution, a CERES rapidity distribution and the Tsallis Blast-Wave  $p_T$  distribution. These  $\phi$  mesons are then decayed into  $\mu^+\mu^-$  pairs. The kinematic requirements of the MTD are then applied to the daughter tracks. The acceptance factor can be expressed as:

$$Acc^\phi = \frac{|y^\phi| < 0.5 \ \& \ |\eta_{1,2}^\mu| < 0.5 \ \& \ p_{T1,2}^\mu > 1.1 \text{ GeV}/c}{|y_\phi| < 0.5} \quad (5.61)$$

The acceptance factor for the  $\phi$  can be seen as a function of its  $p_T$  in Fig. 5.27a. The acceptance factor for the  $J/\psi$  can be seen as a function of its  $p_T$  in Fig. 5.27b. The additional acceptance due to the MTD's limited coverage in the azimuthal direction is combined with its efficiency corrections.

The TPC track reconstruction efficiency is a function of several factors. Naturally, it depends on the track quality, with higher quality tracks (e.g. requiring them to have more hit points) resulting in lower efficiency. The track reconstruction efficiency also depends on multiplicity ranging from better than  $\sim 90\%$  in low multiplicity  $p + p$  events to  $\sim 70\%$  in central Au+Au collisions. While the dimuon triggered events are slightly biased to higher multiplicities than minimum bias events, for tracking purposes the multiplicities are still



extremely low. The TPC track reconstruction efficiency is measured in embedding with:

$$\varepsilon^{\text{TPC}} = \frac{N^{\text{TPC}}}{N^{\text{MC}}} \quad (5.62)$$

where  $N^{\text{MC}}$  is the number of simulated tracks embedded and  $N^{\text{TPC}}$  is the number of simulated tracks properly reconstructed from TPC information. As already mentioned, the total efficiency of the MTD includes several components. The MTD matching algorithm is responsible for matching tracks reconstructed in the TPC with hits in the MTD. The efficiency of the MTD matching ( $\varepsilon^{\text{MTD matching}}$ ) is given by:

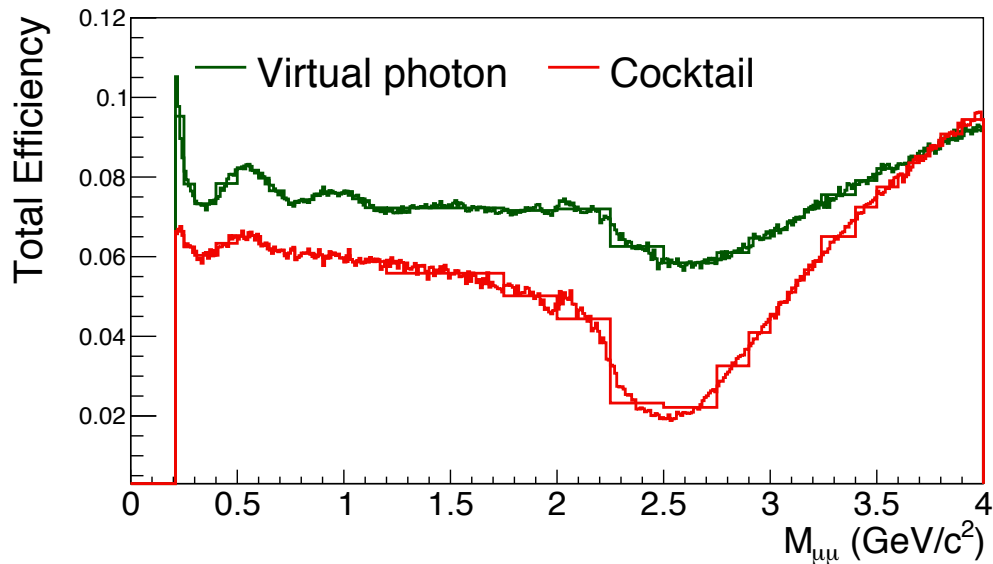
$$\varepsilon^{\text{MTD Matching}} = \frac{N^{\text{MTD Matched}}}{N^{\text{TPC}}} \quad (5.63)$$

where  $N^{\text{MTD matched}}$  are the number of simulated tracks that are reconstructed in the TPC and matched to the correct hit in the MTD. In this case it is implied, in order to not double count factors, that  $N^{\text{TPC}}$  is the number of TPC reconstructed tracks that also fall within the gross acceptance of the MTD. In the simulation software, the active regions of the MTD have a 100% efficiency for producing a hit. However, the real hardware has finite sensitivity and electronic thresholds. For this reason, the response efficiency of the MTD is extracted from cosmic ray data. The efficiency of the MTD electronics can be isolated by using cosmic ray measurements and embedded tracks as:

$$\varepsilon^{\text{MTD electronics}} = \frac{\varepsilon_{\text{cosmics}}^{\text{MTD Matching}}}{\varepsilon_{\text{simulation}}^{\text{MTD Matching}}} \quad (5.64)$$

where  $\varepsilon_{\text{cosmics}}^{\text{MTD Matching}}$  is the MTD matching efficiency defined above extracted from cosmic ray data and  $\varepsilon_{\text{simulation}}^{\text{MTD Matching}}$  is the MTD matching efficiency from simulation where the electronics efficiency is 100%. The dimuon trigger uses a cut on the time of hits to reject background and punch through from slow hadrons at the level of triggering. However, some MTD hits from real muons may not have times that fall into the pre-defined window. These muons will be rejected by the trigger and result in an inefficiency. This affect can be measured using the  $J/\psi$  via:

$$\varepsilon^{\text{Trig}} = \frac{N_{\text{triggered}}^{J/\psi}}{N_{\text{matched}}^{J/\psi}} \quad (5.65)$$



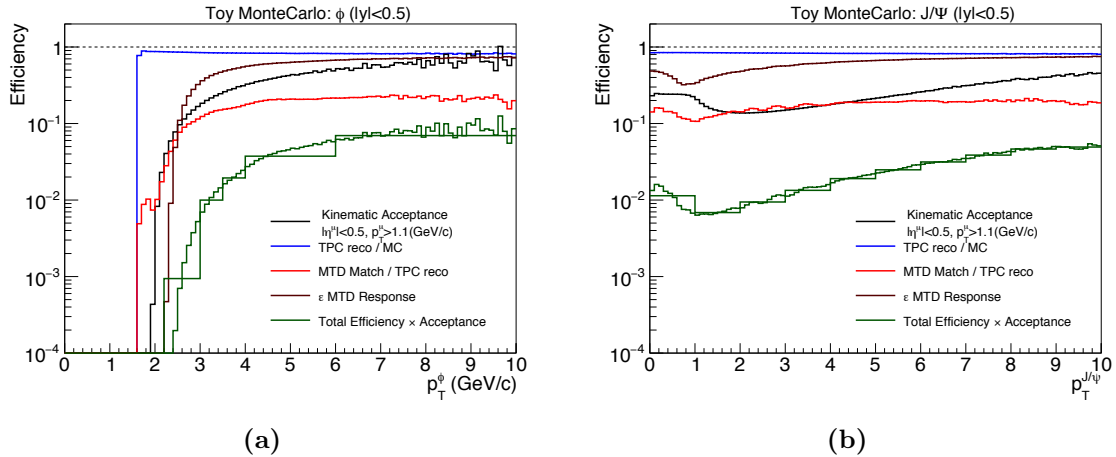
**Figure 5.26 :** The pair efficiency as a function of  $M_{\mu\mu}$  for two scenarios: 1) A  $\gamma^*$  decay with flat mass and  $p_T$  and 2) using the  $M_{\mu\mu}$  and  $p_T^{\mu\mu}$  distributions from the total hadronic cocktail. Note, this includes only the efficiency and acceptance affects for tracks that fall within the gross kinematic acceptance of the MTD (i.e.  $|\eta| < 0.5$  and  $p_T > 1.1$  GeV/ $c$ ).

where  $N_{triggered}^{J/\psi}$  is the number of  $J/\psi$  that are reconstructed from muons passing the trigger conditions and  $N_{matched}^{J/\psi}$  are the number of  $J/\psi$  reconstructed from all MTD matched muon candidates regardless of whether they passed the trigger condition or not. The efficiency of the trigger condition is found to be very high,  $\varepsilon^{Trig} \approx 99\%$  and is taken to be flat since it has very little dependence on the muon's kinematics.

The above efficiencies are computed for individual muons and folded together to determine the pair efficiency as a function of  $M_{\mu\mu}$  and  $p_T^{\mu\mu}$ . This procedure consists of simulating a virtual photon decay to  $\mu^+\mu^-$ . The pair efficiency is considered to be the simple product of the single track efficiencies, i.e:

$$\varepsilon^{pair} = \varepsilon^A(p_T, \eta, \phi) \times \varepsilon^B(p_T, \eta, \phi) \quad (5.66)$$

where  $\varepsilon^A$  and  $\varepsilon^B$  are the efficiencies of the two daughter tracks. Since the pair efficiency is generated from a virtual photon with a flat  $p_T^{\mu\mu}$  distribution the final efficiency as a function of only  $M_{\mu\mu}$  is generated by taking the weighted average over the pair  $p_T$ . The  $p_T$  spectra from the hadronic cocktail is used for the  $p_T^{\mu\mu}$  weighting. The pair efficiency as a function of  $M_{\mu\mu}$  are shown in Fig. 5.26 for two different scenarios: 1) A  $\gamma^*$  decay with flat mass and  $p_T$



**Figure 5.27 :** The total efficiency  $\times$  acceptance for the  $\phi$  (a) and the  $J/\psi$  (b). Each component is shown in addition to the total. The total efficiency  $\times$  acceptance is shown in small bins and in the larger bins used to extract signal.

and 2) using the  $M_{\mu\mu}$  and  $p_T^{\mu\mu}$  distributions from the total hadronic cocktail.

Not all of the pair efficiency factors can be adequately separated into single muon efficiency factors. The MTD trigger unit system is one such case. In order for an event to pass the dimuon trigger and be recorded it must fire two separate trigger units. The requirement that the event fire two separate trigger units imposes an implicit bias on the separation (opening angle) of a pair. For example, a high mass pair ( $M_{\mu\mu} > 3$  GeV/c) with low transverse momentum will produce a  $\mu^+\mu^-$  pair that is nearly back-to-back in the laboratory frame. The daughter tracks from this type of pair will necessarily fall into two separate trigger units and pass the dimuon trigger. However, a low-mass pair with only just enough momentum to produce daughters with  $p_T > 1.1$  GeV/c will produce pairs with a very small opening angle. While the trajectories of the daughter particles will separate somewhat before reaching the MTD, they are much more likely to fall within a single trigger unit and therefore be rejected by the dimuon trigger.

The trigger unit efficiency was computed using the mixed event framework and checked with a sample of simulated  $J/\psi$ . Using mixed-event pairs for this calculation is ideal since the mixed-event pairs can fall into any possible combination of trigger units. The trigger unit efficiency ( $\epsilon^{\text{Trigger Unit}}$ ) is determined by computing the ratio of pairs with daughters that hit

two different trigger units compared to all pairs. That is:

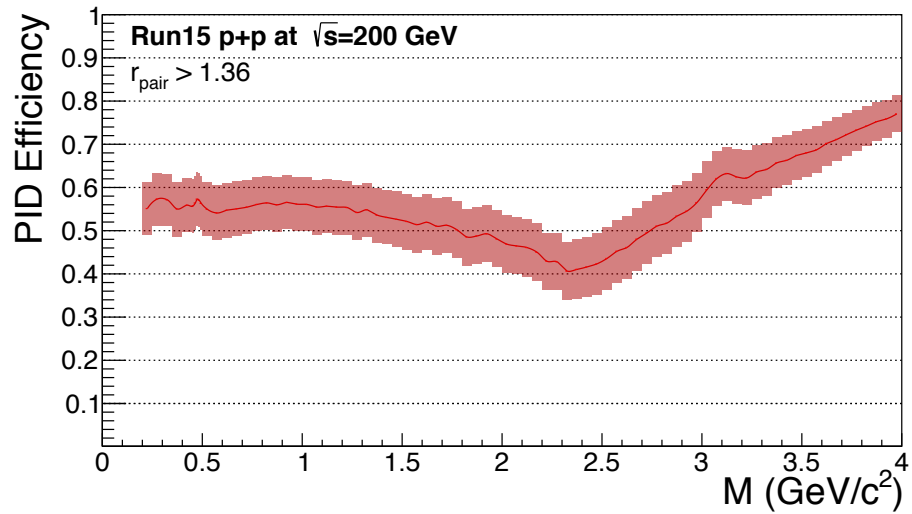
$$\varepsilon^{\text{Trigger Unit}} = \frac{N_{\text{pass}}^{\text{mixed-event}}}{N^{\text{mixed-event}}} \quad (5.67)$$

where  $N_{\text{pass}}^{\text{mixed-event}}$  is the number of mixed-event pairs with daughters that fall into two different trigger units (passing the trigger unit requirement) and  $N^{\text{mixed-event}}$  is all mixed-event pairs including those with daughters that hit the same trigger unit. This efficiency is computed as a function of the pair opening angle and the pair  $p_T$  to fully incorporate the behavior of this effect. The cross-check computation of  $\varepsilon^{\text{Trigger Unit}}$  using the sample of simulated  $J/\psi$  showed an agreement with the mixed-event method at the sub 1% level.

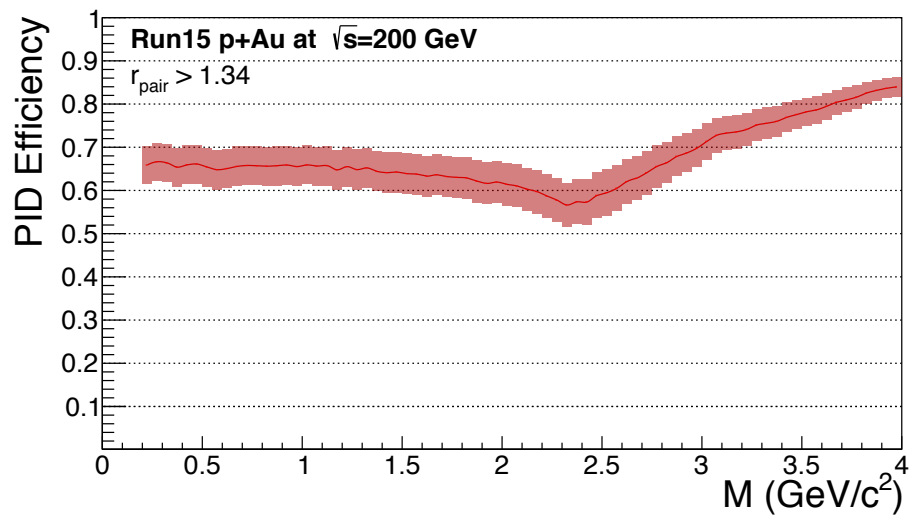
The muon identification is also computed at the pair level since the value of  $r_{\text{pair}}$  is used to select pairs instead of the individual DNN responses for each track. The efficiency for a signal pair is trivial to calculate with the pair templates produced in Sec. 5.5.1. With the signal pair template already computed the PID efficiency is simply:

$$\varepsilon^{\text{PID}}(x) = \frac{\int_x T_{\text{signal}}^{\text{pair}}(r_{\text{pair}}) dr_{\text{pair}}}{\int T_{\text{signal}}^{\text{pair}}(r_{\text{pair}}) dr_{\text{pair}}} \quad (5.68)$$

which gives the signal pair identification efficiency for a cut of  $r_{\text{pair}} > x$ . See Fig. 5.29 for an example showing the signal pair efficiency, purity, S/B, and significance of pairs in the  $J/\psi$  mass window.

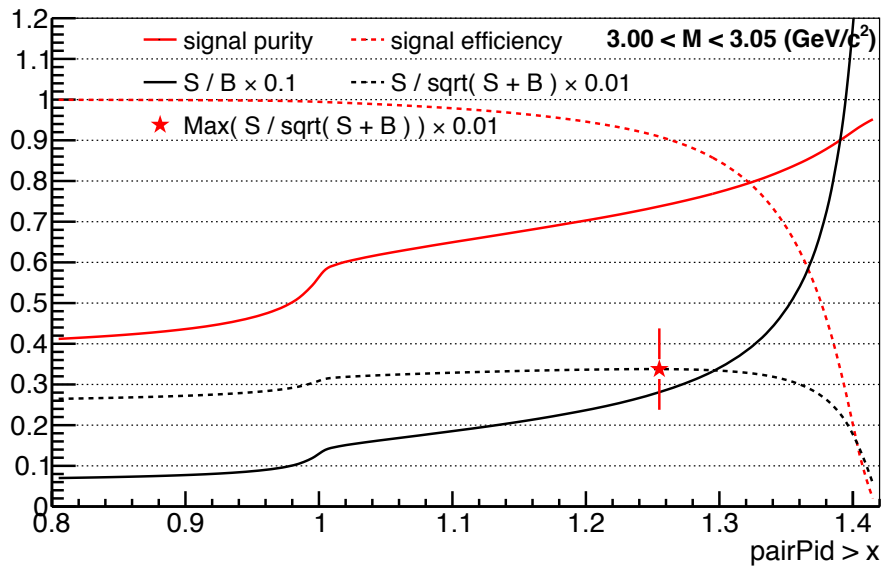


(a)



(b)

**Figure 5.28** : The PID efficiency for  $\mu^+\mu^-$  pairs in  $p + p$  (a) and  $p+\text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV selected with  $r_{\text{pair}} > 1.36$  and  $r_{\text{pair}} > 1.34$  for  $p + p$  and  $p+\text{Au}$  respectively.



**Figure 5.29** : The signal efficiency and purity determined as a function of  $r_{pair}$ . The S/B and significance are also shown with the cut value for maximum significance highlighted.

## 5.8 Systematic Uncertainties

### 5.8.1 Uncertainties on the $\mu^+\mu^-$ Invariant Mass Spectra

The measurement of the  $\mu^+\mu^-$  invariant mass spectra, as described in this chapter, is a complex process including many steps. Each of the assumptions in the procedure may contribute to a bias in the final result. Some parts of the procedure depend on simulations which may not perfectly correspond to the real process they simulate. In this section, a systematic study of the various assumptions and uncertainties inherent in the  $\mu^+\mu^-$  measurement procedure is presented. The goal is to determine the likely range of values which could result if the assumptions and uncertainties in the procedure are changed in reasonable ways. The various sources of systematic uncertainty will be combined and reported in the final result separately from the statistical uncertainties. This practice is common in our field since it helps determine if more data or a better approach (better techniques, improved equipment, different experimental procedure etc.) are needed to produce a higher quality measurement.

Table 5.7 lists the sources of uncertainty that were considered and evaluated in this study. Most of the uncertainties that have been considered are similar between the  $p+p$  and  $p+\text{Au}$  data sets. Therefore, the uncertainties and the technique for evaluating them will be discussed for the  $p+p$  and  $p+\text{Au}$  datasets together, only making special note whenever there is a significant difference between the two cases.

The first set of uncertainties relate to the efficiency and acceptance correction procedure. The single particle efficiency tables were produced using an embedding technique in which simulated tracks are reconstructed alongside real tracks. Often the simulated track distributions (like nHitsFit, DCA, etc.) are not in perfect agreement with the distributions from real tracks. Often, but not always, the simulated tracks are higher quality than real tracks. For instance, the DCA distribution tends to be too narrow, not too wide and the nHitsFit distribution tends to be shifted towards slightly more hits than in data. These differences are taken into account with a conservative 5% uncertainty on the single muon tracking efficiency correction. Likewise, the MTD matching efficiency and the MTD response efficiency are assigned a 5% uncertainty. The single muon trigger efficiency, computed from the data using  $J/\psi$  muons is assigned a 1% uncertainty. These single muon efficiency uncertainties contribute very little deviation to the final measured yield.

The efficiency affects that are more significant relate to the pair efficiency folding pro-

**Table 5.7 :** Sources of systematic uncertainty on the measurement of the  $\mu^+\mu^-$  invariant mass spectra in  $p + p$  and  $p+Au$  collisions at  $\sqrt{s_{NN}} = 200$  GeV.

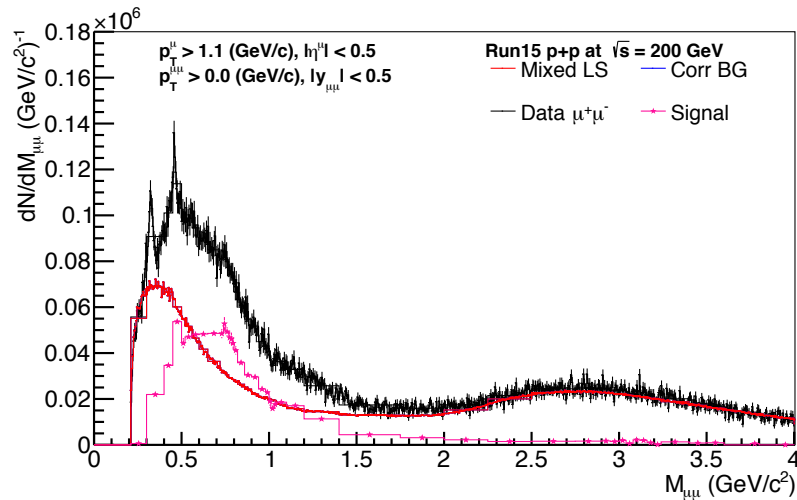
Category	Source Uncertainty	Relative Uncertainty on Yield	
		$p + p$	$p + Au$
Single $\mu$ Tracking Efficiency	5%	1-2%	1-2%
Single $\mu$ MTD Matching Efficiency	5%	1-2%	1-2%
Single $\mu$ MTD Response Efficiency	5%	1-2%	1-2%
Single $\mu$ MTD Trigger Efficiency	2%	<1%	<1%
MTD Trigger Unit Efficiency	50%	40% ( $M_{\mu\mu} < 0.5$ )	45% ( $M_{\mu\mu} < 0.5$ )
Pair Efficiency Folding	Hadronic Cock- tail vs. $\gamma^*$	5-18%	5-20%
Pair PID Efficiency	10%	10%	15%
Total Efficiency $\times$ Acceptance	5-20%	5-25%	15%
Normalization ( $N_{MB}^{equivalent}$ )	5% ( $p + p$ ), 10% ( $p+Au$ )	5%	10%
$r_{\text{pair}}$ Mass Binning	[0.1, 0.05, 0.2] (GeV/ $c^2$ )	<1%	<1%
$r_{\text{pair}}$ Fit Range	$r_{\text{pair}} >$ [0.0, 0.1, 0.2]	6%	6%
$r_{\text{pair}}$ Fit Uncertainty	1.5-3%	3%	3%
DNN Template Uncertainty in LSR method	$\pm 20\%$ ( $p + p$ ), $\pm 30\%$ ( $p+Au$ )	5-40%	5-45%
Event Mixing $p_T^{\mu\mu}$ binning	[0.05, 0.1, 0.2, 0.5]	<1%	<1%
Event Mixing opening angle binning	[0.05, 0.1, 0.2, 0.5]	10% (20–50%, $M_{\mu\mu} < 0.5$ )	10% (20–50%, $M_{\mu\mu} < 0.5$ )
Event Mixing Buffer Size	[all, 9, 36, 81]	1%	1%
Hadron Contamination	1 - purity $\approx 30\%$	30%	40%



cedure, the PID efficiency, and the MTD trigger unit efficiency. The pair efficiencies are computed by folding together the single track efficiencies. The pair efficiency expressed only as a function of  $M_{\mu\mu}$  in this analysis is heavily affected by the  $p_T$  distribution of the pair's daughter particles. The dependence on this effect is investigated by comparing the hadronic cocktail result with the result from virtual photon decay with a flat  $p_T^{\mu\mu}$  distribution. An  $\approx 4\%$  difference between these two is observed in the intermediate mass region with less than 2% difference in other regions of phase space. This in turn results in a 10–15% difference in the yield.

The PID efficiency is computed from the signal template shapes. If the simulation has any systematic bias compared to the data, it will be propagated into the PID efficiency by virtue of a biased template shape. The uncertainty on the template was determined using a tag-and-probe technique with the  $J/\psi$  muons. The PID efficiency was specifically tested by measuring the change in  $J/\psi$  yield as the DNN-based PID cut on one daughter was tightened. In this way the uncertainty on the PID efficiency computed from the templates (averaged over track  $p_T$ ) was found to be  $\sim 10\%$ . Since this efficiency is applied directly at the pair level it has a direct impact on the measured yield. Since the  $\Delta$  TOF distributions were extracted from  $p + p$  data and reused in the  $p$ +Au data, the  $p$ +Au uncertainty is assigned a slightly larger value. A mixed event technique was used to compute the MTD trigger unit efficiency. The trigger unit correction has the largest affects for pairs with a mass just above threshold ( $M_{\mu\mu} < 0.5 \text{ GeV}/c^2$ ). Higher mass pairs that are extremely boosted (e.g.  $J/\psi$  with  $p_T > 10 \text{ GeV}/c$ ) are also affected, since their daughter particles are emitted with a small opening angle in the lab frame. A total uncertainty of 5–20% results from all efficiency and acceptance affects combined with the largest uncertainties found in the intermediate mass region.

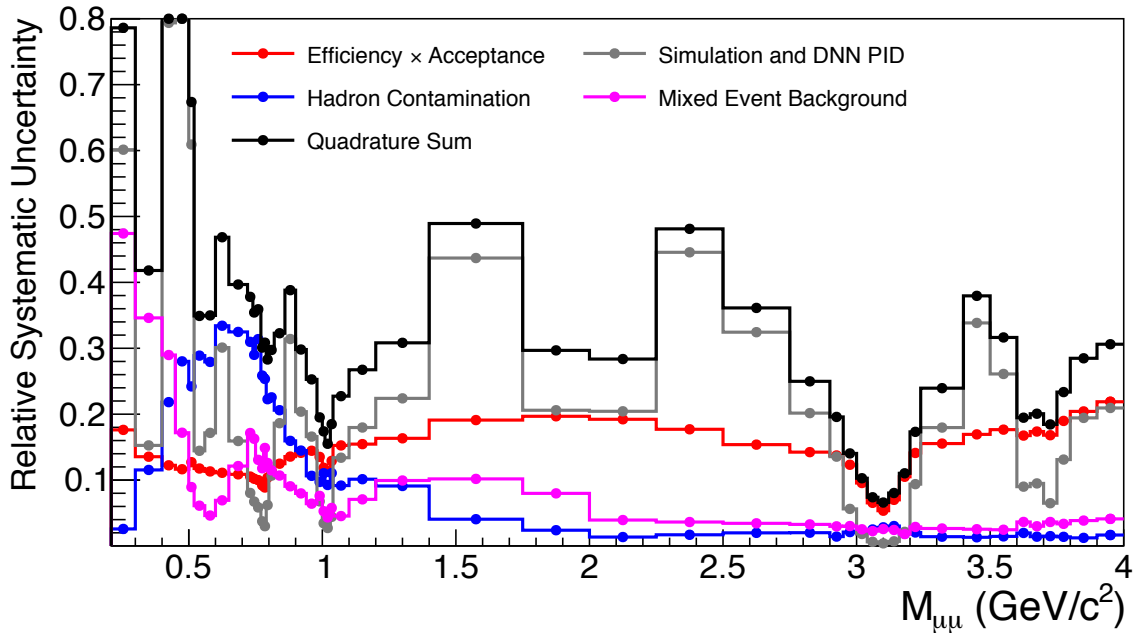
The final mass spectra is reported as the dimuon yield per event. The total number of min-bias events,  $N_{MB}^{equivalent}$ , is used for the normalization. The uncertainty in this value results almost entirely from the variation in performance of the VPD in the two trigger conditions. In the  $p$ +Au dataset the uncertainty is larger partly due to a broader VPD  $v_z$  distribution. An additional uncertainty results from the bias towards higher multiplicities in dimuon triggered events compared to minimum bias triggered events. All of the effects combined contribute the 10% uncertainty on the normalization in  $p$ +Au.



**Figure 5.30** : The invariant mass distribution of the correlated pairs from hadron contamination. The “signal” is extracted from the background only region  $r_{\text{pair}} < 0.9$ . Notice, for instance, the lack of any sign of a  $J/\psi$  peak and the clear  $K_S^0 \rightarrow \pi^+\pi^-$  and  $\phi \rightarrow K^+K^-$  peaks - suggesting that the excess  $\mu^+\mu^-$  pairs are a result of correlated background.

Several sources of uncertainty related to the like-sign ratio background estimation technique are also considered. The effect of changing the mass binning and the  $r_{\text{pair}}$  fit range incurs a negligible impact on the final yield. Additionally, the uncertainty from the fit, computed with the MINOS algorithm is also negligible. However, a variation in the template shapes contributes to a modified correction factor  $R(\text{mass}; \text{signal})$  and can lead to a significant variation in the yield. Since this uncertainty affects the background before being subtracted, it contributes the largest single systematic uncertainty of all sources. This uncertainty takes into account the possible disagreement between the data and the simulation used to train the DNN. Since a change in the DNN template shape induces a change in the  $R$  factor, this uncertainty also represents the maximum possible affect on the yield due to an uncertainty on the background correction factor.

The correlation weighted event mixing technique provides a statistically stable equivalent of the foreground like-sign distribution. In order to test the final yield’s sensitivity to this procedure, several parameters in the event-mixing were varied. The size of the buffer used for track mixing was varied to test the affect of the buffer capacity. Since the total number of tracks is not too large, the nominal setting uses an unlimited buffer size. The final yield is found to be insensitive to the change in the track buffer size. None of the variations to the nominal event mixing technique incurred a large change in the yield except for the binning



**Figure 5.31** : The Systematic uncertainties for the  $\mu^+\mu^-$  invariant mass spectra in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV.

of the opening angle used for the correlation weighting. With bins that are too fine, the mixed-event distribution becomes susceptible to the statistical fluctuations in the original same-event like-sign distribution. If the bins are too large some small angle features are washed out. These effects result in  $\sim 10\%$  uncertainty in the yield but only for low mass pairs which decay with small opening angles (in the lab frame).

Finally, the effect of hadron contamination on the yield is considered. The shape of the hadron contamination uncertainty is obtained from the invariant mass spectra of background pairs ( $r_{\text{pair}} < 0.9$ ). Fig. 5.30 shows that the correlated pairs from hadron contamination are mostly localized to  $M_{\mu\mu} < 1.5$  GeV/ $c^2$ . The magnitude of the shape is normalized by the average impurity ( $1 - \text{purity}$ ). The most significant sources of systematic uncertainty are shown in Fig. 5.31 as a function of  $M_{\mu\mu}$  for the  $p + p$  dataset. The total systematic uncertainty in each mass bin is calculated by summing the components in quadrature.

### 5.8.2 Uncertainty on the $\phi \rightarrow \mu^+ + \mu^-$ Yield

The  $\phi$  yield is extracted via a maximum likelihood fit of the raw  $\mu^+\mu^-$  invariant mass distribution. The final yield is sensitive to several elements in the yield extraction procedure.

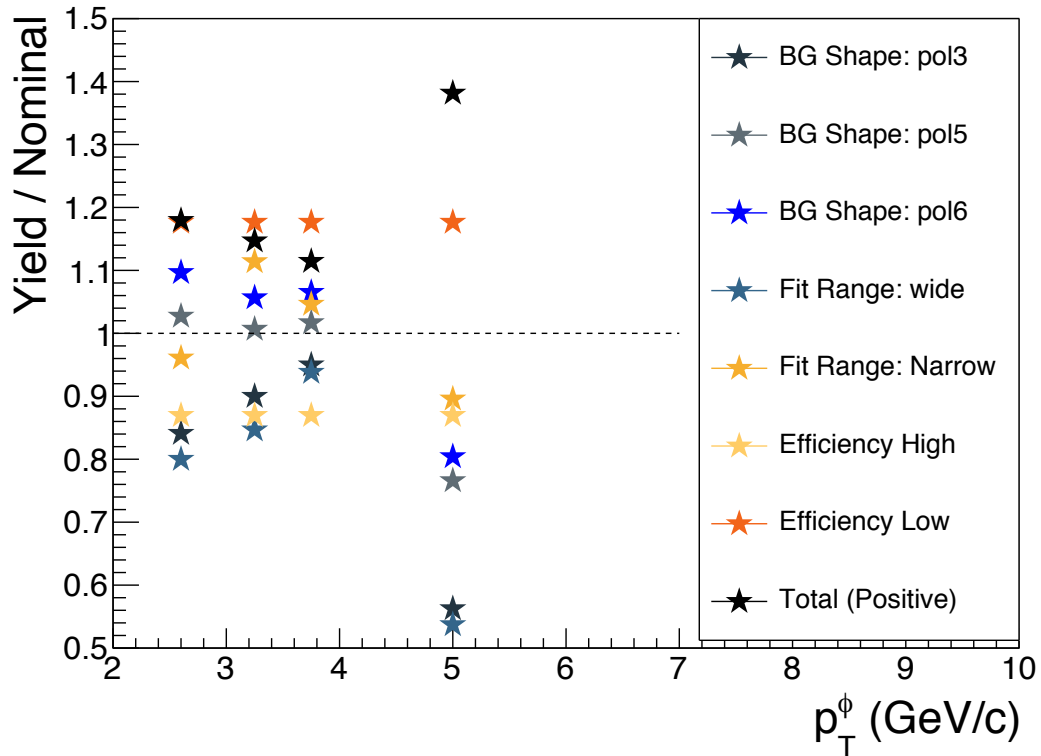
**Table 5.8** : Sources of systematic uncertainty for the  $\phi$  yield extraction.

Category	Source	Relative Uncertainty on Yield
Efficiency	Total	17%
	Tracking Efficiency	2%
	MTD Matching Efficiency	2%
	MTD Response Efficiency	2%
	PID Efficiency	10%
	Pair Efficiency	2.5%
	MTD Trigger Efficiency	1%
	MTD Trigger Unit Efficiency	1%
Normalization	$N_{MB}^{equivalent}$	5%
Background	pol4, pol5, pol6	20–40%
Shape		
Mass Fit Range	narrow, wide	5–50%
Fit Uncertainty	Fit Uncertainty	<2%

Table 5.8 lists the relevant uncertainties with respect to the  $\phi$  yield extraction procedure. The uncertainties are broken in several different categories. The first category relates to the efficiency corrections. The efficiency corrections are determined as described in Sec. 5.7. The single particle efficiency tables calculated from embedding may not perfectly reproduce the efficiency effects present in the data. For this reason, a conservative 5% relative uncertainty is assigned to all of the individual efficiencies. The total efficiency uncertainty in Table 5.8 represents the relative uncertainty on the yield.

There is a 5% global normalization uncertainty resulting from the uncertainty in the determination of  $N_{MB}^{equivalent}$  which gets propagated directly to the final yield. The remaining categories relate to the maximum likelihood procedure for extracting the yield. The nominal fit uses a Gaussian shape for the signal peak and a 4<sup>th</sup> order polynomial for the background shape. The uncertainty is determined by also fitting with a 5<sup>th</sup> and 6<sup>th</sup> order polynomial for the background shape. The mass fit range was varied from the nominal  $0.85 < M_{\mu\mu} < 1.5$  GeV/ $c^2$ , to a narrow range of  $0.95 < M_{\mu\mu} < 1.4$  GeV/ $c^2$ , and a wide range of  $0.8 < M_{\mu\mu} < 1.8$  GeV/ $c^2$ . Finally, the uncertainty from the fit computed asymmetrically by the MINOS algorithm is also included.

The relative uncertainty contribution from each source is shown in Fig. 5.32 for each  $p_T$  bin. Since the  $\phi$  is near the  $\rho$  and  $\omega$  peaks, there is not a large range available for constraining



**Figure 5.32 :** The relative uncertainty on the  $\phi$  yield as a function of  $p_T$  after all corrections due to several different sources of systematic uncertainty.

the background. The sensitivity to the fit range results in the large uncertainty due to small variations in the allowed fit range. The total systematic uncertainty is determined for each  $p_T$  bin by summing the positive contribution from each source in quadrature. The total efficiency is then divided by  $\sqrt{3}$  to approximate a  $\pm 1\sigma$  uncertainty, assuming a Gaussian distribution for the uncertainty.

### 5.8.3 Uncertainties on the Hadronic Cocktail

Each hadron decay channel in the hadronic cocktail is scaled by its  $dN/dy$  in order to predict the dimuon yield for each parent particle. The  $dN/dy$  values, whether measured or extrapolated via the Tsallis Blast-wave fit, have an uncertainty which should be propagated to the final dimuon yield. Table 5.9 shows the nominal  $dN/dy$  and relative uncertainty for each particle. The branching ratios are taken from Ref. [23] and considered to be exact. The PYTHIA event generator is used to determine the kinematics and invariant mass distribution of correlated dimuons from open heavy flavor decays.

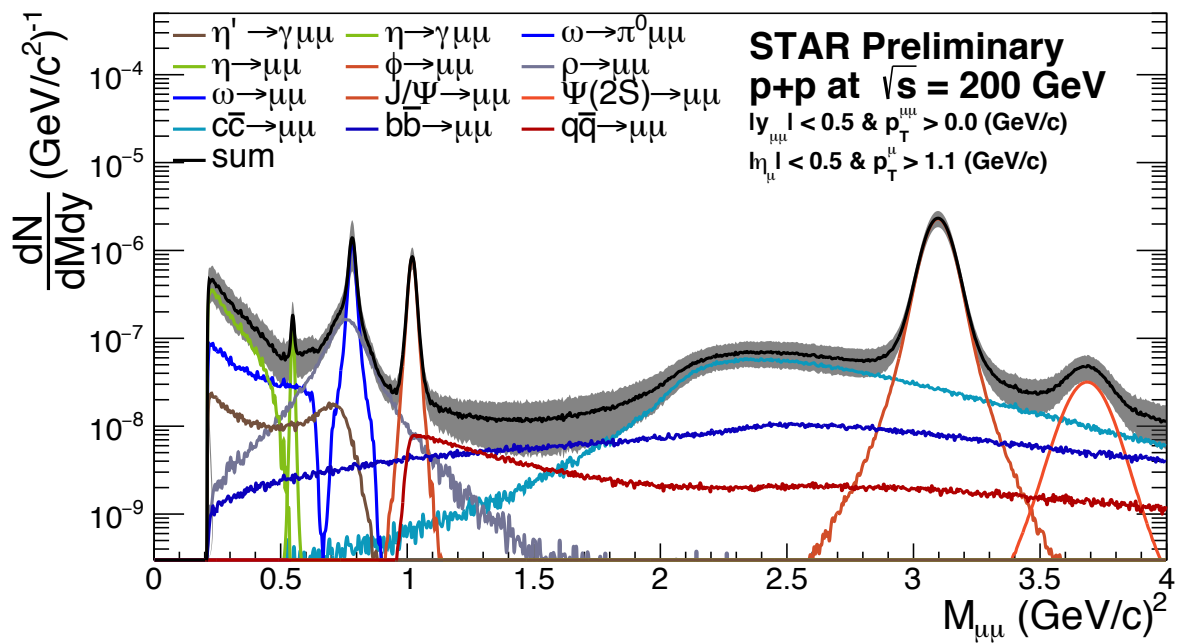
**Table 5.9 :** The nominal value of  $dN/dy$  along with its relative uncertainty is shown for each particle. [147]

Meson	$dN/dy$	Relative Uncertainty
$\eta$	$1.70 \times 10^{-1}$	23%
$\omega$	$1.33 \times 10^{-1}$	21%
$\eta'$	$4.07 \times 10^{-2}$	29%
$\rho^0$	$2.22 \times 10^{-1}$	15%
$\phi$	$1.73 \times 10^{-2}$	20%
$J/\Psi$	$2.44 \times 10^{-5}$	20%
$\Psi(2S)$	$3.38 \times 10^{-6}$	20%

**Table 5.10 :** The nominal value of  $\sigma$  along with its relative uncertainty is shown for  $c\bar{c}$ ,  $b\bar{b}$ , and Drell-Yan. [147]

Channel	$\sigma$	Relative Uncertainty
$c\bar{c}$	0.92 mb	30%
$b\bar{b}$	$3.2 \mu\text{b}$	91%
Drell-Yan	42 nb	15%

The  $\sigma_{c\bar{c}}$ ,  $\sigma_{b\bar{b}}$ , and  $\sigma_{DY}$  values in PYTHIA are not used. Instead, the full phase space cross-sections constrained by recent measurements are used along with their uncertainties [85, 107]. The nominal values of  $\sigma$  are shown with their relative uncertainties in Table 5.10. The assigned uncertainty on  $\sigma_{b\bar{b}}$  is large because the central value of measurements at RHIC are  $+3\sigma$  away from the NLO calculation [70]. In each mass bin the uncertainties from each contributing source are summed in quadrature to determine the total cocktail uncertainty. Fig. 5.33 shows the hadronic cocktail with uncertainties for  $p+p$  collisions at  $\sqrt{s} = 200$  GeV.



**Figure 5.33** : The hadronic cocktail for  $p + p$  collisions at  $\sqrt{s} = 200$  GeV. Each component is drawn separately. The sum is drawn in the black solid curve with the total uncertainty shown in the solid grey band.

## Chapter 6

### Results and Discussion

#### 6.1 Results

##### 6.1.1 First Invariant $M_{\mu\mu}$ Distribution with STAR

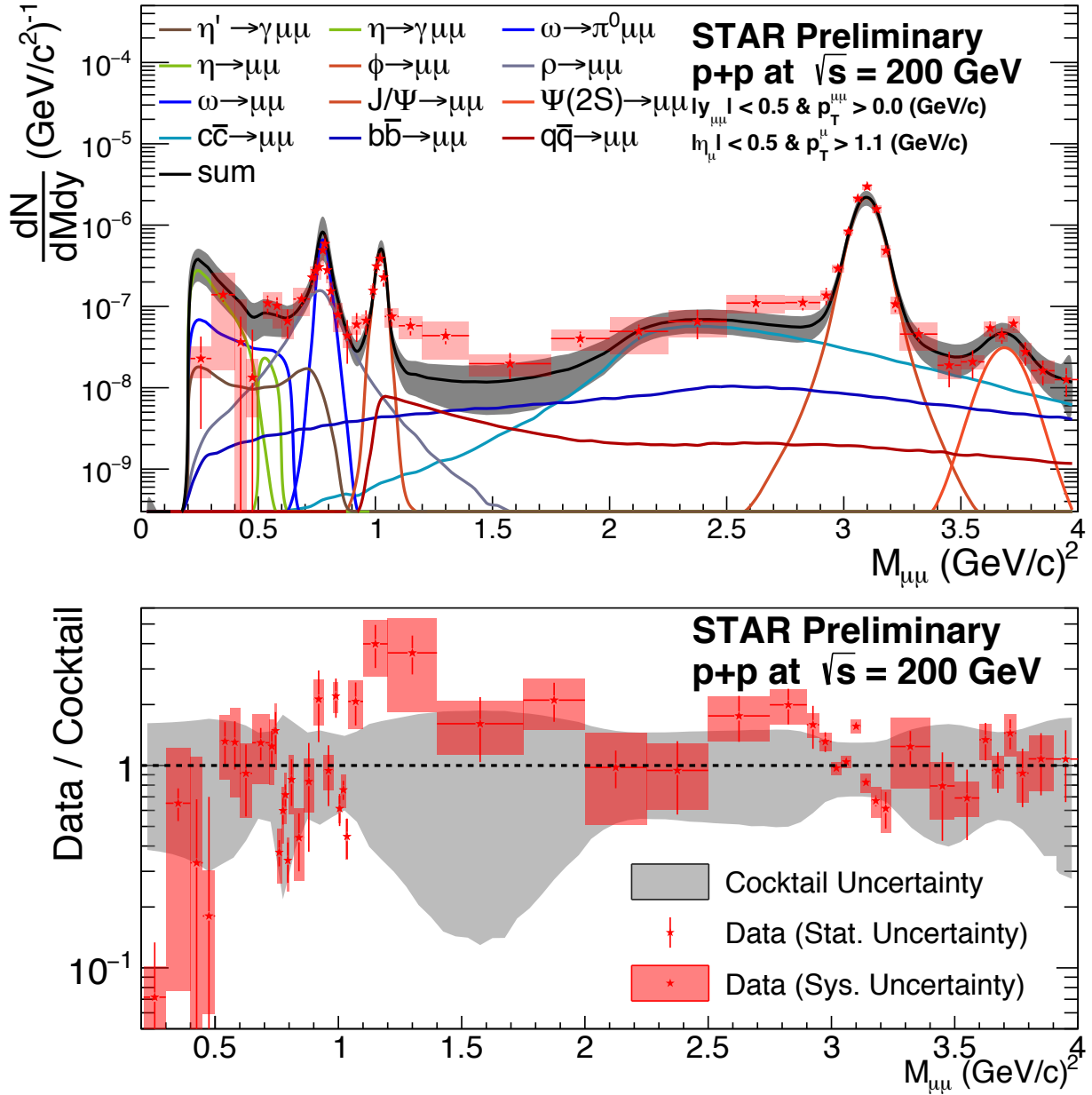
The main purpose of the work in this thesis was to measure the  $\mu^+\mu^-$  invariant mass spectra for the first time with STAR. The challenge of significant background from hadronic punch-through required that the muon identification techniques be optimal. In order to accomplish this, multivariate algorithms were explored for combining the muon identification information from the MTD into an optimal discriminator. Of all the multivariate algorithms investigated, DNNs were found to perform the best in Monte Carlo studies. Using the  $\phi$  meson as a self-analyzing data sample, the DNN-based PID was shown to simultaneously provide better efficiency, signal-to-background ratio, and significance compared to traditional PID techniques.

In addition to providing improved muon identification, the DNN-base PID technique also proved to be crucial for separating signal pairs from background pairs. Unlike NA60's data sets, the data used in these studies did not include a high precision vertex tracker capable of rejecting secondary muons. Instead, a new background estimation technique was developed for *in situ* determination of the correlated backgrounds present in the foreground  $\mu^+\mu^-$  distribution. With these techniques the  $\mu^+\mu^-$  invariant mass spectra was measured in  $p + p$  and  $p+\text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV for the first time with STAR. The fully corrected  $\mu^+\mu^-$  invariant mass spectra are shown in Fig. 6.1 and Fig. 6.2 for the  $p + p$  and  $p+\text{Au}$  data, respectively. The ratio of data to hadronic cocktail is shown in the lower panel of each figure. There is good agreement across the entire invariant mass range within the substantial uncertainties. The agreement with the cocktail demonstrates that the techniques for removing the hadron contamination are largely successful, with the possible exception of the lowest mass region. At masses just above threshold mass, the combination of poor efficiency and

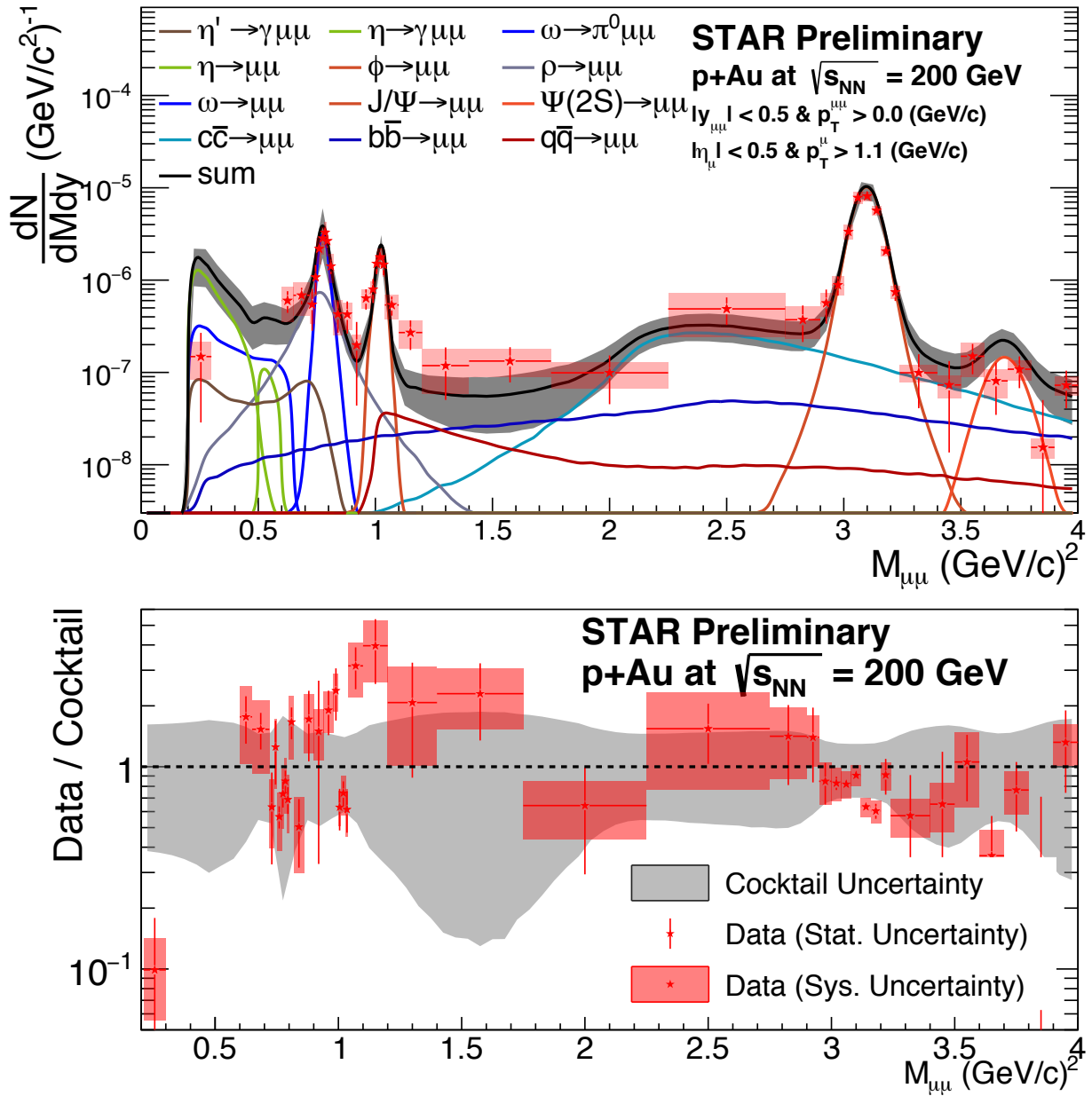


significant, focused hadron contamination from  $K_S^0 \rightarrow \pi^+ + \pi^-$  and  $\phi \rightarrow K^+ + K^-$  makes it difficult to resolve any dimuon signal.

In the intermediate mass region, low  $p_T^{\mu\mu}$  pairs from open heavy flavor decays and Drell-Yan are rejected by the  $p_T > 1.1 \text{ GeV}/c$  requirement enforced by the MTD kinematic acceptance. Due to the mass difference between the  $c$ -quark and the  $b$ -quark, the  $c\bar{c}$  spectra is more heavily affected at  $M_{\mu\mu} < 2 \text{ GeV}/c^2$  which leaves semi-leptonic  $b\bar{b}$  decays as the dominant source in this region. The  $b\bar{b}$  cross section is poorly constrained since recent measurements are not in especially good agreement with the next-to-leading order (NLO) calculations. With the statistics available in the current datasets no further constraint can be made on the  $b\bar{b}$  cross section. However, a higher statistics dataset with the MTD was collected in 2017 from  $p + p$  collisions at  $\sqrt{s} = 510 \text{ GeV}$  that may allow such a study.



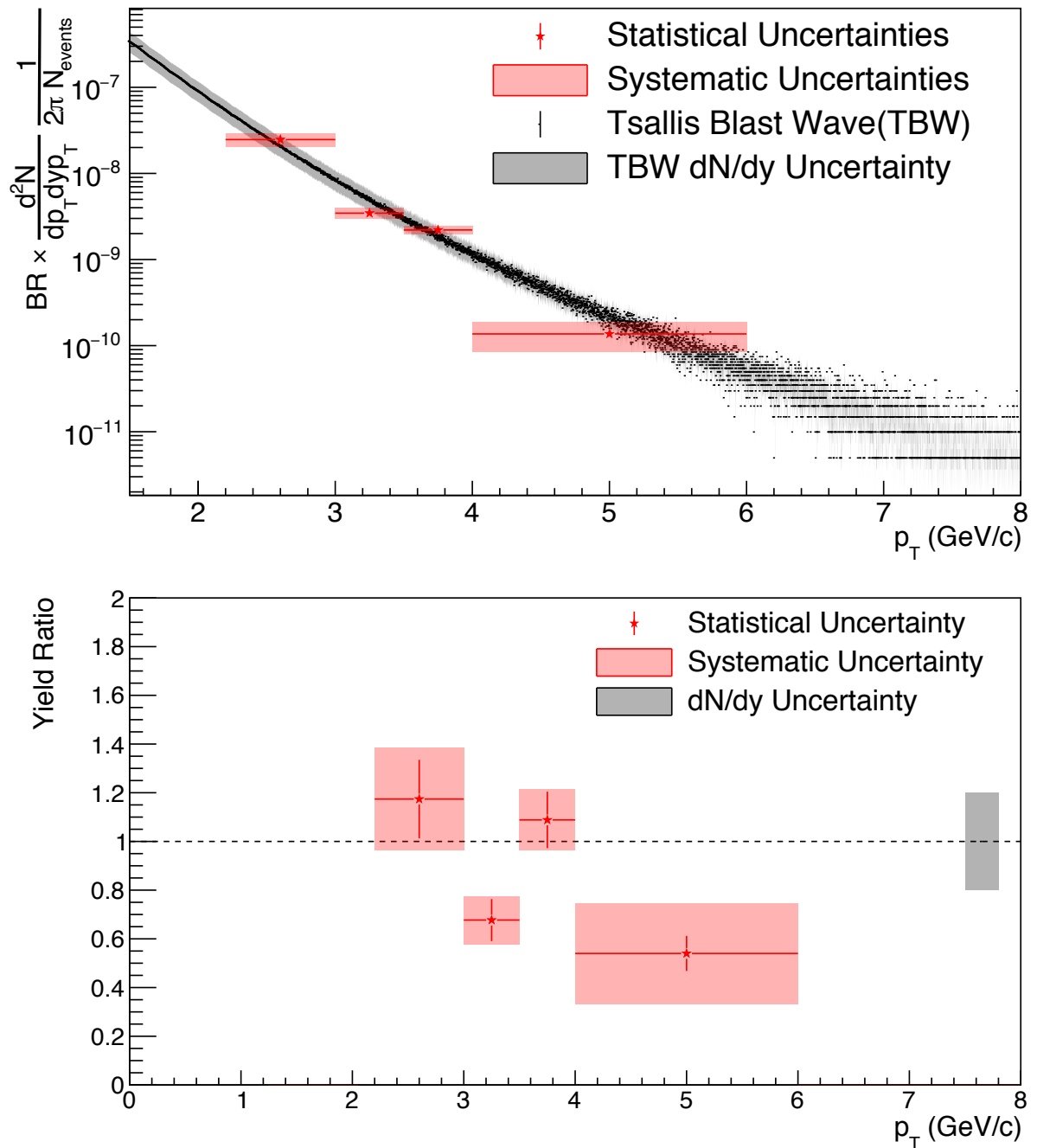
**Figure 6.1** : Top: The  $\mu^+\mu^-$  invariant mass spectra in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV. The statistical uncertainties are shown in vertical bars and the systematic uncertainties are shown in shaded boxes. Bottom: The ratio of data to hadronic cocktail.



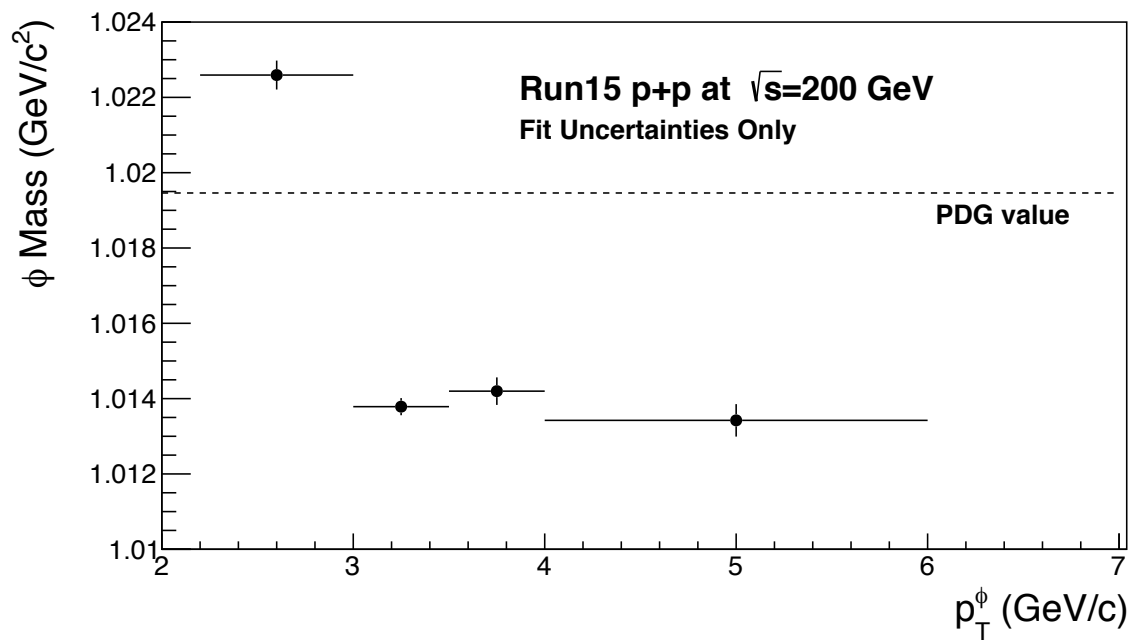
**Figure 6.2** : Top: The  $\mu^+\mu^-$  invariant mass spectra in  $p$ +Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. The statistical uncertainties are shown in vertical bars and the systematic uncertainties are shown in shaded boxes. Bottom: The ratio of data to cocktail.

### 6.1.2 First Measurement of the $\phi$ Spectra via Dimuons with STAR

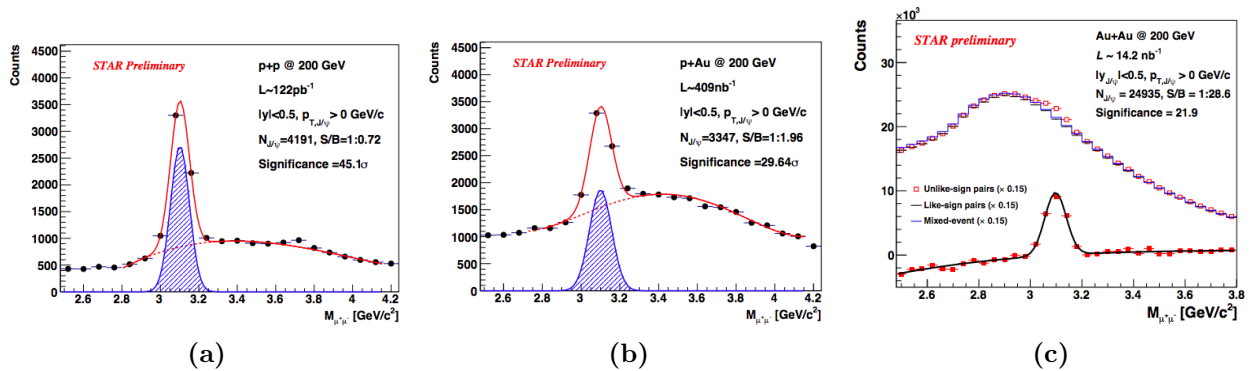
Another primary result of this work is the first measurement of the  $\phi$  meson spectra through the  $\phi \rightarrow \mu^+ + \mu^-$  decay channel with STAR. The measured branching ratio times invariant yield measured in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV is shown in Fig. 6.3. The lower panel of Fig. 6.3 shows the ratio of the measured yield to the the Tsallis Blast-wave shape scaled by the  $dN/dy$  from Ref. [147]. Good agreement is observed within the sizable statistical and systematic uncertainties. The mass of the  $\phi$  measured in each momentum bin is also shown in Fig. 6.4. No attempt is made to correct for energy loss affects which will result in a decrease of the measured mass.



**Figure 6.3 :** Top: The branching ratio times the invariant yield of the  $\phi$  meson measured through the  $\phi \rightarrow \mu^+ + \mu^-$  decay channel. Bottom: The ratio of data to the Tsallis Blast-Wave scaled by the nominal  $dN/dy$  value. The uncertainty on  $dN/dy$  is shown in gray.



**Figure 6.4 :** The mass of the  $\phi$  measured in four  $p_T$  bins. Error bars show only the uncertainties from the maximum likelihood fit. No energy loss correction is included.

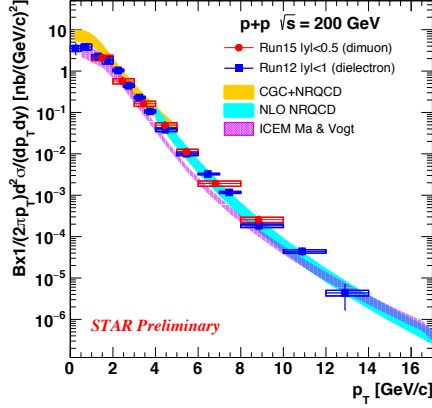


**Figure 6.5 :** The  $J/\psi$  signal measured through the  $\mu^+\mu^-$  channel in  $p + p$ ,  $p+\text{Au}$ , and  $\text{Au}+\text{Au}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV [164].

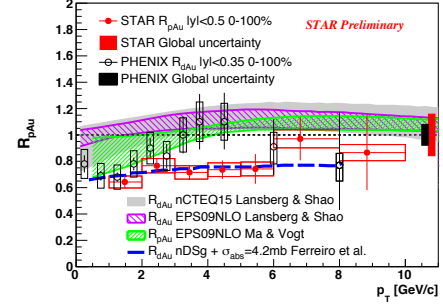
## 6.2 Summary and Discussion

The installation of the Muon Telescope Detector has opened new opportunities to study dimuon production at STAR. The purpose of the MTD was to allow for the identification and measurement of muons at STAR. Muon identification is specifically useful for quarkonia measurements and measurements of the dimuon invariant mass spectra. Since the installation of the MTD was completed  $p + p$ ,  $p+\text{Au}$ , and  $\text{Au}+\text{Au}$  data sets have been collected at  $\sqrt{s_{NN}} = 200$  GeV. These data sets have allowed STAR to augment its dielectron program with additional studies of dilepton production through the  $\mu^+\mu^-$  channel. Figure 6.5 shows the inclusive  $J/\psi$  signals measured in each collision system using the MTD. A comparison of the inclusive  $J/\psi$  cross section in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV measured through the dielectron and dimuons channels is shown in Fig. 6.6. Good agreement is found between the new dimuon measurements and the previous dielectron measurements [164].

In the context of high energy heavy-ion physics, low multiplicity  $p + p$  collisions can be considered a baseline void of any cold or hot nuclear matter effects. Cold nuclear matter effects due, for instance, to the modification of parton distributions within the nucleus can be investigated by comparing  $p+\text{Au}$  yields to  $p + p$  yields. In the absence of nuclear matter effects the  $p+\text{Au}$  and  $\text{Au}+\text{Au}$  yields are expected to be the result of an incoherent collection of  $p + p$  collisions. Nuclear matter effects can be quantified through the nuclear modification



**Figure 6.6 :** The inclusive  $J/\psi$  cross section measured in  $0 < p_T < 14$  GeV/ $c$ . STAR dielectron and dimuon measurements are compared with model calculations. [165–167]. Figure taken from Ref. [168]



**Figure 6.7 :** The nuclear modification factor  $R_{pAu}$  as a function of  $p_T$  for  $J/\psi$  measured through the dimuon channel [168].

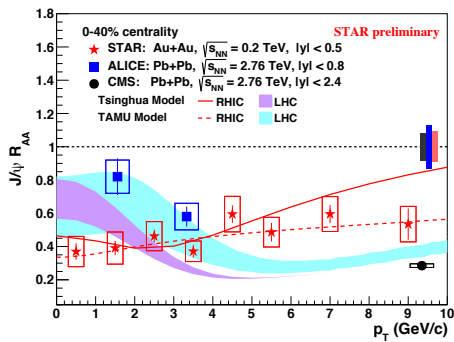
factor which is in the ratio  $R_{pAu}$  ( $R_{AA}$  for Au+Au), where  $R_{pAu}$  is defined as:

$$R_{pAu} = \frac{\sigma_{inel}}{\langle N_{coll} \rangle} \frac{d_{pAu}^N / dy dp_T}{d^2 \sigma_{pp} / dy dp_T} \quad (6.1)$$

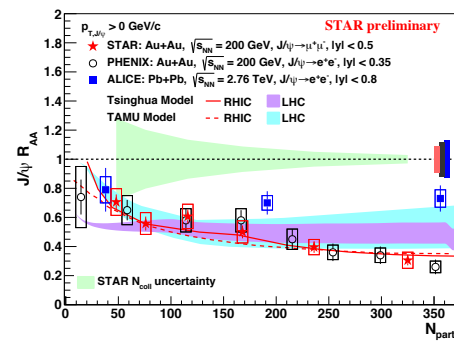
where  $\sigma_{inel}$  is the cross-section for inelastic collisions,  $\langle N_{coll} \rangle$  is the mean number of binary collisions in  $p+Au$  collisions,  $d_{pAu}^N / dy dp_T$  is the yield in  $p+Au$  collisions, and  $d^2 \sigma_{pp} / dy dp_T$  is the cross section measured in  $p + p$  collisions.  $R_{AA}$  results when the yield from Au+Au collisions is used instead. Figure 6.7 shows the  $R_{pAu}$  for  $J/\psi$  as a function of  $p_T^{J/\psi}$  measured through the dimuon channel. It is shown with  $R_{dAu}$  measurements from PHENIX along with several different model calculations [165–167, 169]. The  $R_{pAu}$  result is consistent with the  $R_{dAu}$  at the  $1.4\sigma$  level, suggesting that cold nuclear matter effects are similar between the two systems.

Hot nuclear matter effects, due to the creation of a hot and dense medium can be quantified with the  $R_{AA}$  of  $J/\psi$ . Figure 6.8 shows the  $R_{AA}$  as a function of  $p_T^{J/\psi}$  for 40–80 % central Au+Au collisions. Figure 6.9 shows the centrality dependence of the  $J/\psi$   $R_{AA}$  for measurements at STAR and LHC energies. The strong suppression observed at high  $p_T$  in Fig. 6.8 is evidence of significant  $J/\psi$  disassociation. The  $J/\psi$   $R_{AA}$  as a function of  $p_T$  is compared with the TAMU and Tsinghua transport models. [67, 170–172]. While both models





**Figure 6.8 :** The nuclear modification factor  $R_{AA}$  as a function of  $p_T$  for  $J/\psi$  measured through the dimuon channel [168]. LHC data points from Ref. [68, 173]. Model calculations from Ref. [67, 170, 171]



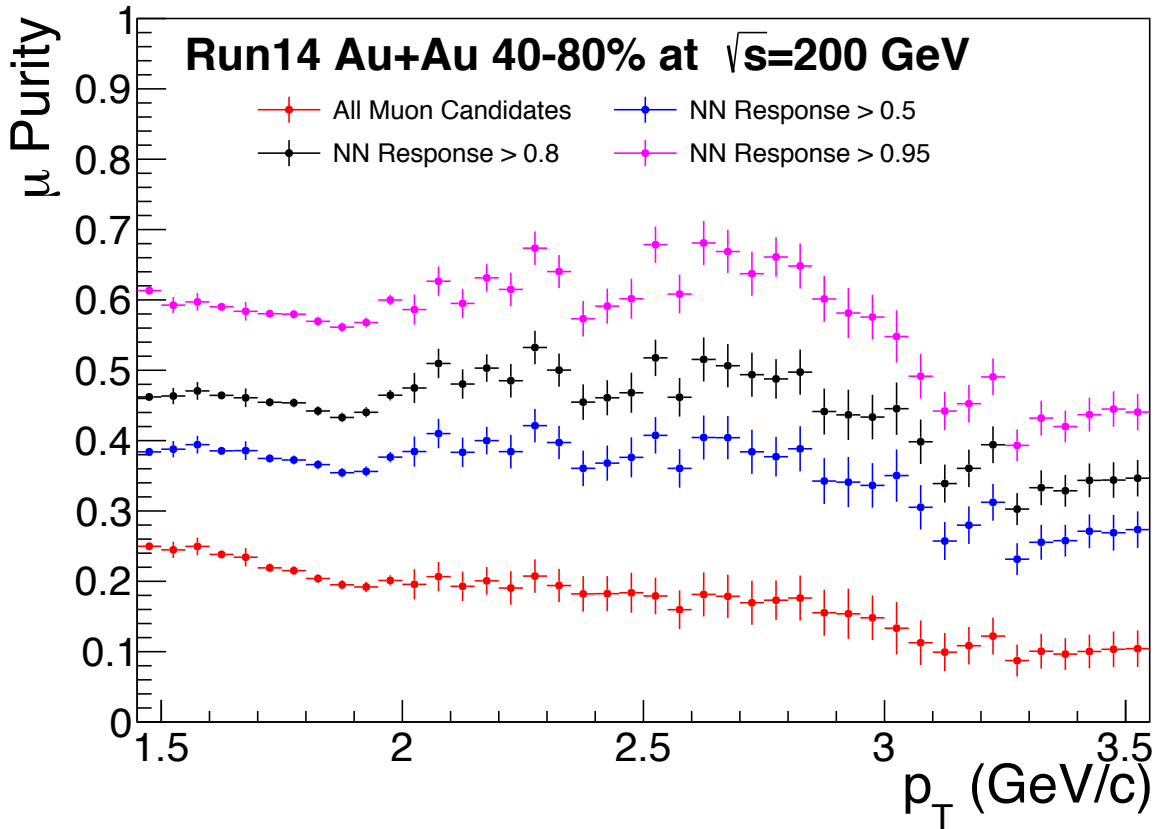
**Figure 6.9 :** The nuclear modification factor  $R_{AA}$  as a function of  $N_{part}$  for  $J/\psi$  measured through the dimuon channel [168]. LHC data points from Ref. [68, 173]. Model calculations from Ref. [67, 170, 171]

describe the data reasonably well, the TAMU model is slightly favored by data. Figure 6.9 shows that the models describe the centrality dependence of the  $J/\psi$   $R_{AA}$  reasonably well at RHIC energies but slightly over estimate the suppression at LHC energies.

The inclusive  $J/\psi$  yield and  $R_{AA}$  measured in Au+Au collisions is being prepared for publication as the first physics results from the MTD. Another paper dedicated to the measurement of the  $\mu^+\mu^-$  invariant mass spectra is in preparation as well. While the MTD has successfully provided muon identification needed for quarkonia measurements in Au+Au collisions, measurement of the  $\mu^+\mu^-$  and  $e-\mu$  spectra has not been possible due to prohibitively low muon purity.

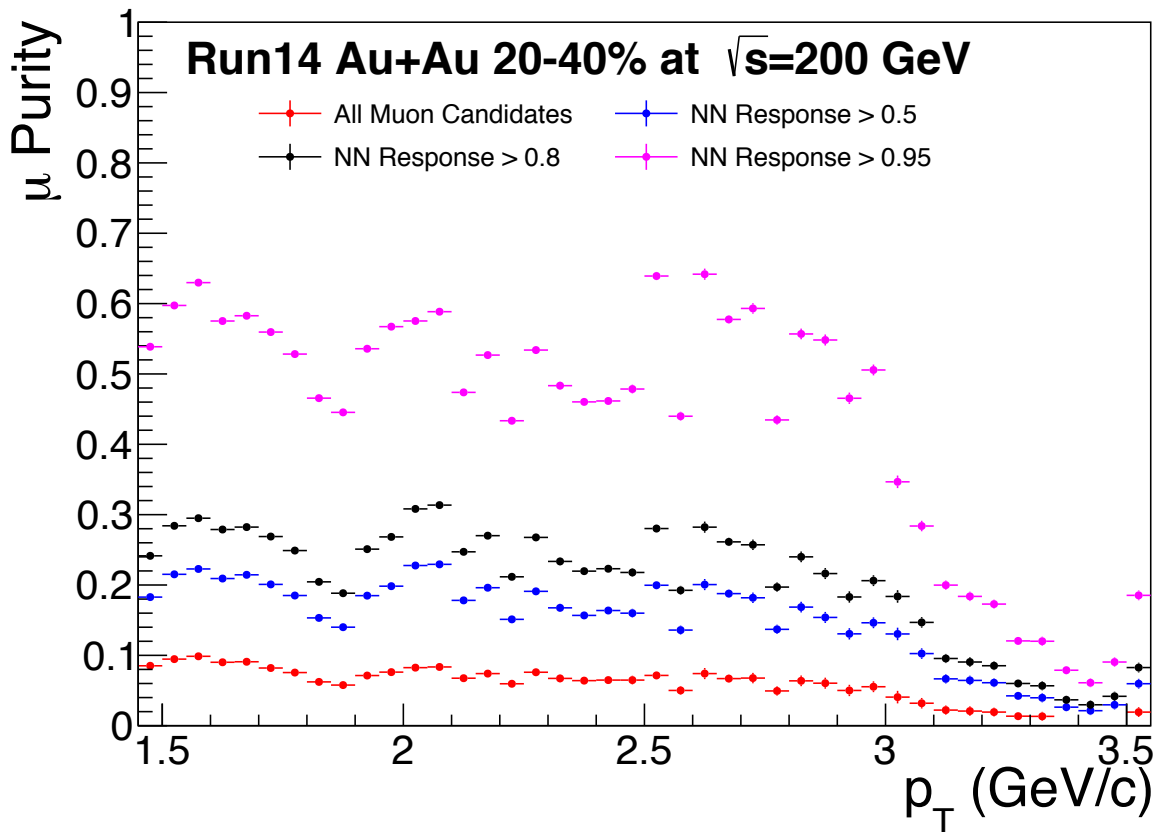
One of the goals of this project was to measure the dimuon invariant mass spectra in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. However, the purity was found to be prohibitively low in the Au+Au datasets. The techniques used for measuring the purity in  $p+p$  and  $p+Au$  collisions was employed for the Au+Au dataset. Despite re-using the DNN model trained for  $p+p$  data, the DNN described the Au+Au data surprisingly well for tracks with  $p_T > 1.5$  GeV/c. Plots summarizing the purity in 40 – 80% central and 20 – 40% central Au+Au collisions are shown in Fig. 6.10 and Fig. 6.11 respectively. The purity fits for tracks with  $p_T > 1.5$  GeV/c are shown in Fig. B.3 – B.11 for  $p+p$ , 40 – 80% Au+Au, and 20 – 40% Au+Au collisions. The  $\chi^2/ndf$  in Au+Au data is worse, largely because the statistics in the data are significantly better than the DNN templates. The uncertainty in the DNN templates

(due to poor statistics) was not included when calculating the  $\chi^2/ndf$ .



**Figure 6.10** : The single muon purity in 40 – 80% central Au+Au collisions for various selection criteria.

While the purity values in 40 – 80% central Au+Au collisions are not substantially worse than in  $p + p$ , the low purity is coupled with a smaller number of events which makes the analysis in peripheral Au+Au collisions impractical. The average purity in the 40 – 80% central Au+Au sample is  $\sim 60\%$  at an DNN cut giving an average signal efficiency of  $\sim 10\%$ . The purity levels in 20 – 40% central Au+Au collisions look surprisingly high at first glance. However, it should be noted that these purity numbers include secondary muons from pion and kaon decays (since they are inseparable). The purity values for inclusive muon candidates (without any DNN cut) are approximately consistent with the values expected for only secondary muon sources - i.e. consistent with zero signal muons. The values listed in Table 5.2 for the fractions of pions and kaons that will decay, on average, within the MTD, are used to come to this conclusion. Since the  $J/\psi$  has been extracted from the Au+Au data set, there must still be a small number of signal muons present. Though even



**Figure 6.11** : The single muon purity in 20 – 40% central Au+Au collisions for various selection criteria.

for the  $J/\psi$  the integrated (over  $p_T$  and centrality) signal to background ratio is  $\sim 1/27$  - less than the significance of the  $J/\psi$  from the dielectron channel, which was  $\sim 1/2$ . Guided by the production predicted by the hadronic cocktail, this would suggest that the signal to background ratio in the low and intermediate mass range is  $\sim 1/300 - 1/3,000$ . This means that the background would need to be measured at better than  $1/30,000$  to achieve a 10% uncertainty. Unfortunately, the like-sign ratio method is not capable of reaching that precision simply due to the effect of statistical fluctuations on the fitting. For this reason, analysis of the low and intermediate dimuon invariant mass spectra in Au+Au is not possible.

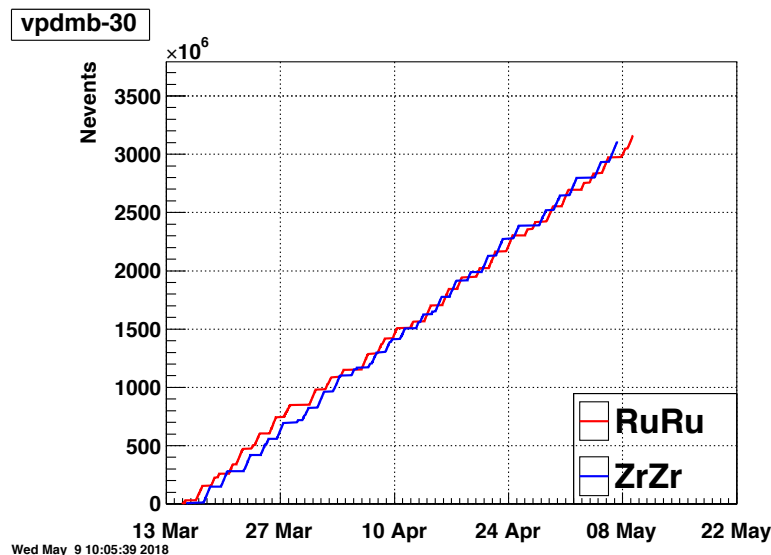
## Chapter 7

### Future $e^+e^-$ Measurements with STAR

The original motivation for this work was to measure the thermally produced dimuons in semi-central and central Au+Au collisions. Towards this end, the measurements in  $p + p$  collisions were used to develop the necessary techniques and methods. Since the hadronic cocktail represents all significant sources of dimuon production in  $p + p$  collisions, it also provided an ideal dataset to demonstrate the precision of the developed techniques. Since low muon purity prevented any measurement of the thermally produced dimuons in Au+Au collisions, the final chapter of this thesis briefly discusses the viability of such a measurement in the recently collected STAR datasets and proposed near future datasets. In 2017 and 2018, STAR collected three datasets that are ideal for measuring the dilepton continuum through the  $e^+e^-$  channel in high energy heavy ion collisions. Additionally, the second phase of the RHIC Beam Energy Scan will commence in 2019 which will provide several additional datasets from Au+Au collisions at  $\sqrt{s_{NN}}$  ranging from 7.7 to 19.6 GeV. The measurements in the recent high energy dataset ( $\sqrt{s_{NN}} = 200$  GeV) will be discussed first followed by a discussion of the proposed measurements in the lower energy datasets.

During 2018 data was collected from  $^{96}_{44}\text{Ru}+^{96}_{44}\text{Ru}$  and  $^{96}_{40}\text{Zr}+^{96}_{40}\text{Zr}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV. Ruthenium and Zirconium were chosen because they are isobars, i.e. they have the same mass number  $A$ , but different number of protons  $Z$ . The primary motivation for this choice was to study effects that are sensitive to  $Z$ , such as magnetic field induced effects and the production of very low momenta dilepton pairs [174, 175]. While neither of these motivations directly relate to the measurement of thermally produced dileptons, the isobar data still provides an ideal dataset for such measurements.

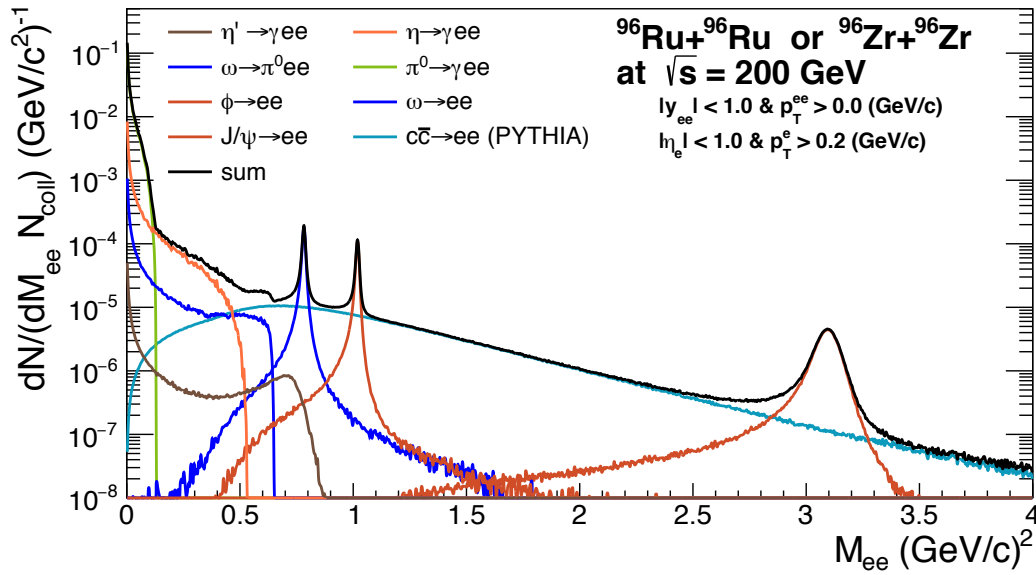
Over 6 billion minimum-bias events were collected from isobar collisions during the 2017 data collection period. Figure 7.1 shows the cumulative number of triggered events collected versus day. Approximately 3.2 billion events were collected for each of  $^{96}_{44}\text{Ru}+^{96}_{44}\text{Ru}$  and  $^{96}_{40}\text{Zr}+^{96}_{40}\text{Zr}$  collisions. The purpose of choosing isobars, and not just any two nuclei with



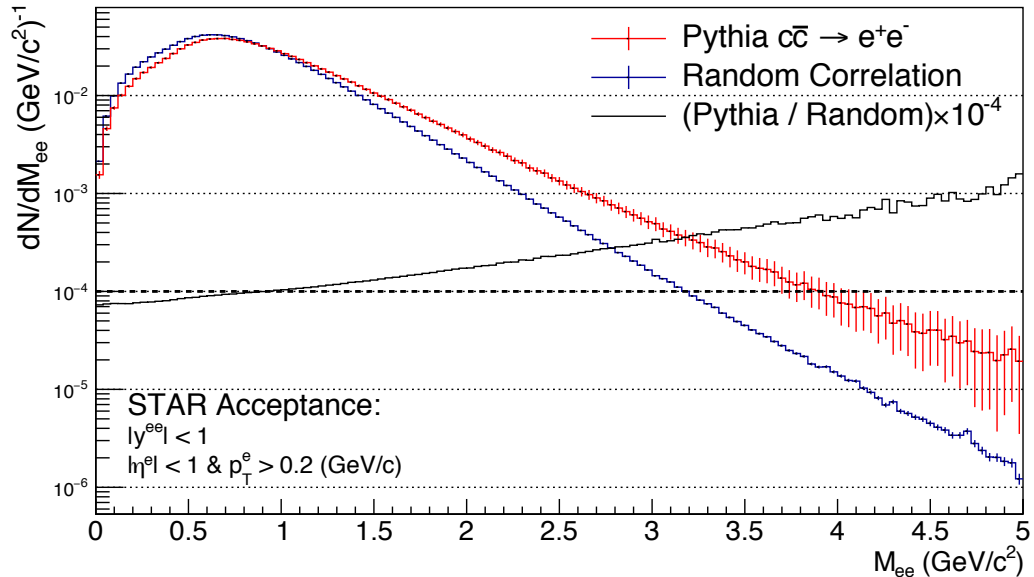
**Figure 7.1** : The cumulative number of events collected versus time through the run.

different  $Z$ , was to remove any variation between the two systems resulting from system size, number of binary collisions, nuclei shape, etc. Measurements of charged particle yields, specifically the  $\pi^\pm$  yields, will be necessary to verify that the gross particle production is similar in the two isobar systems. If the gross particle production is in fact consistent between the two systems, then they can be combined for a measurement of the dielectron invariant mass spectra. Of the 6 billion triggered events a large fraction are expected to pass the quality cuts used for event selection. Even if the two dataset are not combined, the individual datasets including  $\sim 3$  billion events each will still be significantly larger datasets than ever previously collected at STAR for dielectron analyses. The only comparable sized data sets available with STAR are the combined 2014 and 2016 sample of Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. However, since the HFT detector was installed from 2014–2016, the added material budget increases the probability for the formation of conversion electrons. However, the HFT and its support material were removed before the 2017 runs, returning STAR to an ideal state for dielectron analyses.

The expected contributions to  $e^+e^-$  production from hadronic decays and open heavy flavor decays for isobaric collisions at  $\sqrt{s_{NN}} = 200$  GeV can be seen in Fig. 7.2. At  $\sqrt{s_{NN}} = 200$  GeV, semi-leptonic  $c\bar{c}$  decays account for the vast majority of  $e^+e^-$  pairs produced with intermediate masses (between the  $\phi$  and  $J/\psi$ ). Precisely accounting for the dilepton pairs resulting from open heavy flavor decays is the most challenging aspect in iso-



**Figure 7.2 :** The expected hadronic cocktail components for  ${}^{96}\text{Ru}+{}^{96}\text{Ru}$  or  ${}^{96}\text{Zr}+{}^{96}\text{Zr}$  collisions at  $\sqrt{s_{NN}} = 200$  GeV. The contributions from thermal production and the  $\rho^0$  meson are not included. The  $c\bar{c} \rightarrow e^+e^-$  contribution is from PYTHIA and scaled with a cross section of  $800 \mu\text{b}$ .

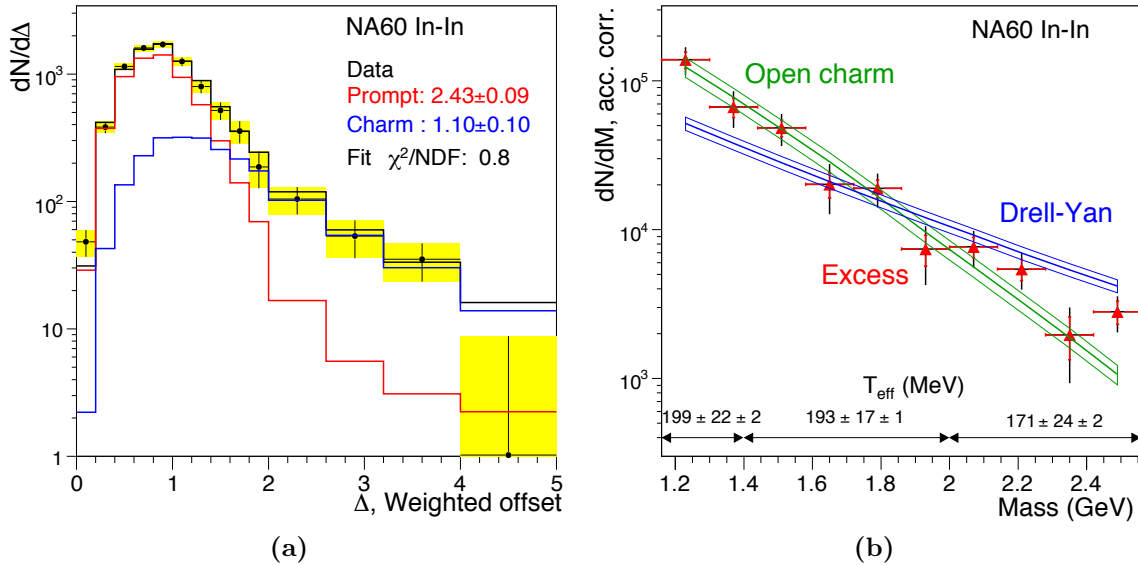


**Figure 7.3 :** The hadronic cocktail component from  $c\bar{c} \rightarrow e^+e^-$  Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. Two extreme scenarios for  $c\bar{c}$  are shown: 1) default PYTHIA i.e. no modification from the medium, 2) fully decorrelated (random correlations).

lating the thermally produced dileptons at RHIC energies. As discussed in Chapter 2, the invariant mass spectra of dileptons from  $c\bar{c}$  decays may be modified by the presence of a dense medium. Specifically, multiple scattering in a dense medium could weaken the correlation between the charmed mesons which later decay into  $e^+e^-$  pairs. A modification in the angular correlation between the parent mesons directly results in a modification of the invariant mass density of produced  $e^+e^-$  pairs. As an example of this, Fig. 7.3 shows a comparison of the  $e^+e^-$   $dN/dM$  distribution from  $c\bar{c}$  as simulated in PYTHIA and for fully decorrelated (random correlations)  $e^+e^-$  pairs. The  $c\bar{c}$  invariant mass distribution from random correlations is much softer (i.e. falls more steeply) and was less favored by the PHENIX data than the default result from PYTHIA.

Another difficulty arises since STAR does not have a precision secondary vertex tracker. The NA60 experiment was able to isolate a prompt thermal radiation component after rejecting  $c\bar{c}$  and Drell-Yan as possible production mechanisms. It was only through fits of the vertex offset that they were able to demonstrate that the observed excess was due to a prompt source and therefore incompatible with an enhanced  $c\bar{c}$  yield. They further note that the prompt excess is not compatible with Drell-Yan because the  $p_T$  spectra is too soft. In fact, the  $p_T$  spectra was compatible with the  $p_T$  spectra for  $c\bar{c}$  within uncertainties. For this reason, the  $p_T$  spectra likely cannot provide distinguishing power between a thermal excess compared to  $c\bar{c}$  and Drell-Yan. For the STAR isobar dataset this means that even with a high statistics dataset, isolating and measuring the thermal dielectrons spectra may not be feasible due to the lack of a precision secondary vertex tracker.

In 2017, STAR collected about 1.5 billion minimum bias events from Au+Au collisions at  $\sqrt{s_{NN}} = 54$  GeV. In 2018, STAR collected about 1.3 billion minimum bias events from Au+Au collisions at  $\sqrt{s_{NN}} = 27$  GeV. Starting in 2019, several more datasets from Au+Au collisions at  $\sqrt{s_{NN}} = 19.6$  GeV and below will be taken as part of the RHIC Beam Energy Scan Phase II. These datasets are invaluable for conducting a systematic study of thermal dilepton production and  $\rho$  meson in-medium broadening as a function of  $\sqrt{s_{NN}}$ , system size, system life time, and  $\mu_B$ . Figure 7.5a shows the  $c\bar{c}$  cross section at these lower collision energies. Measuring dileptons resulting from thermal production becomes significantly more viable at lower energies since the  $c\bar{c}$  cross section falls rapidly with decreasing  $\sqrt{s_{NN}}$ . However, the difficulty of distinguishing the thermal dileptons from the  $c\bar{c}$  decay pairs without a

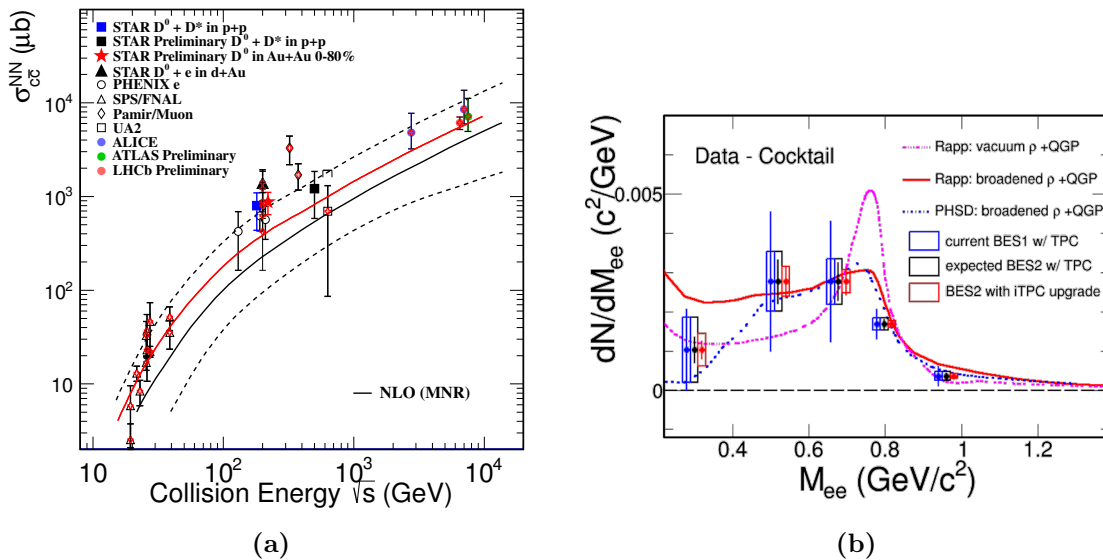


**Figure 7.4 :** The NA60 experiment combined a vertex offset fit (a) with an analysis of the pair  $p_T$  to determine that the excess yield was not from Drell-Yan or  $c\bar{c}$  but instead from a prompt thermal source.

precision secondary vertex tracker will still remain.

The datasets included in the Beam Energy Scan phase II will also allow further investigation of the underlying mechanisms responsible for the observed in-medium broadening of the  $\rho$  meson. In BES I poor statistics prevent the  $e^+e^-$  mass spectra from being measured below  $\sqrt{s_{NN}} = 19.6$  GeV. For each of the energies measured in the range from  $\sqrt{s_{NN}} = 19.6$  GeV to 62.4 GeV, the system's total baryon density and freeze out temperature remained roughly constant [176]. Therefore, the observed excess dielectron yield over the hadronic cocktail contributions resulting from thermal production and  $\rho$  meson decays, primarily probed the changing lifetime of the system. However, at energies below  $\sqrt{s_{NN}} = 19.6$  GeV, the total baryon density begins to increase and the temperature of the system begins to decrease with decreasing  $\sqrt{s_{NN}}$ . Therefore, measurements of the excess dielectron yield will provide stronger distinguishing power between models that respond differently to these effects. Figure 7.5b shows the excess yield measure in Au+Au collisions at  $\sqrt{s_{NN}} = 19.6$  GeV along with two model calculations [177]. It also shows predicted improvements in the size of the uncertainties with the additional statistics and upgraded detectors that will be available for BES II. The reduction in uncertainties may provide the distinguishing power needed to distinguish between the two models shown in Fig. 7.5b. Distinguishing between these models





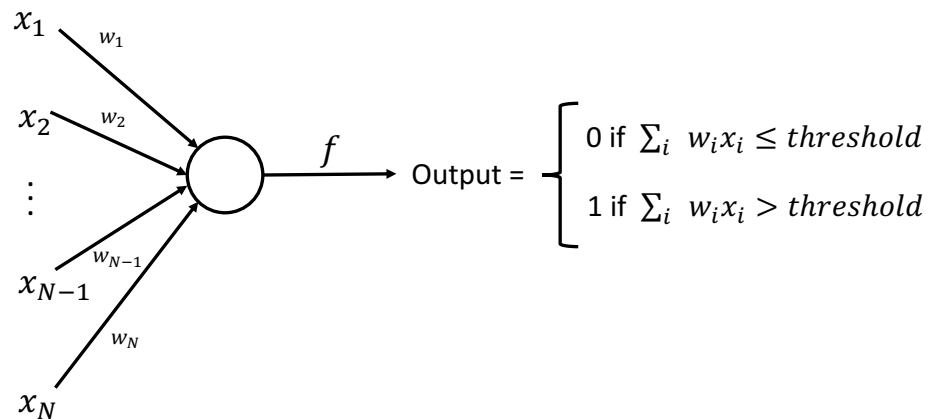
**Figure 7.5 :** The  $c\bar{c}$  cross section versus collision energy ( $\sqrt{s}$ ) measured by several experiments and compared with the next-to-leading order (NLO) calculations (a). The measured excess  $e^+e^-$  spectra near the  $\rho$  meson mass in Au+Au collisions at  $\sqrt{s_{NN}} = 19.6$  GeV from BES 1 compared with various theoretical curves (b). In addition, the expected reduction in uncertainty for BES II is shown for two scenarios, 1) for improved statistics with the current TPC hardware, and 2) for improved statistics with the inner TPC upgrade.

may help to determine whether the broadening of the  $\rho$  spectral function is a result of chiral symmetry restoration or another mechanism. Between the isobar dataset at  $\sqrt{s_{NN}} = 200$  GeV and the various new and upcoming datasets in Au+Au collisions at  $\sqrt{s_{NN}} = 54$  GeV and below, STAR has a multitude of exciting opportunities to study fundamental aspects of QCD and the QGP through dilepton production.

## Appendix A

### Training Neural Networks

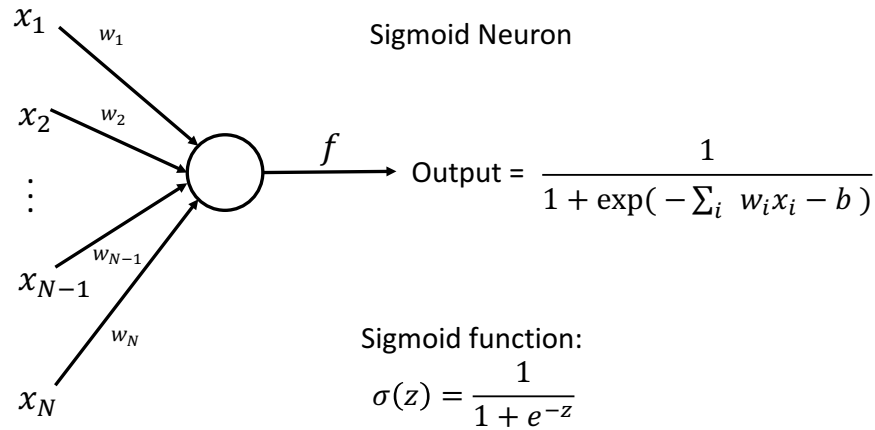
Artificial neural networks (ANN) have made a resurgence in recent years and become one of the most used models for statistical learning applications. In this section I briefly review the concept of ANNs and the back propagation training technique. First, consider the basic building block of the ANNs used in this thesis: the perceptron. The simplest perceptron, shown in Fig. A.1, receives  $N$  inputs and outputs a single value, either 0 or 1. The sigmoid perceptron is more flexible, capable of outputting a continuous value between 0–1. See Fig. A.2 for an example of the sigmoid perceptron.



**Figure A.1 :** A simple perceptron with a binary output value.

The perceptron is a simple building block from which complex networks can be built. In principle, the more perceptrons in an ANN, the more power it has to learn new behaviors. Figure A.3 shows an example network with 3 input variables and two hidden layers between the inputs and the output layer.

During the training phase labeled tuples of data are used to minimize a loss with respect



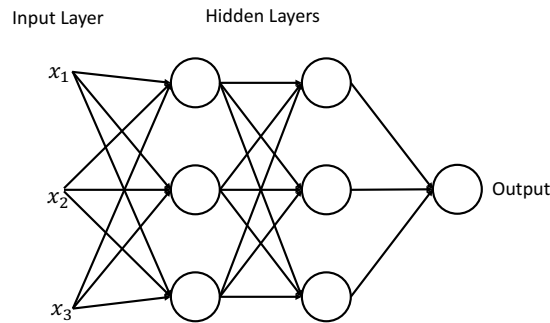
**Figure A.2** : A sigmoid perceptron is capable of outputting a continuous variable.

to the ANN's output value. Labeled data simply means that the tuple includes a one-to-one map between input variables and the *true* output value. These tuples have the form  $(x_1, x_2, \dots, x_{N-1}, x_N, a)$  where  $a$  is the correct output given the inputs  $x_1, x_2, \dots, x_{N-1}, x_N$ . Training the neural network consists of minimizing the loss function given a set of labeled tuples. For this work we used the quadratic loss function:

$$C(w, b) = \frac{1}{2n} \sum_{\text{samples}} \|y(x_1, x_2, \dots, x_{N-1}, x_N) - a\|^2 \quad (\text{A.1})$$

where  $w$  and  $b$  are the weights and bias of a given node and  $y(x)$  is the output of the network. If the network output is correct for all inputs then the loss would be 0. Evaluating the loss function for a given set of nodes with weights and biases on a sample is trivial, but determining the optimal weights and biases is not. However, with the advent of back propagation this problem was definitely solved.

Before describing the back propagation technique it is helpful to define a few terms. The notation used here is based on that used in Ref. [178], a very helpful resource for a thorough discussion of back propagation and its derivation. For an ANN with multiple inputs and multiple layers, the number of weights grows very quickly (combinatorially for dense networks). It is convenient to define  $w_{j,k}^l$  as the weight between the  $j^{\text{th}}$  neuron in layer



**Figure A.3 :** A simple ANN built from multiple perceptrons. Each perceptron has only one output. Multiple exiting lines signify that the output is given as input to many different neurons in the next layer.

$l - 1$  to the  $k^{th}$  neuron in layer  $l$ . Each node has a corresponding bias  $b_j^l$ , for the  $j^{th}$  node in the  $l^{th}$  layer. In practice though, it is common to limit the number of nodes with a bias to only one per layer, since one bias neuron is sufficient to achieve any result. Finally,  $a_j^l$  is the activation of neuron  $j$  by layer  $l$  where:

$$a_j^l = \sigma \left( \sum_k w_{jk}^l a_k^{l-1} + b_j^l \right) \quad (\text{A.2})$$

or more compactly with implicit sums:

$$a^l = \sigma(w^l a^{l-1} + b^l) \quad (\text{A.3})$$

From this relation, it is clear that the activation of each neuron in layer  $l$  depends on all neurons in the previous layer. In order to make the following less verbose we let  $z^l = w^l a^{l-1} + b^l$ . The goal when training the neural network is to compute the partial derivatives of the cost function with respect to the weights and biases, so that we can determine how to update them based on each sample. Back propagation provides a technique for relating  $\delta_j^l = \delta C / \delta z_j^l$  to the quantities that are needed, i.e.  $\delta C / \delta w_{jk}^l$  and  $\delta C / \delta b_j^l$ . Back propagation is then defined by four primary equations. The first equation relates the error in the output layer to its inputs:

$$\delta_j^L = \frac{\delta C}{\delta a_j^L} \frac{\delta \sigma(z_j^L)}{\delta z_j^L} \quad (\text{A.4})$$

This relation is fairly straight forward and easily computed. It simply relates the error in the last layer (output layer  $L$ ) to the activation of the previous layer and its change with respect to its inputs.

The second equation relates the error in layer  $l$  to the error in layer  $l + 1$  as follows:

$$\delta^l = (w^{l+1})^T \delta^{l+1} \circ \frac{\delta\sigma(z^l)}{\delta z^l} \quad (\text{A.5})$$

This equation relates the errors in nearby layers by applying the transpose of the weights in the  $l + 1$  layer. This is essentially propagating the error backwards through the network, hence the need for the transpose of the weights. The third equation relates the change in the cost function to the bias values:

$$\frac{\delta C}{\delta b_j^l} = \delta_j^l \quad (\text{A.6})$$

That is simply that the rate of change of the cost function due to a change in bias values is simply related to the nodes inputs  $z_j^l$ .

Finally, the fourth back propagation equation relates the change in the cost function to the weights :

$$\frac{\delta C}{\delta w_{jk}^l} = a_k^{l-1} \delta_j^l \quad (\text{A.7})$$

This equation defines how to compute the partial derivatives of the cost function with respect to the weights. This partial derivative is a product of the activation function and the cost functions rate of change with respect to the activation function inputs. The four equations define the whole of back propagation. In the end, these are simply the application of the chain rule in an organized way. A pseudo code implementation is given in Listing 6.

```

# evaluate neural network for input data x
a = model.evaluate( x )
# compute the z(l) = w(l) * a(l-1) + b(l)
for l in layers:
    for k in layer_sizes[l-1]:
        for j in layer_sizes[l]:
            z[l] = w[l][j][k] * a[l] + b[l]
            a[l] = sigmoid( z[l] )
# compute the errors of the output layer
delta[:-1] = cost_function_derivative_az( a[:-1], z[:-1] )
# propagate the errors back through the network
for l in range( n_layers - 2, 1 ) :
    for k in layer_sizes[l-1]:
        for j in layer_sizes[l]:
            delta[l] = transpose( w[l+1][k][j] )
            delta[l+1][k][j] sigmoid_derivative( z[l] )
# update the weights
for l in layers :
    w[l] = w[l] - cost_function_derivative_a( transpose(a[l-1]) )
    b[l] = b[l] - cost_function_derivative_z( delta[l] )

```

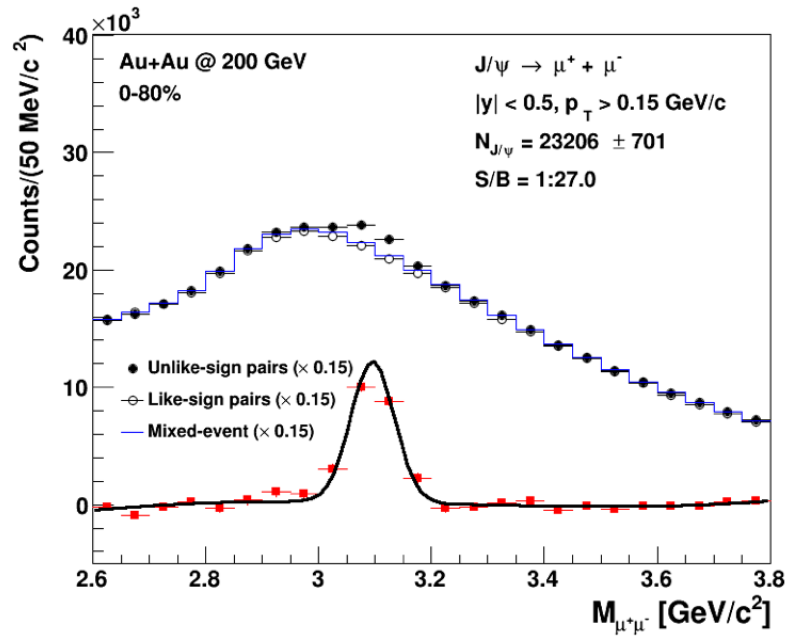
**Listing 6:** A python-like pseudo code implementation of the back propagation algorithm.

## Appendix B

### Muon Purity in Au+Au Collisions at $\sqrt{s_{NN}} = 200$ GeV

One of the goals of this project was to measure the dimuon invariant mass spectra in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. However, the purity was found to be prohibitively low in the Au+Au datasets. The techniques used for measuring the purity in  $p+p$  and  $p$ +Au collisions was employed for the Au+Au dataset. Despite re-using the DNN model trained for  $p+p$  data, the DNN described the Au+Au data surprisingly well for tracks with  $p_T > 1.5$  GeV/ $c$ . Plots summarizing the purity in 40 – 80% central and 20 – 40% central Au+Au collisions are shown in Fig. 6.10 and Fig. 6.11 respectively. The purity fits for tracks with  $p_T > 1.5$  GeV/ $c$  are shown in Fig. B.3 – B.11 for  $p+p$ , 40 – 80% Au+Au, and 20 – 40% Au+Au collisions. The  $\chi^2/ndf$  in Au+Au data is worse, largely because the statistics in the data are significantly better than the DNN templates. The uncertainty in the DNN templates (due to poor statistics) was not included when calculating the  $\chi^2/ndf$ .

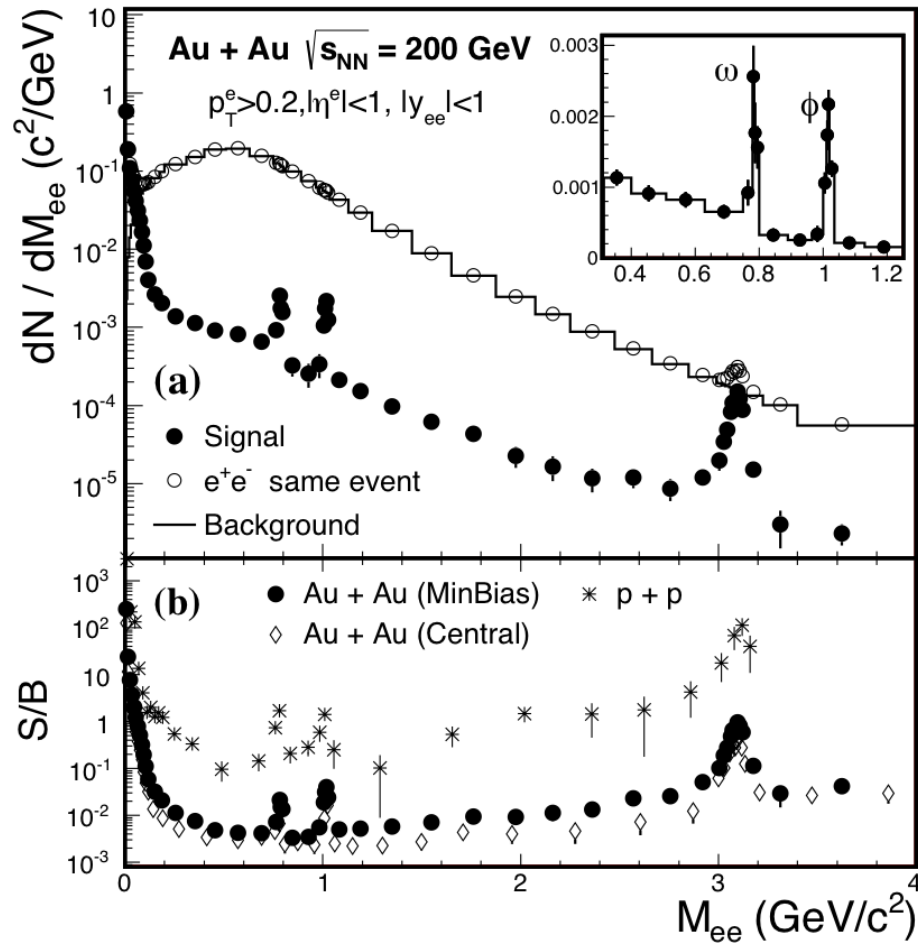
While the purity values in 40 – 80% central Au+Au collisions are not substantially worse than in  $p+p$ , the low purity is coupled with a smaller number of events, makes the analysis in peripheral Au+Au collisions impractical. The average purity in the 40 – 80% central Au+Au sample is  $\sim 60\%$  at an DNN cut giving an average signal efficiency of  $\sim 10\%$ . The purity levels in 20 – 40% central Au+Au collisions look surprisingly high at first glance. However, it should be noted that these purity numbers include secondary muons from pion and kaon decays (since they are inseparable). The purity values for inclusive muon candidates (without any DNN cut) are approximately consistent with the values expected for only secondary muon sources - i.e. consistent with zero signal muons. The values listed in Table 5.2 for the fractions of pions and kaons that will decay, on average, within the MTD, are used to come to this conclusion. Since the  $J/\psi$  has been extracted from the Au+Au data set, there must still be a small number of signal muons present. Though even for the  $J/\psi$  the integrated (over  $p_T$  and centrality) signal to background ratio is  $\sim 1/27$  - less than the significance of the  $J/\psi$  from the dielectron channel, which was  $\sim 1/2$ . Guided by the production predicted by



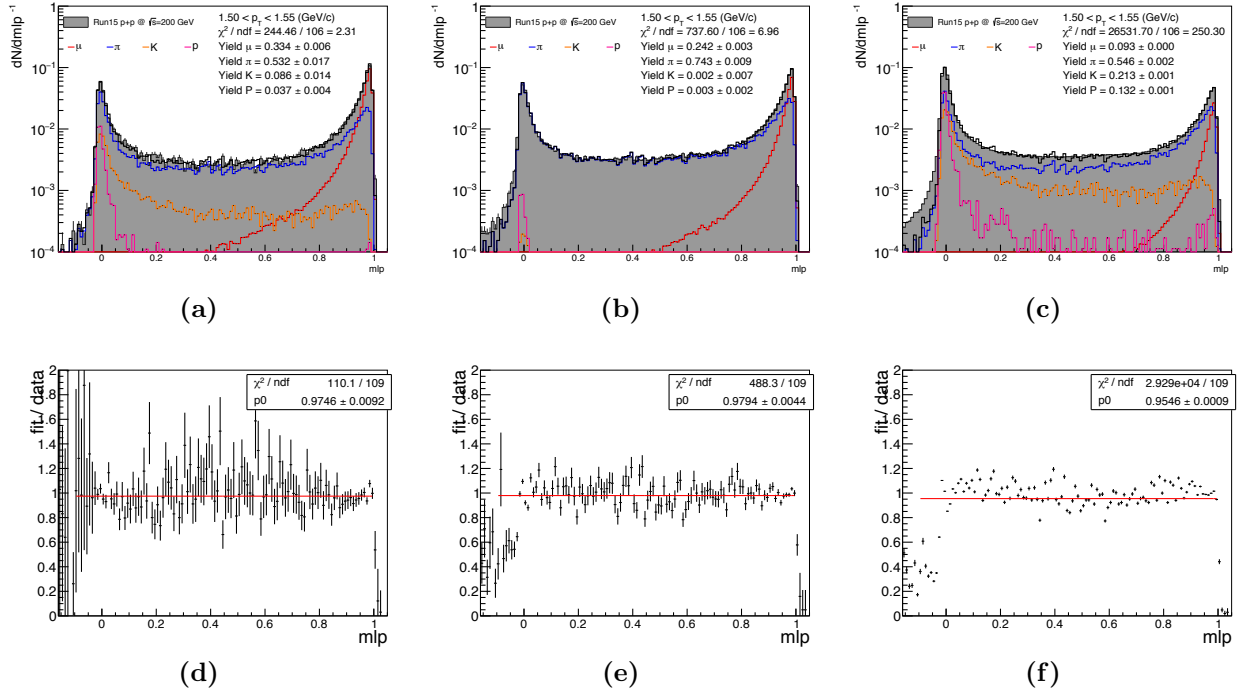
**Figure B.1 :** The  $J/\psi$  yield in 0 – 80% Au+Au collisions measured through the dimuon channel. The S/B ratio is  $\sim 1/27$ .

the hadronic cocktail, this would suggest that the signal to background ratio in the low and intermediate mass range is  $\sim 1/300 - 1/3,000$ . This means that the background would need to be measured at better than  $1/30,000$  to achieve a 10% uncertainty. Unfortunately, the like-sign ratio method is not capable of reaching that precision simply due to the affect of statistical fluctuations on the fitting. For this reason, analysis of the low and intermediate dimuon invariant mass spectra in Au+Au is not possible.

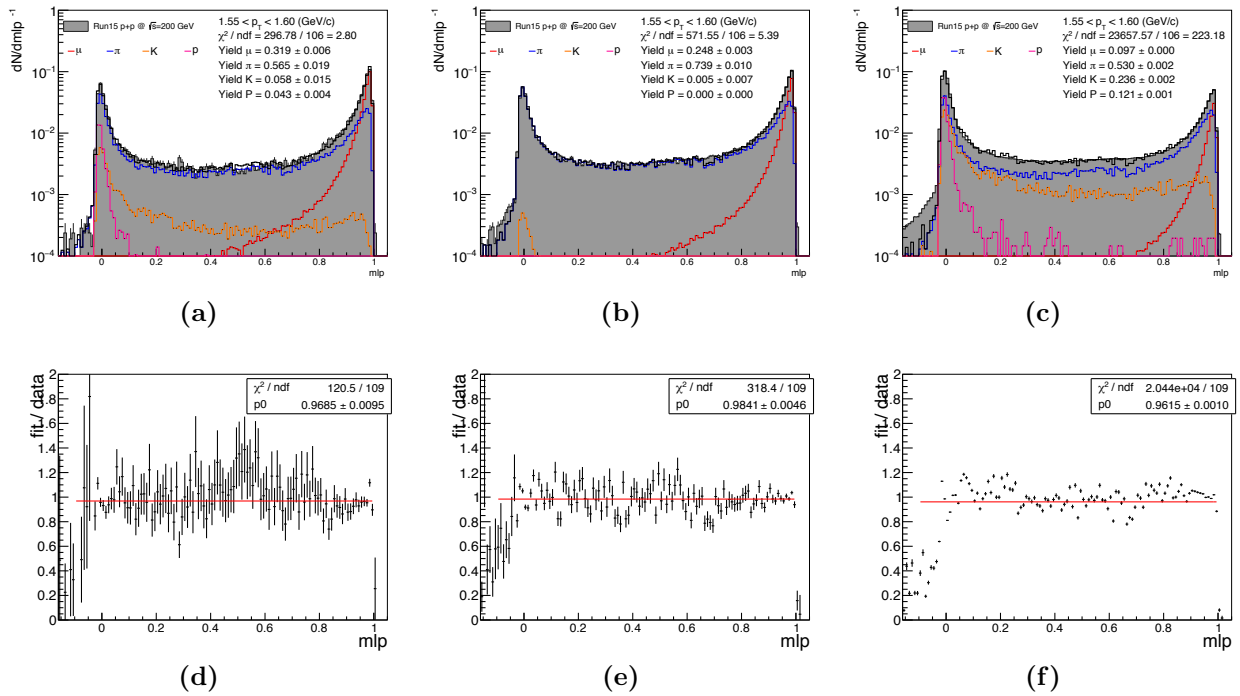




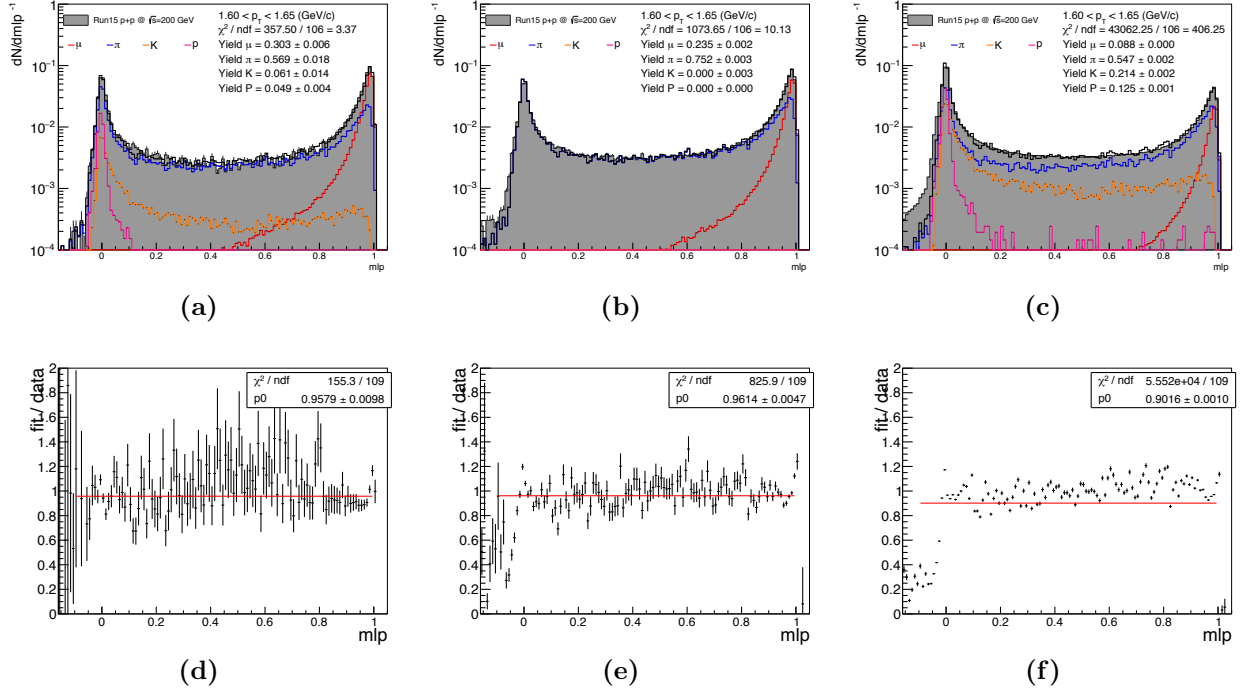
**Figure B.2 :** The  $e^+e^-$  invariant mass spectra in 0 – 80 Au+Au collisions. The lower panel shows the S/B ratio. In the  $J/\psi$  region the S/B reaches  $\sim 1/2 - 1$  [88].



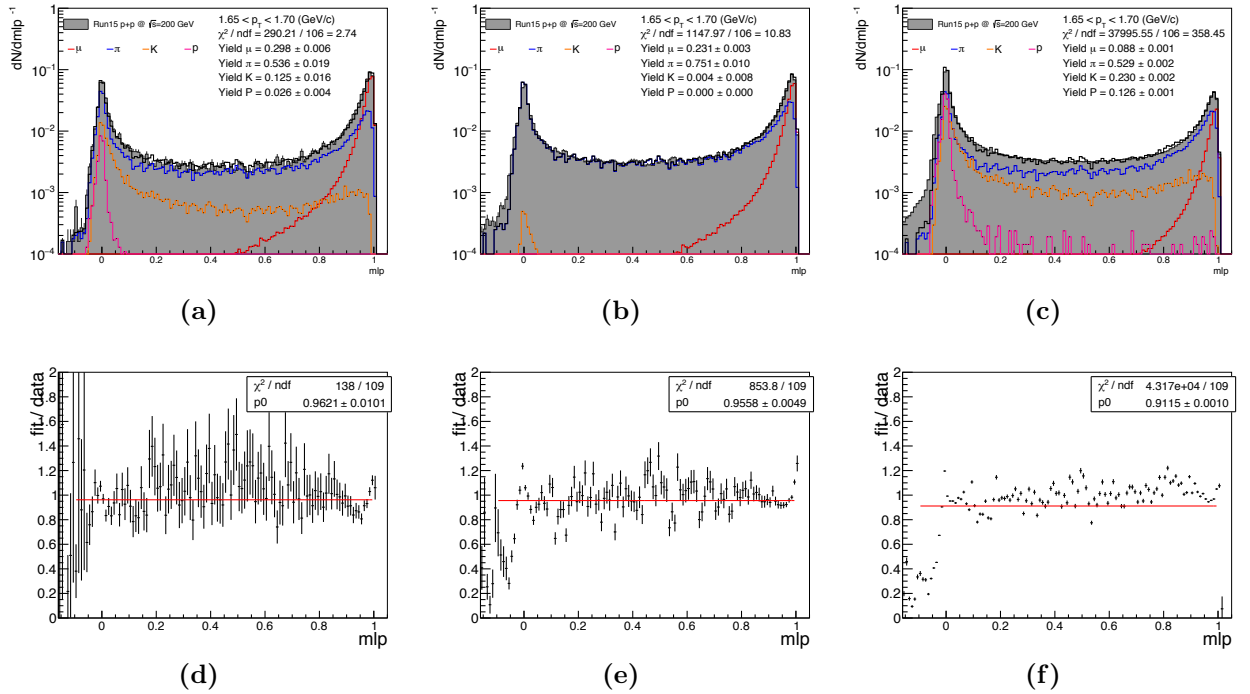
**Figure B.3 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



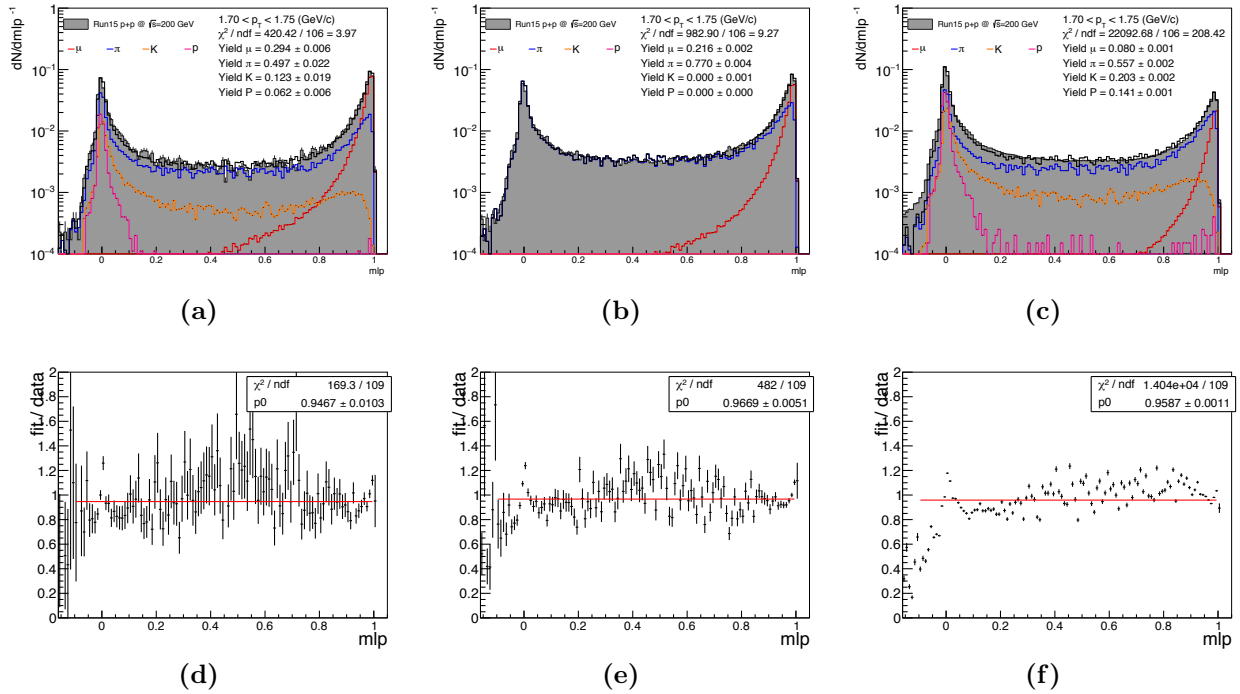
**Figure B.4 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



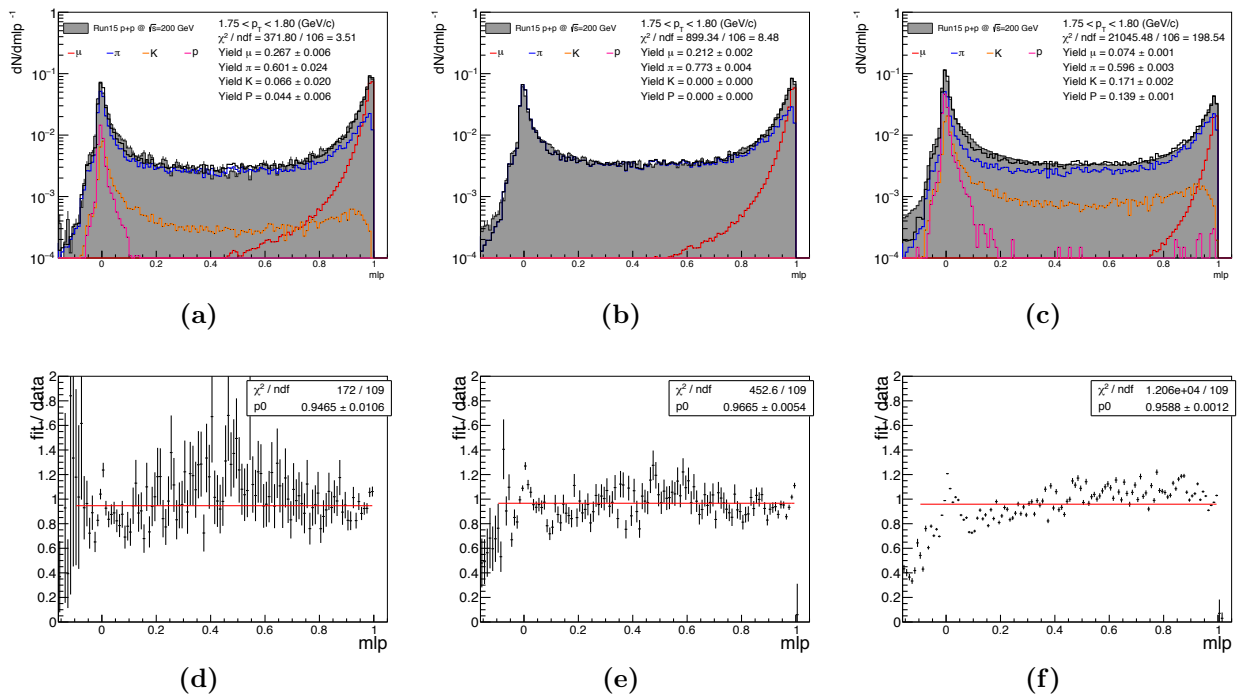
**Figure B.5 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



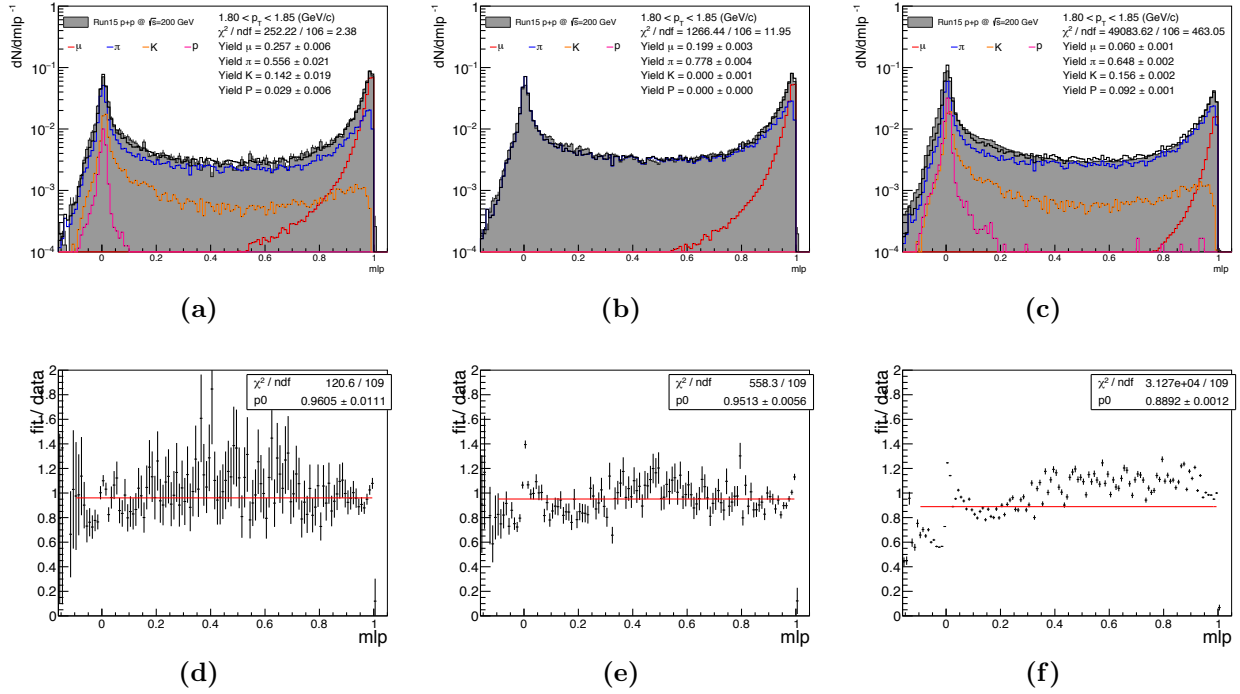
**Figure B.6 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



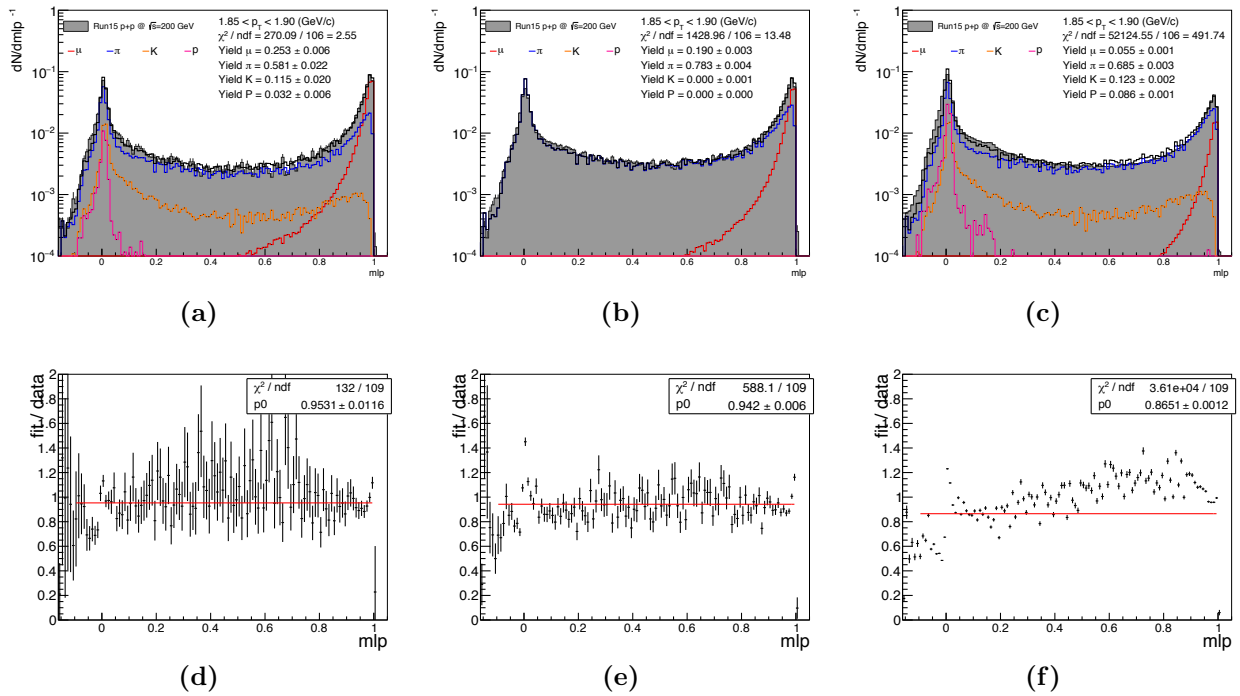
**Figure B.7 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



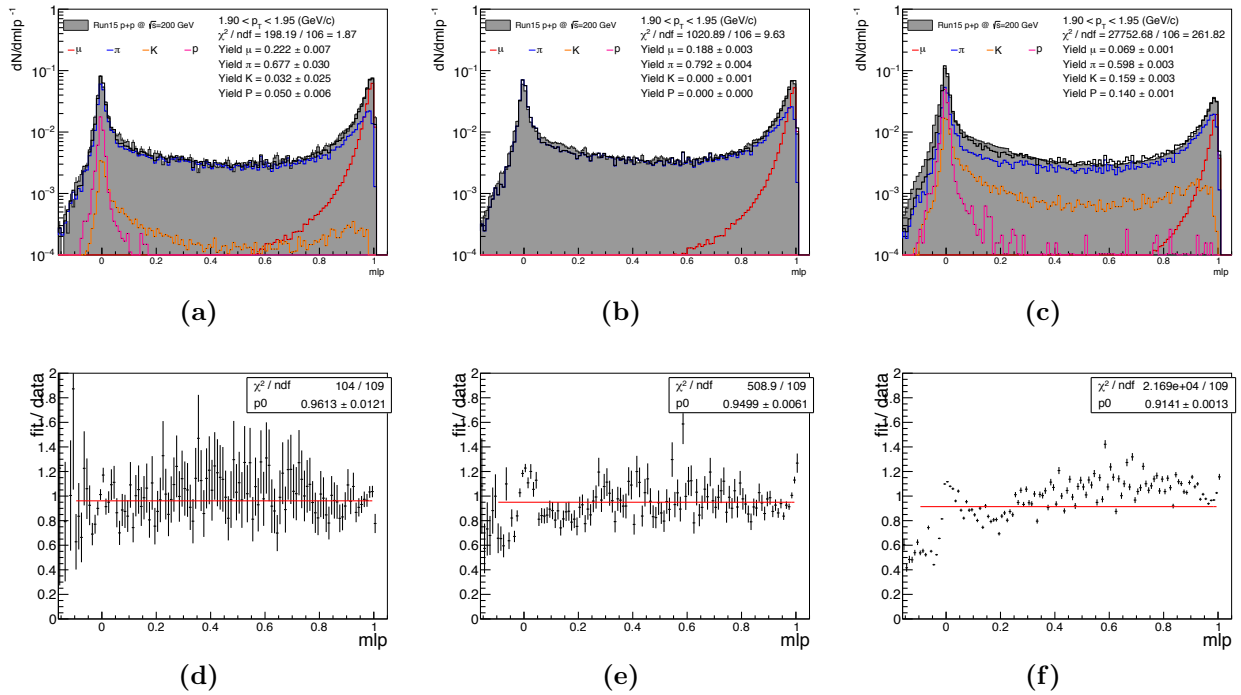
**Figure B.8 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



**Figure B.9 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



**Figure B.10 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p+p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.



**Figure B.11 :** The purity fits are shown for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au in (a), (b), and (c) respectively. The ratio of data over template fit is shown in (d), (e), and (f) for  $p + p$ , 40 – 80% central Au+Au, and 20 – 40% central Au+Au respectively.

## Bibliography

- [1] T. Matsui and H. Satz, “ $J/\psi$  suppression by quark-gluon plasma formation,” *Phys. Lett. B*, vol. 178, pp. 416–422, oct 1986.
- [2] R. Rapp and H. van Hees, “Thermal dileptons as fireball thermometer and chronometer,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 753, pp. 586–590, nov 2016.
- [3] R. Rapp and P. M. Hohler, “Is Rho-Meson Melting Compatible with Chiral Restoration?,” nov 2014.
- [4] P. W. Anderson, “More Is Different,” *Science (80-. )*, vol. 177, pp. 393–396, aug 1972.
- [5] J. C. Maxwell, “A Treatise on Electricity and Magnetism,” *Nature*, vol. 7, pp. 478–480, apr 1873.
- [6] E. Fermi, “Versuch einer Theorie der  $\beta$ -Strahlen. I,” *Zeitschrift für Phys. A Hadron. Nucl.*, vol. 88, pp. 161–177, mar 1934.
- [7] F. L. Wilson, “Fermi’s Theory of Beta Decay,” *Am. J. Phys.*, vol. 36, pp. 1150–1160, dec 1968.
- [8] Royal Swedish Academy of Sciences, “The Nobel Prize in Physics 1979,” 1979.
- [9] S. L. Glashow, “The renormalizability of vector meson interactions,” *Nucl. Phys.*, vol. 10, pp. 107–117, feb 1959.
- [10] S. Weinberger, “A MODEL OF LEPTONS\*,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, nov 1962.
- [11] A. Salam and J. C. Ward, “Weak and electromagnetic interactions,” *Nuovo Cim. Ser. 10*, vol. 11, pp. 568–577, feb 1959.

- [12] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, “The standard model of particle physics,” *Rev. Mod. Phys.*, vol. 71, no. 2, pp. S96–S111, 1999.
- [13] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, oct 1964.
- [14] A. M. Sirunyan *et al.*, “Observation of  $t\bar{t}H$  Production,” *Phys. Rev. Lett.*, vol. 120, p. 231801, jun 2018.
- [15] T. C. Collaboration, “Evidence for the direct decay of the 125 GeV Higgs boson to fermions,” *Nat. Phys.*, vol. 10, pp. 557–560, aug 2014.
- [16] B. Muller and D. K. Srivastava, “How relativistic heavy ion collisions can help us understand the universe,” pp. 1–20, jul 2004.
- [17] S. Frautschi, “Statistical Bootstrap Model of Hadrons,” *Phys. Rev. D*, vol. 3, pp. 2821–2834, jun 1971.
- [18] H. A. Bethe, “An attempt to calculate the number of energy levels of a heavy nucleus,” *Phys. Rev.*, vol. 50, pp. 332–341, aug 1936.
- [19] K. Redlich and H. Satz, “The Legacy of Rolf Hagedorn: Statistical Bootstrap and Ultimate Temperature,” *Melting Hadron. Boil. Quarks - From Hagedorn Temp. to Ultra-Relativistic Heavy-Ion Collisions Cern*, pp. 49–68, 2015.
- [20] W. Broniowski, W. Florkowski, and L. Y. Glozman, “Update of the Hagedorn mass spectrum,” *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, vol. 70, jul 2004.
- [21] M. Gell-Mann, “A schematic model of baryons and mesons,” *Phys. Lett.*, vol. 8, pp. 214–215, feb 1964.
- [22] O. Kaczmarek and F. Zantow, “Static quark-antiquark interactions in zero and finite temperature QCD: I. Heavy quark free energies, running coupling, and quarkonium binding,” *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, vol. 71, p. 114510, jun 2005.
- [23] C. Amsler *et al.*, “Review of Particle Physics,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 667, pp. 1–6, jul 2008.



- [24] R. Aaij *et al.*, “Observation of  $J/\psi$  Resonances Consistent with Pentaquark States,” *Phys. Rev. Lett.*, vol. 115, p. 072001, aug 2015.
- [25] M. Aaboud *et al.*, “Evidence for light-by-light scattering in heavy-ion collisions with the ATLAS detector at the LHC,” *Nat. Phys.*, vol. 13, pp. 852–858, aug 2017.
- [26] D. W. Duke and R. G. Roberts, “Deep inelastic scattering and asymptotic freedom. A detailed analysis and confrontation,” *Nucl. Physics, Sect. B*, vol. 166, pp. 243–283, apr 1980.
- [27] Nobel Media AB 2014, “The Nobel Prize in Physics 2004,” 2004.
- [28] C. Bernard *et al.*, “Continuum limit of lattice QCD with staggered quarks in the quenched approximation: A critical role for the chiral extrapolation,” *Phys. Rev. Lett.*, vol. 81, pp. 3087–3090, oct 1998.
- [29] J. Sadeghi, B. Pourhassan, and S. Heshmatian, “Application of AdS/CFT in quark-gluon plasma,” sep 2013.
- [30] A. Bernamonti and R. Peschanski, “Time-dependent AdS/CFT correspondence and the Quark-Gluon plasma,” *Nucl. Phys. B - Proc. Suppl.*, vol. 216, pp. 94–120, jul 2011.
- [31] R. A. Janik, “AdS/CFT and the Dynamics of Quark-Gluon Plasma,” *Prog. Theor. Phys. Suppl.*, vol. 186, pp. 534–539, oct 2010.
- [32] D. T. Son and A. O. Starinets, “Minkowski-space correlators in AdS/CFT correspondence: recipe and applications,” *J. High Energy Phys.*, vol. 2002, pp. 042–042, sep 2002.
- [33] G. Policastro, D. T. Son, and A. O. Starinets, “From AdS/CFT correspondence to hydrodynamics,” *J. High Energy Phys.*, vol. 2002, pp. 043–043, sep 2002.
- [34] J. Manninen and F. Becattini, “Chemical freeze-out in ultrarelativistic heavy ion collisions at  $\sqrt{s_{NN}}=130$  and 200 GeV,” *Phys. Rev. C - Nucl. Phys.*, vol. 78, no. 5, p. 54901, 2008.
- [35] M. Cheng *et al.*, “QCD equation of state with almost physical quark masses,” *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, vol. 77, p. 014511, jan 2008.

- [36] U. W. Heinz, “Concepts of Heavy-Ion Physics,” jul 2004.
- [37] Y. Akiba *et al.*, “The Hot QCD White Paper: Exploring the Phases of QCD at RHIC and the LHC,” feb 2015.
- [38] B. Müller, “From quark-gluon plasma to the perfect liquid,” in *Acta Phys. Pol. B*, vol. 38, pp. 3705–3729, oct 2007.
- [39] P. K. Kovtun, D. T. Son, and A. O. Starinets, “Viscosity in strongly interacting quantum field theories from black hole physics,” *Phys. Rev. Lett.*, vol. 94, p. 111601, mar 2005.
- [40] STAR Collaboration and J. Adams, “Experimental and Theoretical Challenges in the Search for the Quark Gluon Plasma: The STAR Collaboration’s Critical Assessment of the Evidence from RHIC Collisions,” jan 2005.
- [41] PHENIX Collaboration and K. Adcox, “Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration,” *Nucl. Phys. A*, vol. 757, pp. 184–283, oct 2005.
- [42] B. B. Back *et al.*, “The PHOBOS Perspective on Discoveries at RHIC,” *Nucl. Phys. A* 757, oct 2004.
- [43] I. Arsene *et al.*, “Quark-gluon plasma and color glass condensate at RHIC? The perspective from the BRAHMS experiment,” aug 2005.
- [44] M. A. Stephanov, “QCD phase diagram: an overview,” dec 2006.
- [45] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabó, “The order of the quantum chromodynamics transition predicted by the standard model of particle physics,” *Nature*, vol. 443, pp. 675–678, oct 2006.
- [46] F. R. Brown *et al.*, “On the existence of a phase transition for QCD with three light quarks,” *Phys. Rev. Lett.*, vol. 65, pp. 2491–2494, nov 1990.
- [47] S. Borsanyi *et al.*, “Transition temperature and the equation of state from lattice QCD, Wuppertal Budapest results,” in *Acta Phys. Pol. B, Proc. Suppl.*, vol. 4, pp. 593–601, sep 2011.

- [48] P. De Forcrand, J. Langelage, O. Philipsen, and W. Unger, “Lattice QCD phase diagram in and away from the strong coupling limit,” *Phys. Rev. Lett.*, vol. 113, p. 152002, oct 2014.
- [49] D. H. Rischke, “The quark-gluon plasma in equilibrium,” mar 2004.
- [50] T. M. Schwarz, S. P. Klevansky, and G. Papp, “The phase diagram and bulk thermodynamical quantities in the NJL model at finite temperature and density,” *Phys. Rev. C*, vol. 60, p. 11, oct 1999.
- [51] L. Álvarez-Gaumé and J. Ellis, “Eyes on a prize particle,” jan 2011.
- [52] B. Mueller, “Physics and Signatures of the Quark - Gluon Plasma,” *Reports on Progress in Physics*, vol. 58, pp. 611–636, June 1995. arXiv: nucl-th/9410005.
- [53] M. G. Alford, A. Schmitt, K. Rajagopal, and T. Schäfer, “Color superconductivity in dense quark matter,” *Rev. Mod. Phys.*, vol. 80, pp. 1455–1515, sep 2008.
- [54] J. M. Lattimer and M. Prakash, “Neutron Star Structure and the Equation of State,” feb 2000.
- [55] J. M. Lattimer and M. Prakash, “Neutron star observations: Prognosis for equation of state constraints,” dec 2007.
- [56] R. Snellings, “Elliptic flow: A brief review,” feb 2011.
- [57] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, “Glauber Modeling in High Energy Nuclear Collisions,” *Annu. Rev. Nucl. Part. Sci.*, vol. 57, pp. 205–243, nov 2007.
- [58] C. Shen and B. Schenke, “Dynamical initial-state model for relativistic heavy-ion collisions,” *Phys. Rev. C*, vol. 97, no. 2, p. 24907, 2018.
- [59] R. Glauber and G. Matthiae, “High-energy scattering of protons by nuclei,” *Nucl. Phys. B*, vol. 21, pp. 135–157, jan 1970.
- [60] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, “The Color Glass Condensate,” *Annu. Rev. Nucl. Part. Sci.*, vol. 60, pp. 463–89, feb 2010.

- [61] B. Schenke and S. Schlichting, “3-D Glasma initial state for relativistic heavy ion collisions,” may 2016.
- [62] B. Schenke, P. Tribedy, and R. Venugopalan, “Fluctuating glasma initial conditions and flow in heavy ion collisions,” *Phys. Rev. Lett.*, vol. 108, feb 2012.
- [63] U. W. Heinz, “Concepts of Heavy-Ion Physics,” *arXiv:hep-ph/0407360*, July 2004. arXiv: hep-ph/0407360.
- [64] H. Song and U. W. Heinz, “Causal viscous hydrodynamics in 2+1 dimensions for relativistic heavy-ion collisions,” *Physical Review C*, vol. 77, June 2008. arXiv: 0712.3715.
- [65] B. Hippolyte, “HEAVY-ION SESSION: A (quick) INTRODUCTION,”
- [66] C. Shen, “Sketch of relativistic heavy-ion collisions — iEBE-VISHNU,” 2014.
- [67] K. Zhou, N. Xu, and P. Zhuang, “ production in heavy ion collisions at LHC,” *Nucl. Phys. A*, vol. 931, pp. 654–658, aug 2014.
- [68] S. Chatrchyan *et al.*, “Suppression of non-prompt J/psi, prompt J/psi, and Y(1S) in PbPb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV,” *J. High Energy Phys.*, vol. 2012, p. 63, jan 2012.
- [69] G. T. Bodwin, E. Braaten, and J. Lee, “Comparison of the color-evaporation model and the nonrelativistic QCD factorization approach in charmonium production,” *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, vol. 72, pp. 1–18, apr 2005.
- [70] C. Aidala *et al.*, “Measurements of muon pairs from open heavy flavor and Drell-Yan in p+p collisions at  $\sqrt{s}=200$  GeV,” may 2018.
- [71] P. Ilten, N. L. Rodd, J. Thaler, and M. Williams, “Disentangling heavy flavor at colliders,” *Phys. Rev. D*, vol. 96, p. 054019, sep 2017.
- [72] E. Norrbin and T. Sjöstrand, “Production and hadronization of heavy quarks,” *Eur. Phys. J. C*, vol. 17, pp. 137–161, oct 2000.
- [73] R. Averbek, “Heavy-flavor production in heavy-ion collisions and implications for the properties of hot QCD matter,” may 2013.

- [74] M. D. Cozma, C. Fuchs, E. Santini, and A. Fässler, “Dilepton production at HADES: theoretical predictions,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 640, pp. 170–175, sep 2006.
- [75] G. Agakichiev *et al.*, “Enhanced production of low-mass electron pairs in 200 GeV/nucleon S-Au collisions at the CERN super proton synchrotron,” *Phys. Rev. Lett.*, vol. 75, pp. 1272–1275, aug 1995.
- [76] G. Agakichiev *et al.*, “Systematic study of low-mass electron pair production in p-Be and p-Au collisions at 450 GeV/c,” *Eur. Phys. J. C*, vol. 4, pp. 231–247, jun 1998.
- [77] D. Adamová *et al.*, “Enhanced Production of Low-Mass Electron-Positron Pairs in 40-AGeV Pb-Au Collisions at the CERN SPS,” *Phys. Rev. Lett.*, vol. 91, p. 042301, jul 2003.
- [78] A. Angelis *et al.* and A. A. *et al.*, “Excess of continuum dimuon production at masses between threshold and the  $J/\psi$  in SW interactions at 200 GeV  $/c$  /nucleon,” *Eur. Phys. J. C*, vol. 13, pp. 433–452, apr 2000.
- [79] M. Abreu *et al.*, “Dimuon and charm production in nucleus-nucleus collisions at the CERN-SPS,” *Eur. Phys. J. C*, vol. 14, pp. 443–455, jun 2000.
- [80] M. Abreu *et al.*, “Muon pair and vector meson cross-sections in p-W and SU collisions at 200 GeV/nucleon,” *Phys. Lett. B*, vol. 368, pp. 230–238, feb 1996.
- [81] G. Agakichiev *et al.*, “Dielectron production in C12+C12 collisions at 2AGeV with the HADES spectrometer,” *Phys. Rev. Lett.*, vol. 98, no. 5, 2007.
- [82] T. H. The HADES Collaboration, T. Galatyuk, *et al.*, “Recent results from HADES on electron pair production in relativistic heavy-ion collisions,” nov 2009.
- [83] G. Agakichiev *et al.*, “Origin of the low-mass electron pair excess in light nucleus-nucleus collisions,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 690, pp. 118–122, oct 2010.
- [84] A. Adare *et al.*, “Cross section for bb production via dielectrons in,” *Phys. Rev. C*, vol. 91, p. 014907, may 2015.

- [85] PHENIX Collaboration and A. Adare, “Detailed measurement of the  $e+e-$  pair continuum in p+p and Au+Au collisions at  $\sqrt{s_{NN}}=200$  GeV and implications for direct photon production,” dec 2009.
- [86] L. Adamczyk and Others, “Dielectron Mass Spectra from Au+Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV,” *Phys. Rev. Lett.*, vol. 113, no. July, p. 22301, 2014.
- [87] L. Adamczyk *et al.*, “Di-electron spectrum at mid-rapidity in p+p collisions at  $s=200$  GeV,” *Phys. Rev. C - Nucl. Phys.*, vol. 86, p. 024906, aug 2012.
- [88] L. Adamczyk *et al.*, “Dielectron mass spectra from Au+Au collisions at  $s_{NN} = 200$  GeV,” *Phys. Rev. Lett.*, vol. 113, p. 022301, jul 2014.
- [89] L. Adamczyk *et al.*, “Energy dependence of acceptance-corrected dielectron excess mass spectrum at mid-rapidity in Au+Au collisions at 19.6 and 200 GeV, url = <https://www.sciencedirect.com/science/article/pii/S0370269315006450> <http://linkinghub.elsevier.com/retrieve/pii/S0370269315006450>, volume = 750, year = 2015,” *Phys. Lett. B*, pp. 64–71, nov.
- [90] M. Broz and f. t. A. Collaboration, “Charmonium photoproduction in ultra-peripheral p-Pb and Pb-Pb collisions at the LHC with the ALICE experiment,” *Nucl. Phys. A*, vol. 00, pp. 1–4, nov 2014.
- [91] A. Collaboration, “Dielectron and heavy-quark production in inelastic and high-multiplicity proton-proton collisions at  $\sqrt{s_{NN}} = 13$  tev,” *arXiv:1805.04407 [hep-ex, physics:nucl-ex]*, May 2018. arXiv: 1805.04407.
- [92] T. Ullrich, G. Agakishiev, R. Baur, P. Braun-Munzinger, F. Ceretto, A. Drees, S. Esumi, U. Faschingbauer, Z. Fraenkel, C. Fuchs, E. Gatti, P. Glassel, C. P. de los Heros, P. Holl, C. Jung, H. Kraner, B. Lenkeit, M. Messer, Y. Minaev, Y. Panebratsev, A. Pfeiffer, J. Rak, I. Ravinovich, S. Razin, P. Rehak, M. Sampietro, J. Schukraft, S. Shimansky, E. Sokol, H. J. Specht, J. Stachel, G. Tel-Zur, I. Tserruya, C. Voigt, C. Weber, J. P. Wessels, J. P. Wurm, and V. Yurevich, “First results from CERES/NA45 on low mass electron pair production in Pb Au collisions,” *Nucl. Phys.*, vol. A610, pp. 317C–330C, 1996.

- [93] R. Baur *et al.*, “The CERES RICH detector system,” *Nucl. Inst. Methods Phys. Res. A*, vol. 343, pp. 87–98, apr 1994.
- [94] G. E. Brown and M. Rho, “Chiral restoration in hot matter,” *Nucl. Physics, Sect. A*, vol. 590, pp. 527–530, may 1995.
- [95] R. Rapp and J. Wambach, “Low-mass dileptons at the CERN-SpS: Evidence for chiral restoration?,” *Eur. Phys. J. A*, vol. 6, pp. 415–420, jul 1999.
- [96] M. A. Mazzoni, “Dimuon and vector-meson production in p-W and S-W interactions at 200 GeV/nucleon,” *Nuclear Physics A*, vol. 566, pp. 95–102, Jan. 1994.
- [97] R. Arnaldi *et al.*, “Evidence for the production of thermal muon pairs with masses above 1 GeV/c<sup>2</sup> in 158 AGeV indium-indium collisions : NNNNA60 Collaboration,” *Eur. Phys. J. C*, vol. 59, pp. 607–623, oct 2009.
- [98] R. Arnaldi *et al.*, “First measurement of the  $\rho$  spectral function in high-energy nuclear collisions,” *Phys. Rev. Lett.*, vol. 96, p. 162302, apr 2006.
- [99] M. Abreu *et al.*, “Dimuon and charm production in nucleus-nucleus collisions at the CERN-SPS,” *Eur. Phys. J. C*, vol. 14, pp. 443–455, jun 2000.
- [100] G. Agakishiev *et al.*, “Study of dielectron production in C + C collisions at 1 A GeV,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 663, no. 1-2, pp. 43–48, 2008.
- [101] G. Agakichiev *et al.*, “The high-acceptance dielectron spectrometer HADES,” *Eur. Phys. J. A*, vol. 41, pp. 243–277, feb 2009.
- [102] K. Adcox *et al.*, “PHENIX detector overview,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 499, pp. 469–479, mar 2003.
- [103] W. Anderson *et al.*, “Design, construction, operation and performance of a Hadron Blind Detector for the PHENIX experiment,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 646, pp. 35–58, mar 2011.
- [104] Z. Fraenkel *et al.*, “A Hadron Blind Detector for the PHENIX Experiment at RHIC,” feb 2005.

- [105] L. Aphecetche *et al.*, “PHENIX calorimeter,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 499, pp. 521–536, mar 2003.
- [106] A. Adare *et al.*, “Dielectron production in Au + Au collisions at sNN =200 GeV,” *Phys. Rev. C*, vol. 93, no. 1, p. 51, 2016.
- [107] C. Aidala *et al.*, “Correlations of mumu, emu, and ee pairs in p+p collisions at sqrts=200 GeV and implications for cbarc and bbarb production mechanisms,” may 2018.
- [108] P. Collaboration, “Measurements of  $\mu\mu$  pairs from open heavy flavor and Drell-Yan in  $p + p$  collisions at  $\sqrt{s} = 200$  GeV,” *arXiv:1805.02448 [hep-ex]*, May 2018. arXiv: 1805.02448.
- [109] Y. Guo, “ $\Upsilon$  production at s N N = 200 GeV in p + p and Au + Au collisions at STAR By,” 2008.
- [110] L. Adamczyk *et al.*, “Energy dependence of acceptance-corrected dielectron excess mass spectrum at mid-rapidity in Au+Au collisions at 19.6 and 200 GeV,” *Phys. Lett. B*, vol. 750, pp. 64–71, nov 2015.
- [111] R. Rapp, “Signatures of thermal dilepton radiation at ultrarelativistic energies,” *Phys. Rev. C - Nucl. Phys.*, vol. 63, pp. 549071–5490713, apr 2001.
- [112] R. Rapp, “Dilepton spectroscopy of QCD matter at collider energies,” *Adv. High Energy Phys.*, vol. 2013, pp. 1–17, apr 2013.
- [113] W. Cassing and E. L. Bratkovskaya, “Parton-hadron-string dynamics: An off-shell transport approach for relativistic energies,” *Nucl. Phys. A*, vol. 831, pp. 215–242, jul 2009.
- [114] E. L. Bratkovskaya, W. Cassing, V. P. Konchakovski, and O. Linnyk, “Parton-Hadron-String Dynamics at relativistic collider energies,” *Nucl. Phys. A*, vol. 856, pp. 162–182, jan 2011.
- [115] S. Collaboration, “to be submitted,” *arXiv*, May 2018. arXiv:.



- [116] D. E. Groom, N. V. Mokhov, and S. I. Striganov, “Muon stopping power and range tables 10 MeV–100 TeV,” *At. Data Nucl. Data Tables*, vol. 78, no. 2, pp. 183–356, 2001.
- [117] S. Campbell, “SPHENIX: The next generation heavy ion detector at RHIC,” in *J. Phys. Conf. Ser.*, vol. 832, nov 2017.
- [118] M. Connors, “Fulfilling the RHIC mission with sPHENIX,” *J. Phys. Conf. Ser.*, vol. 736, p. 012027, aug 2016.
- [119] H. Hahn *et al.*, “The RHIC design overview,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 499, pp. 245–263, mar 2003.
- [120] K. H. Ackermann *et al.*, “STAR detector overview,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 499, pp. 624–632, mar 2003.
- [121] A. Schmah, “STAR 3D images.”
- [122] M. Anderson *et al.*, “The STAR time projection chamber: A unique tool for studying high multiplicity events at RHIC,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 499, pp. 659–678, mar 2003.
- [123] H. Bichsel, “A method to improve tracking and particle identification in TPCs and silicon detectors,” jun 2006.
- [124] W. J. Llope, J. Zhou, T. Nussbaum, G. W. Hoffmann, K. Asselta, J. D. Brandenburg, J. Butterworth, T. Camarda, W. Christie, H. J. Crawford, X. Dong, J. Engelage, G. Eppley, F. Geurts, J. Hammond, E. Judd, D. L. McDonald, C. Perkins, L. Ruan, J. Scheblein, J. J. Schambach, R. Soja, K. Xin, and C. Yang, “The STAR vertex position detector,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 759, pp. 23–28, mar 2014.
- [125] W. J. Llope *et al.*, “The TOFp/pVPD time-of-flight system for STAR,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 522, pp. 252–273, aug 2004.

- [126] F. Geurts *et al.*, “Performance of the prototype MRPC detector for STAR,” in *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 533, pp. 60–64, North-Holland, nov 2004.
- [127] F. Sauli, “Principles of Operation of Multiwire Proportional and Drift Chambers,” 1977.
- [128] Y. Haddad, I. Laktineh, G. Grenier, N. Lumb, and S. Cauwenbergh, “High rate resistive plate chamber for LHC detector upgrades,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 718, pp. 424–426, nov 2013.
- [129] W. Riegler and C. Lippmann, “The physics of Resistive Plate Chambers,” *Nucl. Instr. and Meth.*, vol. A518, p. 86, feb 2004.
- [130] A. Collaboration, “ATLAS detector and physics performance: Technical Design Report, 1,” *Design*, vol. I, no. May, pp. 95–144, 1999.
- [131] C. E. Flores, *A Systematic, Large Phase Space Study of Pion, Kaon, and Proton Production in Au+Au Heavy-Ion Collisions from the Beam Energy Scan Program at STAR*. PhD thesis, 2010.
- [132] C. Yang, X. J. Huang, C. M. Du, B. C. Huang, Z. Ahammed, A. Banerjee, P. Bhattarari, S. Biswas, B. Bowen, J. Butterworth, M. Calderón De La Barca Sánchez, H. Carson, S. Chattopadhyay, D. Cebra, H. F. Chen, J. P. Cheng, M. Codrington, G. Eppley, C. Flores, F. Geurts, G. W. Hoffmann, A. Jentsch, A. Kesich, C. Li, Y. J. Li, W. J. Llope, S. Mioduszewski, Y. Mohamed, T. Nussbaum, A. Roy, L. Ruan, J. J. Schambach, Y. J. Sun, Y. Wang, K. Xin, Z. Xu, S. Yang, and X. L. Zhu, “Calibration and performance of the STAR Muon Telescope Detector using cosmic rays,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 762, pp. 1–6, oct 2014.
- [133] L. Ruan *et al.*, “Perspectives of a mid-rapidity dimuon program at the RHIC: A novel and compact muon telescope detector,” *J. Phys. G Nucl. Part. Phys.*, vol. 36, apr 2009.
- [134] Y. J. Sun *et al.*, “New prototype multi-gap resistive plate chambers with long strips,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc.*

- Equip.*, vol. 593, pp. 307–313, may 2008.
- [135] Y. Wang *et al.*, “Performance of a new LMRPC prototype for the STAR MTD system,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 640, pp. 85–90, jun 2011.
- [136] S. Yang, “Dielectron production in  $u + u$  collisions at  $\sqrt{s_{NN}} = 193$  gev at rhic,” May 2016.
- [137] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zancarini, “GEANT3,” 1987.
- [138] G. Contin, “The STAR Heavy Flavor Tracker and Upgrade Plan,” *Nucl. Phys. A*, vol. 956, pp. 858–861, dec 2016.
- [139] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, pp. 359–366, jan 1989.
- [140] C. Debaio, “Degree of approximation by superpositions of a sigmoidal function,” *Approx. Theory its Appl.*, vol. 9, pp. 17–28, dec 1993.
- [141] J. Therhaag, “TMVA - Toolkit for multivariate data analysis,” in *AIP Conf. Proc.*, vol. 1504, pp. 1013–1016, mar 2012.
- [142] B. Efron, “Better bootstrap confidence intervals,” *J. Am. Stat. Assoc.*, vol. 82, pp. 171–185, mar 1987.
- [143] B. Efron, “Bootstrap Methods: Another Look at the Jackknife,” *Ann. Stat.*, vol. 7, pp. 1–26, jan 1979.
- [144] M. Abadi *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems,” 2016.
- [145] T. Huang, R. Ma, B. Huang, X. Huang, L. Ruan, T. Todoroki, Z. Xu, C. Yang, S. Yang, Q. Yang, Y. Yang, and W. Zha, “Muon identification with Muon Telescope Detector at the STAR experiment,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 833, pp. 88–93, oct 2016.
- [146] F. James, “MINUIT User’s Guide (Old),” *Unpublished*, p. 50, 1994.

- [147] Z. Tang, Y. Xu, L. Ruan, G. Van Buren, F. Wang, and Z. Xu, “Spectra and radial flow in relativistic heavy ion collisions with Tsallis statistics in a blast-wave description,” *Phys. Rev. C - Nucl. Phys.*, vol. 79, p. 051901, may 2009.
- [148] M. Shao, L. Yi, Z. Tang, H. Chen, C. Li, and Z. Xu, “Examination of the species and beam energy dependence of particle spectra using tsallis statistics,” *J. Phys. G Nucl. Part. Phys.*, vol. 37, dec 2010.
- [149] Z. B. Tang *et al.*, “The statistical origin of constituent-quark scaling in QGP hadronization,” *Chinese Phys. Lett.*, vol. 30, jan 2013.
- [150] K. Gallmeister, B. Kämpfer, O. P. Pavlenko, and C. Gale, “A unique parametrization of the shapes of secondary dilepton spectra observed in central heavy-ion collisions at CERN-SPS energies,” *Nucl. Phys. A*, vol. 688, pp. 939–955, oct 2001.
- [151] PHENIX Collaboration and A. Adare, “Inclusive cross section and double helicity asymmetry for  $\pi^0$  production in p+p collisions at  $\sqrt{s}=200$  GeV: Implications for the polarized gluon distribution in the proton,” *Phys. Rev. D*, vol. 76, p. 051106, sep 2007.
- [152] S. S. Adler *et al.*, “Nuclear effects on hadron production in d+Au collisions at  $\sqrt{s_{NN}}=200$  GeV revealed by comparison with p+p data,” *Phys. Rev. C - Nucl. Phys.*, vol. 74, mar 2006.
- [153] PHENIX Collaboration and S. S. Adler, “High transverse momentum eta meson production in p+p, d+Au and Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV,” pp. 1–38, nov 2006.
- [154] S. S. Adler *et al.*, “Production of  $\omega$  mesons at large transverse momenta in p+p and d+Au collisions at  $\sqrt{s_{NN}}=200$  GeV,” *Phys. Rev. C - Nucl. Phys.*, vol. 75, nov 2007.
- [155] A. Adare *et al.*, “Dilepton mass spectra in p + p collisions at  $\sqrt{s} = 200$  GeV and the contribution from open charm,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 670, pp. 313–320, feb 2009.
- [156] A. Adare *et al.*, “ $J/\psi$  Production vs Transverse Momentum and Rapidity in p+p Collisions at  $s = 200$  GeV,” *Phys. Rev. Lett.*, vol. 98, nov 2007.

- [157] P. Collaboration, “Suppressed  $\pi^0$  Production at Large Transverse Momentum in Central Au+Au Collisions at  $\sqrt{s_{NN}}=200$  GeV,” *Phys. Rev. Lett.*, vol. 91, pp. 0–5, apr 2003.
- [158] PHENIX Collaboration and S. S. Adler, “Identified Charged Particle Spectra and Yields in Au+Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV,” jul 2003.
- [159] N. M. Kroll and W. Wada, “Internal pair production associated with the emission of high-energy gamma rays,” *Phys. Rev.*, vol. 98, pp. 1355–1359, jun 1955.
- [160] A. Uras, “Measurement of the  $\eta$  and  $\omega$  Dalitz decays transition form factors in p-A collisions at 400 GeV/c with the NA60 apparatus,” in *J. Phys. Conf. Ser.*, vol. 270, aug 2011.
- [161] L. G. Landsberg, “Electromagnetic decays of light mesons,” nov 1985.
- [162] C. Terschlüsen and S. Leupold, “Electromagnetic transition form factors of light vector mesons,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 691, pp. 191–201, aug 2010.
- [163] S. P. Schneider, B. Kubis, and F. Niecknig, “transition form factors in dispersion theory,” *Phys. Rev. D*, vol. 86, p. 054013, sep 2012.
- [164] T. Todoroki and f. t. S. collaboration, “Quarkonium measurements via the di-muon decay channel in  $p + p$  and au+au collisions with the STAR experiment,” *Journal of Physics: Conference Series*, vol. 779, p. 012040, Jan. 2017. arXiv: 1612.02499.
- [165] Y. Q. Ma and R. Venugopalan, “Comprehensive description of  $J/\psi$  production in proton-proton collisions at collider energies,” *Phys. Rev. Lett.*, vol. 113, aug 2014.
- [166] H. S. Shao, H. Han, Y. Q. Ma, C. Meng, Y. J. Zhang, and K. T. Chao, “Yields and polarizations of prompt  $J/\psi$  and  $\psi(2S)$  production in hadronic collisions,” *J. High Energy Phys.*, vol. 2015, pp. 1–18, nov 2015.
- [167] Y. Q. Ma and R. Vogt, “Quarkonium production in an improved color evaporation model,” *Phys. Rev. D*, vol. 94, p. 114029, dec 2016.

- [168] T. Todoroki, “Measurements of charmonium production in p+p, p+Au, and Au+Au collisions at  $\sqrt{s_{NN}}=200$  GeV with the STAR experiment,” *Nucl. Phys. A*, vol. 967, pp. 572–575, apr 2017.
- [169] A. Adare *et al.*, “Transverse-Momentum Dependence of the J/psi Nuclear Modification in d+Au Collisions at  $\sqrt{s_{NN}}=200$  GeV,” *Phys. Rev. C*, vol. 87, p. 034904, mar 2013.
- [170] X. Zhao and R. Rapp, “Charmonium in medium: From correlators to experiment,” *Phys. Rev. C - Nucl. Phys.*, vol. 82, aug 2010.
- [171] X. Zhao and R. Rapp, “Medium modifications and production of charmonia at LHC,” *Nucl. Phys. A*, vol. 859, pp. 114–125, feb 2011.
- [172] Y. Liu, Z. Qu, N. Xu, and P. Zhuang, “J /  $\psi$  transverse momentum distribution in high energy nuclear collisions,” *Phys. Lett. Sect. B Nucl. Elem. Part. High-Energy Phys.*, vol. 678, pp. 72–76, jan 2009.
- [173] H. Probes, V. Feuillard, C. E. A. Saclay, and A. Collabora, “Charmonium production in Pb-Pb collisions measured by ALICE at the LHC,” pp. 1–31, mar 2016.
- [174] V. V. Skokov, a. Y. Illarionov, and V. D. Toneev, “Estimate of the Magnetic Field Strength in Heavy-Ion Collisions,” *Int. J. Mod. Phys. A*, vol. 24, pp. 5925–5932, jul 2009.
- [175] J. Adam *et al.*, “Measurement of an Excess in the Yield of J/psi at Very Low pT in Pb-Pb Collisions at  $\sqrt{s_{NN}}=2.76$  TeV,” *Phys. Rev. Lett.*, vol. 116, sep 2016.
- [176] STAR Collaboration, “Bulk Properties of the Medium Produced in Relativistic Heavy-Ion Collisions from the Beam Energy Scan Program,” *Physical Review C*, vol. 96, Oct. 2017. arXiv: 1701.07065.
- [177] STAR Collaboration, “Energy dependence of acceptance-corrected dielectron excess mass spectrum at mid-rapidity in au+au collisions at 19.6 and 200 gev,” *Physics Letters B*, vol. 750, pp. 64–71, Nov. 2015. arXiv: 1501.05341.
- [178] M. A. Nielsen, “Neural Networks and Deep Learning,” 2015.