

DRAFT

be precise--

true

signal events, even with the improved triggering. The number of $W \rightarrow \mu$ events produced over the 2013 data set number in the hundreds, while the total number of recorded events is approximately 15 billion.

This leads to the substantial problem of fishing out the appropriate physics events from the 15 billion event haystack. Why, I hear you ask, don't we just have a detector which can record only the W Boson events of interest?

Because, as a multipurpose spectrometer, PHENIX must be ready to take all kinds of data, and satisfy many experimental requirements, in addition to fitting a lot of functionality into a relatively tight budget, as is common for federally funded research. Although this measurement would have been made much simpler with a forward calorimeter, we can't simply build and install a calorimeter the moment that an analysis would benefit from its presence.

So, instead, we must rely on our ingenuity and deep understanding of the data set, to tease out the results we want to measure.

Think about
this..

how will THIS
measurement benefit
for a calorimeter?
 $N \rightarrow M$?

5.2 Raw Data to Reconstructed Parameters

Any time a PHENIX trigger condition is satisfied, all of the information recorded by the PHENIX spectrometer are read out from temporary on-detector memory, and fed into a data stream that eventually is archived as a 'PHENIX Raw Data File Format' or PRDFF.

PRDFF data is hierarchical, first being organized by event-type, and then organized by packet-type. There are many event types - 'DATAEVENTS' typically carry the information relevant to a physics analysis, whereas other event-types carry very important QA information for determining the status of the RHIC apparatus, the beam, polarization, and PHENIX performance.

Every packet has a header, which contains general information such as what the packet contains, and in what order that packet was received. Every packet recorded can be associated with a unique event-sequence number, which specifies roughly the order in which the event owning the packet was received by the DAQ. Within a given run number, an event-number is guaranteed to be unique. The complexity of the packet is limited by the bandwidth available to move data off PHENIX onto other storage, and the buffers/reconstruction ability of the front end electronics modules built onto PHENIX sub-

DRAFT

systems. PHENIX archives data from the DAQ at a rate of approximately 700 Megabytes per second - or one compact disk.

Generally, raw PHENIX data is too complex to use straight-away, because minimal to no reconstruction of physical properties for a certain event is done, due to hardware limitations and time limitations - some of this raw data is often directly used in triggering decisions, which must be made once every 106 nanoseconds or faster (the bunch crossing frequency).

The raw data collected from PHENIX undergoes a process called "Data Production", where physical parameters are reconstructed from the simpler raw data. Raw data could take any form - for example - which cathode strips were activated in an event in the muon tracker, or the number of photons counted in a photomultiplier tube. This information is often combined with extensive survey information about the geometry of a given detector, the known magnetic field in a detector, to reconstruct quantities such as momentum, or deposited energy.

Once reconstruction has finished in a Data Production, the data are then repackaged into ROOT files, often times internally structured into custom output objects which are associated with a ~~various~~ ^{specific} detector. These output objects are simply custom ~~written~~ C++ classes which have a serialization scheme, which have libraries and dictionaries compiled that allow for them to be serialized into ROOT's file format.

For the purposes of this analysis, all data has been reconstructed and serialized into a specific type of output object called a 'picoDST' or even more concisely, 'pDST'. This name, like many others in PHENIX has historical context: DST stands for 'Data Summary Tape' hearkening back to the days when data was stored primarily on magnetic tape (it is still archived on magnetic tape!), and 'pico' because of its relatively small disk-space requirement, compared to 'nanoDST' files or simply 'DST' files. I'm not making this up, I swear!

↳ which contain more
grainy information

5.3 Choosing Analysis Variables

Even data reduced to the point of a pDST still is much more complicated and comprehensive than what is needed for this analysis - there are thousands of variables relating to reconstruction parameters. We only need a handful of variables for this analysis, summarized on Tables 5.1, 5.2, 5.3 and 5.4. When Cartesian coordinates are referenced,

↳ You miss (and perhaps missed) an important point here.
the thousands of variables sounds like we don't know what we are doing. The reason is we didn't in the beginning and all these variables were used to see if they were useful, or check the "used" variables in the analysis. That's why there are 1000s...

DRAFT

implicitly, the reference frame is the PHENIX Coordinate system (Figure 5.2).

As you can probably guess, the only variables which are truly relevant to this analysis need to be relevant to understanding two questions:

1. Is this reconstructed muon track the result of a real W -Boson Decay?
2. What is the polarization of the two colliding protons for every recorded collision?

Incident To properly answer these questions, we need to comprehensively understand what processes are capable of producing muons, as well as whether or not our detector can be 'tricked' by signals which look like muons, but really aren't. Secondly, we need a means of recovering the proton spin polarization for each colliding bunch-pair.

Polarization recovery is straight-forward - we already have mechanisms in the data stream which number and track the colliding bunch pairs. We additionally have well defined spin patterns which are applied to the 120 bunches, the same way every time this pattern is applied to the fill. As discussed previously, we have good QA apparatuses in place to ensure the advertised spin pattern is the same as the delivered spin pattern. Since polarization patterns do not typically change in a standard physics beam fill (if they do, alarms are raised and the data is typically discarded), all that is needed is to associate a PHENIX run number, with a RHIC fill number, and then look up the spin pattern, which in effect, is stored in a database call. Of course, the overall beam polarization percentage is an important factor, which dilutes any spin asymmetry, but this is taken into account in the final spin database QA analysis [66]. *which is discussed in Section X ... or detail here*

This leaves us with the first question, and the difficulty in answering this question is essentially that it is challenging to differentiate signal $W \rightarrow \mu$ events from other $X \rightarrow \mu$ events, or even events which look like muons, but are really due to incorrect track reconstruction.

Therefore, the thrust of the Data Analysis portion of this work is really just to tease apart the real W -genic muons, from all other muon candidates. To this requires some substantial feature engineering, and creating some statistical models, as well as a means of evaluating the performance of these statistical models, which is difficult because validating any statistical differentiation technique (aka machine learning technique) requires a labeled

Polarization recovery is relatively straightforward. Each event is uniquely mapped to a specific colliding bunch (out of 120).⁸⁸ In turn, known 'spin patterns' are applied to each fill, which maps polarization directly to each bunch.

DRAFT

*Appendix, just use picture to
describe*

Name	Description
Run_Number	A unique number identifying a run in a RHIC fill for PHENIX
Evt_Number	A unique number within a single run identifying the approximate order an event was taken.
Evt_bbcZ	The event z-vertex calculated by the BBC
triggerbit	The result of a bit-wise 'OR' applied to all 32-bit trigger bits which fired
clockcross	The bunch number of the two colliding bunches [0 – 119]. Required to look up the spin polarization, along with Run_Number

Table 5.1: Variables characterizing events overall

Name	(Unit) Description
Evt_Nmu	The number of muon tracks reconstructed for a given event
charge	($\pm e$) The charge associated with a reconstructed muon track
p_z	(GeV) The z-momentum associated with the muon track
p	(GeV) The total momentum of a charged track
χ^2	The result of the Kalman fitter reconstructing the track
lastGap	The last gap in the Muon Tracker which was activated (there are 4)
η	The rapidity of the track
ϕ	(rad) The azimuthal position angle the track makes relative to the x-axis
DG0	(cm) A Track matching variable (matching between MuID and MuTR) associated with the MuID road, at MuID station 3.
DDG0	(degree) The opening angle between the MuID track road, and the MuTr projection onto the MuID
x_{Sta_i}	(cm) The x-coordinate of the track at Station i , $i \in 1, 2, 3$ of the MuTr
y_{Sta_i}	(cm) The y-coordinate of the track at Station i , $i \in 1, 2, 3$ of the MuTr
ϕ_i	(rad) The angle the track makes with Station i , $i \in 1, 2, 3$, i.e.: $\phi_i = \tan^{-1} \left(\frac{x_i}{y_i} \right)$
θ	(rad) Azimuthal angle of track, $\tan^{-1} \left(\frac{p_T}{p_z} \right)$
DCA _z	(cm) Distance of closest approach between the z-vertex positions extracted by projecting the MuTR track z-vertex back to the BBC z-vertex
DCA _r	(cm) Distance of closest approach between the track and beam axis

Table 5.2: Muon tracker variables. Generally, this data set is indexed on a subevent level, where one event will contain all reconstructed muon tracks seen for that event.

DRAFT

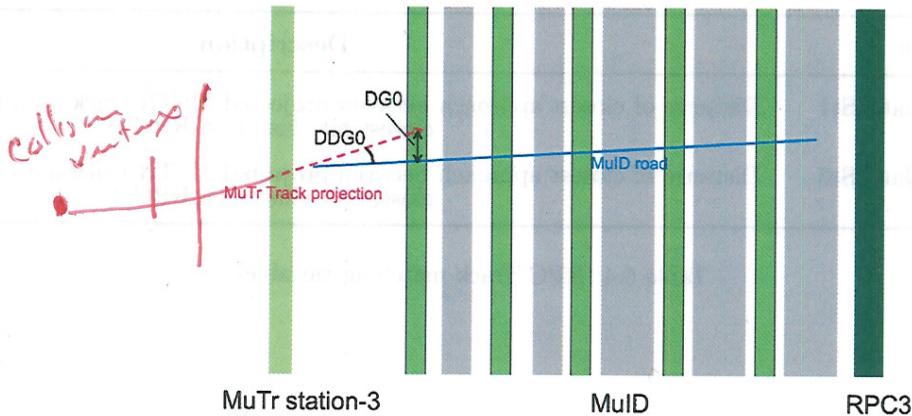


Figure 5.1: A schematic representation of the matching variables, DG0 and DDG0 at the intersection between the Muon Tracker and Muon Identifier [67]

Name	Description
$f_{vtx_{d\phi}}$	The ϕ residual between MuTR track and FVTX track
$f_{vtx_{d\theta}}$	The θ residual between the MuTR track and FVTX track
$f_{vtx_{dr}}$	The radial residual between the MuTR track and the FVTX track
$f_{vtx_{conebits}}$	The number of FVTX clusters inside a cone around the track defined by: $0.04\text{rad} < dR < 0.52\text{rad}$ where $dR = \sqrt{d\eta^2 + d\phi^2}$

Table 5.3: A summary of the variables reconstructed from FVTX raw data [68].

Appendix
data set, and we intrinsically do not possess this, since otherwise, this analysis would not need to be done.

The engineered variables in this analysis are dw_{13} , dw_{23} , $d\phi_{13}$, and $d\phi_{23}$. These variables are calculated from reconstructed physics data. They play an important role in our extraction of the signal to background ratio. The $d\phi_{ij}$ variables represent the difference in azimuthal angle observed at the MuTR station i and j respectively. dw_{ij} is constructed from $d\phi_{ij}$ as follows:

$$dw_{ij} = p_T \times \sin(\theta) \times d\phi_{ij} \quad (5.1)$$

DRAFT

Name	Description
RpcMatchSt1	Distance of closest approach between projected MuTR track onto the RPC 1 and the closest hit cluster on RPC 1
RpcMatchSt3	Distance of closest approach between projected MuTR track onto the RPC 3 and the closest hit cluster on RPC 3

Table 5.4: RPC Track matching variables

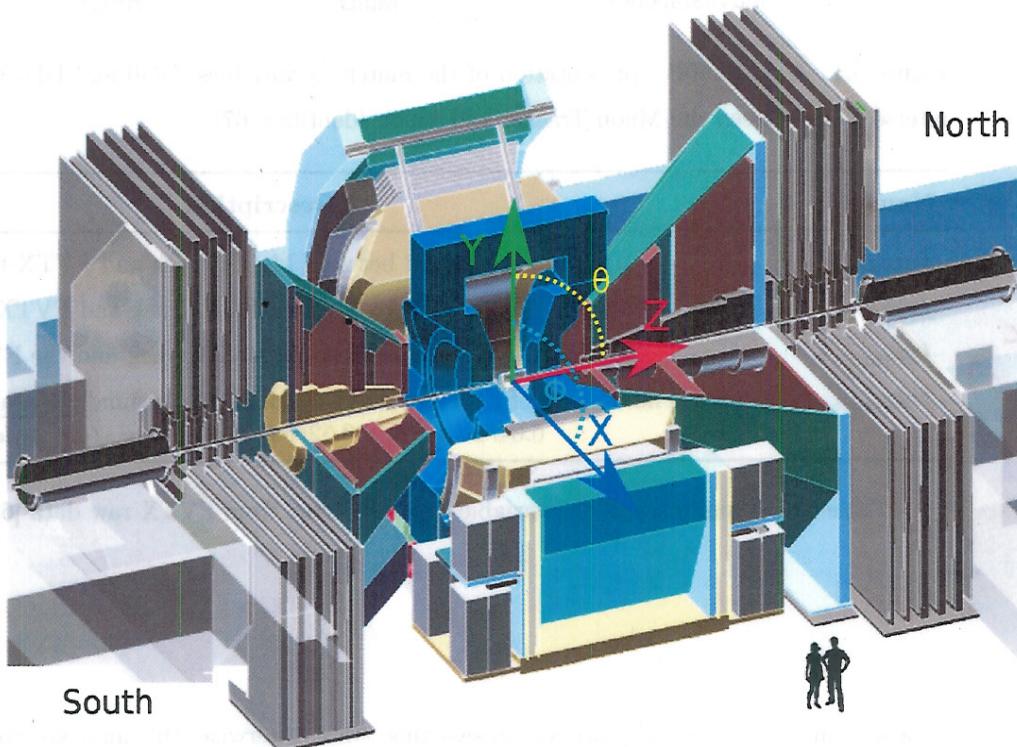


Figure 5.2: The PHENIX coordinate system is shown (RGB arrows) at the center of the nominal interaction point within PHENIX, the origin, in this quarter-cutaway drawing. The small black figures are actually miniaturized human beings, the PHENIX detector is very small - this is a full scale drawing of PHENIX. Shown: the x, y, and z coordinates, as well as the azimuthal coordinate, θ and polar coordinate ϕ [69]

Halie

DRAFT

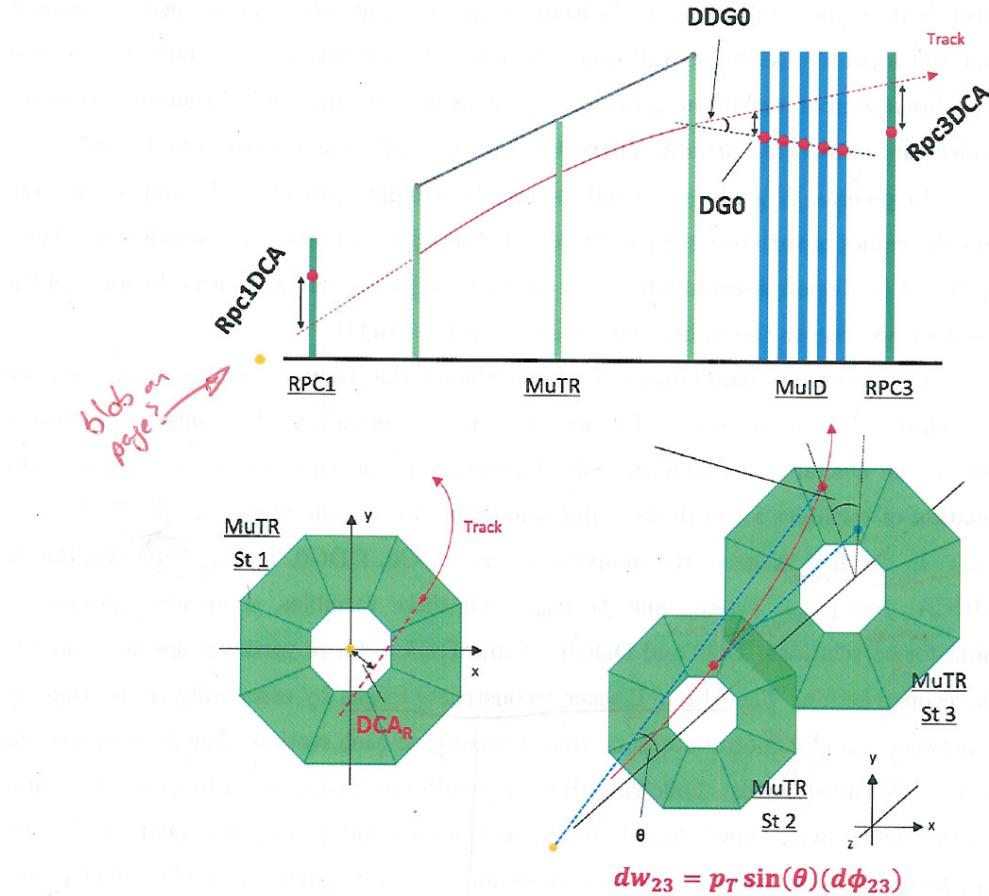


Figure 5.3: A nice summary of discriminating kinematic variables reproduced with permission from Dr. Chong Kim. We see the MuTR tracking planes in green, and a muon track penetrating the planes in red, and reference coordinates in the lower right-hand corner. The geometric relationship between the roads, reconstructed track are shown in the annotations.

ref

It's all backwards, start with simplest variables
(DGO, DDG0, ... the derived variables)

DRAFT

While ϕ_i is calculated from the x and y coordinates at Station i:

$$\phi_i = \tan^{-1} \left(\frac{y_{Sta_i}}{x_{Sta_i}} \right) \quad (5.2)$$

A common theme amongst these variables is that they should help us distinguish between high momentum muon tracks from W-Bosons, and other muon tracks. The hope is that the muon tracks from W-Bosons are kinematically restricted to have a relatively narrow momentum distribution in the forward kinematic regime, and so therefore, tracking variables can be used to partially differentiate between signal and background events.

In general, W-genic events will be mostly straight, geometrically, and so this constrains the values of variables such as DCA_r substantially, and other variables less so. Thus, dw_{23} should be a good discriminator, as it depends on p_T and the azimuthal bending of the charged tracks, due to the radial magnetic field in the MuTR.

Our secondary requirement of our variables is that they are relatively uncorrelated with each-other, to leave plenty of room for statistical modeling. Ultimately, we chose a subset of the available tracking variables to carry out the analysis, in two stages. The correlation of variables for both data and simulation are summarized in Table ??.

In the first stage of the analysis, we use: DGO, DDG0, DCA_r , χ^2 , Rpc1DCA, Rpc3DCA, $fvtx_{dr \times d\theta}$, $fvtx_{d\phi}$, and $fvtx_{cone}$. Of these variables, some were grouped to account for correlations: DGO and DDG0, χ^2 and DCA_r . These variables are all related to track reconstruction. The Muon Tracker reconstructs tracks by essentially connecting the dots between x and y coordinate 'hits' that it records at each station. The lines connecting these hits are called 'roads'. Following this, the roads and hits are used to generate a curve fit to the data, given knowledge of the muon tracker's radial magnetic field. From this curve, we extrapolate the charge and momentum, and we construct variables which codify the difference between the reconstructed curve, and the 'connect the dots' roads. The smaller these differences are, the more straight the track is, and as discussed, straightness points to higher momentum, which ultimately leads to labeling as a W-genic particle, if the momentum is in the correct range.

In the second phase of the analysis, we use dw_{23} and η primarily. Both stages of the analysis are discussed in the following sections. dw_{23} is related to track straightness as well, and is referred to as "reduced azimuthal bending". Since we're interested in forward muons, η is used as our second variable.

The muon tracker does NOT reconstruct tracks!!

This makes no sense, why would
be use it we are
forward? what about if we are
at midrapidity??

DRAFT

Generally, we are interested in recovering forward μ^+ , forward μ^- , backward μ^+ and backward μ^- . As the muon arms do not have the same rapidity coverage, we separate the data into these four categories - forward positive charged tracks, forward negatively charged tracks, backwards positively charged tracks and backward negatively charged tracks. Due to the geometry of the muon arms, the North Arm will always correspond with forward positive rapidity, whereas the South Arm will always correspond with backward, negative rapidity. I will use 'forward and backward' interchangably with 'North and Sout'.
We perform all calculations with our data set in parallel between these four conditions.

(could be written much more concisely)

The data is further subdivided based on the available track matching variables for a given event, but these subdivisions are not kept separate from the overall arm-charge separation. Some variables, such as the RPC track matching variables and the FVTX track matching variables exist for some events, but not others. We will discuss how this is managed in later sections, but this data is not generally partitioned in this way.

Yes because it's defined that way, not "due to geometry"

Charge to such as: (ie remove " $\otimes \circ$ ")
The calculations are performed on all these datasets in parallel.

Yeah, we forgot to switch on the detectors ...

Why? Kinematic range is different \Rightarrow Did you have a picture to illustrate this?

DRAFT

Chapter 6

Feature Engineering

The ultimate goal of Feature Engineering is to clean the data and transform the data with heuristic strategies that tag events coming from signal sources, as separate from events which come from our background, so that we can proceed with the calculation of the physics asymmetry.

Even with the Forward Upgrade (Section ??), our data set is still composed mostly of background events (Figure 6.1). The primary constituents of the data set are muons from the following sources:

- Hadronic Background

- The hadronic background is composed of hadrons which are produced from at the primary event vertex, and then travel into the muon arms. The hadrons then decay into muons with the hit pattern for such decay. It is a very small fraction of decays or misreconstructed as high- p_T muons, originating at the event vertex. Although the probability of this confusion is small, the large number of hadrons creates a large background.

- Muon Background
- W Signal

- The muon background is composed of processes which produce real muons which fall into a similar kinematic regime of the W-genuine muons. They are denoted by ...
- These are muons we are looking for. They come from the W boson decay, and carry information about the proton spin.

DRAFT

In the first step a likelihood event selector
is used

We will differentiate these types of data in two different steps. In the first step, we will use likelihood event selection, where we lump hadronic background and muon background together, and merely distinguish it from W-genic events. In the second stage, we will further differentiate our model for the background, drawing heavily on the data to understand the hadronic background, and providing simulations of the W-genic events as well as simulation cocktail for other muon background events. Before we discuss this differentiation, it is important to discuss the simulations, as the rest of the analysis hinges on the simulation of both the W genic events and the Muon Background events.

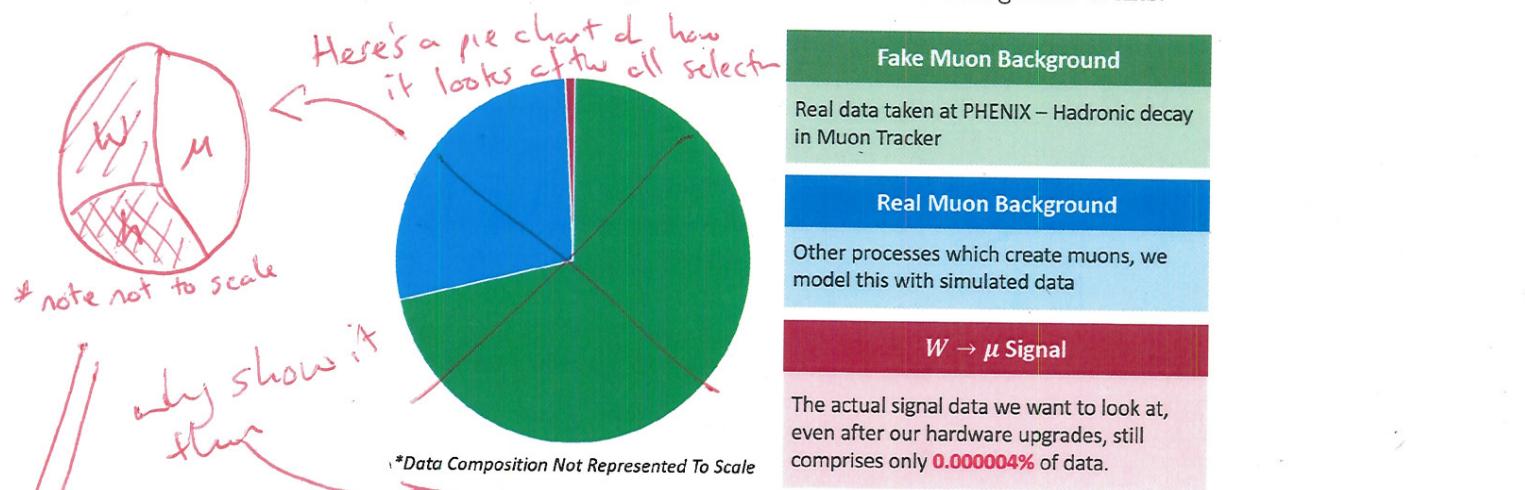


Figure 6.1: A cartoon of the dataset composition. The data, even after the Forward Upgrade, is mostly composed of hadronic background, which has tricked our Muon Tracker.

In subsequent sections, I will discuss what we do with the variables which we have chosen to use to identify W-genic events. Because our data set is so dominated by background sources, we must rely heavily on simulations to estimate what our signal events might look like. As of the time of this writing, the analysis has not yet incorporated the simulation of hadronic background, which is quite difficult, as there are a lot of effects at play - particles which interact with the material of PHENIX itself to produce secondary and tertiary vertices, for example. However, if we can simulate accurately the signal process (the W Boson cross section and its scaling with energy is known to excruciating precision), and the muon background processes (known to similarly high precision), we can approach the problem from a standpoint of using our data set as a relatively good model for what

See why it's
important to
have real stuff!

"what we do"
"how we use"

in the
detector

the analysis is relies

As the % of hadronic cuts
which mimic W muons is very
small billions & units need
to be generated. a millions

as these
we both
know to
high preci

DRAFT

To form the basis to separate signal and background, the data itself is used as a proxy for what the hadronic background' looks like, and use simulations to fit the portions of the data which cannot come from this hadronic background.

The one thing that we do know is that the w signal must be in the data sample.

DRAFT

6.1 The Basic Cut

The basic cut aims to remove all obviously bad events from our event mix. The cut approaches this from two premises. The first, is that if track reconstruction variables simply cannot have resulted from a ~~W~~^{muon} genic event, then we remove the track. Secondly, if the track corresponds to a reconstructed energy which is larger than what is physically allowable for a W-genic track, we remove it. ~~is removed~~

The you solved
all problems!!

The "Basic Cut" is defined:

Variable	Lower Bound	Upper Bound
MuID lastGap	*	Gap 4
χ^2 what	0	20
$DG0$	0	20
$DDG0$	0	9
μ candidate	*	1

Table 6.1: The Basic Cuts used in the Run 13 analysis. lastGap refers to the last gap in the MuID which saw a μ candidate event. The fourth gap is the furthest penetration possible, therefore suggesting a high energy muon. Other parameters are described in Tables 5.1, 5.2, 5.3, and 5.4

With this cut, we have removed quite a large fraction of background events from ~~the~~ our dataset, without worry of removing any events in that fall within the kinematic range of W-Boson production.

No it doesn't it
removes bad track
candidates for w analysis

How could this
have been possible?

HOW?
describe

why don't we just use
this, makes life easier...

10%? 50%? ...
are removed

DRAFT

You should have said at the beginning:

I did not build PHENIX

I was not responsible for software

I did not sit every shift, I did not produce the DST

→ Who cares, this is YOUR thesis

6.2 Simulations

I did not do any work to produce the simulations used in this analysis, all of that credit goes to Dr. Ralf Seidl. I will generally describe how the simulations were produced, and paraphrase from the analysis note which we co-wrote with contributions also from Dr. Giordano, Abraham Meles and Daniel Jumper: [70].

PHENIX has a rather well developed simulation framework, which uses the in-house built "PHENIX Integrated Simulation Application" (PISA) [?] custom simulation framework. The simulation framework models in great detail the entire 12m \times 18m \times 18m volume of the PHENIX apparatus, as well as all the various material properties of the apparatus. The software package originated from the GEANT geometry and tracking packages. PISA encompasses more than this, though, it additionally encapsulates event-generators, a standalone geometry verification package, and the PHENIX offline analysis shell, to generate data that is completely compatible with PHENIX's data packaging framework. PISA has since been integrated into a simulation work-flow with the popular PYTHIA event generation system.

The simulations were created by selecting the biggest sources of muon background known to be produced at PHENIX as well as the W-Boson event and producing many events to generate good statistics. The primary purpose of simulating the muon background and W-Signal is to ultimately generate probability distribution functions for the variables which have the largest analyzing power - i.e. ability to differentiate between signal and background. The simulation and data both are ultimately described by the same variables. If they weren't it wouldn't work.

The data are added together, when combined to generate a 'muon background pdf' or 'W-Signal pdf' according to the cross-section of the process and the number of generated events, so as to not add these ingredients into the cocktail in the wrong amounts. This process is described in the next section, but the simulations used in this analysis are summarized here, in Table 6.2. When adding the simulations together, one must be careful to scale the final-yields by a correction factor (called k-factor) such that events which produce boosted dimuons are properly accounted for.

PHENIX uses some somewhat exclusive jargon when describing the various quark bound states which contribute to the muon background, and signal events. Open charm or charmonium refer to the bound state of the $c\bar{c}$ quarks. Onium generally refers to any process where a particle is in a bound-state with its own antiparticle (without of course

Like $p\bar{p}$?
or $u\bar{u}$?

Is this the
"exclusive" jargon? which
is used by everyone else?

DRAFT

There are some gaps in your
physics knowledge

NO not even
close...
do you
mean
"virtual
photon"
in this work.

double-counting open charm/charmonium). Direct photon, or alternatively $d\gamma$ (sometimes also written as DY) refers to photons that are produced as an immediate result of an inelastic scattering process, not from secondary decays. Open bottom refers to the bound state of $b\bar{b}$ quarks. $Z/d\gamma$ refers to the production and decay of the mixing between the Z-boson and virtual photons. ONLY Z refers to Z production and decay. W is naturally the signal event. W tau refers to the production of tau leptons, which can decay weakly, producing electrons or muons. W had refers to the production of W bosons from hadronic processes, rather than as the primary event vertex. All of these processes are summarized in Table 6.2.

Reference Run 393888

Process	k factor	σ (mb)	# Events Generated	\mathcal{L} (fb^{-1})
$c\bar{c}$	2.44	5.71e-01	5.85e+11	1.02
onium	0.415	1.35e-01	1.5e+11	1.11
$d\gamma$	0.0	5.32e-02	5.84e+10	1.10
$b\bar{b}$	1.83	7.30e-03	7.36e+09	1.01
ONLY Z	1.25	3.37e-07	1.73e+08	577.0
W	1.5	1.66e-06	3.38e+08	198.9
W tau	0.0	1.66e-06	3.43e+08	201.8
W had	0.0	1.66e-06	3.42e+08	201.2
Z	1.25	1.02e-06	2.93e+08	61.2

Table 6.2: Simulated sub processes in Run 13 including their generated event numbers as well as the corresponding luminosity and cross sections. Dr. Sanghwa Park has done an extensive analysis of the simulated data to determine an appropriate k-factor. Process which contribute very little to the muon background include W had, W tau, and $d\gamma$; they are scaled to zero. *so why did we bother... explain*

The simulations must additionally be weighted for trigger efficiency. To accomplish this, we weight events for each arm and charge with the associated trigger efficiency when constructing probability density functions representing the muon background. The trigger efficiencies generally manifest as η dependent functions *O* thus we bin the data into 20

DRAFT

separate η bins and calculate the efficiency associated with each bin. The bin ranges, and efficiency corrections are summarized in Table 6.4 for the North arm, and Table 6.3 for the South arm. *in the appendix*

One We can visualize the composition of the simulated data set by stacking the relative distributions of these variables. By looking the cross-sections of these variables as a function of p_T , for each arm and charge combination, we can get a feeling for how the data set *allows one to see* ~~background~~ *process contribute as a fraction of* *background* varies with p_T (Figure 6.2).



and other information such as the number of tracks and the number of photons. These variables are used to calculate the likelihood of the event occurring under the standard model hypothesis. The likelihood is calculated for each event, and the probability of the event occurring under the standard model hypothesis is determined. This probability is then compared to the observed probability of the event occurring under the standard model hypothesis.

defining the likelihood of the event occurring under the standard model hypothesis. The likelihood is calculated for each event, and the probability of the event occurring under the standard model hypothesis is determined. This probability is then compared to the observed probability of the event occurring under the standard model hypothesis.

South Arm			
η_{min}	η_{max}	$\mu^- \pm stat \pm sys$	$\mu^+ \pm stat \pm sys$
1.10	1.17	$0.27912 \pm 0.00297 \pm 0.10243$	$0.30607 \pm 0.00423 \pm 0.01108$
1.17	1.25	$0.40422 \pm 0.01642 \pm 0.04811$	$0.43125 \pm 0.01717 \pm 0.26702$
1.25	1.32	$0.27958 \pm 0.00056 \pm 0.05539$	$0.36619 \pm 0.00925 \pm 0.07316$
1.32	1.40	$0.26563 \pm 0.00542 \pm 0.02485$	$0.25312 \pm 0.00349 \pm 0.04927$
1.40	1.48	$0.39802 \pm 0.00497 \pm 0.07770$	$0.34295 \pm 0.00306 \pm 0.03127$
1.48	1.55	$0.43156 \pm 0.00633 \pm 0.17060$	$0.37567 \pm 0.00248 \pm 0.03644$
1.55	1.62	$0.34831 \pm 0.00309 \pm 0.03720$	$0.40246 \pm 0.00546 \pm 0.04605$
1.62	1.70	$0.33043 \pm 0.00280 \pm 0.09227$	$0.40219 \pm 0.00472 \pm 0.05637$
1.70	1.77	$0.33152 \pm 0.00318 \pm 0.11668$	$0.30805 \pm 0.00360 \pm 0.03644$
1.77	1.85	$0.34710 \pm 0.00633 \pm 0.00918$	$0.38565 \pm 0.00439 \pm 0.04295$
1.85	1.92	$0.32448 \pm 0.00404 \pm 0.14670$	$0.30118 \pm 0.00418 \pm 0.10071$
1.92	2.00	$0.31461 \pm 0.00714 \pm 0.01799$	$0.31263 \pm 0.00545 \pm 0.01643$
2.00	2.07	$0.64632 \pm 0.01161 \pm 0.23329$	$0.63252 \pm 0.01040 \pm 0.10507$
2.07	2.15	$0.60582 \pm 0.00565 \pm 0.05569$	$0.67335 \pm 0.01245 \pm 0.05630$
2.15	2.22	$0.45058 \pm 0.00697 \pm 0.45101$	$0.69619 \pm 0.01247 \pm 0.65623$
2.22	2.30	$0.45185 \pm 0.01358 \pm 0.36032$	$0.51436 \pm 0.01288 \pm 0.43781$
2.30	2.38	$0.43890 \pm 0.07336 \pm 0.34632$	$0.61623 \pm 0.06221 \pm 0.62209$
2.38	2.45	$0.00000 \pm 0.25000 \pm 0.00000$	$0.00000 \pm 0.25000 \pm 0.00000$
2.45	2.52	$0.00000 \pm 0.25000 \pm 0.00000$	$0.00000 \pm 0.25000 \pm 0.00000$
2.52	2.60	$0.00000 \pm 0.25000 \pm 0.00000$	$0.00000 \pm 0.25000 \pm 0.00000$

Table 6.3: η dependent trigger efficiencies are calculated for the South arm in 20 η bins. Each correction has both systematic and statistical error accounted for.

		North Arm	
η_{min}	η_{max}	$\mu^- \pm stat \pm sys$	$\mu^+ \pm stat \pm sys$
1.10	1.17	$0.56285 \pm 0.03834 \pm 0.32882$	$0.52850 \pm 0.01938 \pm 0.36163$
1.17	1.25	$0.67803 \pm 0.02249 \pm 0.13431$	$0.49546 \pm 0.00261 \pm 0.16304$
1.25	1.32	$0.69537 \pm 0.01551 \pm 0.03465$	$0.63287 \pm 0.01285 \pm 0.08350$
1.32	1.40	$0.39864 \pm 0.00724 \pm 0.02330$	$0.38435 \pm 0.00762 \pm 0.11954$
1.40	1.48	$0.52102 \pm 0.00750 \pm 0.05014$	$0.49573 \pm 0.00698 \pm 0.03733$
1.48	1.55	$0.48068 \pm 0.00498 \pm 0.11579$	$0.48874 \pm 0.00357 \pm 0.08063$
1.55	1.62	$0.54113 \pm 0.00860 \pm 0.04895$	$0.50041 \pm 0.00659 \pm 0.05165$
1.62	1.70	$0.45140 \pm 0.00822 \pm 0.05718$	$0.46948 \pm 0.00755 \pm 0.09718$
1.70	1.77	$0.43203 \pm 0.00547 \pm 0.04976$	$0.40722 \pm 0.00546 \pm 0.07957$
1.77	1.85	$0.42141 \pm 0.00815 \pm 0.04366$	$0.44450 \pm 0.00628 \pm 0.04575$
1.85	1.92	$0.37946 \pm 0.00620 \pm 0.01766$	$0.37183 \pm 0.00700 \pm 0.01848$
1.92	2.00	$0.37499 \pm 0.00782 \pm 0.05026$	$0.40156 \pm 0.00678 \pm 0.02291$
2.00	2.07	$0.51268 \pm 0.00547 \pm 0.10416$	$0.60041 \pm 0.00973 \pm 0.21212$
2.07	2.15	$0.56990 \pm 0.00614 \pm 0.14507$	$0.58276 \pm 0.01392 \pm 0.25179$
2.15	2.22	$0.60527 \pm 0.01524 \pm 0.10354$	$0.60766 \pm 0.00425 \pm 0.23618$
2.22	2.30	$0.70200 \pm 0.01678 \pm 0.25233$	$0.45067 \pm 0.01008 \pm 0.24192$
2.30	2.38	$0.48294 \pm 0.00294 \pm 0.12663$	$0.54157 \pm 0.02109 \pm 0.06230$
2.38	2.45	$0.47814 \pm 0.02338 \pm 0.42026$	$0.42606 \pm 0.03092 \pm 0.25031$
2.45	2.52	$0.61788 \pm 0.14438 \pm 0.61788$	$0.29673 \pm 0.06686 \pm 0.04941$
2.52	2.60	$0.00000 \pm 0.25000 \pm 0.00000$	$0.15630 \pm 0.15630 \pm 0.18223$

Table 6.4: η dependent trigger efficiencies are calculated for the North arm in 20 η bins. Each correction has both systematic and statistical error accounted for.

DRAFT

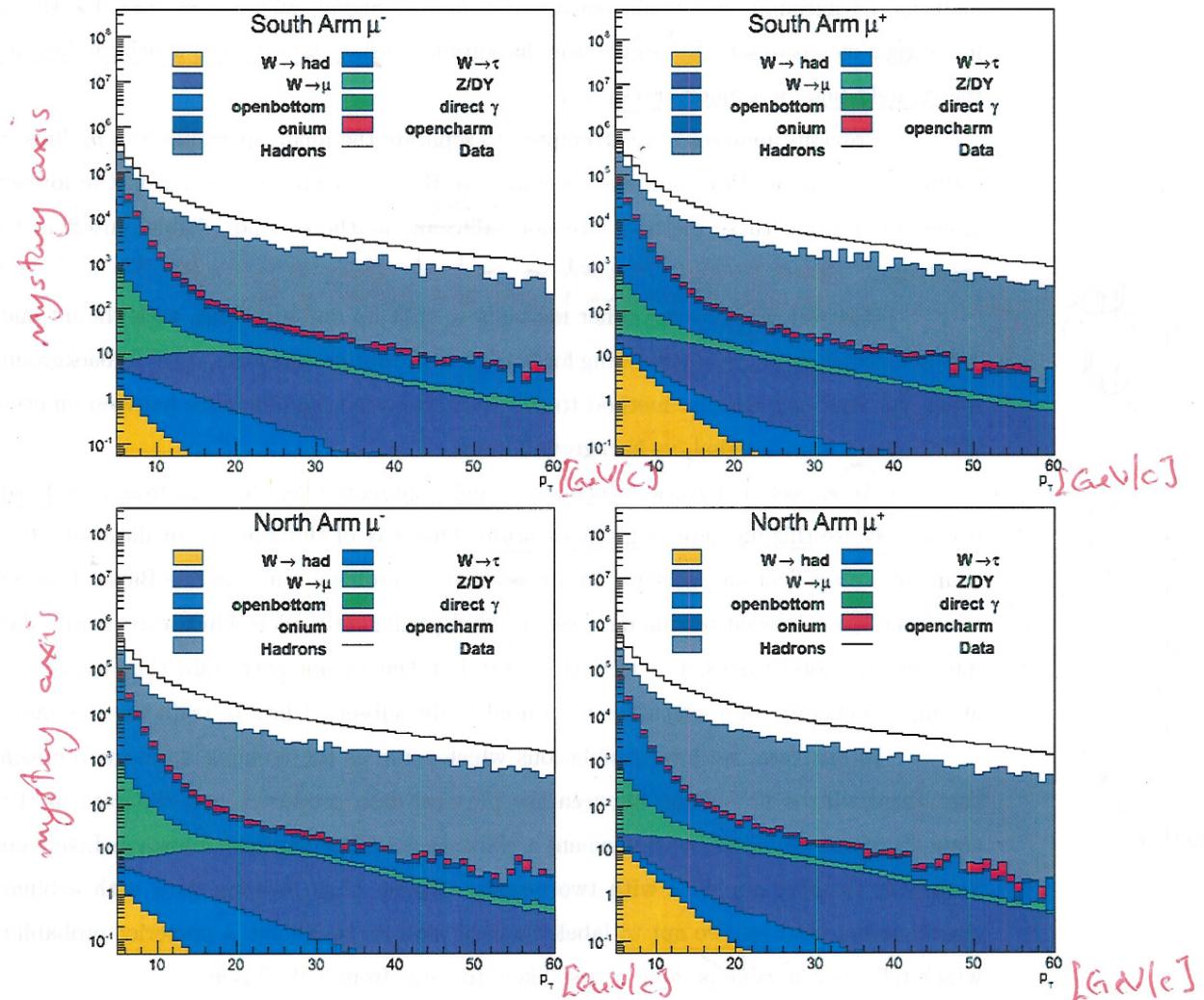


Figure 6.2: Here, we see the stacked cross-sections of all simulated processes as a function of p_T . All data shown has been created from the PISA+PYTHIA framework. Top Left: South μ^- , Top Right: South μ^+ , Bottom Left: North μ^+ , Bottom Right: North μ^- . Figure reproduced from my analysis note. Dr. Ralf Seidl produced the original [70].

DRAFT

6.3 W_{ness} : Likelihood Event Tagging

Recalling that we have already split the dataset into three main contributions: hadronic background, real muon background, and W-Signal, we are now tasked with formulating a means to separate signal from background, using the variables which can indicate the straightness of a muon track.

Previous analyses have attempted to separate the muon spectrum into p_T bins, to estimate the composition, however, because the $W \rightarrow \mu$ signal is so small in the forward kinematic regime, these methods are not sufficient, as there is no 'visible' cutoff in the spectrum.

However, we may use other methods to split up our spectrum, with the ultimate goal of calculating A_L , and correcting for background dilution using the signal to background ratio. We must use another method to effectively describe the difference between an event which comes from a signal vs background event.

We expect that tracks which are straight are more likely to come from a W-Boson decay, because this indicates high momentum. One way of thinking of our data set can in terms of a classification problem. In a classification problem, one can use Bayes' Theorem when one has a labeled testing data set to build predictive models which can classify data into two or more classes, provided that care is taken to not over-train the classifier, or attempt to classify data which has been used in the subset of data to train the classifier.

In our case, we have simulations which serve as the training data, guaranteeing that there will be no overlap between the physical data produced, and the data used to train the classifier. Thus, we implement a Naive Bayes Classifier (also known as Likelihood Selection) to label our data with two classes. Rather than labeling data with a binary classification, however, we opt to label the data with its likelihood, a posterior probability which tells us if a value is more or less likely to come from a W-Boson.

6.3.1 Naive Bayes Classification

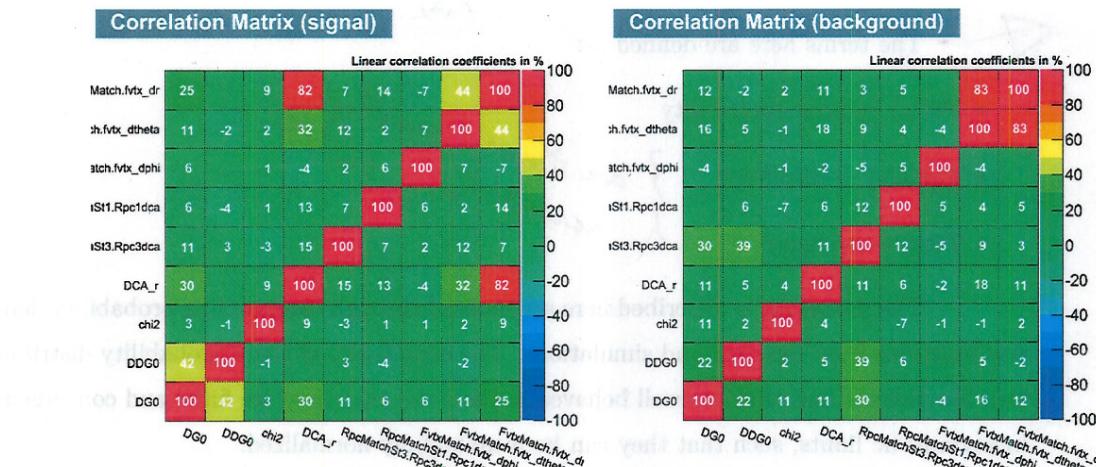
There are many techniques available for classifying a collection of variables (a feature set) into categories. Naive Bayes classification is an excellent candidate for classification, in cases where we have two classifications with distributions of feature sets which are uncorrelated. Naive Bayes even works when feature sets are slightly correlated. It is a robust, fast, scalable machine learning technique. Traditionally used for classification of

DRAFT

text documents, Naive Bayes is also able to handle numeric features whose distributions are known [71].

In our analysis, we begin with a Naive Bayes classifier which is trained to classify signal muons or background muons. We combine both Real Muon Background muons and Fake Muons (Hadronic Background Muons) in the label of "Background Muons" at this stage, though, later, we will separate out the muons further.

In order to obtain the best performance from our classifier, without over-training, we need to ensure that the variables (or feature set) used to determine a class are maximally uncorrelated. The variables which match this criteria are: DG0, DDG0, χ^2 , f_{vtx} variables, Rpc1DCA, Rpc3DCA, DCA_r, and DCA_z. The Linear Correlations between these variables are shown for both the data, and the simulated W-Signal in Figure 6.3.



(a) Correlations between kinematic variables, produced from simulated data.

(b) Correlations between kinematic variables, produced from the data, which is composed mostly of hadronic background

Figure 6.3: Low correlations between the signal variable distributions (from simulation), and the background variable distributions make this data set a good candidate for classification using Naive Bayes

As one can see from Figure 6.3, DG0 and DDG0 are slightly correlated, as are χ^2 and DCA_r. A Naive Bayes classifier may be constructed from the core of the familiar Bayes Theorem from probability and statistics. In our case, we understand Naive Bayes as a

chi
or
chi^2
DCA_r → 106

DRAFT

conditional probability. Concretely, we consider a vector of features (i.e. our discriminating kinematic variables):

No indent

$$\mathbf{x} = (x_1, \dots, x_n) \quad (6.1)$$

and assume independence between each feature x_n . We then define the probability of a given classification, C_k given a set of features x_n :

No indent

$$\mathcal{P}(C_k|x_1, \dots, x_n) \quad (6.2)$$

This conditional probability is defined in terms of Bayes Theorem:

Diagram

$$\mathcal{P}(C_k|\mathbf{x}) = \frac{\mathcal{P}(C_k) \mathcal{P}(\mathbf{x}|C_k)}{\mathcal{P}(\mathbf{x})} \quad (6.3)$$

The terms here are defined as:

- $\mathcal{P}(C_k) \rightarrow$ prior probability
- $\mathcal{P}(\mathbf{x}|C_k) \rightarrow$ likelihood
- $\mathcal{P}(\mathbf{x}) \rightarrow$ evidence

*are these usually
named like this?*

The probabilities described here are realized through constructing probability density functions from the data and simulations. The constraints of these probability distribution functions is that they are well behaved in the sense that they are finite and convergent in asymptotic limits, such that they can be meaningfully normalized.

In our case, we construct a likelihood ratio, using the posterior probability for each classification, which is defined as W_{ness} :

$$\lambda_{sig} = \prod_k \mathcal{P}(\mu_{sig}|C_k)$$

$$\lambda_{bak} = \prod_k \mathcal{P}(\mu_{bak}|C_k)$$

where

Where λ_{sig} and λ_{bak} represent the total likelihoods that a given track is either signal, or background, constructed from the product of likelihoods calculated from each probability density function. The λ 's are *luckily* combined to calculate the W_{ness} :

$$W_{ness} = \frac{\lambda_{sig}}{\lambda_{sig} + \lambda_{bak}} \quad (6.4)$$

DRAFT

Thus, we must construct probability density functions representing the likelihood of an event being W-genic or from the combined hadronic+muon background. Our data set after the basic cut has approximately 1 million events. Based on the cross-section of the $W \rightarrow \mu$ decay, we expect the final population of W-Bosons in the data set to be on the order of 1 thousand. Therefore, we can confidently use the data set as is, in order to generate PDFs representing the hadronic+muon background, as any effect from the signal would be only one part in a thousand. Thus, we loop over the data set, and the W-Simulation set, and filter the data into probability distribution functions. Because some events do not have archived data from all subsystems, we construct a variety of PDFs, selecting the appropriate PDF cocktail based on whether or not the requisite variables were archived for that given track, Figure 6.4.

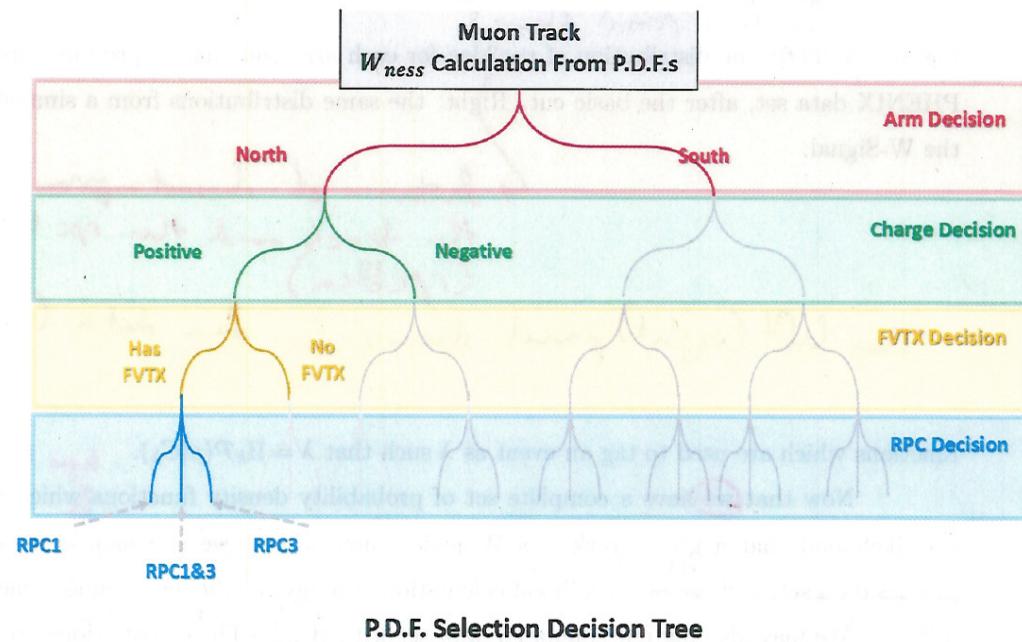
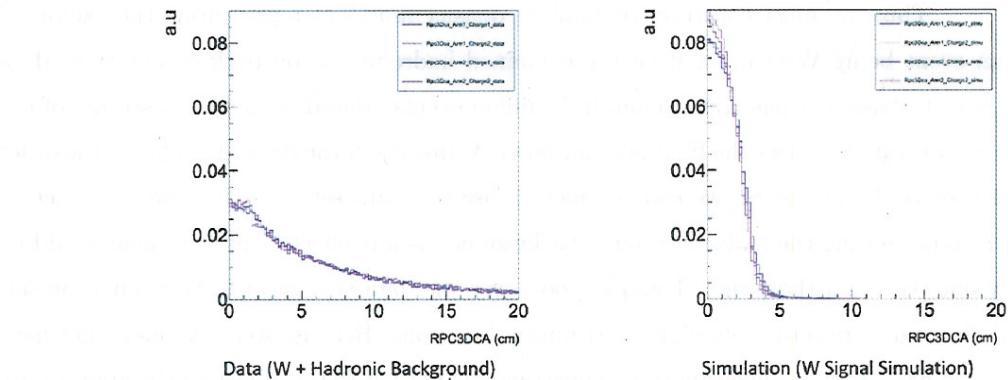


Figure 6.4: A cartoon of the decision tree to determine the PDF cocktail to use for quantifying the W_{ness} of a given track. The track's properties are used to traverse the tree, and select the cocktail contents.

In figures 6.5-6.10, we can see the various distributions which are used to create probability distribution functions. In the figures, we represent the product of all probability

DRAFT

Order these
figures from
left to right
is DGO, R,



$$\lambda = p(DG0, DDG0)p(\chi^2)p(DCA_r)p(Rpc1/3dca)p(fvtx_{dr})p(fvtx_{d\theta})p(fvtx_{d\phi})$$

The left panel shows

Figure 6.5: Left: the distribution of rpc3dea for each arm and charge, produced from the PHENIX data set, after the basic cut. Right: the same distributions from a simulation of the W-Signal.

↳ distance of closest approach between the track and the rpc hits at RPC3DCA

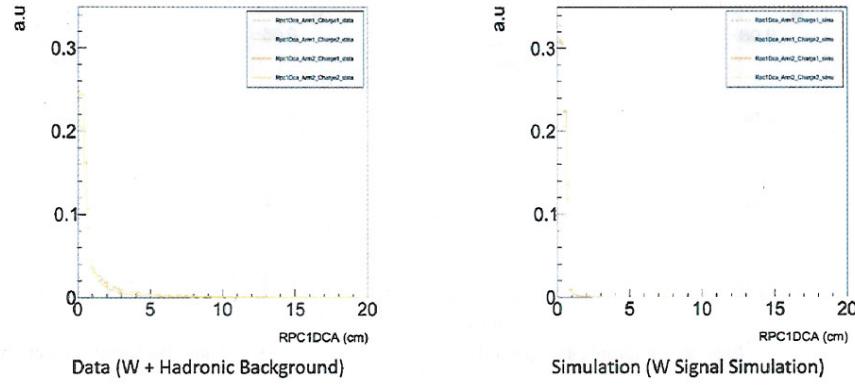
The left (right) panel shows ... for data (W simulation) ...

functions which are used to tag an event as λ such that $\lambda = \prod_k \mathcal{P}(\mu | C_k)$. *has been produced*

Now that we have a complete set of probability density functions which predicts the likelihood that a given track is a W-genic muon or not, we can loop over the real physics data set, and use our likelihood calculation strategy to label every muon track with a W_{ness} . We may also tag our simulated data set with W_{ness} . The distributions are shown in Figure 6.11.

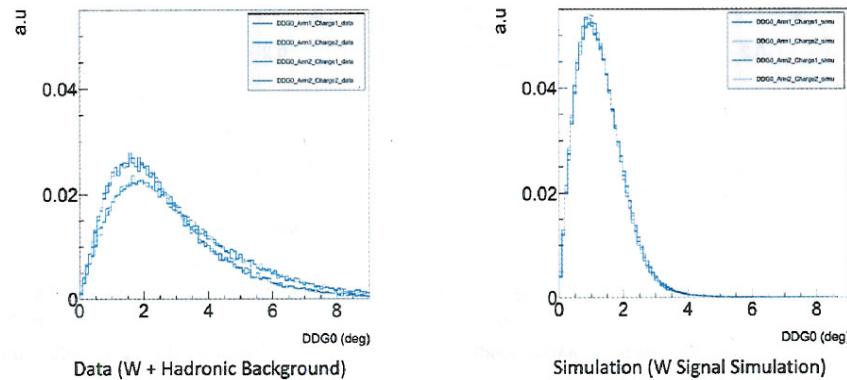
As we can see from Figure 6.11, most of the simulated data falls in the high W_{ness} range while most of the physics data falls in the low W_{ness} range. The goal of the likelihood analysis is to tag the data with W_{ness} such that we can apply a cut on the data based on the parameter's value. We wish to apply the cut in a way that minimally removes any signal, and we may calculate the efficiency of this cut, summarized in Figure 6.12.

DRAFT



$$\lambda = p(DG0, DDG0)p(\chi^2)p(DCA_r)p(Rpc1/3dca)p(fvtx_{dr})p(fvtx_{d\theta})p(fvtx_{d\phi})$$

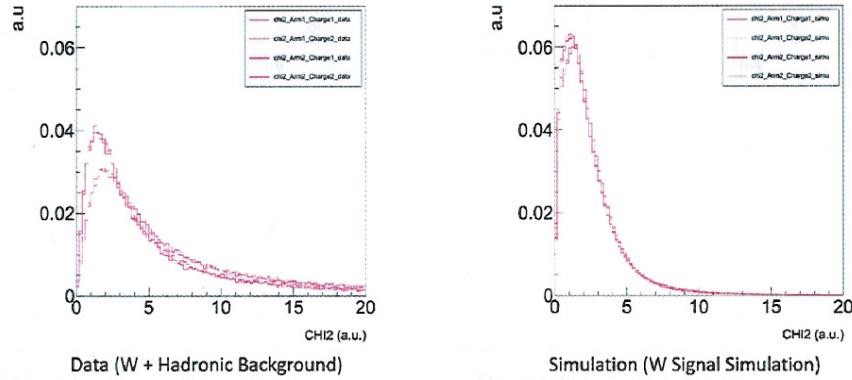
Figure 6.6: Left: the distribution of rpc1dca for each arm and charge, produced from the PHENIX data set, after the basic cut. Right: the same distributions from a simulation of the W-Signal.



$$\lambda = p(DG0, DDG0)p(\chi^2)p(DCA_r)p(Rpc1/3dca)p(fvtx_{dr})p(fvtx_{d\theta})p(fvtx_{d\phi})$$

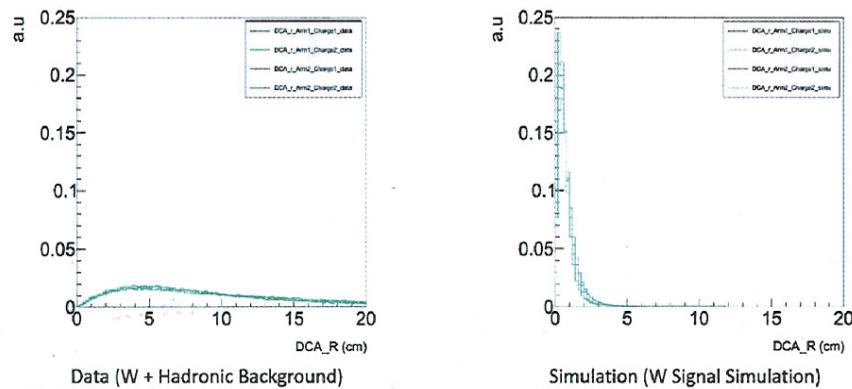
Figure 6.7: Left: the distribution of DDG0 for each arm and charge, produced from the PHENIX data set, after the basic cut. Right: the same distributions from a simulation of the W-Signal.

DRAFT



$$\lambda = p(DG0, DDG0)p(\chi^2)p(DCA_r)p(Rpc1/3dca)p(f vtx_{dr})p(f vtx_{d\theta})p(f vtx_{d\phi})$$

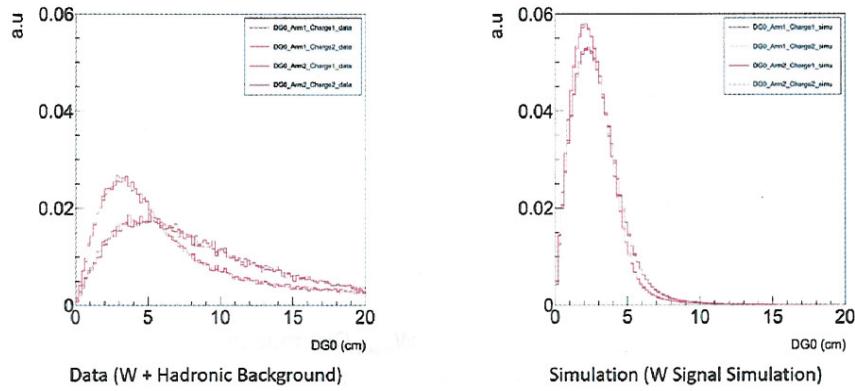
Figure 6.8: Left: the distribution of chi2 for each arm and charge, produced from the PHENIX data set, after the basic cut. Right: the same distributions from a simulation of the W-Signal.



$$\lambda = p(DG0, DDG0)p(\chi^2)p(DCA_r)p(Rpc1/3dca)p(f vtx_{dr})p(f vtx_{d\theta})p(f vtx_{d\phi})$$

Figure 6.9: Left: the distribution of dcar for each arm and charge, produced from the PHENIX data set, after the basic cut. Right: the same distributions from a simulation of the W-Signal.

DRAFT



$$\lambda = p(DG0, DDG0)p(\chi^2)p(DCA_r)p(Rpc1/3dca)p(fvtx_{dr})p(fvtx_{d\theta})p(fvtx_{d\phi})$$

Figure 6.10: Left: the distribution of DG0 for each arm and charge, produced from the PHENIX data set, after the basic cut. Right: the same distributions from a simulation of the W-Signal.

As we make successive cuts in W_{ness} , we find that the optimum cutoff is at $W_{ness} < 0.95$. We can throw out all data below this threshold, and maintain a good portion of our signal data.

Note that now with this reduced data set, we could simply assume that all remaining data is signal, and calculate an asymmetry, however, there is clearly still a lot of background present. Any background that is still present will dilute our main observable, A_L . Therefore, we employ the unbinned maximum likelihood fit to a three dimensional data set, composed of W_{ness} , η , and dw_{23} .

(Sectn X)

Something more like this:
To obtain the optimal dress cut-off, successive cuts in W_{ness} are made. The fraction of signal and background are compared at each point. It is found that $W_{ness} < 0.95$...

to estimate the residual background cuts

112
You should probably explain what this means (the 0.95 number)
is it 95% ...?

DRAFT

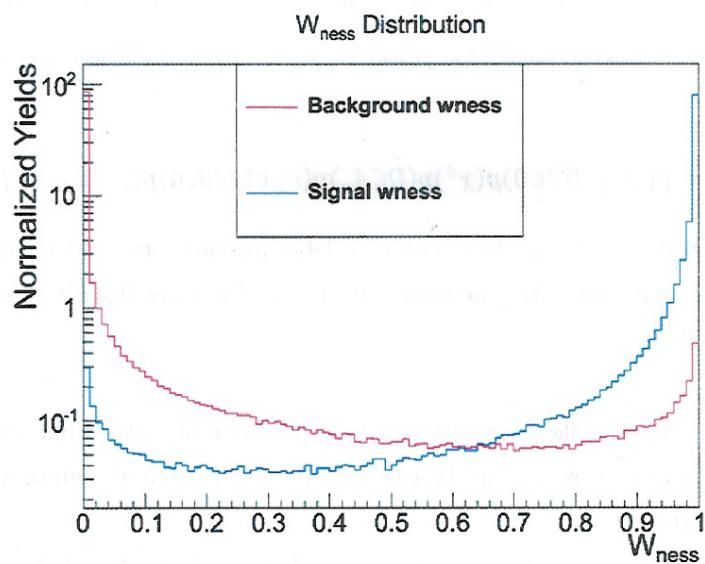


Figure 6.11: After W_{ness} tagging, we can visualize the classification of signal from background by comparing the distribution of W_{ness} in physics data, and the simulated data data. Note that the vertical is plotted on a log scale. The two distributions have been normalized prior to plotting.

by area

usually colors are
not allowed

DRAFT

6.1.2. Cuts and the fraction of background remaining vs. signal efficiency

Let's consider the effect of a cut on the total fraction of signal and background remaining. If we want to keep a certain fraction of signal, we can choose a cut that removes the lowest fraction of background. This is the most conservative choice, and it is often called the "quadratic" choice. For example, if we want to keep 90% of the signal, we can remove 10% of the background.

Let's look at the effect of a cut on the total fraction of signal and background remaining. We can choose a cut that removes the lowest fraction of background. This is the most conservative choice, and it is often called the "quadratic" choice.

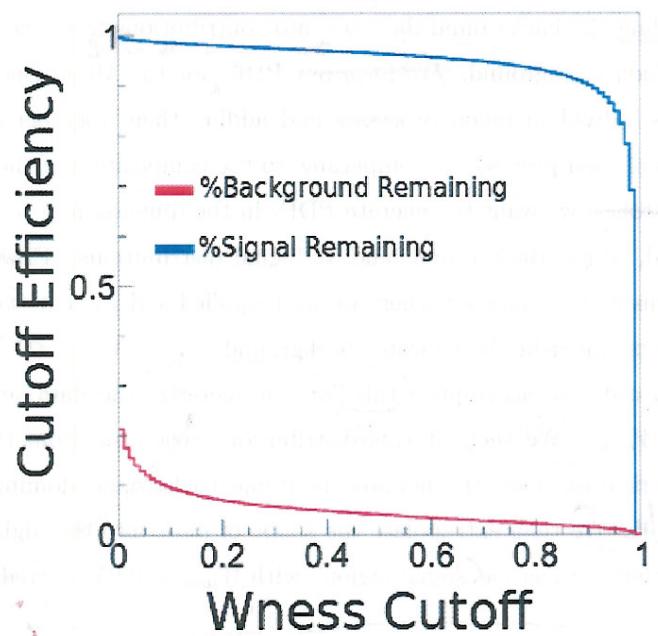
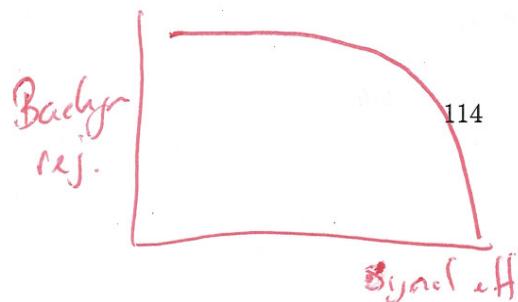


Figure 6.12: We look at the fraction of signal and background remaining in the total data set as we make successively higher cuts in W_{ness} . At the turning point of the blue distribution (the fraction of remaining signal) is where we choose to cut the data, corresponding to removing data with a W_{ness} value of less than 0.95.

I prefer the other style figure



DRAFT

6.4 Extended Unbinned Maximum Likelihood Selection: The Signal to Background Ratio

The goal of the Extended Unbinned Maximum Likelihood Fit (EULMF) is to calculate the signal to background ratio, so that we can calculate A_L and correct for the dilution from the background. The EULMF is another statistical method which relies on creating Probability Density Functions to represent the likelihood of given track to originate from a known source. However, at this stage in the analysis, we are interested in subdividing the background data set into contributions from the Hadronic Background and the Muon Background. We form our PDFs for the Muon Background by weighting the various individual muon processes and adding them together such that the relative frequencies of each process are comparable to the composition of the real physics data set. In broad strokes, we want to generate PDFs in the dimension of η and dw_{23} for Hadronic Background, Muon Background, and W-Signal distributions. However, since we will be applying this fit to a data set where we have applied a W_{ness} cut, we must be very careful to not over or under-fit the hadronic background.

In order to accomplish this, we parameterize the data set as a 2D function in dw_{23} and W_{ness} . We then fit this distribution, generated from the physics data, with a parameterization, over the nominal hadronic background dominated region from $0 < W_{ness} < 0.95$. We then extrapolate the shape of dw_{23} into the high $W_{ness} > 0.95$ region, hereafter referred to as the 'signal region', with $W_{ness} < 0.95$ referred to as the 'background region'.

Similarly to any analysis which uses probability density functions, the PDFs representing η and dw_{23} must be uncorrelated so as to not over-fit the data.

The purpose of the EULMF is to essentially scale the PDFs for each arm and charge for η and dw_{23} so as to obtain yields for W-genic muons, Muon Background muons, and hadronic background fake muons.

To use this method, we must construct the likelihood function (Equation ??) and maximize it. We write down the likelihood function in as a product of the individual likelihoods:

An unbinned maximum likelihood fit can then be performed to extract the number of events for each process: n_{sig} , n_μ , n_{had} . The likelihood function is defined accounting for a Poisson distribution of the events x_i :

DRAFT

$$\mathcal{L}(\theta|X) \equiv \frac{n^N e^{-n}}{N!} \prod_{x_i \in X} \sum_c \frac{n_c}{n} p_c(x_i), \text{ ; with } n = \sum_c n_c \quad (6.5)$$

no indent

where X is the sample of N total events $x_i = (\eta_i, dw_{23i})$, and θ gives the parameters of the fit $\theta = (n_{sig}, n_\mu, n_{had})$. To reduce the number of parameters, we fixed the number of muon background events n_μ to the expected yield according to the cross section of muon background processes, and then extracted the remaining parameters (n_{sig}, n_{had}) by minimizing the $-\log(\mathcal{L}(\theta|X))$. With Run 13 data we have enough statistics to divide the data sample in three η region: $1.10 < \eta < 1.40$, $1.40 < \eta < 1.80$ and $1.80 < \eta < 2.60$.

6.4.1 Hadronic Background PDFs

The main analysis challenge for the EULMF is obtaining an adequate description of the dw_{23} and η distributions for the hadronic background. They are shown, along with the W_{ness} distribution for the data, for the background region, in Figure 6.13, and the simulated W-genic data in Figure 6.14.

Two features to note from Figure 6.13 and Figure 6.14 is that the distribution for dw_{23} should be expected to be quite narrow for W-genic events, whereas η is more broad.

We have good statistic for η over all orders of magnitude, so we can directly construct PDFs for this variable from a binned dataset. However, with dw_{23} we need to be more careful, as this variable will offer us the analyzing power.

We create a model for dw_{23} , to fully parameterize the event distribution when viewed as a function of dw_{23} vs W_{ness} . We model this by assuming that the dw_{23} vs W_{ness} distribution can be separated into two parts: R_j ...

$$F(W_{ness}, dw_{23}) = f(W_{ness}) \times g(W_{ness}, dw_{23}) \quad (6.6)$$

no indent

$f(W_{ness})$ is modeled simply as a fourth-degree polynomial (the third column) of Figures 6.13 and 6.14. The polynomial fit is summarized in Equation 6.7 and Figure 6.15:

$$f(W_{ness}) = P_8 + P_9 W_{ness} + P_{10} W_{ness}^2 + P_{11} W_{ness}^3 + P_{12} W_{ness}^4 \quad (6.7)$$

\$P_{W_ness} \sim \Sigma^{43\%}\$

how is it
complexly modelled, is it
better? (careful with wording)

why why why
you cannot just throw in stuff, explain...

DRAFT

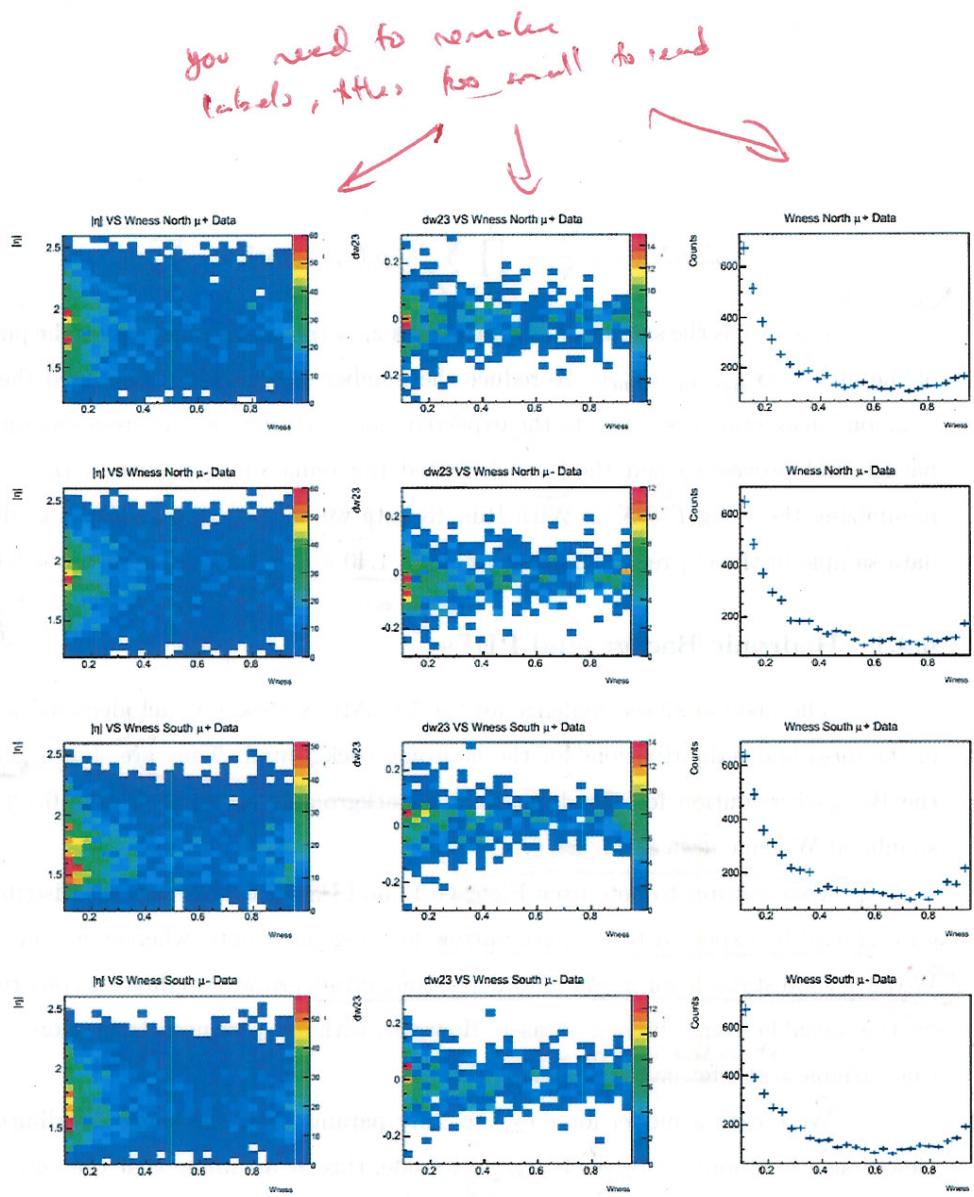


Figure 6.13: The first column of plots is η plotted as a function of W_{ness} where we see a 2D histogram of the even distribution. The middle column is dw_{23} as a function of W_{ness} , and the right column is a simple histogram of W_{ness} . The rows all correspond to the same arm and charge. From top to bottom: North, μ^+ , North μ^- , South μ^+ , North μ^- . Distributions shown here are all from the physics data set.

shows
why simple?

DRAFT

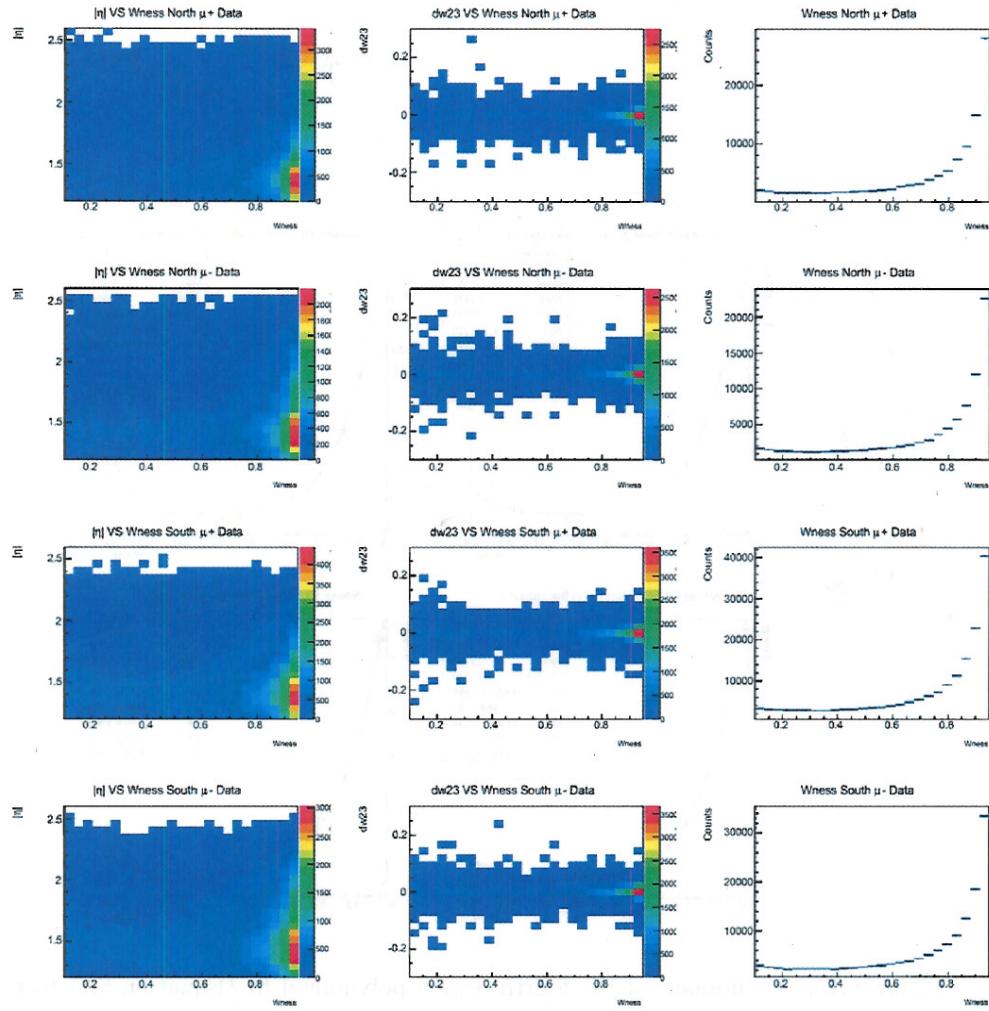


Figure 6.14: The first column of plots is η plotted as a function of W_{ness} where we see a 2D histogram of the even distribution. The middle column is dw_{23} as a function of W_{ness} , and the right column is a simple histogram of W_{ness} . The rows all correspond to the same arm and charge. From top to bottom: North, $\mu+$, North $\mu-$, South $\mu+$, North $\mu-$. Distributions shown here are all from simulated W-genic data set.

sel
6/13

DRAFT

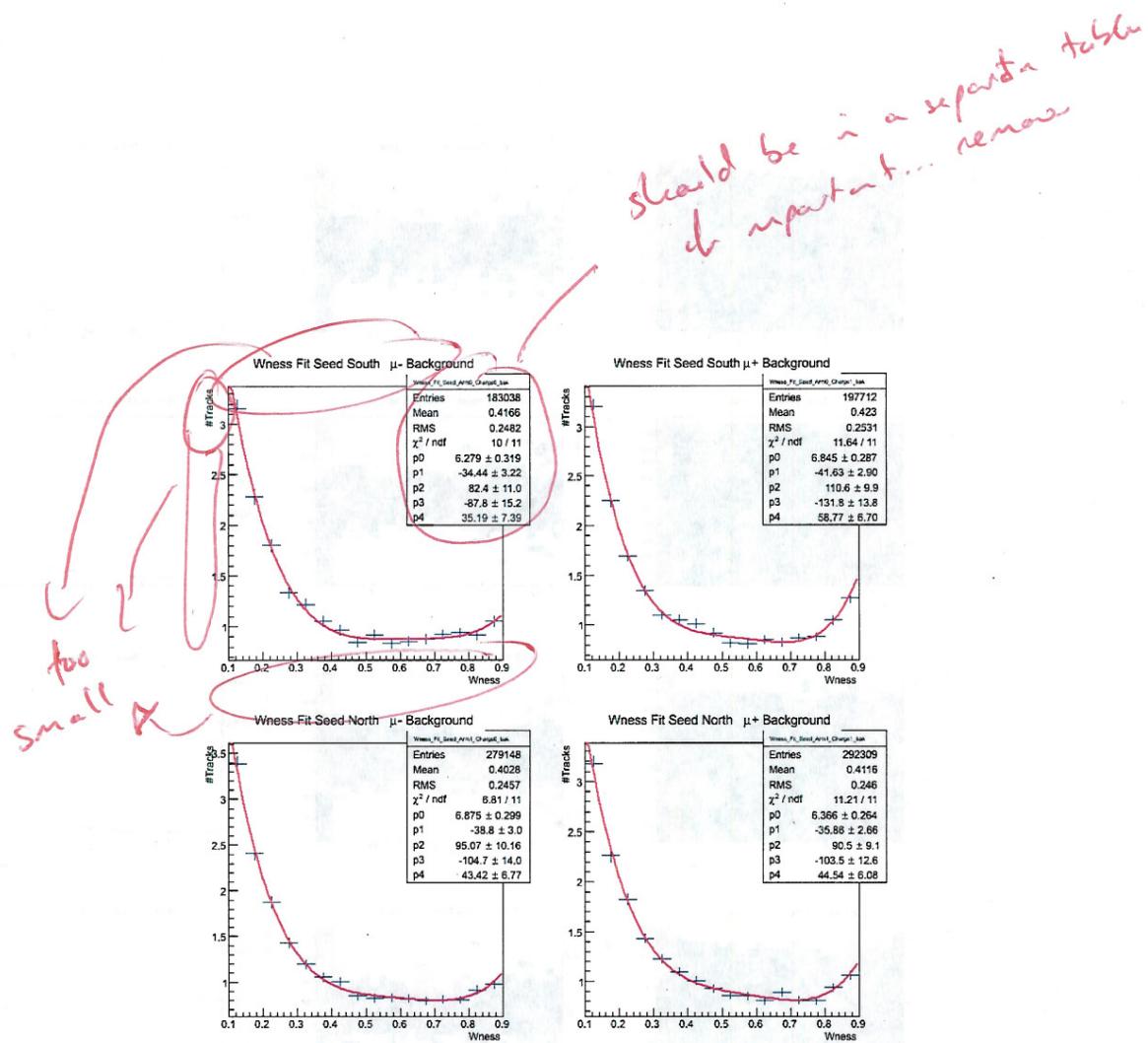


Figure 6.15: A summary of the fourth degree polynomial fit (Equation 6.7) to the W_{ness} distribution from the physics dataset in the background region.

for real data, fit over the "background" region.

DRAFT

Next the second component ($g(W_{ness}, dw)$) is modeled as

We then model the other element of the distribution $g(W_{ness}, dw_{23})$ as a co-axial double Gaussian. We allow the Parameters of the co-axial double Gaussian to vary linearly with W_{ness} , as seen below:

where ...?

$$\sigma_1 = P_1 + P_3 \times W_{ness}$$

$$C_g = P_6 + P_7 \times W_{ness} \quad (6.8)$$

$$\sigma_2 = P_4 + P_5 \times W_{ness}$$

$$\mu = P_0 + P_1 \times W_{ness} \quad (6.9)$$

$$g(W_{ness}, dw_{23}) = C_w \times \left(\left(\frac{1}{\sqrt{2\pi}\sigma_1 + C_g \sqrt{2\pi}\sigma_2} \right) \times \left(e^{\frac{1}{2} \left(\frac{dw_{23}-\mu}{\sigma_1} \right)^2} + C_g e^{\frac{1}{2} \left(\frac{dw_{23}-\mu}{\sigma_2} \right)^2} \right) \right) \quad (6.10)$$

We seed these linearized parameters by taking slices of dw_{23} in W_{ness} and then fitting this slice with a co-axial double Gaussian. The parameters of the results of these fits are then plotted against the value of the W_{ness} slice, and fit with a line. These parameters are then used to seed the fit of Equation 6.6 to the physics data set. Fits to the individual slices of dw_{23} are summarized in Figure 6.16. The results of the co-axial double Gaussian parameters as functions of W_{ness} slice are shown in Figure 6.17.

please
a
little more
effort...

The results of this fitting procedure are summarized in Figure 6.18.

Finally, the extrapolation of dw_{23} was reproduced independently by four separate analyzers, Daniel, Abraham, Ralf and Myself, with distributions lining up very closely, Figure 6.20 *WHY?* because \rightarrow it's critical for...

This is NOT
scientific

The PDF for η was obtained by creating a histogram of the variable for events tagged with $W_{ness} < 0.9$.

No names allowed

How important is the
seed how did you vary it?

Meson?

or

distortion...

... for the η distribution...

DRAFT

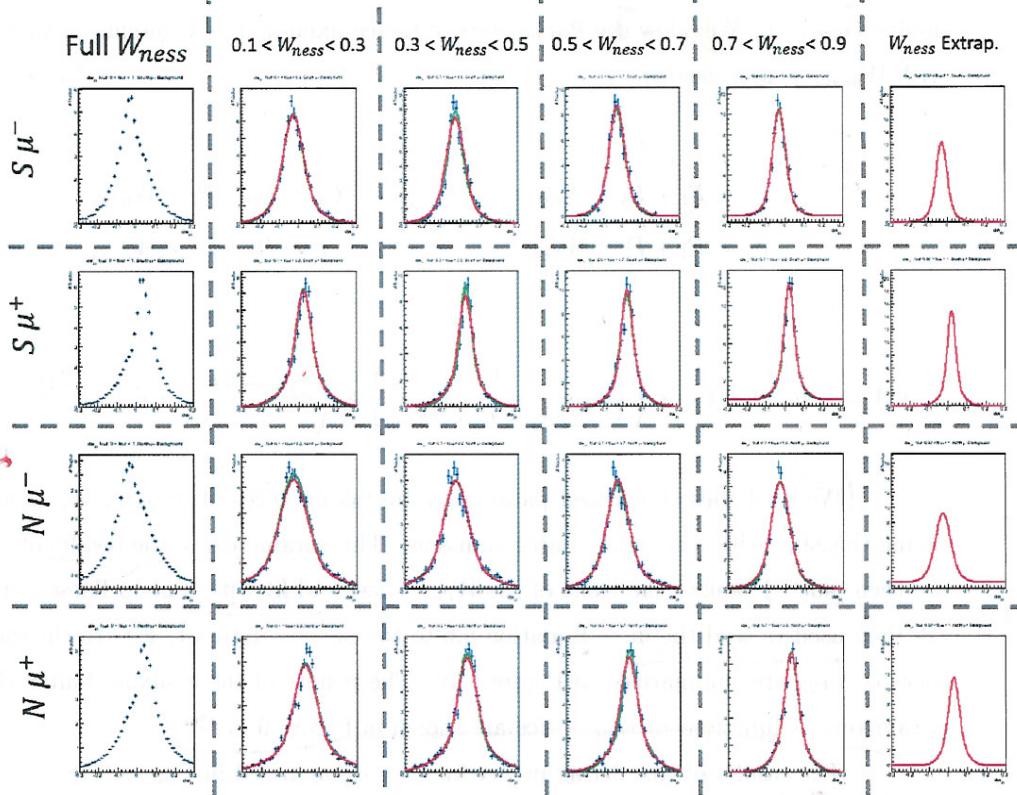


Figure 6.16: From left to right the columns are: ~~dw_{23}~~ for the full W_{ness} range, $0.1 < W_{ness} < 0.3$, $0.3 < W_{ness} < 0.5$, $0.5 < W_{ness} < 0.7$, $0.7 < W_{ness} < 0.9$, and finally the extrapolated shape for $W_{ness} > 0.95$. In red, we see the 1D projection of the 2D distribution to the slice. This overlays a green curve, which is a fit done independently to a slice. The rows are labeled with the Arm and charge corresponding to the subdivided dataset. As you can see, the matching is often exact, between green and red curves. As the final column is the extrapolation, there is no slice-fit.

BUT
there is data
why not show
the comparison so
we can get a feeling
for "sync" & "bljd"
fit ...?

why it was not
mentioned ... independent
how about kind of fit¹²¹ same?
explain, explain, ...

The red curve shows a 1D projection
of the 2D distribution made over
the W_{ness} slice indicated.

What is really matching ...
the curves of the fits...
(same as before, just did not do recos...)

DRAFT

REMAKE
MEET for at least
(readable)

review

too small

σ_1

σ_2

G_{factor}

G_{offset}

before???

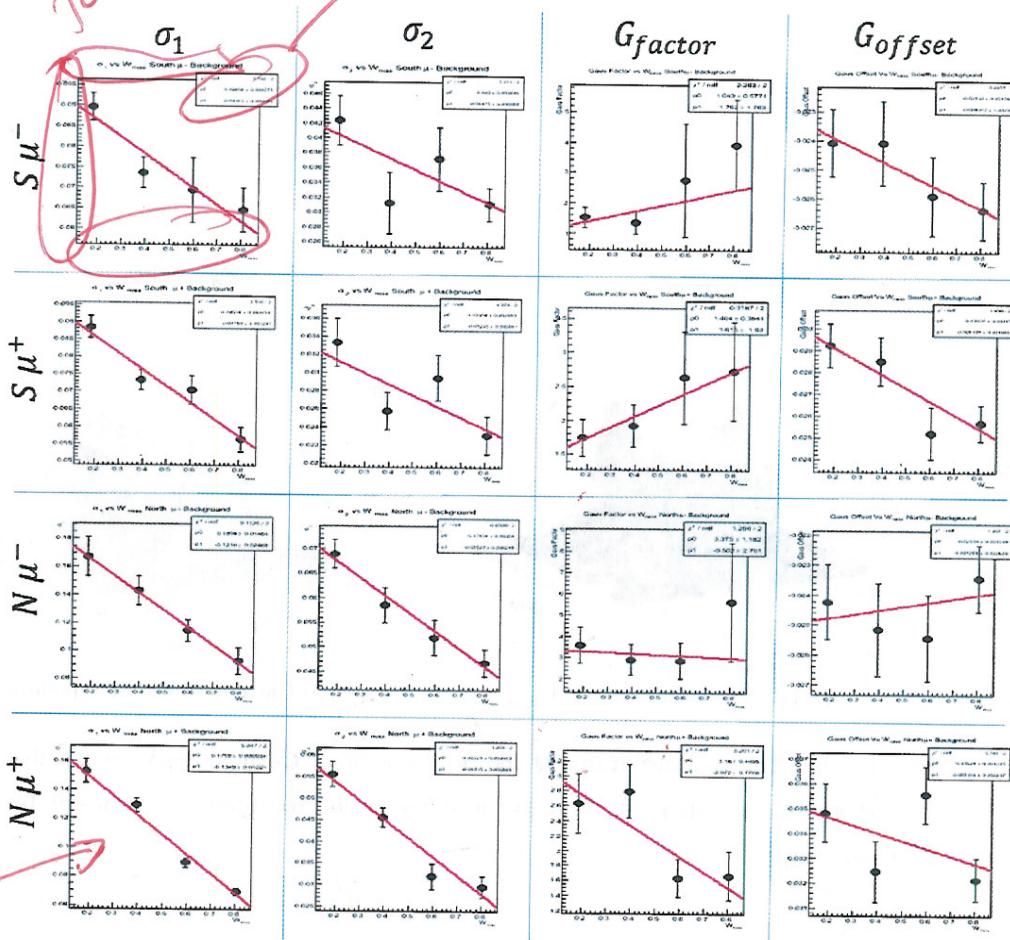
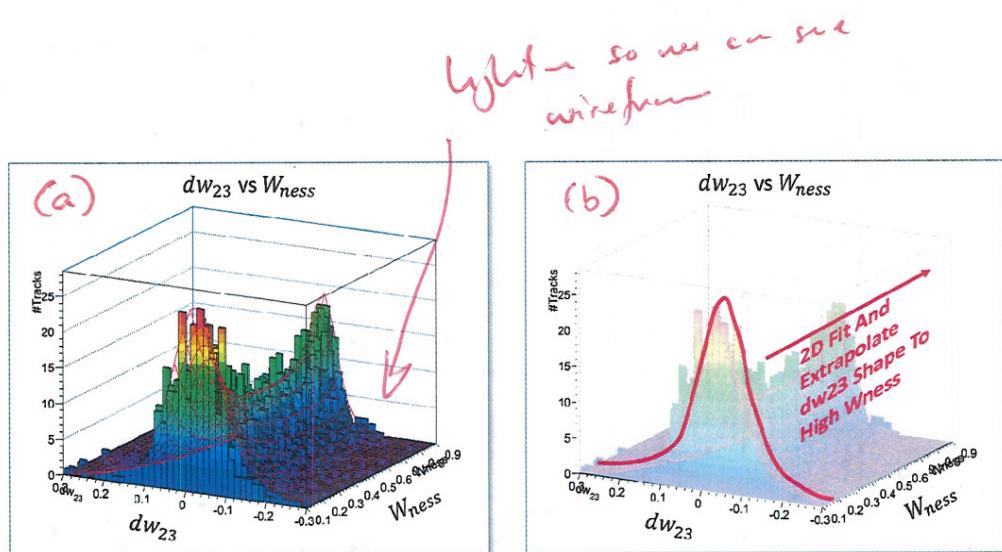


Figure 6.17: The four parameters from the co-axial Gaussian parameterization of dw_{23} as a function of W_{ness} . Though some parameters ($G_{\text{factor}}, N\mu^-$) may appear to be non-linear, note that the uncertainty on some bins is quite large. Rows are arm/charge, labeled on the left, while columns are co-axial Gaussian parameters, summarized in Equation 6.10

why did you
use? did you
draw that..

for text, not caption

DRAFT



(a) The final fit to dw_{23} vs W_{ness}

(b) Cartoon of the extrapolation

Figure 6.18: The red wire-frame is the resultant fit of to the dw_{23} vs W_{ness} distribution. We extrapolate the shape of dw_{23} to the signal region to obtain the hadronic background PDF for dw_{23} .

Panel (a) shows Panel (b) shows ...

already
send in text...

DRAFT

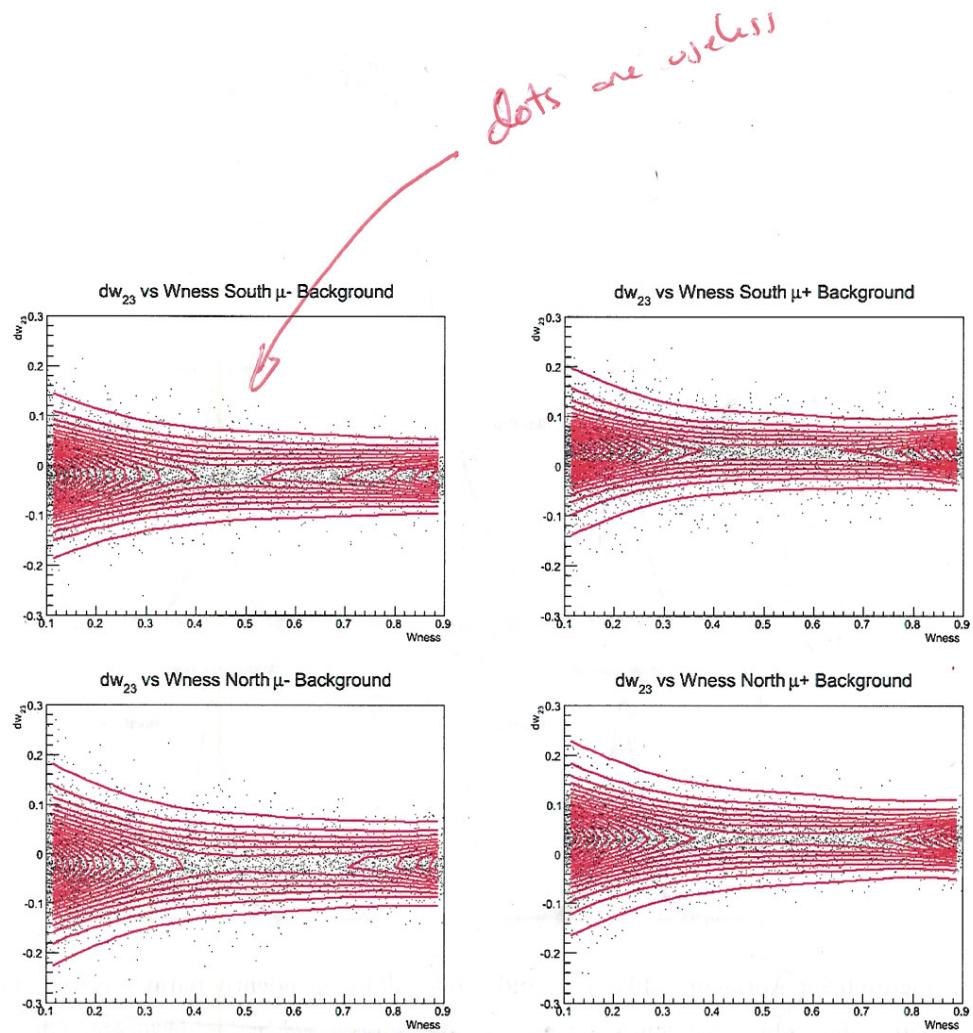


Figure 6.19: An overhead view of the various results of the dw_{23} vs W_{ness} fit for each arm and charge combination.

Obviously there's no point to this
figure... not ref'd in text
Remove

DRAFT

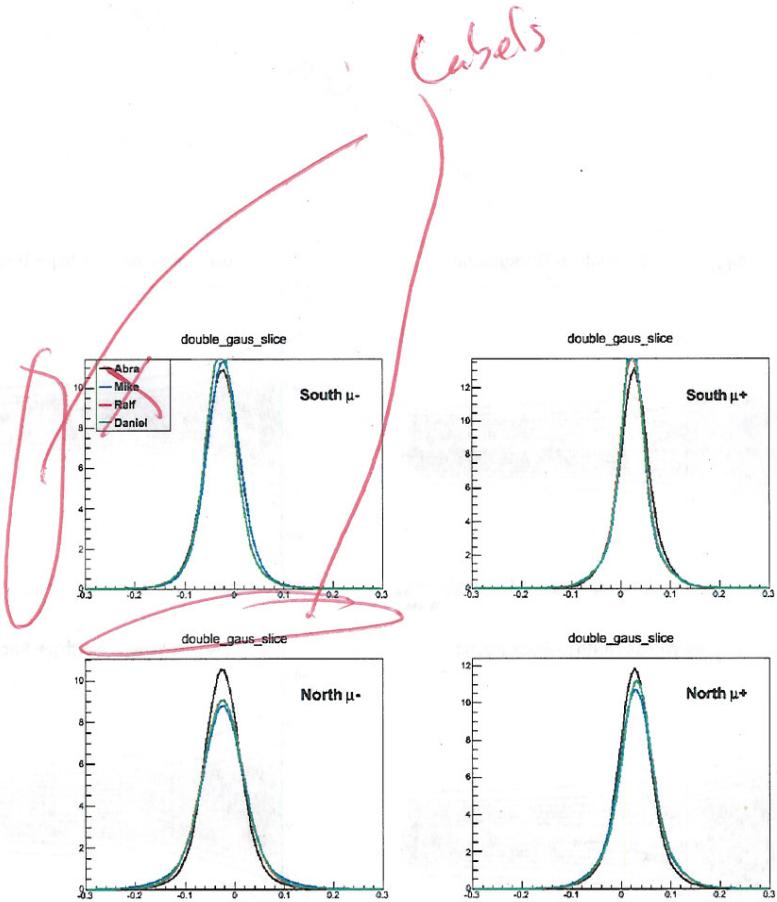


Figure 6.20: Abraham, Mike, Ralf and Daniel all independently parameterized and extrapolated dw_{23} obtaining consistent results. Figure prepared by Dr. Francesca Giordano [70]

no names allowed
you all are same fuck, same data, ...
how could they be diffent?

DRAFT

I stopped here... but I hope you got
the gist... no "we", no names, explain,
more scientific, sharper language.

Finally, you must find someone else to read it who
has either read or (pref) written Thesis...

6.4.2 Muon Background and W-Signal PDFs

The muon background probability density functions and the W-Signal probability density functions must be carefully summed from simulations so as to match the likely composition of the data set. This is done by using the well known cross-section of each of the processes which are simulated and normalizing with the integrated luminosity delivered to PHENIX during the 2013 run of RHIC. This luminosity was found to be 277pb^{-1} .

One caveat is that the minimum bias trigger of PHENIX is easily fooled by effects such as pile-up and multiple collisions. Concretely, this occurs when there is more than one collision in a single bunch crossing. This is typically not a problem when PHENIX operates at lower energies and beam luminosities, but for this data set, it was a real factor, that must be corrected for, else all the ingredients in the muon background cocktail will be present in the wrong amounts and we'll get the wrong answer from using them. Pile up refers to the process where some events aren't read out quickly enough, and so one recorded event will contain information for two actual beam crossings. Unaccounted for, pile-up and multiple collisions both have the affect of lowering the measured luminosity.

The 277pb^{-1} luminosity figure has been corrected for pile-up and multiple collisions. This is a separate analysis, done by my colleagues working on this analysis, so it will not be described in detail in this thesis, however it is described in detail in our analysis note: [70].

Finally, the PDFs used in the EULMF are shown in Figures 6.21-6.24.

With all PDFs prepared, we can perform the extended unbinned maximum likelihood fit, and extract the yields for the number of signal muons, and the number of fake hadronic background muons (recalling that the number of muon background muons are fixed).

The signal to background ratio extraction is summarized by each analyzer in Table ??, for the South Arm $\mu-$ (the canonical cross check).

DRAFT

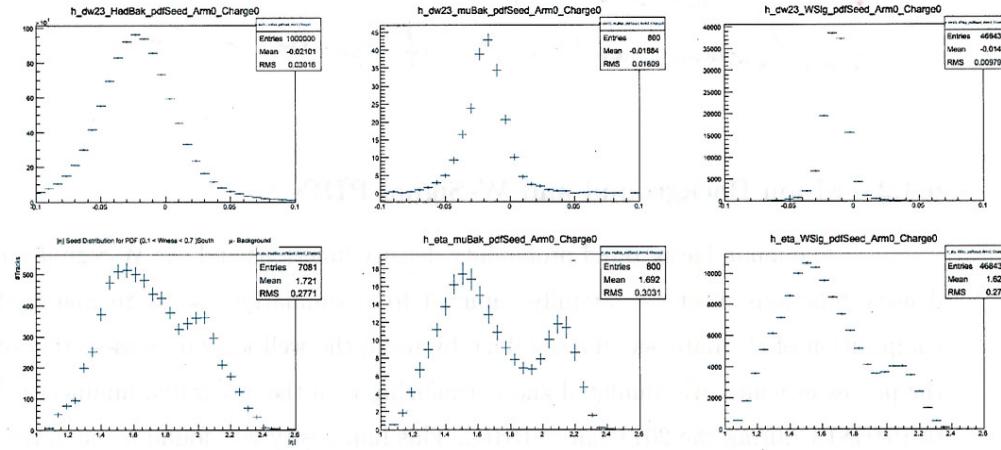


Figure 6.21: Left Column: The hadronic background PDFs, Middle Column: The Summed Muon Background PDFs, Right Column: The W-Signal PDF. For South Arm, $\mu+$

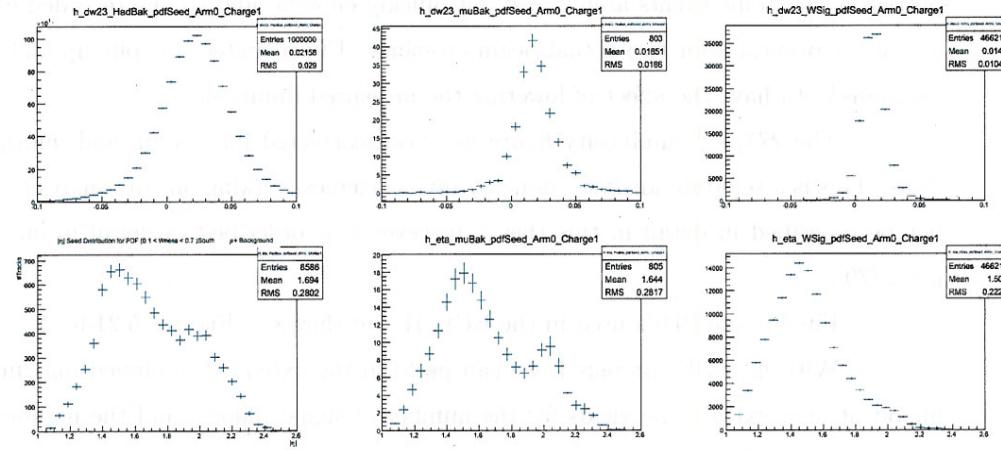


Figure 6.22: Left Column: The hadronic background PDFs, Middle Column: The Summed Muon Background PDFs, Right Column: The W-Signal PDF. For South Arm, $\mu-$