Search for Weak Scale supersymmetric particles in compressed scenarios

Justin Anguiano

B.S. Engineering Physics, University of Kansas, 20XX

M.S. Computational Physics and Astronomy, University of Kansas, 20XX

Submitted to the graduate degree program in Department of Physics and Astronomy and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Committee members

_____
Member Name 1, Chairperson

_____
Member Name 2, Co-Chair

_____
Member Name 3

_____
Member Name 4, The One Who Never Answers Email

_____
Member 5 has a long name

_____
Member 6, External Reviewer

Date defended: _____ July 02, 2019

The Dissertation Committee for Justin Anguiano certifies
that this is the approved version of the following dissertation :

Search for Weak Scale supersymmetric particles in compressed scenarios

_____
Member Name 1, Chairperson

Date approved: _____ August 06, 2019 _____

Abstract

This is the abstract

Acknowledgements

I would like to thank all of the little people who made this thesis possible. Sleepy, Dopey, Grumpy, you know who you are.

Contents

List of Figures

List of Tables

Chapter 1

The Standard Model and Supersymmetry

Abstract

blah blah c1 abstract

1.1 Introduction

c1 introduction section

1.2 The Standard Model

1.3 Supersymmetry

Chapter 2

The CMS Experiment at the Large Hadron Collider

Abstract

This is a chapter absract

2.1 The Large Hadron Collider

here is the body of the chapter

2.2 The CMS Experiment

Chapter 3

Compressed SUSY Search

Abstract

blah blah c3 abstract

3.1 analysis overview

3.2 reference frames

3.3 signal models

Chapter 4

The Tag-and-Probe

Abstract

The tag-and-probe is a method used to measure the selection efficiencies of an object using data. In the context of this compressed SUSY analysis, the Tag-and-probe measures the efficiencies separately of each light lepton selection critera. The total lepton selection efficiency is then computed by combining factorized efficiency components. The same general method is used for both electrons and muons, however, Muons utilize the J/ψ di-muon trigger which allow direct efficiency measurements from data at lower p_T.

4.1 Introduction and Methodology

An important element of a lepton based search for new physics is properly modeling the efficiency of selected leptons. A purely Monte-Carlo driven approach is inadequate in perfectly describing nuances in data due to imperfections in modeling. Instead of trying to model exactly all physics and detector effects with simulation, the efficiencies can be directly measured from data by using the Tag-and-Probe method.

The Tag-and-Probe method is used to measure a selection criteria by using a well known resonance such as a Z, J/ψ, or Υ and counting the number of probes that pass that criteria. Each counted instance of the Tag-and-Probe consists of two selected leptons. One of the selected leptons is the tag and the other is the probe. The tag passes tight selection require-

ment to give high confidence that it isn't a fake lepton. Fake leptons fall into two possible categories: reducible and irreducible. A reducible fake lepton is a particle that fakes the signature of a lepton such as a charged pion. An irreducible fake lepton is an actual lepton which coincidentally passes some selection criteria but is not the targeted leptons of interest e.g. an isolated muon from a jet accompanying a lepton Z decay of interest. The second lepton in the Tag-and-Probe is the probe. The probe is subjected to the selection criteria whose efficiency is being measured. The invariant mass of pair of leptons is calculated and required to fall within a defined range around the resonance. A particular event may have multiple lepton pairs but the tag and the probe are not allowed to switch positions and be counted twice, this would lead to a bias in the efficiency measurement ?AN 111). To avoid bias the tag and probe are required to be oppositely charged and same flavor with the tag being randomly selected. If multiple pairs can formed in a single event no arbitration is used thus the pair that is not from the resonance will then contribute as combinitorial background. The selected probes can either pass or fail their selection leading to the formation of three distributions, one with a passing probe, one with a failing probe, and one with all probes. he probability of observing k passing probes in n Tag-and-Probe pair trials is dependent on the selection efficiency ε and can be expressed as a binomial likelihood $P(k|\varepsilon, n) = \binom{n}{k}\varepsilon^k(1 - \varepsilon)^{n-k}$. From this the MLE estimator for efficiency is the fraction of passing probes from the total number of pairs, or $\varepsilon = k/n$. Technical documentation for the Tag-and-Probe in CMS is scarce, but, an early strategy for fitting efficiency is defined in Berryhill et al. (2010). The legacy code base as of `CMSSW_10_6_X` uses a binned maximum likelihood between the observed passing probes and failing probes wher the efficiency extracted is an explicit fit parameter. The two simultaneously fit functions are:

$$N^{\text{Pass}} = N_{\text{Total}}(\varepsilon \cdot f_{\text{All}}^{\text{sig}}) + \varepsilon_{\text{bkg}} \cdot (1 - f_{\text{All}}^{\text{sig}})) \tag{4.1}$$

$$N^{\text{Fail}} = N_{\text{Total}}((1 - \varepsilon) \cdot f_{\text{All}}^{\text{sig}} + (1 - \varepsilon_{\text{bkg}}) \cdot (1 - f_{\text{All}}^{\text{sig}})) \tag{4.2}$$

$N^{\mathrm{Pass/Fail}}$ is the total number of observed probes that either pass or fail the selection criteria while N_{Total} is the total number of Tag-and-Probe pairs. The binomial estimator for efficiency, ε, enters the fit functions as the first term but is accompanied by a second term that describes the background contribution with its own efficiency $\varepsilon_{\mathrm{bkg}}$. The term $f_{\mathrm{All}}^{\mathrm{sig}}$ is the fraction of background subtracted signal events over the allowed dilepton mass range. $f_{\mathrm{All}}^{\mathrm{sig}}$ depends on the defined signal and background pdfs. The nominal pdfs chosen for reported fits uses a Voigtian+Voigtian signal model combined with an Exponential background model.

4.2 Lepton Object Definitions

Leptons are selected according to the minimium requirment 'VeryLoose', the criteria defining VeryLoose is defined in Table 4.1. The set of VeryLoose leptons is further subdivided by quality into three mutually exclusive categories: Gold, Silver, and Bronze. Each category has a measure of three main quantities: the quality of the pre-determined ID (citation about ID here), the "promptness" or distance of the lepton production point from a primary interaction point, and the isolation. Promptness is measured by the significance of the 3D impact parameter whereas the isolation is a measure of the density of particles in a cone around the lepton. Two similar but complimentary, in terms of background rejection, isolations are used: PFIso and MiniIso. Both isolations are an energy sum of neighboring particles inside a cone. PFIso has a fixed cone size of 0.5? and miniIso cone sizes varies inversely with lepton p_{T} (cite MINIISO)

For miniIso the EA corrections are applied from a centrally provided look up table in bins of pt and eta in the CMSSW Producer/Ntuplizing stage. CITE EA stuff?

The explicit formulas for both electrons and muons can be generalized with the three fac-

torized components previously mentioned:

$$\epsilon_{\text{Gold}} = \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID}} \times \epsilon_{\text{Prompt}|(\text{ID}\cap\text{Isolated})}$$

$$\epsilon_{\text{Silver}} = \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID}} \times \left(1 - \epsilon_{\text{Prompt}|(\text{ID}\cap\text{Isolated})}\right) \quad (4.3)$$

$$\epsilon_{\text{Bronze}} = 1 - \left(\epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID})}\right)$$

Table 4.1: The criteria define the minimum requirements for a lepton accepted. The electron and muon requirements are equivalent in terms of pseudorapidity, vertexing, and isolation but vary in p_{T} threshold and the MVA VLooseFO working point. The MVA VLooseFO ID is also dependent on year. TODO: REMAKE CAPTION

Criteria	Electron	Muon
p_{T}	≥ 5 GeV	≥ 3 GeV
$|\eta|$	< 2.4	< 2.4
$\text{IP}_{3D}/\sigma_{\text{IP}_{3D}}$	< 8	< 8
$|d_{xy}|$	< 0.05 cm	< 0.05 cm
$|d_z|$	< 0.1 cm	< 0.1 cm
$\text{PFIso}_{\text{abs}}$	$< 20 + (300/p_{\text{T}})$ GeV	$< 20 + (300/p_{\text{T}})$ GeV
MVA VLooseFO ID	✓	−

4.3 Electron Tag-and-Probe

The electron tag and probe is done by using Z

4.4 Muon Tag-and-Probe

For the muon tag and probe we measure efficiencies above 20 GeV using the Z and below 20 GeV using J/ψ there are central and forward η bins which are divided at the endcaps at $\eta = 2.1$ The bins edges follow the recommendation from the muon object group and are as follows:

$J^{Low}|p_T \in [3.0, 4.0, 5.0, 6.0, 7.0, 9.0, 14.0, 20.0]$

$J^{Low}||\eta| \in [0, 1.2, 2.4]$

$Z^{High}|p_T \in [10, 20, 30, 40, 60, 100]$

$Z^{High}||\eta| \in [0, 1.2, 2.4]$

The exact criteria chosen for the tag and probe also follows the recommendations from the muon group.

Tag:

$J/\Psi : isGlobalMuon, \ p_T > 5\text{GeV}, \ numberOfMatchedStations > 1,$ Matches trigger object in hltIterL3

Probe:

$J/\Psi :$ Matches trigger object in hltTracksIter or hltMuTrackJpsiEffCtfTrackCands

Tag:

$Z : p_T > 15\text{GeV},$ passes tightID, $(\sum p_T^{chg-hadron})/p_T < 0.2$

Probe:

$Z : No requirement$

For MC there is an inherit difference in the objects selected. Data will also have an implicit selection due to triggering. To account trigger bias and better simulate the behavior of data, trigger bits are added to the denominator of efficiency for all criteria. The triggers available vary from year to year and are as follows:

$2016 JMu7p5Tk2$

$2017 JMu7p5Tk2$

$2018 JMu7p5Tk2$

$2016 ZIsoTkMu22$

$2017 ZisoMu24eta2p1$

$2018 ZisoMu24eta2p1$

The muon efficiencies are derived from two separate processes with the Tag-and-Probe method. For the range of muons with $3 \leq p_T < 20$ the Medium Id is measured directly

from J/ψ. For higher momentum muons above 20 GeV, Z is used to measure all the efficiency components of Equation **??**. The production of J/ψ in b-jets inhibits directly measuring efficiencies with vertexing or isolation, so, the vertexing and isolation efficiencies are extrapolated from Z fits to low p_{T}. Two different binnings and ranges are defined for the Z measurements, they are denoted by $Z_H \in [20, 100]$ GeV, which is used for the higher momentum direct efficiency measurements, and $Z_L \in [6, 50]$ GeV , which is used for extrapolating efficiencies into the lower p_{T} range. The overall efficiency and scale factors for the three muon tiers are obtained by factorizing the efficiency into mutually exclusive components from the subsets defined in Equation **??**. The efficiency factors depend on the range of p_{T} and are defined as:

$p_{\mathrm{T}} \in [3, 20)$

$$
\begin{aligned}
\epsilon_{\mathrm{Gold}} &= \epsilon_{\mathrm{ID}}^{J/\psi} \times \epsilon_{\mathrm{Isolated|ID}}^{Z_L} \times \epsilon_{\mathrm{Prompt|(ID \cap Isolated)}}^{Z_L} \\
\epsilon_{\mathrm{Silver}} &= \epsilon_{\mathrm{ID}}^{J/\psi} \times \epsilon_{\mathrm{Isolated|ID}}^{Z_L} \times \left(1 - \epsilon_{\mathrm{Prompt|(ID \cap Isolated)}}^{Z_L}\right) \\
\epsilon_{\mathrm{Bronze}} &= 1 - \left(\epsilon_{\mathrm{ID}}^{J/\psi} \times \epsilon_{\mathrm{Isolated|ID}}^{Z_L}\right)
\end{aligned}
\tag{4.4}
$$

$p_{\mathrm{T}} \in [20, 100]$

$$
\begin{aligned}
\epsilon_{\mathrm{Gold}} &= \epsilon_{\mathrm{ID}}^{Z_H} \times \epsilon_{\mathrm{Isolated|ID}}^{Z_H} \times \epsilon_{\mathrm{Prompt|(ID \cap Isolated)}}^{Z_H} \\
\epsilon_{\mathrm{Silver}} &= \epsilon_{\mathrm{ID}}^{Z_H} \times \epsilon_{\mathrm{Isolated|ID}}^{Z_H} \times \left(1 - \epsilon_{\mathrm{Prompt|(ID \cap Isolated)}}^{Z_H}\right) \\
\epsilon_{\mathrm{Bronze}} &= 1 - \left(\epsilon_{\mathrm{ID}}^{Z_H} \times \epsilon_{\mathrm{Isolated|ID}}^{Z_H}\right)
\end{aligned}
\tag{4.5}
$$

The order of the conditional efficiencies is important, by applying the vertexing efficiency last the gold, silver, and bronze efficiencies can be parameterized with the same three efficiency components. When the efficiency conditional sequence uses (Prompt | ID ∩ Isolated) last, this efficiency and inefficiency can be directly used in gold and silver respectively. Although the isolation and vertexing requirements are physically uncorrelated, intersection between these sets of criteria exist, thus, the prompt efficiency should be conditionally dependent on both the ID and the isolation. The ID efficiency with statistical errors for both data and MC

9

are shown in Figure 4.1. The other efficiencies for each year for all p_T ranges are included in the appendix. The overlapping bins between J/ψ and Z do not all match within statistical uncertainties. However, the average deviation of the efficiency central values are 0.02% for MC and 1% for data.

The extrapolation of the vertexing and isolation efficiencies below 20 GeV is done by fitting a quadratic polynomial to the efficiencies on the Z_L interval. Both data and MC are shown in Figure 4.2. The errors for each bin are the combined statistical and systematic errors from Table **??** and are adjusted before the polynomial fit. Any efficiencies below 20 GeV are then reported from the fit model. The fit errors are the 68% confidence interval combined with the systematic errors. The worst observed P-value is $\approx 2\%$ but most of the fits are high quality.

The very loose and the efficiency components combined into Gold, Silver, and Bronze are summarized in Figure 4.3, the other years are included in the appendix. The tool used to store/calculate efficiencies can be found at `https://github.com/Jphsx/LepTool`. The cumulative efficiency for a muon also includes the efficiency of the VeryLoose selection and is defined as:

$$\epsilon_\mu = \epsilon_\text{VeryLoose} \times \epsilon_\text{Gold/Silver/Bronze} \tag{4.6}$$

4.5   Lepton Systematics

The systematic error for the muon efficiencies is derived by varying the Tag-and-Probe signal and background models, as well as the mass window and number of bins used in the fit. The systematic error is defined as the maximum spread in efficiencies between the modeling variations. The analysis is not sensitive to muon systematics, so, a simplified approach is used in extracting and applying errors in which we apply systematics from 2017 to all three years. The errors are separated by efficiency and the shape of failed probe backgrounds distribution, leading to separations at $|\eta| = 1.2$ and $p_\text{T} = 20$ GeV, respectively. The low-
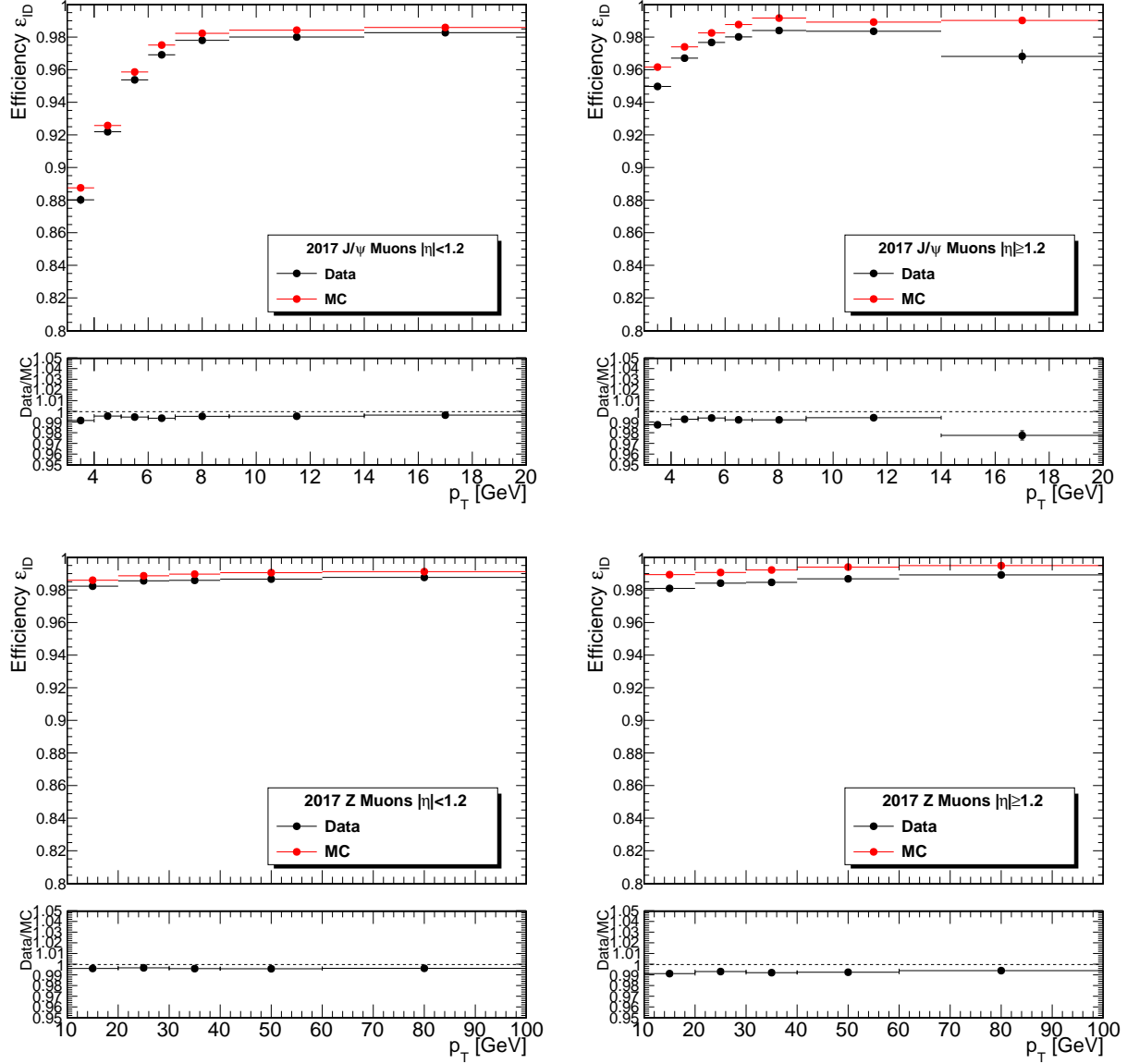
Figure 4.1: Tag-and-Probe efficiencies for the Medium Id in 2017. The left plots show the barrel while the right plots show the endcaps. The top fits use J/ψ resonance while the bottom use the Z resonance.
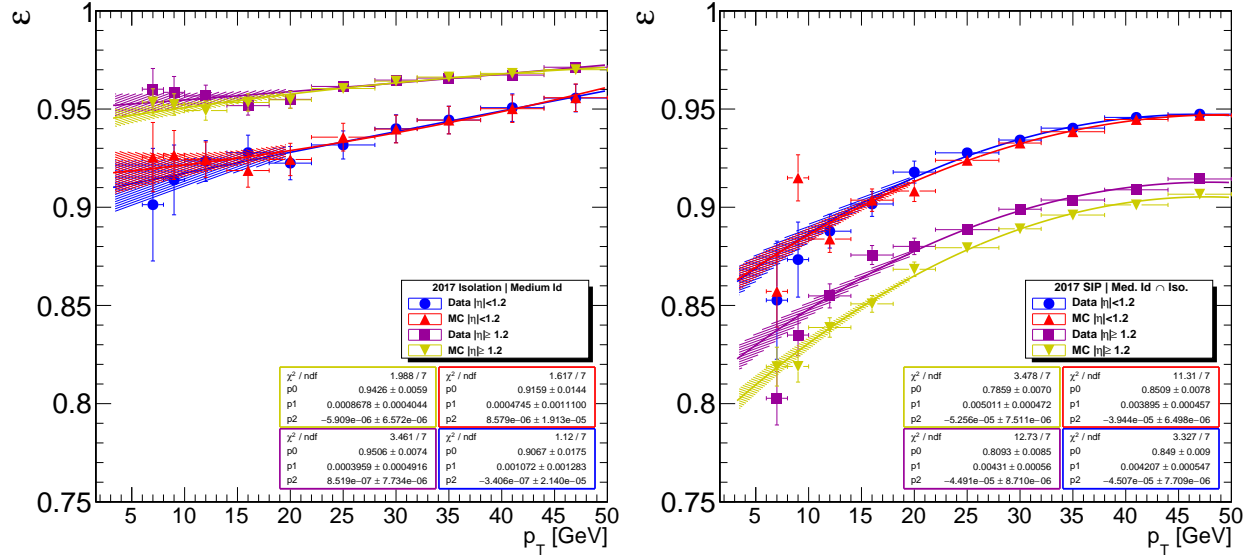
Figure 4.2: The fitted muon isolation and SIP3D efficiencies for 2017. Includes both data and MC which are separated between barrel and endcap.

and high-p_T regions included muons with [3,20) GeV and [20,45] GeV. The efficiency shapes between the years are similar and are consistent with centrally available muon systematics for all three years which advocates a streamlined single year systematics approach. The derived systematics are for 2017 data are shown in Table ??.

2017 Gold Eff.
Data MC Data/MC

0.884 +/- 0.009 0.878 +/- 0.009
1.007 +/- 0.015

0.861 +/- 0.009 0.863 +/- 0.009
0.998 +/- 0.015

0.892 +/- 0.009 0.894 +/- 0.009
0.998 +/- 0.014

0.873 +/- 0.009 0.872 +/- 0.009
1.002 +/- 0.015

0.876 +/- 0.009 0.878 +/- 0.009
0.998 +/- 0.014

0.859 +/- 0.009 0.859 +/- 0.009
1.000 +/- 0.015

0.854 +/- 0.009 0.855 +/- 0.009
0.999 +/- 0.015

0.843 +/- 0.009 0.840 +/- 0.009
1.004 +/- 0.015

0.816 +/- 0.015 0.821 +/- 0.015
0.994 +/- 0.026

0.814 +/- 0.015 0.804 +/- 0.015
1.011 +/- 0.026

0.701 +/- 0.017 0.716 +/- 0.016
0.979 +/- 0.032

0.754 +/- 0.016 0.747 +/- 0.015
1.008 +/- 0.029

p_T [GeV] / η / Data/MC

2017 Silver Eff.
Data MC Data/MC

0.056 +/- 0.006 0.063 +/- 0.006
0.900 +/- 0.119

0.095 +/- 0.006 0.097 +/- 0.006
0.978 +/- 0.086

0.050 +/- 0.005 0.051 +/- 0.005
0.973 +/- 0.147

0.084 +/- 0.006 0.093 +/- 0.006
0.909 +/- 0.082

0.055 +/- 0.005 0.057 +/- 0.005
0.961 +/- 0.130

0.092 +/- 0.006 0.100 +/- 0.006
0.921 +/- 0.076

0.066 +/- 0.005 0.070 +/- 0.005
0.943 +/- 0.106

0.103 +/- 0.006 0.112 +/- 0.006
0.913 +/- 0.068

0.089 +/- 0.006 0.089 +/- 0.006
0.996 +/- 0.099

0.127 +/- 0.007 0.140 +/- 0.006
0.912 +/- 0.063

0.105 +/- 0.009 0.100 +/- 0.008
1.051 +/- 0.118

0.154 +/- 0.008 0.166 +/- 0.008
0.926 +/- 0.068

2017 Bronze Eff.
Data MC Data/MC

0.060 +/- 0.008 0.060 +/- 0.008
1.006 +/- 0.184

0.044 +/- 0.008 0.040 +/- 0.008
1.086 +/- 0.288

0.058 +/- 0.008 0.055 +/- 0.008
1.062 +/- 0.204

0.043 +/- 0.008 0.036 +/- 0.008
1.190 +/- 0.335

0.069 +/- 0.008 0.065 +/- 0.008
1.062 +/- 0.174

0.049 +/- 0.008 0.042 +/- 0.008
1.183 +/- 0.287

0.080 +/- 0.008 0.075 +/- 0.008
1.061 +/- 0.147

0.054 +/- 0.008 0.048 +/- 0.008
1.129 +/- 0.243

0.095 +/- 0.015 0.090 +/- 0.015
1.058 +/- 0.250

0.059 +/- 0.016 0.056 +/- 0.016
1.057 +/- 0.411

0.194 +/- 0.017 0.184 +/- 0.016
1.055 +/- 0.128

0.093 +/- 0.016 0.087 +/- 0.016
1.070 +/- 0.273

2017 V.L. Eff.
Data MC Data/MC

1.000 +/- 0.005 0.999 +/- 0.005
1.000 +/- 0.008

0.997 +/- 0.005 0.991 +/- 0.006
1.005 +/- 0.008

1.000 +/- 0.005 1.000 +/- 0.005
1.000 +/- 0.008

0.997 +/- 0.005 0.992 +/- 0.005
1.005 +/- 0.008

1.000 +/- 0.005 0.999 +/- 0.005
1.001 +/- 0.008

0.996 +/- 0.005 0.988 +/- 0.005
1.009 +/- 0.008

0.999 +/- 0.005 0.998 +/- 0.005
1.001 +/- 0.008

0.995 +/- 0.005 0.983 +/- 0.005
1.013 +/- 0.008

0.998 +/- 0.006 0.995 +/- 0.006
1.002 +/- 0.009

0.994 +/- 0.006 0.974 +/- 0.006
1.021 +/- 0.009

0.997 +/- 0.008 0.991 +/- 0.008
1.005 +/- 0.011

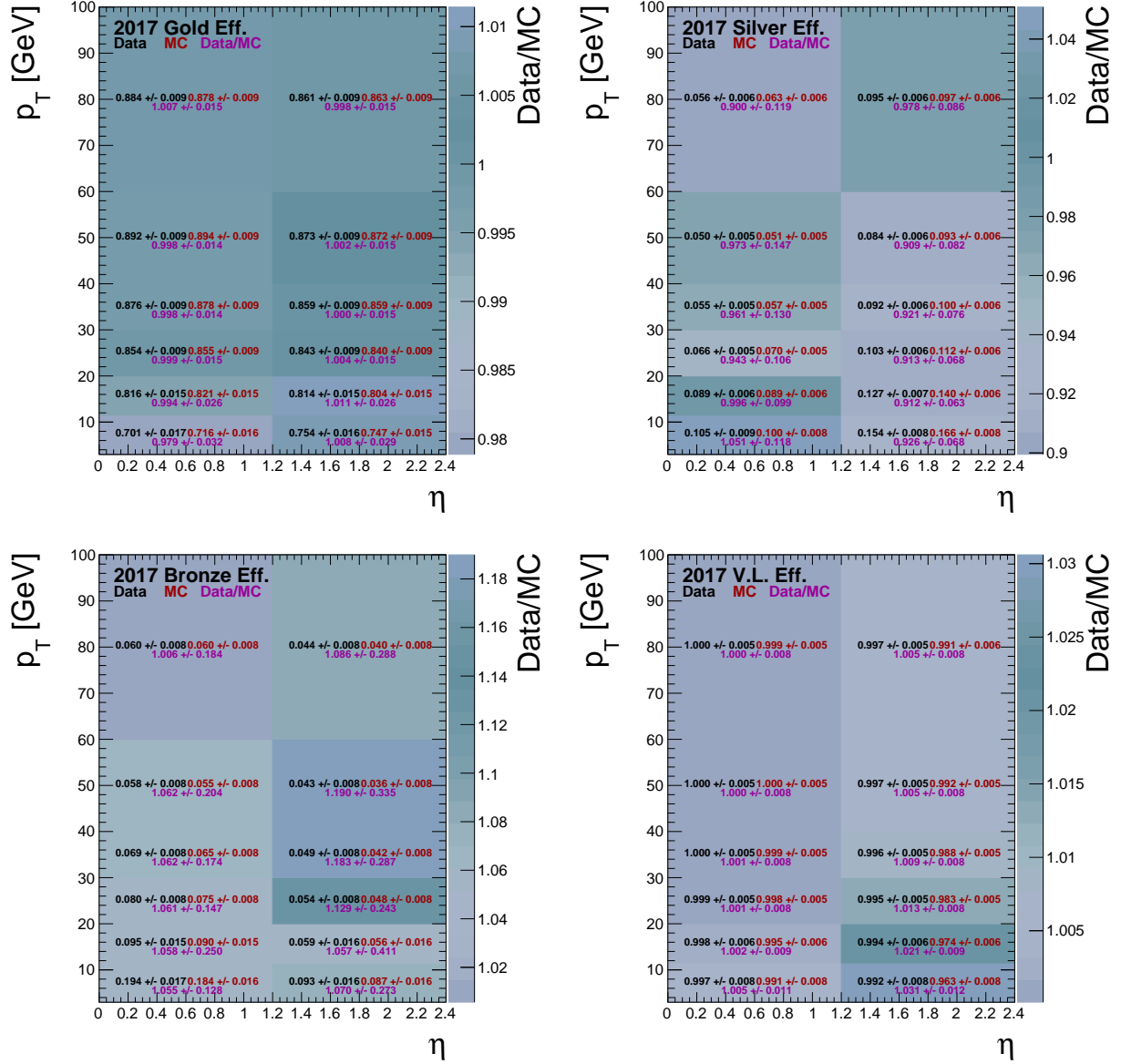0.992 +/- 0.008 0.963 +/- 0.008
1.031 +/- 0.012

Figure 4.3: The combined efficiency components from equations 4.4 and 4.5 and Very Loose for 2017. The low-p_T region (< 20 GeV) includes the contributions from J/ψ as well as the isolation and SIP3D extrapolations. Propagated errors are treated as uncorrelated.

13

References

Berryhill, J., Daskalakis, G., Halyo, V., Werner, J., & Xie, S. (2010). Electron efficiency measurements with 2.88 pb^{-1} of pp collision data at $\sqrt{s} = 7$ tev.

Appendix A

My Appendix, Next to my Spleen

There could be lots of stuff here