

# Search for Weak Scale Supersymmetric Particles in Compressed Scenarios

©2022

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.

---

Graham Wilson, Chairperson

---

Alice Bean, Co-Chair

Committee members

---

Christopher Rogan

---

Ian Lewis

---

Zsolt Talata, 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

---

Graham Wilson, Chairperson

Date approved: August 06, 2019

# Abstract

This is the abstract

# Acknowledgements

Thanks everybody

# Contents

<b>1</b>	<b>The Standard Model and Supersymmetry</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	The Standard Model . . . . .	1
1.3	Supersymmetry . . . . .	3
<b>2</b>	<b>Motivating the Search for SUSY</b>	<b>6</b>
2.1	Introduction . . . . .	6
2.2	Motivating SUSY . . . . .	6
2.2.1	Stabilizing the Higgs mass . . . . .	6
2.2.2	Muon anomalous magnetic moment . . . . .	7
2.2.3	W mass measusremnt . . . . .	9
2.3	The current status of SUSY . . . . .	10
<b>3</b>	<b>The Tag-and-Probe</b>	<b>11</b>
3.1	Introduction and Methodology . . . . .	11
3.2	Lepton Object Definitions . . . . .	14
3.3	Electron Tag-and-Probe . . . . .	17
3.4	Muon Tag-and-Probe . . . . .	21
3.5	Lepton Systematics and Scale Factors . . . . .	27

# List of Figures

1.1	particles figure cite wiki . . . . .	3
3.1	Example Tag-and-Probe Z di-muon fits for passing,failing, and all probes with the Medium Id, $ \eta  < 1.2$ , and $p_T < 20$ GeV . . . . .	13
3.2	Gold (Top-Left), Silver (Top-Right) and Bronze (Bottom) MC truth matching in TTJets sample 2017. Signal is defined here as prompt electrons from a $W$ decay. . . . .	16
3.3	Gold, Silver, and Bronze efficiency on truth matched prompt electrons as signal and secondary electrons as Fakes. . . . .	17
3.4	2017 efficiencies . . . . .	19
3.5	2017 electron GSB efficiency and SF . . . . .	20
3.6	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 $J/\psi$ resonance while the bottom use the Z resonance. . . . .	24
3.7	The fitted muon isolation and SIP3D efficiencies for 2017. Includes both data and MC which are separated between barrel and endcap. . . . .	25
3.8	The combined efficiency components from equations 3.5 and 3.6 and Very Loose for 2017. The low- $p_T$ region ( $< 20$ GeV) includes the contributions from $J/\psi$ as well as the isolation and SIP3D extrapolations. Propagated errors are treated as uncorrelated. . . . .	26
3.9	Tag-and-Probe di-muon mass distributions for both passing and failing probes. The top set of plots consist of probes below 20 GeV and the bottom set are about 20 GeV. . . . .	27

3.10	Example systematic spread from various fit models and binnings for muons.	
	Includes the four combinations of regions either low or high pt and central	
	and forward eta. . . . .	28

# List of Tables

3.1	The criteria that define the minimum requirements for an accepted lepton. 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 also varies between years. . . . .	15
3.2	Data and MC samples for each year used for the electron Tag-and-Probe. . .	17
3.3	selection . . . . .	18
3.4	. . . . .	21
3.5	muon binning . . . . .	22
3.6	The electron systematic error derived from the Tag-and-Probe for 2017 data and split into $p_T$ and $ \eta $ regions. . . . .	29
3.7	The muon systematic error derived from the Tag-and-Probe data and split into $p_T$ and $ \eta $ regions. . . . .	29



# Chapter 1

## The Standard Model and Supersymmetry

### Abstract

Introduce SM, talk about electroweak origins, susy origins

### 1.1 Introduction

c1 introduction section

### 1.2 The Standard Model

The standard model is a collection of an amalgamation of adhoc theories which are used to predict or reproduce experimental data. The theory itself incorporates Quantum Field theory, the Dirac equation, the gauge principle, and the Higgs mechanism. These principles are constrained by physical data and describe the set of elementary particles and interactions between them. The SM consists of 26 free parameters which are input by hand. These parameters are: the masses of the 12 fermions, 3 coupling constants that describe gauge interactions:  $g, g', g_s$ ; 2 parameters to describe the Higgs potential  $\mu, \lambda$ , and 9 mixing angles which describe the PMNS and CKM matrices. The 12 fermion parameters are subdivided by three neutrinos  $m_\nu^i$ , three charged leptons  $m_\ell^\pm$ , and six quarks  $m_q$ ;

Quantum field theory (QFT) provides a description for both known and theoretical particles and interactions by combining quantum theory, the field, and relativity (cite Peskin?).

The gauge theory aspect describes the exact nature of QFT interactions and provides the mechanisms for the electromagnetic, strong, weak forces. We know of four gauge fields:  $G_\mu^C$  which transforms under SU(3) and governs strong interactions,  $W_\mu^{1-3}$  and  $B_\mu$  which transform under SU(2)xU(1) and governs electromagnetic and weak interactions. The combination of the gauge fields and fermion fields with the Dirac equation yields eigenstates that represent fermionic matter particles. In general these particles would be massless if not for the inclusion of the complex scalar Higgs field. The Yukawa coupling to the Higgs field causes spontaneous symmetry breaking creating a non zero vacuum expectation value. The symmetry breaking is responsible for generating masses of the electroweak gauge bosons and. Additionally the interaction between the fermionic fields and the non zero vev generates the masses of SM fermions.

The set of standard model elementary particles can be divided into two subgroups: fermions and bosons. The fermions consist of leptons and quarks. There are three flavors of charged leptons, the electron ( $e$ ), the muon ( $\mu$ ), and the tau ( $\tau$ ). Each charged particle and pairing neutral neutrino. The  $e$  and  $\mu$  are also generally considered as "light" leptons for their small mass relative to the  $\tau$  and the term lepton, depending on context, often refers to only the charged particles. As for the quarks, there are three generations in mass in 3 pairs of quarks. The quarks have complementary fractional electric charges. The lightest set of quarks are the up ( $u$ ) and down ( $d$ ) quarks, followed by the charm ( $c$ ) and strange ( $s$ ), then the bottom ( $b$ ) and extremely massive top quark ( $t$ ). The bosons are the force carrying particles which represent the gauge fields. They are comprised of the vector bosons - the photon ( $\gamma$ ), gluon ( $g$ ), the  $W^\pm$ , and the  $Z^0$  - along with the singular scalar boson the Higgs ( $h$ ). The elementary particles are summarized in Figure 1.1.

The SM is an asymmetric chiral theory, meaning that spinors with different chiralities can transform differently under the Lorentz group and the EW gauge group SU(2)<sub>L</sub>\*U(1) (cite slides). The consequence of chiral properties is limiting the possible helicity combinations between interaction vertices (cite Thompson?). Helicity is defined by the projection of

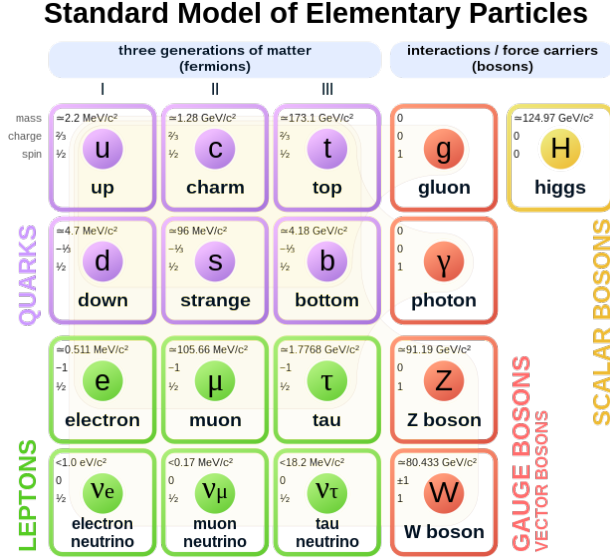


Figure 1.1: particles figure cite wiki

a particles spin ontion its direction of motion (thompson). A particle is considred right-handed if the direction of its spin and motion is parallel. It is left-handed if spin and motion antiparallel. (cite wikipedia helicity). This peculiar property shows up with the W boson, which only couples to left handed particles or right handed antiparticles.

### 1.3 Supersymmetry

Supersymmetry is an extension of the standard model. It adds an generator that rotates the spin between bosons and fermions. This then introduces a bosonic degree of freedom for every fermionic degree of freedom (cite run2 susy paper) and generates super partners for each particle which differ in spin by a half integer. Thus there is a mirrored symmetry of elementary particles with a super symmetric partner and are refered to as sparticles. Each fermionic sparticle carries the same as its fermion partner but with an "s" prefix e.g. sfermion, squark, selectron. As for the bosonic sparticles, the gauge  $B$  and  $\vec{W}$  fields are accompanied by the the super fields - Higgsino  $\tilde{H}$ , Bino  $\tilde{B}$ , and Wino  $\tilde{W}$ . The B and  $\vec{W}$  SM fields mix and yield mass eigenstates for  $\gamma, Z, W^\pm$ . The Higgsino, Bino, and Wino mix

to produce four neutral and two charged eigenstates, the neutralinos ( $\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$ ) and charginos ( $\tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm$ ). SUSY also requires an additional higgs doublet to give mass to up-type and down-type fermions which leads to five higgs boson states with 2 Charged higgs and 3 neutral higgs with the lightest neutral higgs representing the SM higgs boson. The second higgs doublet also introduces another vev and the ratio of the two is commonly denoted as  $v_1/v_2 = \tan \beta$ . With the introduction to a new set of particles we denote a new bookkeeping parameter similar to lepton number or baryon number conservation which is called R-parity where the total number of SM particles (+1) and sparticles (-1) are conserved. R-parity conservation then requires sparticles to be produced in pairs. If R-parity is violated, the common consequence is that the lightest supersymmetric particle (LSP) is unstable.

Supersymmetry is an extremely expansive model and intractable to experimentally test without significant well motivated simplifications. The most experimentally common is the Minimally Super Symmetric Standard Model (MSSM). The MSSM contains the smallest number of new particle states and new interactions which are consistent with phenomenology (cite howie direct weak scale book). The MSSM is still experimentally inaccessible due to the presence of over 100 parameters, a small change in parameter space could completely change the overall structure of masses and couplings significantly morphing the experimental signature. To tackle such a vast parameter space further simplification is needed, an example of a simpler model the phenomenological MSSM (pMSSM). This contains 19 parameters which includes mass parameters for each generation of squark and slepton, parameters to control the mixing of super fields, and dials for the higgs. More commonly the MSSM is boiled down into 5 parameters in Gravity -mediate supersymmetry breaking models or (minimal supergravity mSUGRA). It can also be reduced, which is what our MC is made from, into 4 parameters  $M_1, M_2, \mu, \tan \beta$ . These come from Realistic simplified gaugino-higgsino models in MSSM (cite benjamin Fuks papaer). A direct quote from the paper describing the 4 parameters: "that are the off-diagonal Higgs(ino) mass parameter, the ratio of the vacuum expectation values of the neutral components of the two doublets of Higgs fields and the two

soft supersymmetry-breaking electroweak gaugino mass parameters, respectively."

Include a plot with mass hierarchies. Since there are so many possible parameters, varying sets of parameters can produce significant differences in experimental signatures and topologies. Typically for a model we decouple specific sectors when generating monte carlo, For instance if we are searching for sleptons, the squark or electroweakino sector will be chosen to be significantly heavier (out of current experimental range) effectively decoupling it from the slepton sector. Then from a simplified model with everything else decoupled we scan various topologies with particular mass values.

SUSY Chirality!

MSSM how many parameters, what about msugra and all of that, what are the main interesting parameters??

talk about the higgsino double and the particles that arise from this model

Talk about and define r-parity, what is rparity pair production

simplified models

pmssm

what are the important parameters and specific particles?

## Chapter 2

# Motivating the Search for SUSY

### Abstract

First im opening up with the basic motivations for susy, like solving the hierarchy problem and dark matter candidate, next we look at a theoretical motivation for SUSY via the higgs mass. Motivate simplified models with naturalness etc, talk about how susy needs to be at a few TeV scale to work out. Then we visit two recent experimental measurements which strongly motivate the search for susy and more speficically this body of work.

### 2.1 Introduction

susy is dark matter candidate, susy solve hierarchy problem

### 2.2 Motivating SUSY

blah summarize higgs mass,  $g-2$  , and w mass measurement

#### 2.2.1 Stabilizing the Higgs mass

UV problem with the higgs and large scale corrections to the mass

### 2.2.2 Muon anomalous magnetic moment

g-2 overview, measurement, could point to susy or could deal with QCD lattice and weird diagrams

g-2 is an experiment designed to measure the anomalous magnetic dipole moment of the muon. The spin magnetic moment of a charged, spin-1/2 particle that does not possess any internal structure (a Dirac particle) is given by (wiki direct quote [https://en.wikipedia.org/wiki/G-factor\\_\(physics\)](https://en.wikipedia.org/wiki/G-factor_(physics)))  $\mu = g \frac{e}{2m} \mathbf{S}$ . where  $g$  is the particles g-factor,  $\mu$  is the magnetic moment,  $m$  is the particle mass and  $S$  is the spin. The g-factor in quantum electrodynamics is close to 2 so typically the reported measurement is the difference from 2 or  $g-2$  or as a significance  $a_\mu = g - 2/2$ . The difference from 2 arises from higher order contributions in quantum field theory

What is the g-factor The quantity  $g$  is a factor in the electromagnetic coupling of charged particles to a photon. The factor largely depends on the tree level lepton-photon coupling but gets small quantum corrections from higher order loops, the largest being the single photon loop or Schwinger term. There are three types of corrections applied when determining the SM g-factor for light leptons – QED, Electroweak, and Hadronic. Corrections due to the Higgs are neglected due to the mass disparity  $m_h \gg m_{e,\mu}$  and the mass dependence in the Higgs coupling. This causes the Higgs contributions to be unnecessary at our current level of precision. If the g-factor is calculated to first order in QED the g-factor is exactly 2, when accounting for quantum corrections the g-factor deviates slightly from 2. Experimentally the deviations from 2 are studied, which are written in the form  $a_\ell = \frac{g-2}{2}$  and referred to as  $(g - 2)$ . These small contributions are interesting because if they were to deviate from the SM prediction, it would be an indication of new particles interacting with the SM leptons. There are three possible g-factors we can measure, one for each generation of lepton:  $a_e, a_\mu, a_\tau$ . The current best candidate to both test the SM and search for new physics is by measuring  $(g - 2)_\mu$  or  $a_\mu$  because of experimental precision potential. The electron measurement is already known to the highest precision and is expected to have

the smallest contributions from new physics (cite). The  $(g - 2)_\tau$  is not yet experimentally tractable to measure with precisions that are competitive with  $e$  and  $\mu$ , so  $(g - 2)_\mu$  has been measured at both at Brookhaven National Lab (BNL) and again at Fermi National Accelerator Laboratory (FNAL).

What is the current calculation and current status of  $(g-2)_\mu$  The currently accepted best SM prediction of  $a_\mu$  from (CITE) includes QED, Electroweak(EW) and Hadronic contributions and is reported as  $a_\mu^{SM} = a_\mu^{QED} + a_\mu^{EW} + a_\mu^{Hadronic} = 116591810(43) \times 10^{-11}$ . For each of the  $a_\mu$  components, the QED component enters at the  $O(10^{-3})$  and is known to  $O(10^{-11})$ . the EW component enters the sum at  $O(10^{-9})$  and is known to  $O(10^{-10})$ . Finally the most complicated component, hadronic, contributes at  $O(10^{-8})$  and is known up to  $O(10^{-9})$ , the main sub components that contribute to the  $a_\mu^{Hadronic}$  is the Hadronic vacuum polarization and light by light scattering, diagrams illustrated in Figure X. The hadronic precision is constrained by data driven measurements and computation approaches – QCD lattice theory, this error dominates the overall uncertainty of  $a_\mu$ . The BNL measurement of  $a_\mu$  yields a difference with the SM prediction of  $\Delta a_\mu := a_\mu^{BNL} - a_\mu^{SM} = 279(76) \times 10^{-11}$  which is a significance of  $3.7\sigma$ . The most recent  $a_\mu$  measurement from FNAL confirms the BNL measurement within  $1\sigma$  and the combined experimental average increases the SM deviation with a significance of  $4.2\sigma$ .

What could this deviation mean? The  $4.2\sigma$  significance is a compelling sign for potential new physics, but can be somewhat explained by improvements in QCD lattice calculations of the HVP and LBL contributions. There is also a calculation which resolves this tension up to  $1.2\sigma$  (cite BMW). The more interesting explanation is that the tension could be due to effects of new particles. There are several models which can quantify the  $a_\mu$  SM deviation, those being Super Symmetry, Dark Matter (DM mediator Dark photon), Lepto quarks, 2 Higgs doublet models.



### 2.2.3 W mass measusremnt

the most recent w mass measurement yielded a heavy W, this higher mass is more favorable for light higgsino and compressed susy models

What is the W boson The W boson is an very important and peculiar particle, it is the electrically charged boson that mediates flavor change but only with left handed particles. The mass of the W-boson underpins many important parameters in the SM. In fact,  $m_W$  is related to the vacuum expectation value  $v$  of the Higgs, this implies that scale of mass depedent coupling of the higg field to all particles is tuned by the mass of the W. Similary the W mass is related to the  $g$  factor from  $(g - 2)_\ell$

What is the current status of the W boson? Since there is interdependence of many paramters such as  $v$ ,  $m_z$  there is no exact SM prediction of the W mass, rather a value that is constrained by experimatally measured parameters. at tree level  $m_W$  can be parameterized with  $m_Z$  which can be precisely measured in the visible leptonic and hadronic modes,  $G_F$  or the fermi constant which is also related to  $v$  and can be measured precisely with muon lifetime, and  $\alpha$  the fine structure constant (cite sm w paper). The most recent expereimental measurment of  $m_W$  was performed by CDF II at the tevatron. The mass was obtained by fitting the kinematic distributions from light leptonic decays recoiling against a system of jets. This measurement is 50% more precise than the previous measurement by ATLAS (cite atlas) and heavier than the SM prediction. The combination deviation and precision results in a  $7\sigma$  significance with the SM (cite CDFII).

What could this deviation mean? Possible early sign of new physics would be slight inconsistencies between measurements and different SM obervables, and the  $7\sigma$  deviation of the W mass, if correct, is a very strong indicator of new physics. There are a slew of SUSY models that could explain the excessive mass of the W boson. (CDF lightsusy) In general, a slightly heavier W favors light SUSY models. This means electro weak scale SUSY particles in the cases of light wino/bino and higgsino models, many of which favor compressed mass

scenarios. Due to the interdependence of  $m_W$  and  $(g-2)_\mu$ , these parameters both constraint compressed SUSY and spotlight a critical area to search. An example of model points of sleptons and gaugino models which satisfy the newest  $g-2$  and  $W$  mass constraints is shown in FIGURE Z (cite Wmass and g-2 sven paper)

## 2.3 The current status of SUSY

drop the most recent limits here, start with multi TeV excluded gluino and squark models which leaves the a good place to search in the weak scale sector with electoweakinos. Talk about electroweak limits and how alot of these are excluded already one of the remaining places to search is the compressed corridor where mass splittings are small. link this limit motivation with how both g-2 and  $W$  mass favor compressed scenarios

## Chapter 3

### 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( $e/\mu$ ) selection criteria. 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  $J/\psi$  di-muon trigger which allow more precise efficiency measurements from data at lower  $p_T$ .

#### 3.1 Introduction and Methodology

An important element of a lepton based search 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  $Z$ ,  $J/\psi$ , or  $\Upsilon$  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 leptonic 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 the 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, as double counting would lead to a bias in the efficiency measurement [? ]. To avoid bias, the tag and probe are required to be the opposite charge and same flavor where the tag is randomly selected. If multiple same flavor lepton pairs occur in single event i.e. there are multiple probes to a single tag, the treatment for selecting the pairs differs between electrons and muons. There is no specific study which led to justifying the differing arbitration approaches in flavors, only that the choice reflects the default choices implemented in the existing code bases. For muons, no arbitration is used, all pairs are utilized which means an additional pair not truly from the resonance will then contribute as combinatorial background in a single event. For electrons, only a single probe is selected per event which has the highest  $p_T$ . The selected probes can either pass or fail their selection which leads to the formation of three distributions, one with a passing probe, one with a failing probe, and one with all probes. An example of all three distributions is shown in Figure 3.1. The probability of observing  $k$  passing probes in  $n$  Tag-and-Probe pair trials is dependent on the selection efficiency  $\varepsilon$  and can be expressed as a likelihood from the binomial probability density  $P(k|\varepsilon, n) = \binom{n}{k} \varepsilon^k (1 - \varepsilon)^{n-k}$ . The MLE estimator for efficiency is then the fraction of passing probes to 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 [? ]. The legacy code base as of CMSSW\_10\_6\_X uses a binned maximum likelihood between the

observed passing probes and failing probes where the efficiency extracted is an explicit fit parameter. The two simultaneously fit functions are:

$$N^{\text{Pass}} = N_{\text{Total}}(\varepsilon \cdot f_{\text{All}}^{\text{sig}}) + \varepsilon_{\text{bkg}} \cdot (1 - f_{\text{All}}^{\text{sig}}) \quad (3.1)$$

$$N^{\text{Fail}} = N_{\text{Total}}((1 - \varepsilon) \cdot f_{\text{All}}^{\text{sig}} + (1 - \varepsilon_{\text{bkg}}) \cdot (1 - f_{\text{All}}^{\text{sig}})) \quad (3.2)$$

$N^{\text{Pass/Fail}}$  is the total number of observed probes that either pass or fail the selection criteria while  $N_{\text{Total}}$  is the total number of Tag-and-Probe pairs. The binomial estimator for efficiency,  $\varepsilon$ , enters the fit functions as the first term but is accompanied by a second term that describes the background contribution with its own efficiency  $\varepsilon_{\text{bkg}}$ . The term  $f_{\text{All}}^{\text{sig}}$  is the fraction of background subtracted signal events over the allowed dilepton mass range.  $f_{\text{All}}^{\text{sig}}$  depends on the defined signal and background pdfs. The nominal pdfs chosen for reported fits uses a 5 parameter Voigtian+Voigtian signal model which share a common mean but use independent  $\Gamma$  and  $\sigma$ . The signal model is combined with an Exponential background model.

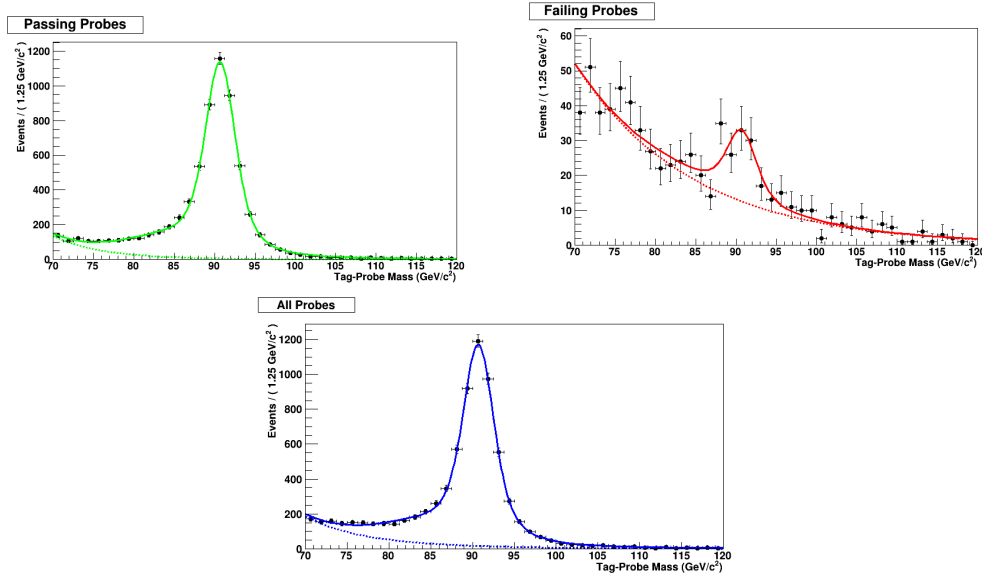


Figure 3.1: Example Tag-and-Probe Z di-muon fits for passing,failing, and all probes with the Medium Id,  $|\eta| < 1.2$ , and  $p_T < 20$  GeV

## 3.2 Lepton Object Definitions

Leptons are selected according to the minimum requirement “VeryLoose” which depend kinematic and topological quantities which are shown in Table 3.1. The electrons use an additional loose MVA requirement: MVA VLooseFO ID [? ]. The set of VeryLoose leptons are further subdivided by quality into three mutually exclusive categories: Gold, Silver, and Bronze. Each category has a measure of three main quantities, the first being the quality of the pre-determined Id. The Id’s differ per flavor and are the standard working points defined by the corresponding physics object group. The muons use the Medium Id [? ] and electrons use a more strict selection, due to their messy nature, with the Tight Id [? ]. The second quantity is the “promptness” or distance of the lepton production point from the primary vertex. Promptness is measured by the significance of the 3D impact parameter (SIP3D) which is defined as the impact parameter normalized by its measured error. A  $SIP3D > 1$  is associated with a secondary particle which is not produced at the primary vertex. The last component is the isolation, a measure of the density of particles in a cone around the lepton. Two similar but complimentary absolute isolations are used: PFIso [? ] and MiniIso [? ]. Both isolations are an energy sum of neighboring particles inside a cone, but, PFIso has a fixed cone size of  $R = 0.4$  cm and miniIso cone sizes varies inversely with lepton  $p_T$  as shown in 3.3.

$$R_{\text{miniIso}} = \begin{cases} 0.2 & p_T^\ell < 50\text{GeV} \\ \frac{10}{p_T^\ell} & 50\text{GeV} \leq p_T^\ell \leq 200\text{GeV} \\ 0.05 & p_T^\ell > 200\text{GeV} \end{cases} \quad (3.3)$$

Mini isolation also includes effective area pile-up corrections provided in a look up table of bins of  $p_T$  and  $\eta$  in the CMSSW Producer/Ntuplizing stage. The implementation of mini-isolation and their corrections utilize the same IsoValueMap producer as used in NANO AOD as of CMSSW\_10\_6\_X.

The explicit flavor independent formulas for Gold, Silver, and Bronze can be generalized

by the product of three components which are the measured efficiencies of the three previously mentioned quantities. The efficiencies take the form of conditional probabilities to be measured independently in sequence relative to each other:

$$\begin{aligned}
\epsilon_{\text{Gold}} &= \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID}} \times \epsilon_{\text{Prompt}|\text{(ID}\cap\text{Isolated)}} \\
\epsilon_{\text{Silver}} &= \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID}} \times (1 - \epsilon_{\text{Prompt}|\text{(ID}\cap\text{Isolated)}}) \\
\epsilon_{\text{Bronze}} &= 1 - (\epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID}})
\end{aligned} \tag{3.4}$$

The subscript for an efficiency, e.g.  $\epsilon_{\text{Prompt}|\text{(ID}\cap\text{Isolated)}}$ , reads as the efficiency to pass the SIP3D requirement given the lepton passes the Id and Isolation requirements. From equation 3.4 the Gold, Silver, and Bronze efficiencies can be read off as Gold passes all criteria, Silver fails only the SIP3D requirement, and Bronze fails either the Id or isolation and is agnostic to SIP3D. While isolation and vertexing requirements are physically uncorrelated, there is an intersection between the two, meaning a lepton can be both prompt and isolated. This intersection then demands the necessity for conditional efficiencies. The order of the conditional efficiencies is also chosen to minimize the number of measured efficiencies by reusing efficiencies across Gold, Silver, and Bronze.

Table 3.1: The criteria that define the minimum requirements for an accepted lepton. The electron and muon requirements are equivalent in terms of pseudorapidity, vertexing, and isolation but vary in  $p_{\text{T}}$  threshold and the MVA VLooseFO working point. The MVA VLooseFO ID also varies between years.

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

The advantage of having various lepton quality categories allows for robust sensitivity to a wide range of signal processes. This strategy boosts the overall modeling statistics and

provides control regions for multiple scenarios. The populations of different truth selected objects are shown in Figure 3.2 and the overall efficiency for Gold, Silver, and Bronze on truth matched objects are shown in Figure 3.3. The gold region is mainly populated by prompt and isolated leptons that are produced within the primary vertex. This region also coincides with the signature of many targeted electroweakino models. The silver selection accommodates both leptonically decaying taus, providing an ideal region for stau's, and assists in recovering efficiency of isolated b decays in stop production. The bronze selection is rich in fake leptons and provides the best regions to extract overall fake rates for other regions as well as a surplus of events to anchor the fit.

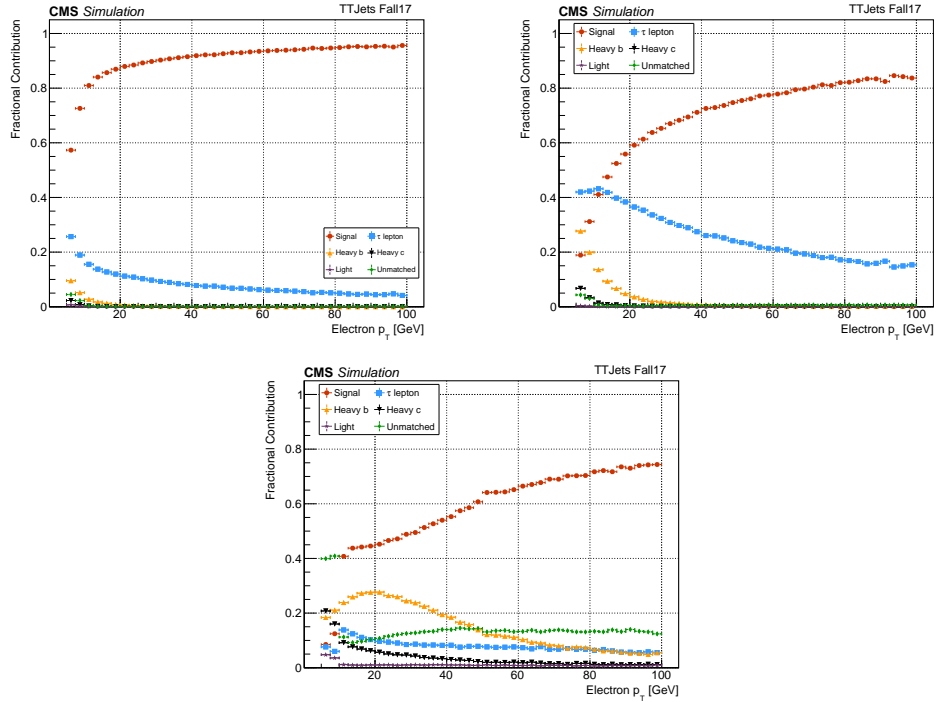


Figure 3.2: Gold (Top-Left), Silver (Top-Right) and Bronze (Bottom) MC truth matching in TTJets sample 2017. Signal is defined here as prompt electrons from a  $W$  decay.



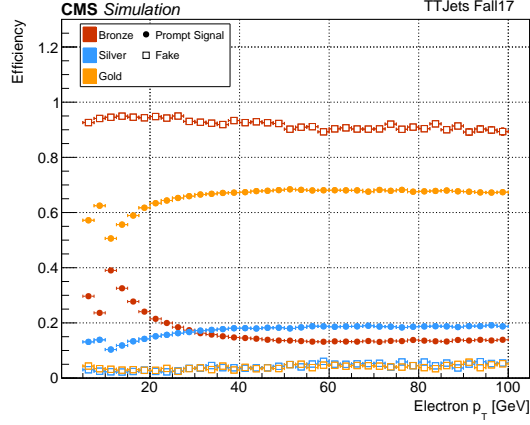


Figure 3.3: Gold, Silver, and Bronze efficiency on truth matched prompt electrons as signal and secondary electrons as Fakes.

### 3.3 Electron Tag-and-Probe

The electron tag and probe is done by using the Z resonance over the entire  $p_T$  range of selected electrons. The selected binnings follow the  $p_T$  and  $\eta$  binning conventions from the electron physics object group and are  $p_T \in [5, 10, 20, 30, 40, 70, 100]$  and  $|\eta| \in [0, 0.6, 1.4, 2.4]$ . The electron Tag-and-Probe tools uses a centrally curated CMSSW PhysicsTools in `CMSSW_10_2_X`. The software pipeline consists of two steps, an ntuplizing stage and a fitting stage. The Ntuplizing stage selects Tag-and-Probe pairs along with all potential variables of interest and loads them onto an ntuple using `TnPTreeProducer` [? ]. The samples used in the Ntuplizing stage are listed in Table 3.2. In the fitting stage, a random subset of of TnP pairs are sampled with `TnPTreeAnalyzer` [? ]. The analyzer performs all of the fitting and efficiency measurements according to the specified selection criteria.

A general selection is applied for electron TnP candidates. The selection for electrons dif-

Table 3.2: Data and MC samples for each year used for the electron Tag-and-Probe.

Type	Year	Sample Name
Data	2016	/SingleElectron/Run2016B-17Jul2018_ver2-v1/MINIAOD
Data	2017	/SingleElectron/Run2017C-31Mar2018-v1/MINIAOD
Data	2018	/EGamma/Run2018A-PromptReco-v1/MINIAOD
MC	2016	/DYJetsToLL_Pt-100To250_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3_ext5-v2/MINIAODSIM
MC	2017	/DYJetsToLL_Pt-100To250_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14-v1/MINIAODSIM
MC	2018	/DYJetsToLL_Pt-100To250_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM

Table 3.3: selection

Tag-and-Probe Electron Candidate Selection Criteria			
Tag	Probe	Super Cluster	Pair
$ \eta_{SC}  \leq 2.1$ veto $1.4442 \leq  \eta_{SC}  \leq 1.566$ $p_T \geq 30.0$ GeV Passes Tight Id	$ \eta_{SC}  \leq 2.5$ $E_{ECAL} \sin(\theta_{SC}) > 5.0$ GeV	$ \eta  < 2.5$ $E_T > 5.0$ GeV	$50\text{GeV} < m_{ee} < 130\text{GeV}$

fers between the tag and probe, but, both depend on super cluster (SC) kinematics. The super clusters are expected to fall within the calorimeter acceptance which includes vetoing super clusters in the endcap gaps. The invariant mass of the electron of the pair also is required to fall within a specified Z-window. The selection specifics are listed in Table 3.3. The tag is also required to pass a trigger requirement to reflect the inherit trigger bias which is not applied in simulation by default. The triggers selected are HLT electron collections and are grouped by specific paths and filters. The electrons are matched to trigger objects in the path/filter combination and passed based on the OR of triggers in the collection. The probes are not subjected to trigger matching. The chosen trigger combinations are HLT\_Ele27\_eta2p1\_WPTight\_Gsf\_v\*, HLT\_Ele32\_WPTight\_Gsf\_L1DoubleEG\_v\*, HLT\_Ele32\_WPTight\_Gsf\_v\* for 2016 through 2018 respectively.

The measurments of the gold silver and bronze efficiencies components, based on Equations 3.4, are shown in Figure 3.4. The relative efficiencies per component range from approximately 75% to 95% with a slight dependence on  $|\eta|$  which is the strongest lower  $p_T$ . The largest combined systematic and statistical errors are  $O(4\%)$  and occur in data with the lowest  $p_T$  bins. The data and MC agreement is within a few percent for both the Id and Isolation but the average data and MC agreement in SIP3D averages closer to  $O(10\%)$  with the highest  $p_T$  bins discrepancies about 20% and a consistent deficit in data efficiency. The product of the efficiency components into their corresponding Gold, Silver, and Bronze category is shown in Figure 3.5. The efficiency for Very Loose is also included separately but

is factored into the denominator efficiencies components, so, the Gold, Silver, and Bronze efficiencies represent the overall electron efficiency for that particular lepton ranking. The range of efficiencies for each ranking are  $(50 - 70)\%$ ,  $(10 - 20)\%$ , and  $(10 - 30)\%$  for Gold, Silver, and Bronze respectively. The component combined agreement for all three ranks ranges around 10% to 20% but large discrepancies can be seen at the highest and lowest  $p_T$  bins for Silver and Bronze. Better measurements could be obtained by using a different resonance such as  $J/\psi \rightarrow ee$  to measure the lower  $p_T$  ranges, however, data triggers with electrons for  $J/\psi$  are not available.

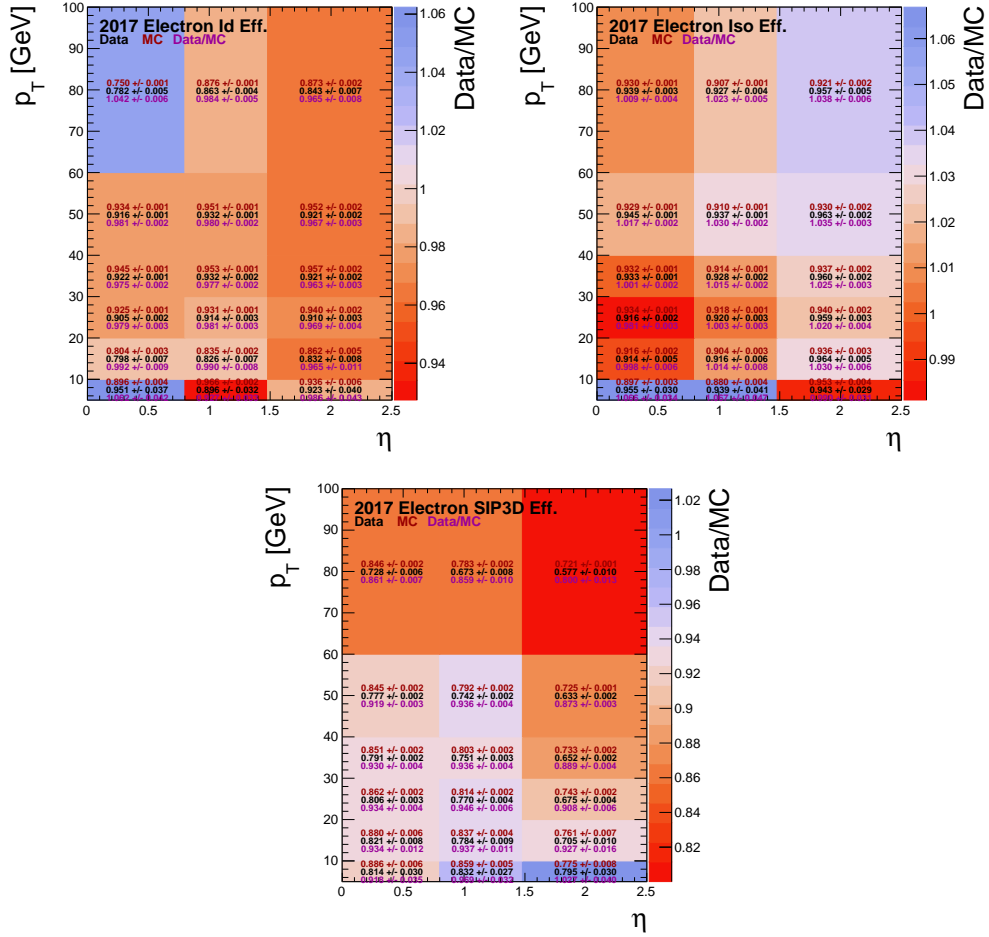


Figure 3.4: 2017 efficiencies

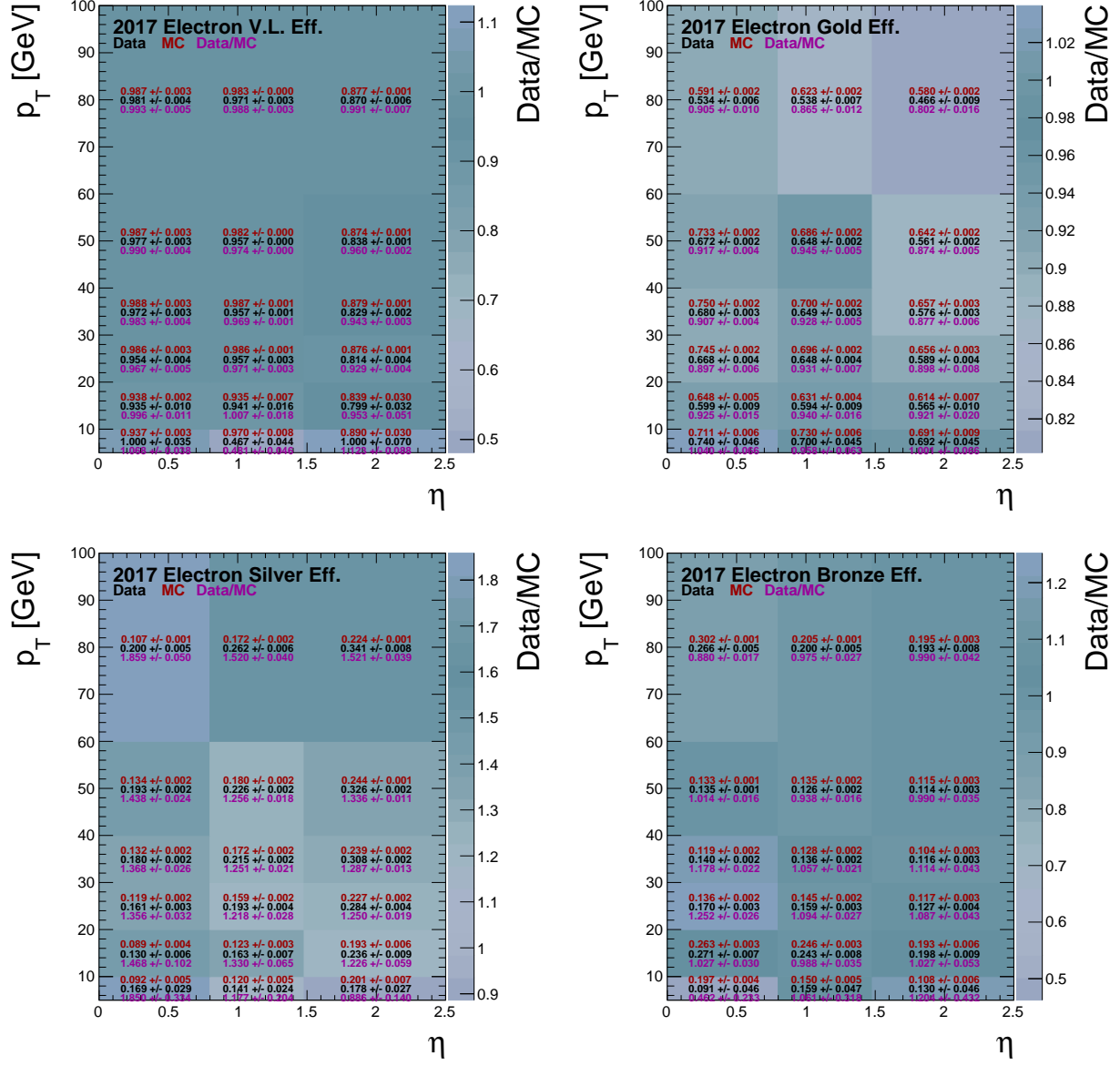


Figure 3.5: 2017 electron GSB efficiency and SF

### 3.4 Muon Tag-and-Probe

The muon Tag-and-Probe tools also uses a centrally curated CMSSW PhysicsTools in `CMSSW_10_6_X`. The software pipeline is identical to electrons in that it consists of an ntuplizing [?] and fitting [?] stage. The code bases for muons and electrons are separate but functionally identical. The samples chosen for  $Z$  measurements are shown in Table 3.4. The  $J/\psi$  ntuples are available from a central repository of standard Tag-and-Probe selection variables which use the pre-ultra legacy samples for each year [?]. The muon Tag-and-Probe efficiencies are measured above 20 GeV using the  $Z$  boson while below 20 GeV benefits from the  $J/\psi$  meson for Id measurements. The  $\eta$  bins are divided into a central and forward regions around the endcaps at  $|\eta| = 2.1$ . In total there are three sets of binnings: The low  $p_T$   $J/\psi$  binning  $J/\psi^L$  for muon Id below 20 GeV, the high  $p_T$   $Z$  binning  $Z^H$  above 20 GeV, and the low  $p_T$   $Z$  binning  $Z^L$  used to extrapolate isolation and impact parameter efficiencies down to 3 GeV. The explicit bin edges for each range are defined in Table 3.5.

Topological dependencies for isolation and impact parameters prevent measurement using the  $J/\psi$ . About 30% of prompt  $J/\psi$  are produced from higher mass states  $\chi_c$  and  $\Psi(2S)$  thus  $J/\psi$  will be produced from a cascade inside jets and likely be unisolated [?]. Similary another 10% of all  $J/\psi$  are produced within b-jets and leading to non-prompt unisolated events [?].

The exact criteria chosen for the tag and probe vary between physics processes but are identical across the two  $Z$  ranges. The selections follow the standards defined from the centrally produced muon Tag-and-Probe efficiencies.

Table 3.4

Type	Year	Sample Name
Data	2016	/SingleMuon/Run2016C-17Jul2018-v1/MINIAOD
Data	2017	/SingleMuon/Run2017C-31Mar2018-v1/MINIAOD
Data	2018	/SingleMuon/Run2018A-17Sep2018-v2/MINIAOD
MC	2016	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3_ext2-v2/MINIAODSIM
MC	2017	/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIIFall17MiniAODv2-PU2017RECOStep_12Apr2018_94X_mc2017_realistic_v14_ext1-v1/MINIAODSIM
MC	2018	/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM

Table 3.5: muon binning

Muon Binning		
Range	$p_T$ GeV	$ \eta $
$J/\psi^L$	[3.0, 4.0, 5.0, 6.0, 7.0, 9.0, 14.0, 20.0]	[0, 1.2, 2.4]
$Z^H$	[10, 20, 30, 40, 60, 100]	[0, 1.2, 2.4]
$Z^L$	[6,8,10,14,18,22,28,32,38,44,50]	[0, 1.2, 2.4]

Tag-and-Probe Muon Candidate Selection Criteria		
$J/\psi$		
Tag	Probe	Pair
isGlobalMuon numberOfMatchedStations > 1 $p_T > 5$ GeV Matches hltIterL3MuonCandidates	Matches hltTracksIter OR Matches hltMuTrackJpsiEffCtfTrackCands	$2.8\text{GeV} < m_{\mu\mu} < 3.4\text{GeV}$ $ z_{\mu_1} - z_{\mu_2}  < 1$ cm
$Z$		
passes tightID $\sum p_T^{ch}/p_T < 0.2$ $p_T > 15$ GeV	No requirement	$m_{\mu\mu} > 60$ GeV $ z_{\mu_1} - z_{\mu_2}  < 4$ cm

The muon data will also have an implicit selection due to triggering. To reflect this selection in MC, the tag is required to pass a chosen trigger in the efficiency denominator in addition to HLT object matching. The triggers available vary from year to year for  $Z$  using IsoTkMu22 in 2016 and isoMu24eta2p1 in 2017 and 2018. A single  $J/\psi$  triggers is available for all years which is Mu7p5Tk2.

The Gold, Silver, and Bronze efficiency definitions are split based on  $p_T$  and reflect the high and low binning separations shown in Table 3.5. The low  $p_T$  muons include the Id measured by  $J/\psi$  as well as the extrapolated efficiencies from SIP3D and isolation fits in  $Z_L$ . The high  $p_T$  muons are composed of all the factors directly measured in  $Z_H$ .

$$p_T \in [3, 20)$$

$$\begin{aligned}
\epsilon_{\text{Gold}} &= \epsilon_{\text{ID}}^{J/\psi} \times \epsilon_{\text{Isolated|ID}}^{Z_L} \times \epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}}^{Z_L} \\
\epsilon_{\text{Silver}} &= \epsilon_{\text{ID}}^{J/\psi} \times \epsilon_{\text{Isolated|ID}}^{Z_L} \times (1 - \epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}}^{Z_L}) \\
\epsilon_{\text{Bronze}} &= 1 - (\epsilon_{\text{ID}}^{J/\psi} \times \epsilon_{\text{Isolated|ID}}^{Z_L})
\end{aligned} \tag{3.5}$$

$$p_T \in [20, 100]$$

$$\begin{aligned}\epsilon_{\text{Gold}} &= \epsilon_{\text{ID}}^{Z_H} \times \epsilon_{\text{Isolated}|\text{ID}}^{Z_H} \times \epsilon_{\text{Prompt}|\text{(ID}\cap\text{Isolated)}}^{Z_H} \\ \epsilon_{\text{Silver}} &= \epsilon_{\text{ID}}^{Z_H} \times \epsilon_{\text{Isolated}|\text{ID}}^{Z_H} \times (1 - \epsilon_{\text{Prompt}|\text{(ID}\cap\text{Isolated)}}^{Z_H}) \\ \epsilon_{\text{Bronze}} &= 1 - (\epsilon_{\text{ID}}^{Z_H} \times \epsilon_{\text{Isolated}|\text{ID}}^{Z_H})\end{aligned}\tag{3.6}$$

The 2017 Id efficiency with statistical errors for both data and MC are shown in Figure 3.6. The other efficiencies for each year for all  $p_T$  ranges are included in the appendix. The overlapping bins between  $J/\psi$  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 relative efficiencies per component range from approximately 88% to 98% and are fairly uniform between the central tracker and endcaps. The efficiencies for the isolation ranges from (90 – 95)% where the encaps generally are about 5% more efficient. As for SIP3D, the efficiency ranges from about (80 – 93)% with another 5%  $|\eta|$  based efficiency gap, however, in the SIP3D case, the central tracks are more efficient as opposed to isolation. 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 3.7. The errors for each bin are the combined statistical and systematic errors from Table 3.7 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 right tail P-value from all fits is  $\approx 2\%$ , the median P-value from the Figure 3.7 is 84%. The fits in each year behave qualitatively the same as 2017. The product of the efficiency components into their corresponding Gold, Silver, and Bronze category is shown in Figure 3.8. Similar to electrons, the efficiency for Very Loose is also included separately but is factored into the denominator efficiencies components, so, the Gold, Silver, and Bronze efficiencies represent the overall electron efficiency for that particular lepton ranking. The range of efficiencies for each ranking are (70–80)%, (5–15)%, and (4 – 20)% for Gold, Silver, and Bronze respectively. The Data and MC agreement for all three ranks is better than electrons with the largest discrepancy in Gold being 2% and

the average deviation in Silver and Bronze begin approximately  $(5 - 10)\%$ .

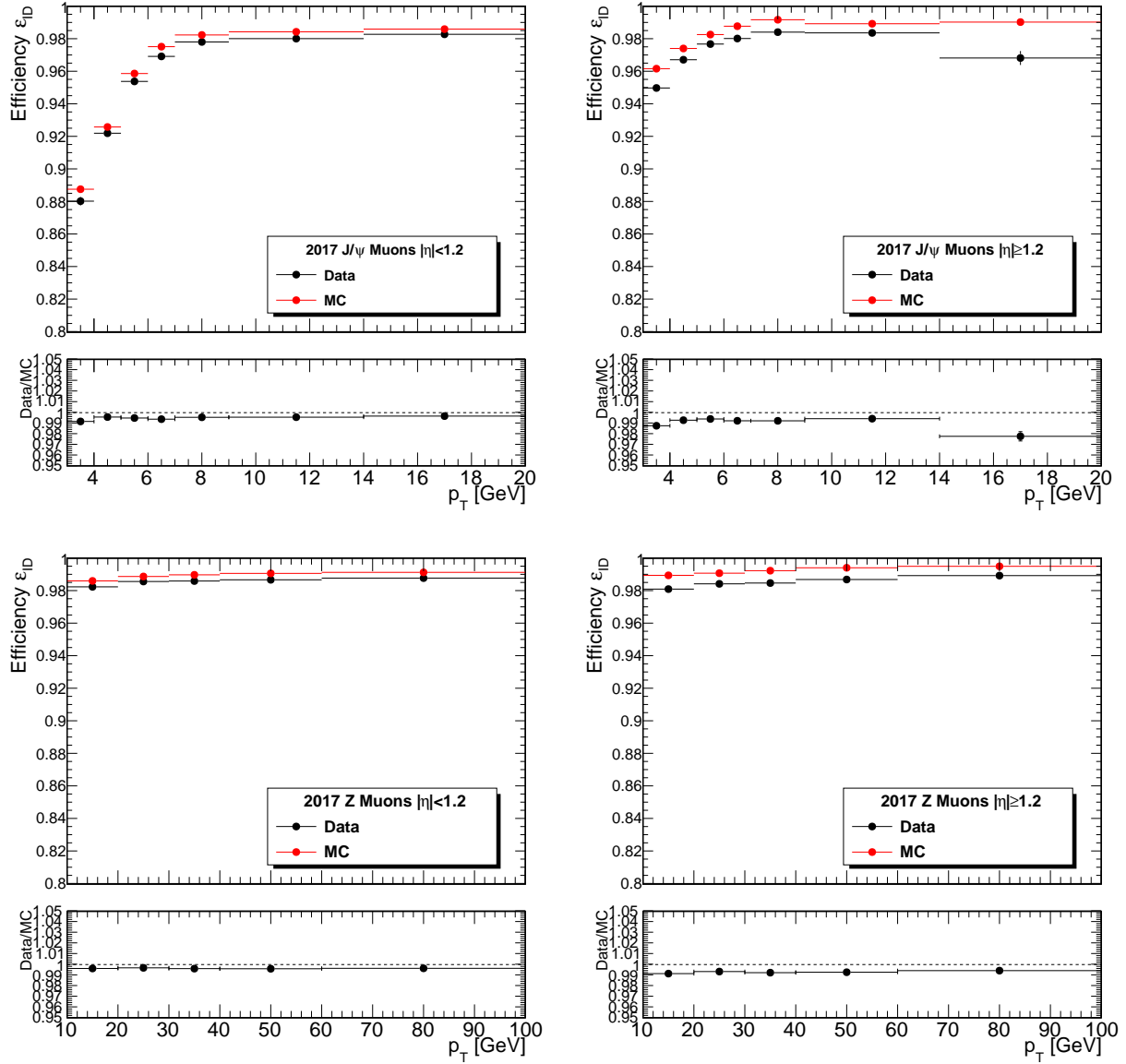


Figure 3.6: Tag-and-Probe efficiencies for the Medium Id in 2017. The left plots show the barrel while the right plots show the endcaps. The top fits use  $J/\psi$  resonance while the bottom use the  $Z$  resonance.



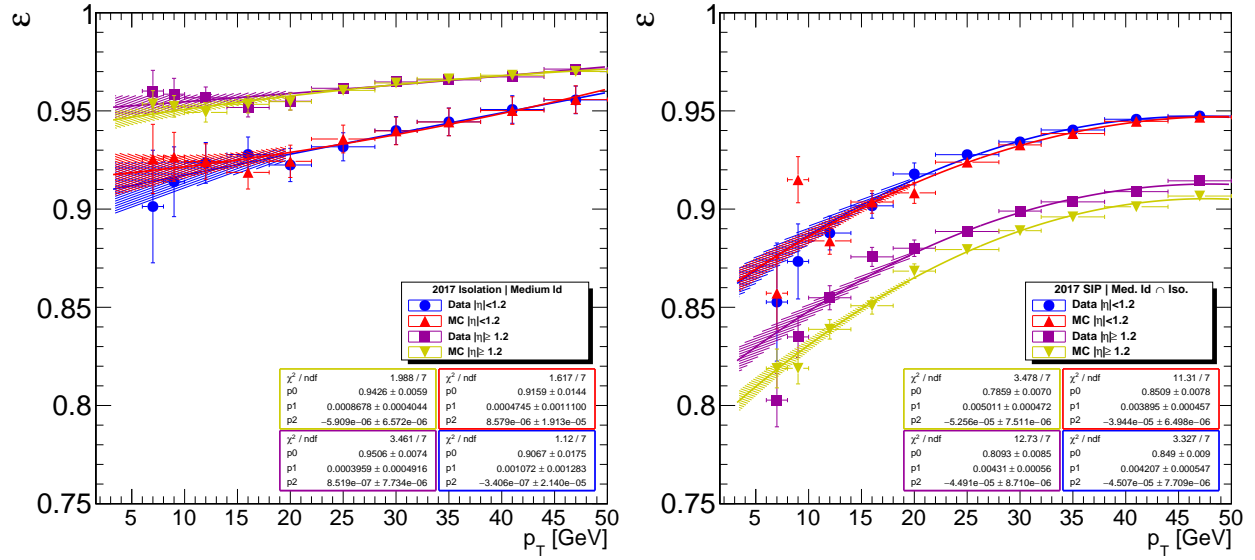


Figure 3.7: The fitted muon isolation and SIP3D efficiencies for 2017. Includes both data and MC which are separated between barrel and endcap.

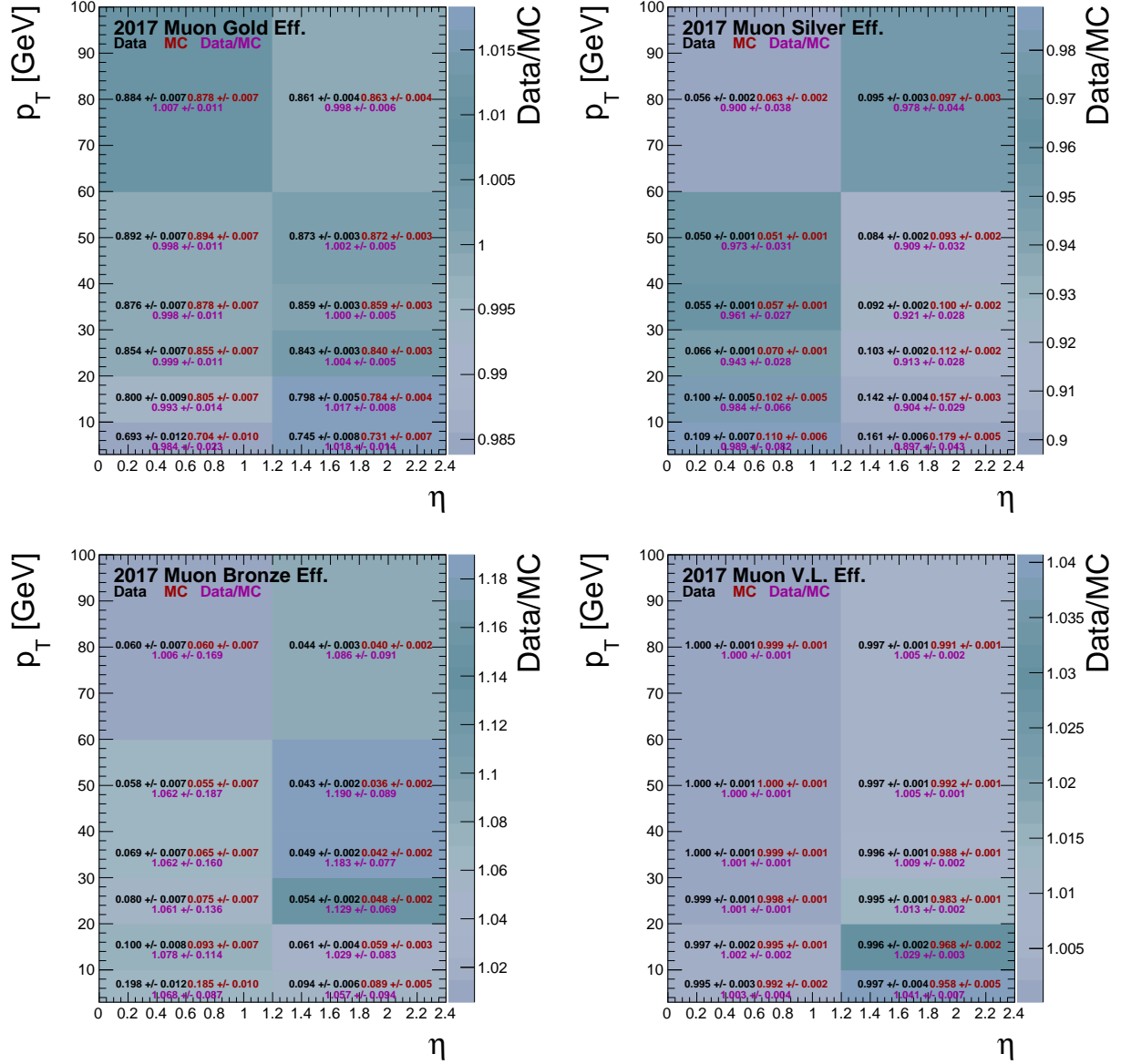


Figure 3.8: The combined efficiency components from equations 3.5 and 3.6 and Very Loose for 2017. The low- $p_T$  region ( $< 20$  GeV) includes the contributions from  $J/\psi$  as well as the isolation and SIP3D extrapolations. Propagated errors are treated as uncorrelated.

### 3.5 Lepton Systematics and Scale Factors

The systematic error for the electron and muon efficiencies are derived by varying the Tag-and-Probe signal and background models, slimming and widening the mass window, and increasing and decreasing the number of bins used in the fit. The systematic error is defined as the maximum spread in efficiencies between the modeling variations with an example spread shown in Figure 3.10. Rather than compute the systematic error for every bin, similarities between neighboring bins motivates using a simplified bin approach which was chosen qualitatively by the background shape. The shape of the  $p_T$  based mass distributions is illustrated in Figure 3.9. The same  $\eta$  bins are utilized according to lepton flavor, but the  $p_T$  bins are consolidated into a high and low bin pivoting on 20 GeV. A high and low systematic is derived for each selection criteria per flavor per year and is applied to the efficiencies that fall within the corresponding  $p_T$  and  $\eta$  range.

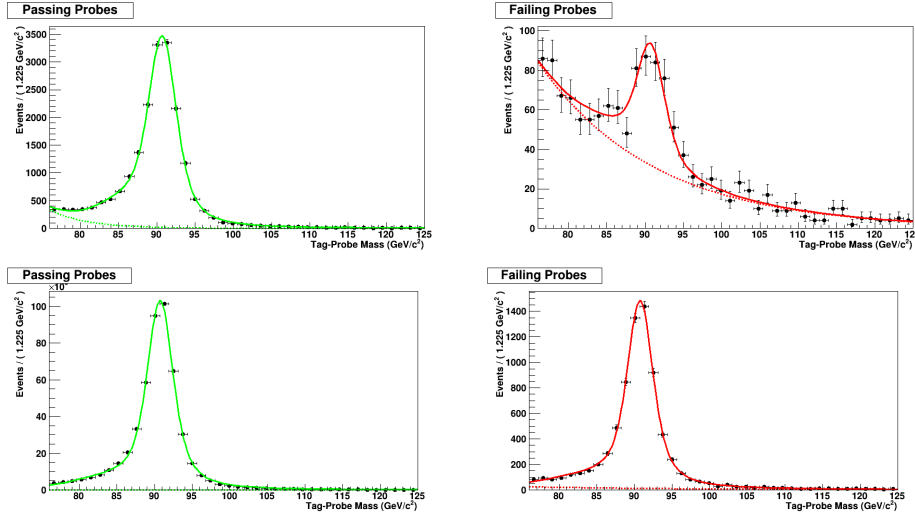


Figure 3.9: Tag-and-Probe di-muon mass distributions for both passing and failing probes. The top set of plots consist of probes below 20 GeV and the bottom set are about 20 GeV.

Scale factors are derived bin by bin for each criteria per flavor per year by finding the ratio of efficiencies in data to Monte Carlo. The scale factor variance is propagated by combining both the statistical error from the Tag-and-Probe in quadrature with the systematic error. The full 2017 set of systematics electrons and muons is shown in Table 3.6 and Table 3.7.

Additional scale factors are also needed adjusting the differences between samples which are either created with a full simulation or fast simulation. The Fast to Full factor is obtained by extracting the criteria efficiency ratio between full and fast sim  $t\bar{t}$  samples.

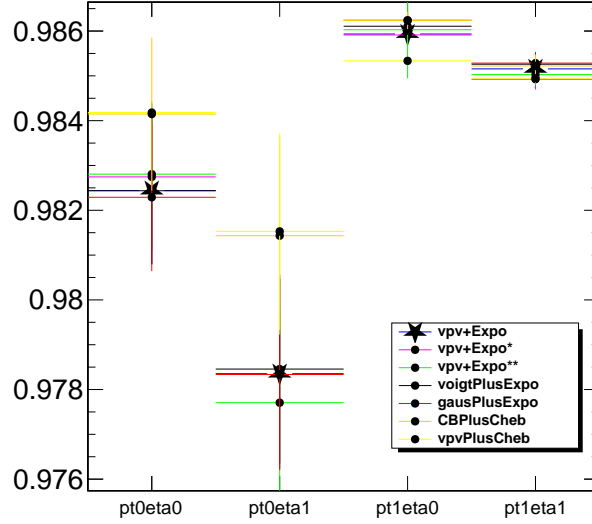


Figure 3.10: Example systematic spread from various fit models and binnings for muons. Includes the four combinations of regions either low or high  $p_t$  and central and forward  $\eta$ .

Table 3.6: The electron systematic error derived from the Tag-and-Probe for 2017 data and split into  $p_T$  and  $|\eta|$  regions.

ID	$0 \leq  \eta  < 0.8$	$0.8 \leq  \eta  < 1.479$	$ \eta  \geq 1.479$
$p_T < 20$ [GeV]	0.003	0.001	0.005
$p_T \geq 20$ [GeV]	0.001	0.001	0.002
Iso   ID			
$p_T < 20$ [GeV]	0.002	0.003	0.003
$p_T \geq 20$ [GeV]	0.001	0.001	0.002
SIP   Iso $\cap$ ID			
$p_T < 20$ [GeV]	0.006	0.004	0.007
$p_T \geq 20$ [GeV]	0.002	0.002	0.0006
VeryLoose			
$p_T < 20$ [GeV]	0.002	0.007	0.03
$p_T \geq 20$ [GeV]	0.003	0.0001	0.0007

Table 3.7: The muon systematic error derived from the Tag-and-Probe data and split into  $p_T$  and  $|\eta|$  regions.

ID	$ \eta  < 1.2$	$ \eta  \geq 1.2$
$p_T < 20$ [GeV](J)	0.001	0.001
$p_T \geq 20$ [GeV](Z)	0.001	0.0003
Iso   ID		
$p_T < 20$ [GeV]	0.007	0.004
$p_T \geq 20$ [GeV]	0.007	0.002
SIP   Iso $\cap$ ID		
$p_T < 20$ [GeV]	0.005	0.003
$p_T \geq 20$ [GeV]	0.001	0.002
Very Loose		
$p_T < 20$ [GeV]	0.001	0.0003
$p_T \geq 20$ [GeV]	0.001	0.001