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

Graham Wilson, Chairperson

Alice Bean, Co-Chair

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

List of Figures

List of Tables

Chapter 1

The Standard Model and Supersymmetry

Abstract

Here we outline the fundamental concepts of particle physics, we introduce the set of fundamental particles, fields, and interactions which are described by the standard model. Since the standard model does not solve all of the problems of particle physics I introduce supersymmetry, a well motivated extension to the standard model. The implications of SUSY model space on various phenomoglly of processes as well as experimental observations. The motivation for the SUSY extension is reserved for the following chapter.

1.1 Introduction

c1 introduction section maybe a little history here

1.2 The Standard Model

The Standard Model(SM) is a collection of adhoc theories used to predict and reproduce experimental data. The theory itself incorporates four major concepts: Quantum Field theory, the Dirac equation, the gauge principle, and the higgs mechanism. These four principles are constrained by physical data and describe the set of elementary particles, known as fermions and bosons. The SM generally refers to the SM Lagrangian, an equation with different

sectors that describe different subsets of particles, fields, and their interactions. The SM Lagrangian itself 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 desribe the Higgs potential μ, λ which are the higgs mass m_h and the vacuum expectation value (vev), and 9 mixing angles which describe the PMNS and CKM matrices i.e. the mixing of different fermionic fields. The 12 fermion paramters are subdivided by three neutrinos m_ν^i, three charged leptons $m_\ell^\pm$, and six quarks m_q;

Quantum field theory (QFT) provides a descriptions for both known and theoretical particles by combining quantum theory, the field concept, 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: $\vec{G}$ which transforms under $SU(3)$ and governs strong interactions, $\vec{W}$ and B which transform under $SU(2)_L \times U(1)$ and governs electromagnetic and weak interactions. The combination of the gauge fields and fermion fields along 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 spontaneous symmetry breaking of the Higgs field, due to the Yukawa coupling, creates a non zero vev which is responsible for generating the masses of the electroweak gauge bosons. 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 is divided into two subgroups: fermions and bosons. The fermions consist of both charged and neutral leptons as well as fractionally charged quarks. There are three flavors of charged leptons (ℓ), the electron (e), the muon (μ), and the tau $\tau)$. Each charged particle and pairing neutral neutrino ν_ℓ. The e and μ are also generally considered as "light" leptons for there small mass relative to the τ. The term lepton, depending on context, often refers to only the charged particles. As for the quarks, there are also three generations of pairs of quarks.. The lighest set of quarks are the up (u) and down (d) quarks, followed by the charm (c) and strange (s), and lastly 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 (γ), gluon (g), the $W^{\pm}$, and the Z^0 - along with the singular scalar boson the Higgs (h). The elementary pariticles are summarized in Figure 1.1.

Standard Model of Elementary Particles
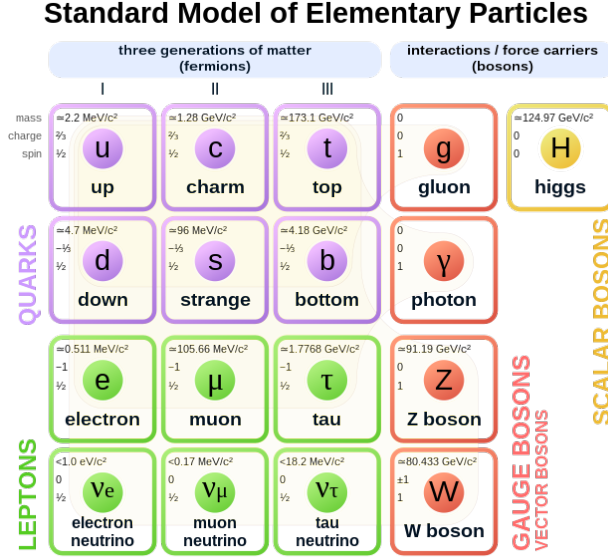
Figure 1.1: particles figure cite wiki

The SM is a asymmetric chiral theory, combining three groudps $SU(3)_L \times SU(2)_L \times U(1)$. The L, or left handed, subscript indicates that mirrored fields (with different chiralities) transform differently under the Lorentz group and the EW gauge group (cite slides). The consequence of chirality is that the possible combinations between interaction vertices is limited(cite thompson?). This peculiar property shows up with the W boson, which only couples to left handed particles or right handed antiparticles. Extensions of the standard model also often include chiral properties.

1.3   Supersymmetry

Supersymmetry (SUSY) is an extension of the standard model. It adds a 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) which generates a super partner for each particle differiring in spin by a half integer. The resulting set of mirrored elementary particles are referred to as sparticles. Each bosonic sparticle carries the same name as its fermion partner but with an "s" prefix e.g. sfermion, squark, selectron. As for the bosons, with the gauge fields B and $\vec{W}$, these are accompanied by three super symmetric fields - the Higgsino $\tilde{H}$, Bino $\tilde{B}$, and Wino $\tilde{W}$. The B and $\vec{W}$ SM fields mix and are represented by particle matrix, the diagnolization of this matrix and yield mass eigenstates for $\gamma, Z, W^{\pm}$ while the Higgsino,Bino, and Wino mix to produce four neutral and two charged eignestates, the neutralinos ($\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$) and charginos ($\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^{\pm}$) (cite erich's 43 susy matrix eigenstates). SUSY also requires an additional Higgs doublet to give mass to up-type and down-type fermions, (cite this run 2 paper reference maybe) leading to five higgs boson states inlcuding two charged higgs and thre neutral Higgs. The lightest neutral higgs in the two higgs doublet represents the SM Higgs boson. The full set of SM and SUSY partilcles are featured in Figure **??**. An additional higgs doublet also introduces a second vev. The ratio between the two vev's is commonly denoted as $v_1/v_2 = \tan\beta$ and is an important parameter in experimental searches. We also introduce a new parameter similar to lepton number or baryon number conservation which is called R-parity. This parameter tallies the total number of SM particles (+1) and sparticles (-1) and expects the net total of particles to be conserved between initial and final states. 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 simplified SUSY model 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 inaccesible due to the presence of over 100 parameters, meaning small differences in parame-
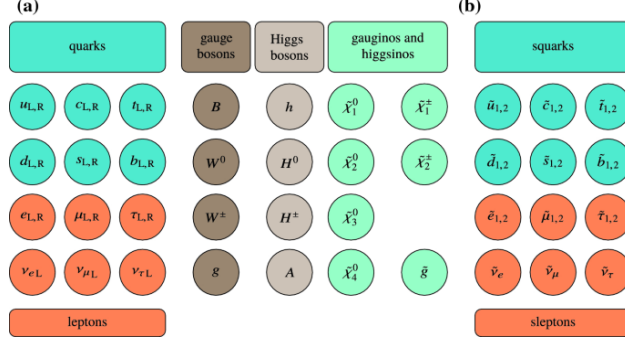
Figure 1.2: stolen from this springer thesis book, probably make my own figure later `https://link.springer.com/chapter/10.1007/978-3-030-25988-4_4`

ter space can completely morph the model structure and experimental signatures. To tackle the vast parameter space further simplification is needed, the most popular simplified model is the phenomological MSSM (pMSSM). The pMSSM contains 19 parameters which include the masses of each generation of squark and slepton, parameters to control the mixing of $\tilde{H}, \tilde{W}, \tilde{B}$, and dials for the higgs doublet. The pMSSM is still borderline too complicated to attack directly, so the MSSM is boiled down into four parameters M_1, M_2, μ, and $\tan\beta$ which are the gaugino mass parameters, the Higgsino mass parameter, and the vev ratio respectively (cite Fuks paper). A model point from this four parameter space is referred to as Realistic simplified gaugino-higgsino model, these target specific and well motivated regions of MSSM phase space and experimental topologies.

To effectively grasp the structure of SUSY and various models either in the pMSSM or simplified models there are a couple key elements to condsider. The first being the mass scale of the relative susy sectors meaning how massive are the gauginos versus sleptons versus squarks. In an electroweak SUSY search with a simplified model the model is further simplified by assuming squarks are at the several TeV scale while the targeted electroweakinos are at detectable mass scale at the TeV and sub-TeV scale. By decoupling sectors outside the sector-of-interest we remove the interaction between these groups, so, if sleptons are decoupled from the gauginos we avoid complicated dependencies between the two. The key element is the composition of the LSP, typically $\tilde{\chi}_1^0$, each unique model point yields a specific

mixing of $\tilde{H}, \tilde{W}, \tilde{B}$ with an LSP that reflects that mixing. The model point is denoted by the field that dominates the overall mix, so a Higgsino model has an LSP composed of mostly $\tilde{H}$ (cite mixing altas paper?). The take away from simplified model types is that H,W,B can characterize the nature of the model by governing the overall cross sections for sparticles, the topological infrastructure, and how these sparcticles specicially interact and decay amongst themselves and SM particles. Two pMSSM examples comparing the mass structure between two arbitrary mass points for a Wino model versus Higgsino model is shown in Figure 1.3. For both models the Higgs and slepton sectors are decoupled at a multi-TeV scale while the squark and gaugino sectors are at an accesible TeV and sub-TeV scale. Note that with small changes pMSSM model space results differing LSP content and large variations in the relative mass structure and orderings. The difference in cross sections between the same two model points for gaugino pair production combinations is showin in Figure 1.4. This illustrates that the same tweak in parameter space can induce order of magnitude changes sparticle production.
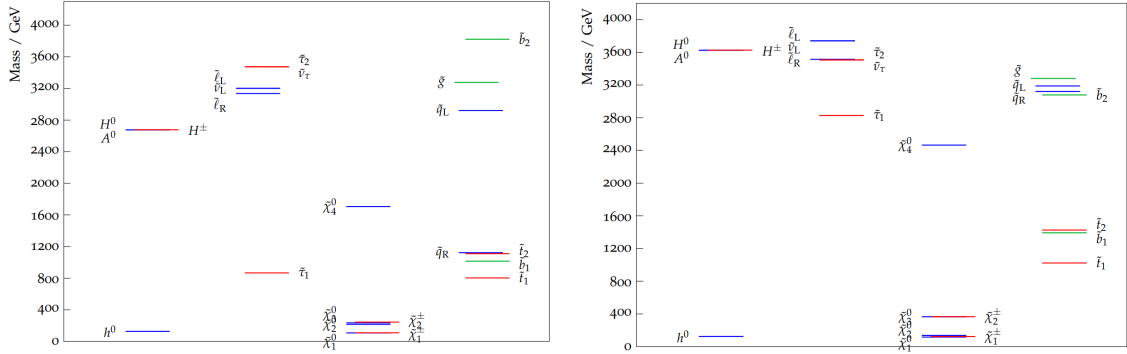
Figure 1.3: mass structure winno vs higgsino modelpoints

Example of differing model points

In addition to the mass structure and cross sections the decay nature of H/W/B models also varies which has a significant impact on the experimental channels and signatures of interest. In an experimental search we would expect the lighter sparticles to decay to both SM particles along with the LSP, if the LSP is close in mass to the heavier sparticle of interest say O(100) GeV or less the model would be considered as a compressed scenario. Between
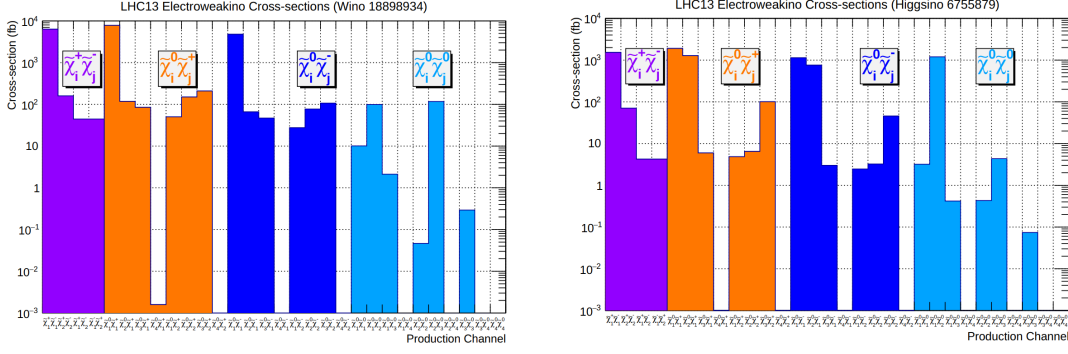
Figure 1.4: xsec strucutre wino vs higgsino modelpoints

the 3 types of models the most likely candidates for compression Higgsino-like and Bino-like models. The Wino type models by far have the largest cross sections but are the least likely to have compressed states. Particular interesting topologies involve decay signatures of processes like N2->ZstarN1, N2->C1N1, C1->W*N1, stop->tN1. For instance, the decay pattern of N2 is both dependent on the H/W/B nature of the model but also the level of compression Figure 1.5 shows the various average decay modes for H W or B from a selection of pMSSM from (cite atlas pmssm paper).

Figure 1.5: N2 BFs

7
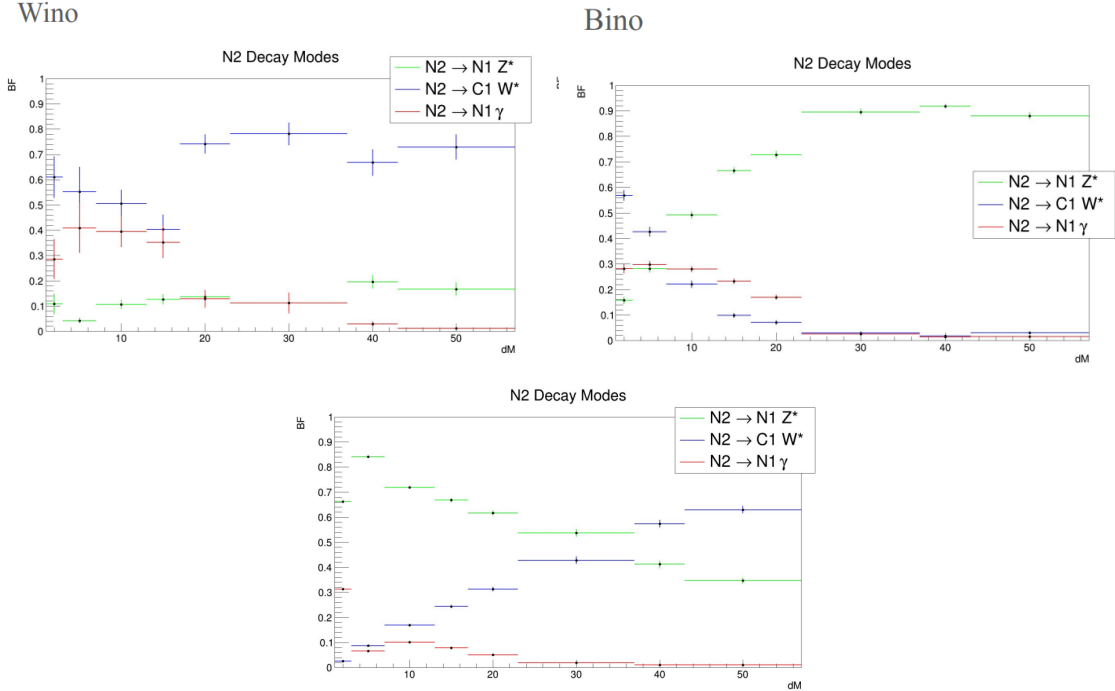
Notice between each model the Z* and W* modes can be highly suppressed or enhanced depending on model specifics. In some cases, not shown, specific modes like C1W* can be forbidden or excluded in simplified models to streamline MC production and enhance the statistical power of the targeted final states i.e. Wtolnu or Ztoll. In addition to decay specific complicattions, the phase space of the final state particles is subject to model specifics. For instance, in N2->Z*N1, dilepton mass distribution $m_{\ell\ell}$ from the Z* is dependent on the sign of the gaugino eigenstates, essentially again, the interpretation of the model being either Wino/Bino or Higgsino. The $m_\ell\ell$ distribution is shown in Figure 1.6 and illustrates the experimentally observable differences in a single model point compared under two different interpretations. Overall, with the complications of model dependent decays and observables for even a single spariticle in addition to inherently rare production then add in varying structures and coupled versus decoupled sectors like sleptons and squarks this yields an extraordinary challenge to design a search to tackle a large and unwieldy model space with many well motivated corners and processes to choose from.
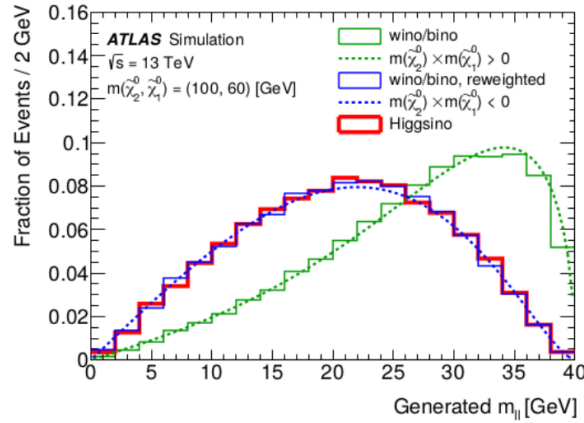
Figure 1.6: mll reweight with w/b or H interpretations from altas paper in grahams talk

Chapter 2

Motivating the Search for SUSY

Abstract

Introduce some of the issues of the SM and opening up with the basic motivations for susy, like solving the hierarchty 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 spefically this body of work.

2.1 Introduction

The standard model is a consistent and remarkable theory that has held up to test to many orders of magnitude in energy. However it is not a perfect theory in that there are a few problems both experimental and theoretical that the SM can not yet explain. The current SM doesn't have an explanation for neutrino mass which is implied by observations of flavor oscillations (cite solar and atmo neutrino data). There also is not a theoretically proven understanding of the all particle masses and mixing including generations of mass ordering and neutrino mass orderings. Observations of from the relic microwave background suggest the existence of cold dark matter however there is no discovered particle which could yet represent dark matter. SUSY specifically can handle the DM problem because it offers a DM candidate such as the LSP χ_1^0. In fact the expected relic DM density of the universe can

constrain the model space in which to search for SUSY particles. Three big motivations which will be discussed in the following sections is the major theoretical improvement that SUSY offers stabilization divergent terms in the higgs mass, an explanation to the signficantly SM deviating measurment of the muon $(g-2)$ factor, as well as an explanation to the deviation observed in the W mass at CDF. Both of the latter experimental measurements not only motivate searching for SUSY but point to compressed SUSY scenarios.

2.1.1 Stabilizing the Higgs mass

A commonly pursued attribute of theoretical models is naturalness, typically we expect a model to function "naturally" if the ratio of free parameters in a model are of $O(1)$, large swings in parmeters is considered "fine tuning". Fine tuning is only an aesthetic problem, but, could indicate new physics which would balance out fine tuned paramters. One such fine tuning problem arises in the higgs mass due to the Higgs self interaction terms. The SM Higgs Lagrangian that represents the higgs self interaction is shown in equation X

$$\mathcal{L} = \frac{gm_h}{4M_W}H^3 - \frac{g^2m_h^2}{32M_W^2}H^4 \tag{2.1}$$

where H represents the scalar higgs field, m_h the higgs mass, and m_W the W mass. A correction to the higgs mass can be calculated using standard perturbation theory by evaluating the second term of the lagrangian from equation X.

$$\begin{aligned}
\Delta m_h^2 &= \langle H| \frac{g^2m_h^2}{32M_W^2}H^4|H\rangle = 12\frac{g^2m_h^2}{32M_W^2}\int \frac{d^4k}{(2\pi)^2}\frac{i}{k^2 - m_h^2} \\
&= 12\frac{g^2m_h^2}{32M_W^2}\frac{1}{16\pi^2}\left(\Lambda^2 - m_h^2\log\frac{\Lambda^2}{m_h^2} + O(\frac{1}{\Lambda^2})\right)
\end{aligned} \tag{2.2}$$

Here the intergal term is the propagator for the exchange of a virtual higgs and integrated over four momenta, and we introduce the scale cutoff parameter Λ which is a large number which can be interpreted as the scale at which the SM breaks down, possibly near the planck

scale $O(10^{19})$GeV. Notice the leading term Λ^2 indicates a quadratic divergence which means there needs to be extremely large finely tuned cancellations to maintain $\Delta m_h \propto O(m_h)$. This is known as the hierarchy problem, SUSY offers a solution by protecting the Higgs from divergence by introducing a symmetry. The additional fermionic/bosonic from SUSY introduce paramters to equation Y which cancel the λ^2 term and provide a natural structure for the higgs boson.

2.1.2 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 anomolaus 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)`) $\boldsymbol{\mu} = g\dfrac{e}{2m}\mathbf{S}$. where g is the particles g-factor, μ is the magetic 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 signficance $a_\mu = g-2/2$. 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 $m_h >> m_{e,\mu}$ 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 contribtuions 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_μ, a_τ. The current best candidate to both test the SM and search for new physics is by measuring $(g-2)_\mu$ or a_μ 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 μ, 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 currrent status of (g-2)mu The currently accepted best SM prediction of a_μ from (CITE) includes QED, Electroweak(EW) and Hadronic contributions and is reported as $a_\mu^{SM} = a_\mu^{QED} + a_\mu^{EW} + a_\mu^{\text{Hadronic}} = 116591810(43) \times 10^{-11}$. For each of the a_μ components, the QED compenent 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_μ^{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_μ. The BNL measurement of a_μ 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 3.7σ. The most recent a_μ measurement from FNAL confirms the BNL measurement within 1σ and the combined experimental average increases the SM deviation with a significance of 4.2σ. The standard model is a consistent and remarkable theory that holds over many orders of magnitude in energy

What could this deviation mean? The 4.2σ 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 1.2σ (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_μ SM deviation, those being Super Symmetry, Dark Matter (DM mediator Dark photon), Lepto quarks, 2 Higgs doublet models.

2.1.3   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 α 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 7σ signficance 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 7σ 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 scenraios. 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 gaug-ino models which satisfy the newest $g-2$ and W mass constraints in shown in FIGURE Z (cite Wmass and g-2 sven paper)

2.2 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/μ) 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 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/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 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 combinitorial 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 ε and can be expressed as a likelihood from the binomial probabilty 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:

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

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

$N^{\text{Pass/Fail}}$ 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 ε_{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 Γ and σ. The signal model is combined with an Exponential background model.
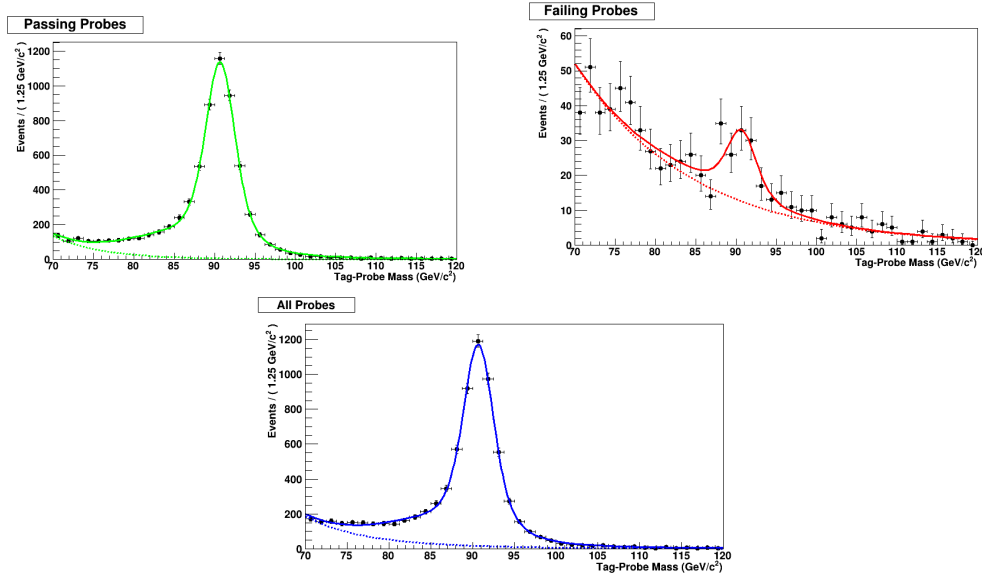
Figure 3.1: Example Tag-and-Probe Z di-muon fits for passing,failing, and all probes with the Medium Id, $|\eta| < 1.2$, and $p_T < 20$ GeV

3.2 Lepton Object Definitions

Leptons are selected according to the minimium 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.

$$
R_{\mathrm{miniIso}} = \begin{cases} 0.2 & p_\mathrm{T}^\ell < 50 \mathrm{GeV} \\ \frac{10}{p_\mathrm{T}^\ell} & 50 \mathrm{GeV} \leq p_\mathrm{T}^\ell \leq 200 \mathrm{GeV} \\ 0.05 & p_\mathrm{T}^\ell > 200 \mathrm{GeV} \end{cases} \tag{3.3}
$$

Mini isolation also includes effective area pile-up corrections provided in a look up table of bins of p_T and η 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 efficiences of the three previously mentioned quantities. The efficiencies take the form of conditional probabilities to be measured independently in sequence relative to each other:

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

$$\epsilon_{\text{Silver}} = \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated}|\text{ID}} \times \left(1 - \epsilon_{\text{Prompt}|(\text{ID}\cap\text{Isolated})}\right) \tag{3.4}$$

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

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 effiencies 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_{T} threshold and the MVA VLooseFO working point. The MVA VLooseFO ID also varies between years.

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
PFIso$_{\text{abs}}$	$< 20 + (300/p_{\text{T}})$ GeV	$< 20 + (300/p_{\text{T}})$ GeV
MVA VLooseFO ID	$\checkmark$	$-$

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 accomodates 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.
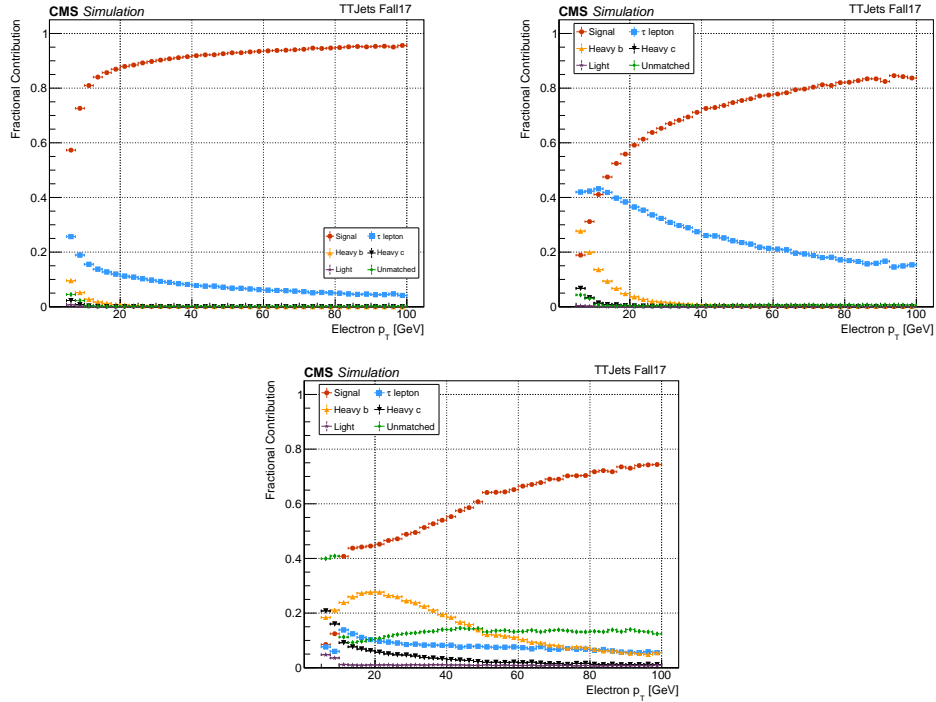
Figure 3.2: Gold (Top-Left), Silver (Top-Right) and Bronze (Bottom) MC truth matching in TTJets sample 2017. Signal is defined here as prompt electrons from a W decay.
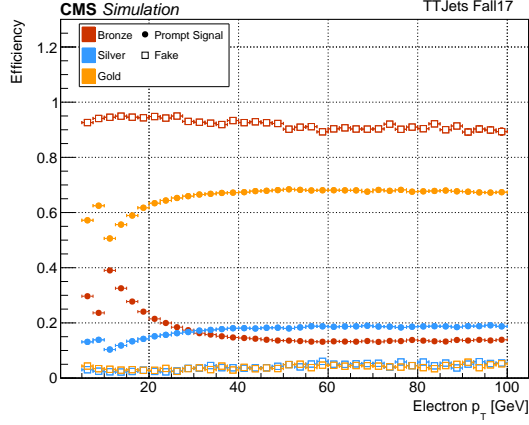
Figure 3.3: Gold, Silver, and Bronze efficiency on truth matched prompt electrons as signal and secondary electons as Fakes.

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 η 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 Ntupilizing 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 $1.4442 \leq |\eta_{SC}| \leq 1.566$ $p_\text{T} \geq 30.0$ 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	50GeV $< m_{ee} <$ 130GeV

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 $J/\psi \to ee$ to measure the lower p_{T} ranges, however, data triggers with electrons for J/ψ are not available.

Figure 3.4: 2017 efficiencies

23

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

0.987 +/- 0.003   0.983 +/- 0.000   0.877 +/- 0.001
0.981 +/- 0.004   0.971 +/- 0.003   0.870 +/- 0.006
0.993 +/- 0.005   0.988 +/- 0.003   0.991 +/- 0.007

0.987 +/- 0.003   0.982 +/- 0.000   0.874 +/- 0.001
0.977 +/- 0.003   0.957 +/- 0.000   0.838 +/- 0.001
0.990 +/- 0.004   0.974 +/- 0.000   0.960 +/- 0.002

0.988 +/- 0.003   0.987 +/- 0.001   0.879 +/- 0.001
0.972 +/- 0.003   0.957 +/- 0.001   0.829 +/- 0.002
0.983 +/- 0.004   0.969 +/- 0.001   0.943 +/- 0.003

0.986 +/- 0.003   0.986 +/- 0.001   0.876 +/- 0.001
0.954 +/- 0.004   0.957 +/- 0.001   0.814 +/- 0.004
0.967 +/- 0.005   0.971 +/- 0.003   0.929 +/- 0.004

0.938 +/- 0.002   0.935 +/- 0.007   0.839 +/- 0.030
0.935 +/- 0.010   0.941 +/- 0.016   0.799 +/- 0.032
0.996 +/- 0.011   1.007 +/- 0.018   0.953 +/- 0.051

0.937 +/- 0.003   0.970 +/- 0.008   0.890 +/- 0.030
1.000 +/- 0.035   0.467 +/- 0.044   1.000 +/- 0.070
1.068 +/- 0.018   0.481 +/- 0.046   1.128 +/- 0.088

2017 Electron Gold Eff.
Data MC Data/MC

0.591 +/- 0.002   0.623 +/- 0.002   0.580 +/- 0.002
0.534 +/- 0.006   0.538 +/- 0.007   0.466 +/- 0.009
0.905 +/- 0.010   0.865 +/- 0.012   0.802 +/- 0.016

0.733 +/- 0.002   0.686 +/- 0.002   0.642 +/- 0.002
0.672 +/- 0.002   0.648 +/- 0.002   0.561 +/- 0.002
0.917 +/- 0.004   0.945 +/- 0.005   0.874 +/- 0.005

0.750 +/- 0.002   0.700 +/- 0.002   0.657 +/- 0.003
0.680 +/- 0.003   0.649 +/- 0.003   0.576 +/- 0.003
0.907 +/- 0.004   0.928 +/- 0.005   0.877 +/- 0.006

0.745 +/- 0.002   0.696 +/- 0.002   0.656 +/- 0.003
0.668 +/- 0.004   0.648 +/- 0.004   0.589 +/- 0.004
0.897 +/- 0.006   0.931 +/- 0.007   0.898 +/- 0.008

0.648 +/- 0.005   0.631 +/- 0.004   0.614 +/- 0.007
0.599 +/- 0.009   0.594 +/- 0.009   0.565 +/- 0.010
0.925 +/- 0.015   0.940 +/- 0.016   0.921 +/- 0.020

0.711 +/- 0.006   0.730 +/- 0.006   0.691 +/- 0.009
0.740 +/- 0.046   0.700 +/- 0.045   0.692 +/- 0.045
1.040 +/- 0.066   0.958 +/- 0.063   1.001 +/- 0.086

2017 Electron Silver Eff.
Data MC Data/MC

0.107 +/- 0.001   0.172 +/- 0.002   0.224 +/- 0.001
0.200 +/- 0.005   0.262 +/- 0.006   0.341 +/- 0.008
1.859 +/- 0.050   1.520 +/- 0.040   1.521 +/- 0.039

0.134 +/- 0.002   0.180 +/- 0.002   0.244 +/- 0.001
0.193 +/- 0.002   0.226 +/- 0.002   0.326 +/- 0.002
1.438 +/- 0.024   1.256 +/- 0.018   1.336 +/- 0.011

0.132 +/- 0.002   0.172 +/- 0.002   0.239 +/- 0.002
0.180 +/- 0.002   0.215 +/- 0.002   0.308 +/- 0.002
1.368 +/- 0.026   1.251 +/- 0.021   1.287 +/- 0.013

0.119 +/- 0.002   0.159 +/- 0.002   0.227 +/- 0.002
0.161 +/- 0.003   0.193 +/- 0.004   0.284 +/- 0.004
1.356 +/- 0.032   1.218 +/- 0.024   1.250 +/- 0.019

0.089 +/- 0.004   0.123 +/- 0.003   0.193 +/- 0.006
0.130 +/- 0.006   0.163 +/- 0.009   0.236 +/- 0.009
1.468 +/- 0.102   1.330 +/- 0.065   1.226 +/- 0.059

0.092 +/- 0.005   0.120 +/- 0.005   0.201 +/- 0.007
0.169 +/- 0.029   0.141 +/- 0.024   0.178 +/- 0.027
1.850 +/- 0.314   1.177 +/- 0.204   0.886 +/- 0.140

2017 Electron Bronze Eff.
Data MC Data/MC

0.302 +/- 0.001   0.205 +/- 0.001   0.195 +/- 0.003
0.266 +/- 0.005   0.200 +/- 0.005   0.193 +/- 0.008
0.880 +/- 0.017   0.975 +/- 0.027   0.990 +/- 0.042

0.133 +/- 0.001   0.135 +/- 0.002   0.115 +/- 0.003
0.135 +/- 0.001   0.126 +/- 0.002   0.114 +/- 0.003
1.014 +/- 0.016   0.938 +/- 0.016   0.990 +/- 0.035

0.119 +/- 0.002   0.128 +/- 0.002   0.104 +/- 0.003
0.140 +/- 0.002   0.136 +/- 0.002   0.116 +/- 0.003
1.178 +/- 0.022   1.057 +/- 0.021   1.114 +/- 0.043

0.136 +/- 0.002   0.145 +/- 0.002   0.117 +/- 0.003
0.170 +/- 0.003   0.159 +/- 0.003   0.127 +/- 0.004
1.252 +/- 0.026   1.094 +/- 0.027   1.087 +/- 0.043

0.263 +/- 0.003   0.246 +/- 0.003   0.193 +/- 0.006
0.271 +/- 0.007   0.243 +/- 0.008   0.198 +/- 0.009
1.027 +/- 0.030   0.988 +/- 0.035   1.027 +/- 0.053

0.197 +/- 0.004   0.150 +/- 0.005   0.108 +/- 0.006
0.091 +/- 0.046   0.159 +/- 0.047   0.130 +/- 0.046
0.462 +/- 0.213   1.061 +/- 0.318   1.208 +/- 0.432
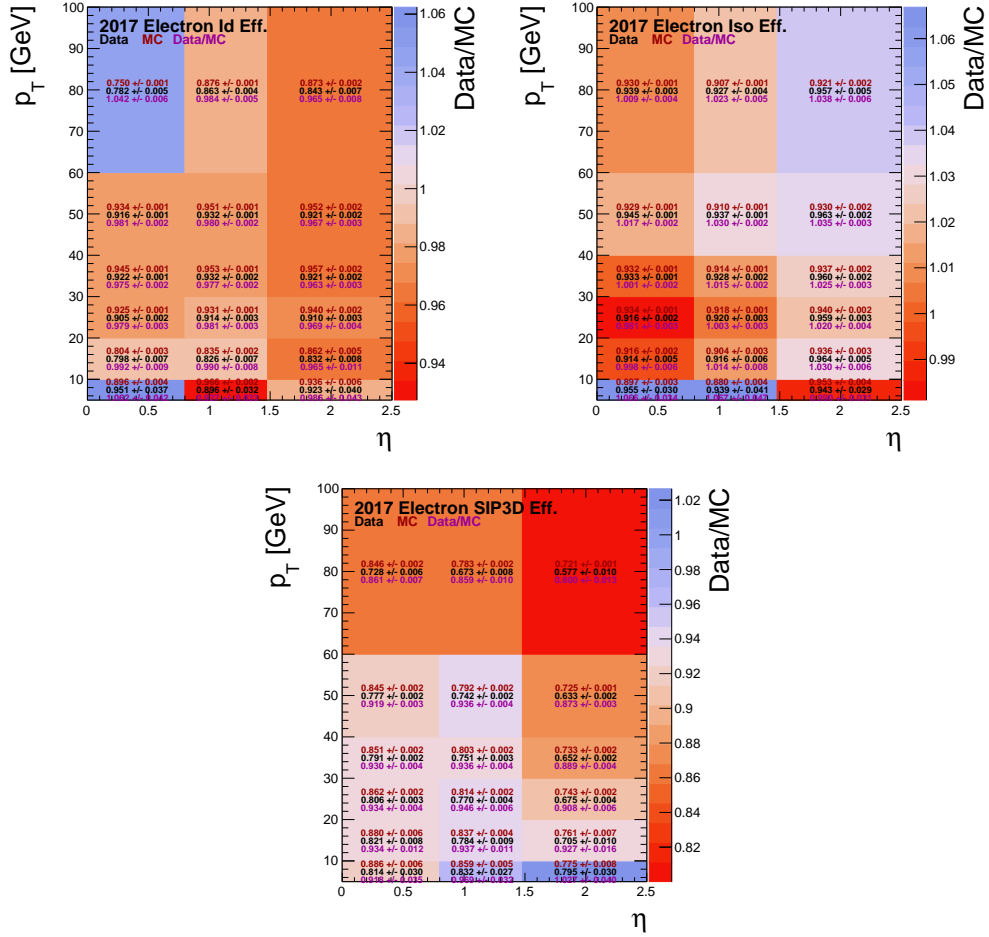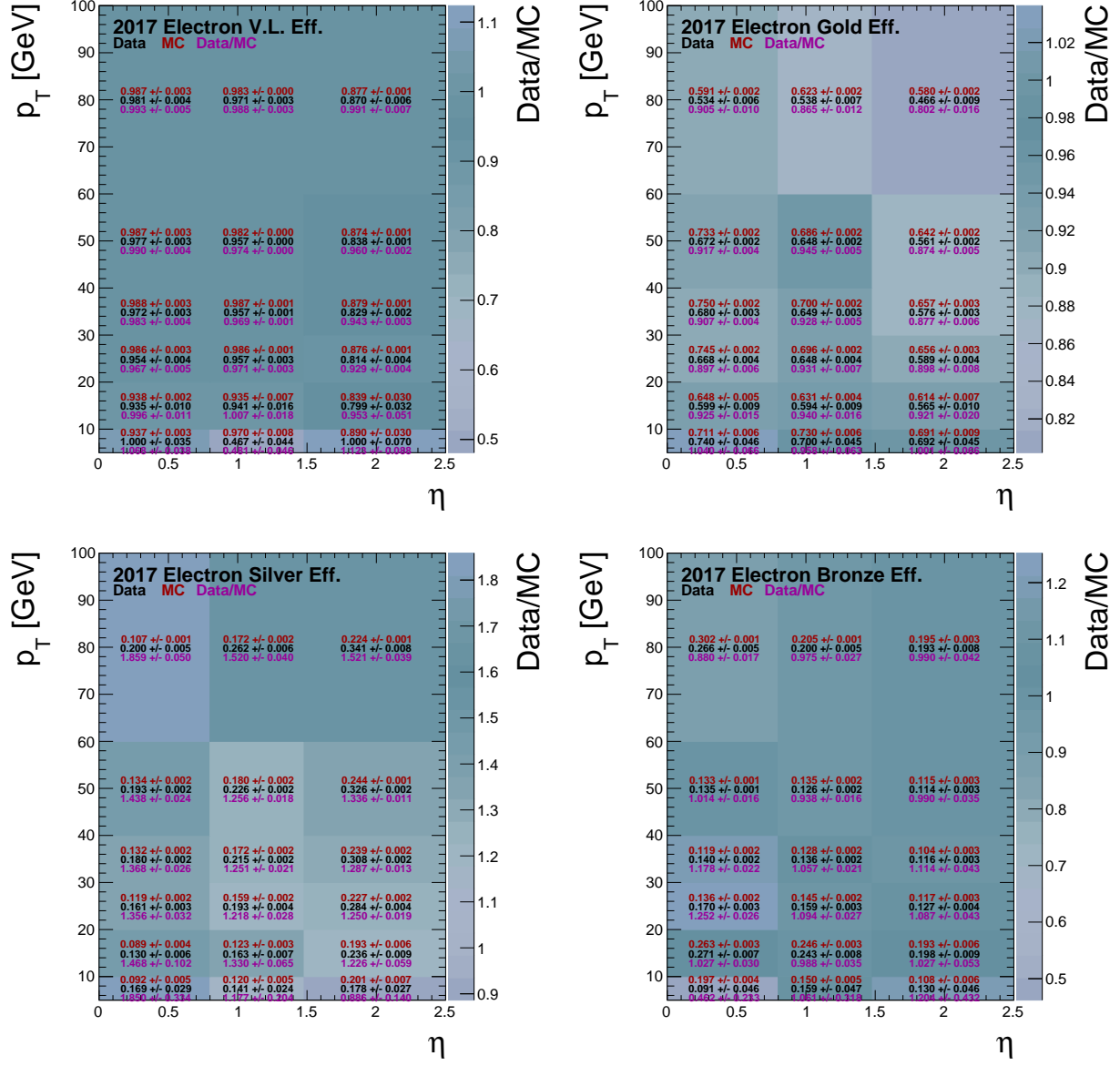
Figure 3.5: 2017 electron GSB efficiency and SF

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 electons 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/ψ 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/ψ meson for Id measurements. The η 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/ψ binning J/ψ^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 dependecies for isolation and impact parameters prevent measurement using the J/ψ. About 30% of prompt J/ψ are produced from higher mass states χ_c and $\Psi(2S)$ thus J/ψ will be produced from a cascade inside jets and likely be unisolated [?]. Similary another 10% of all J/ψ 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-PU2017RECOSIMstep_12Apr2018_94X_mc2017_realistic_v14_ext1-v1/MINIAODSIM
MC	2018	/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM

Muon Binning		
Range	p_T GeV	$|\eta|$
J/ψ^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/ψ		
Tag	Probe	Pair
isGlobalMuon numberOfMatchedStations> 1 $p_{\rm T} > 5$ GeV Matches hltIterL3MuonCandidates	Matches hltTracksIter OR Maches hltMuTrackJpsiEffCtfTrackCands	2.8GeV $< m_{\mu\mu} <$ 3.4GeV $|z_{\mu_1} - z_{\mu_2}| < 1$ cm
Z		
passes tightID $\sum p_{\rm T}^{ch}/p_{\rm T} < 0.2$ $p_{\rm T} > 15$ 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/ψ triggers is available for all years which is `Mu7p5Tk2`.

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

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

$$\epsilon_{\rm Gold} = \epsilon_{\rm ID}^{J/\psi} \times \epsilon_{\rm Isolated|ID}^{Z_L} \times \epsilon_{\rm Prompt|(ID \cap Isolated)}^{Z_L}$$

$$\epsilon_{\rm Silver} = \epsilon_{\rm ID}^{J/\psi} \times \epsilon_{\rm Isolated|ID}^{Z_L} \times \left(1 - \epsilon_{\rm Prompt|(ID \cap Isolated)}^{Z_L}\right)$$

$$\epsilon_{\rm Bronze} = 1 - \left(\epsilon_{\rm ID}^{J/\psi} \times \epsilon_{\rm Isolated|ID}^{Z_L}\right)$$

(3.5)

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

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

$$\epsilon_\text{Silver} = \epsilon_\text{ID}^{Z_H} \times \epsilon_\text{Isolated|ID}^{Z_H} \times \left(1 - \epsilon_\text{Prompt|(ID$\cap$Isolated)}^{Z_H}\right) \tag{3.6}$$

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

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/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 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 $(5-10)\%$.
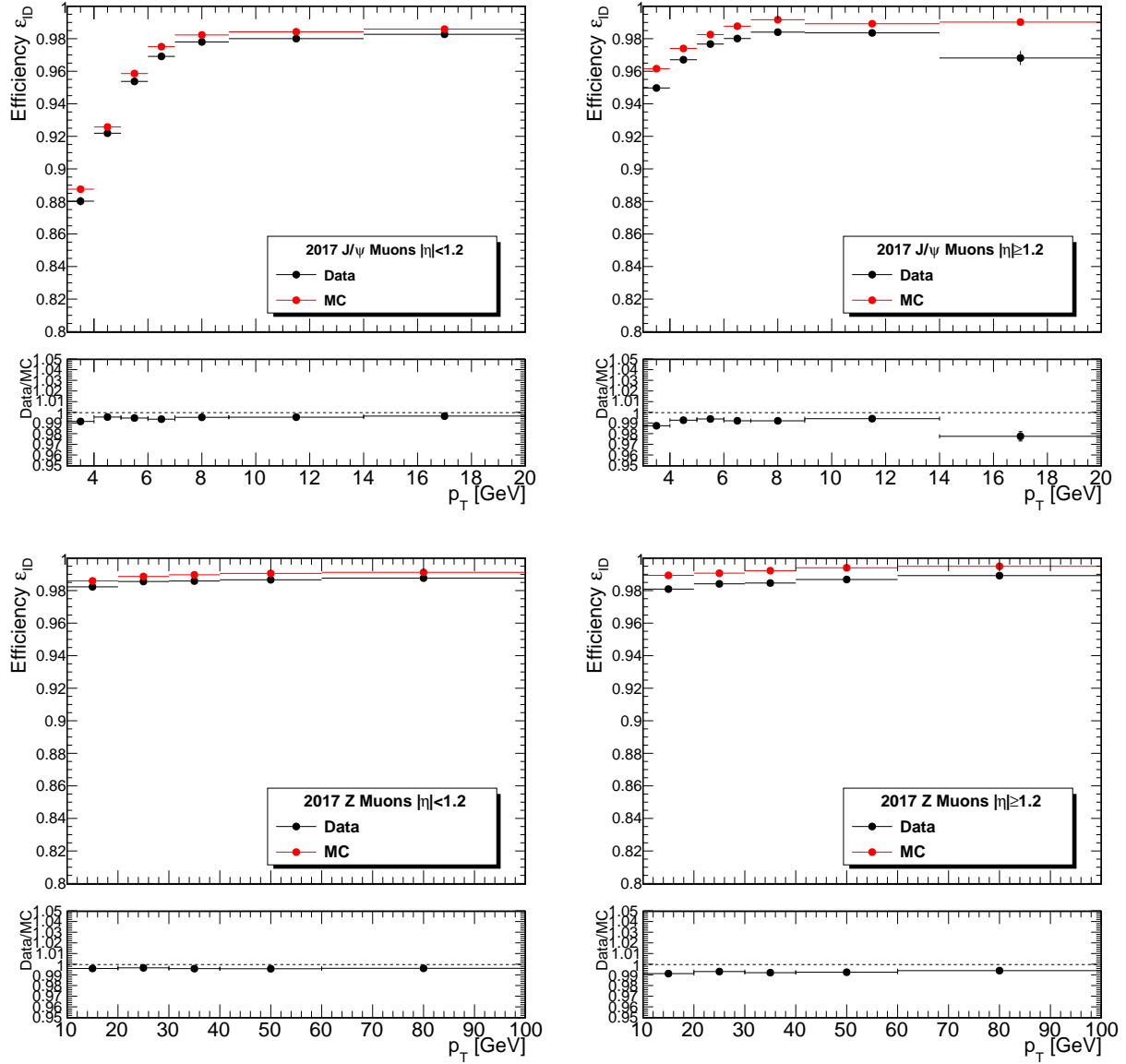

Figure 3.6: Tag-and-Probe efficiencies for the Medium Id in 2017. The left plots show the barrel while the right plots show the endcaps. The top fits use J/ψ resonance while the bottom use the Z resonance.
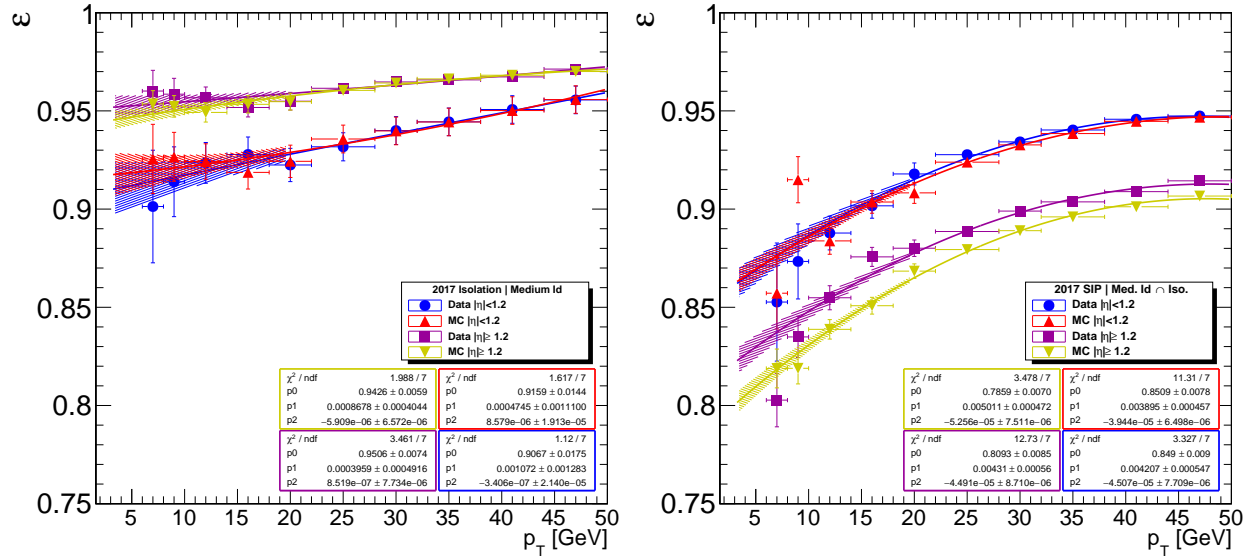
Figure 3.7: The fitted muon isolation and SIP3D efficiencies for 2017. Includes both data and MC which are separated between barrel and endcap.
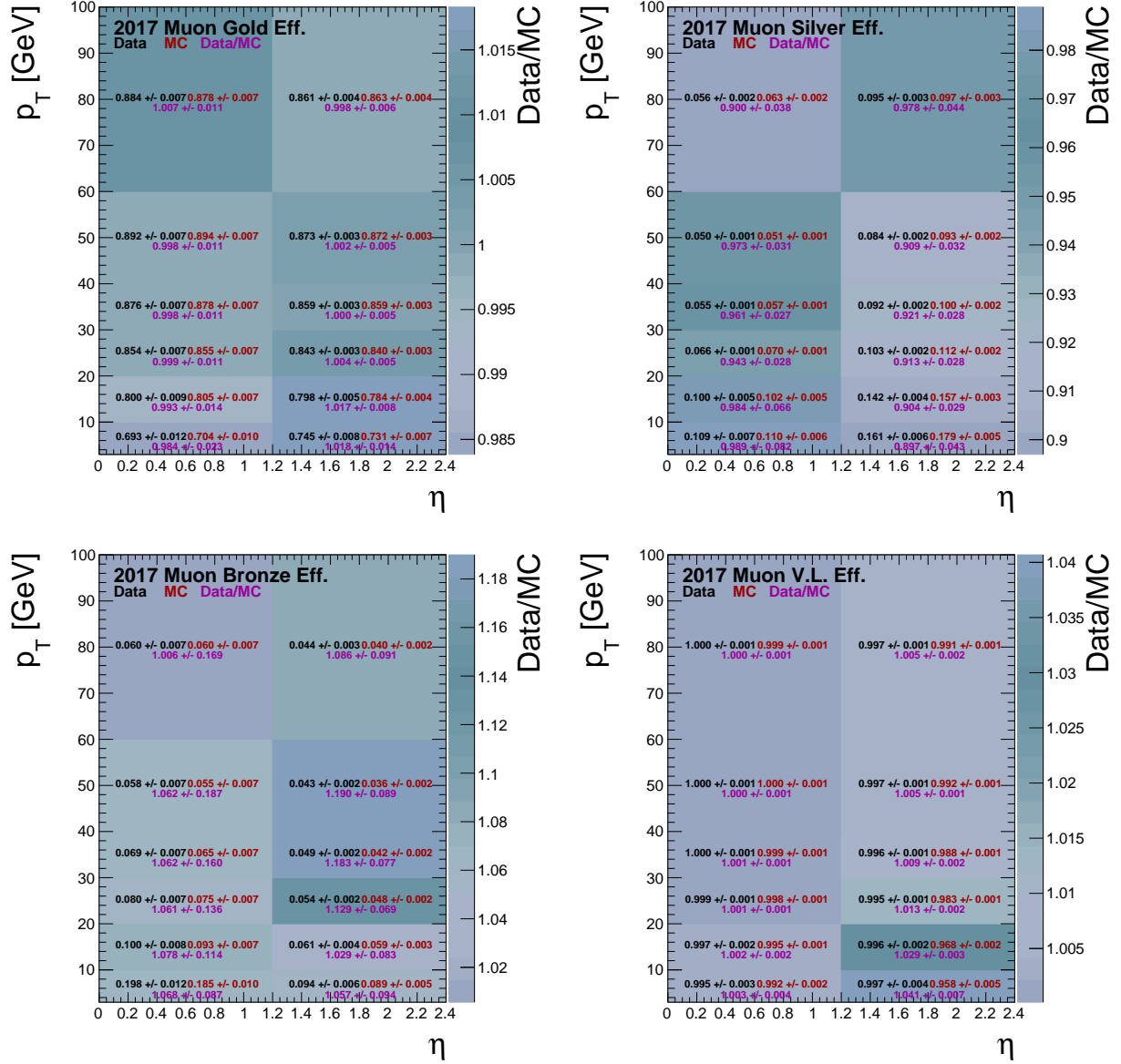
Figure 3.8: The combined efficiency components from equations 3.5 and 3.6 and Very Loose for 2017. The low-p_T region (< 20 GeV) includes the contributions from J/ψ as well as the isolation and SIP3D extrapolations. Propagated errors are treated as uncorrelated.

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 η 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 η range.
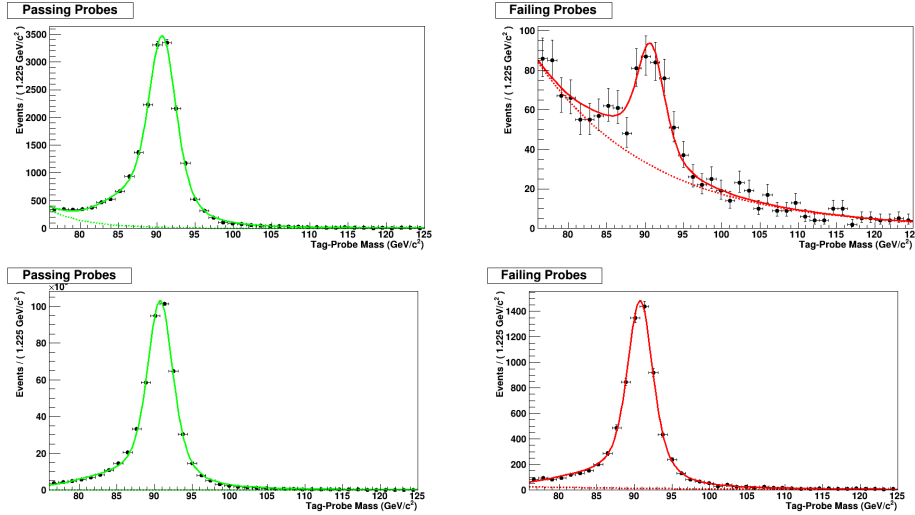

Figure 3.9: Tag-and-Probe di-muon mass distributions for both passing and failing probes. The top set of plots consist of probes below 20 GeV and the bottom set are about 20 GeV.

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 ttbar samples.
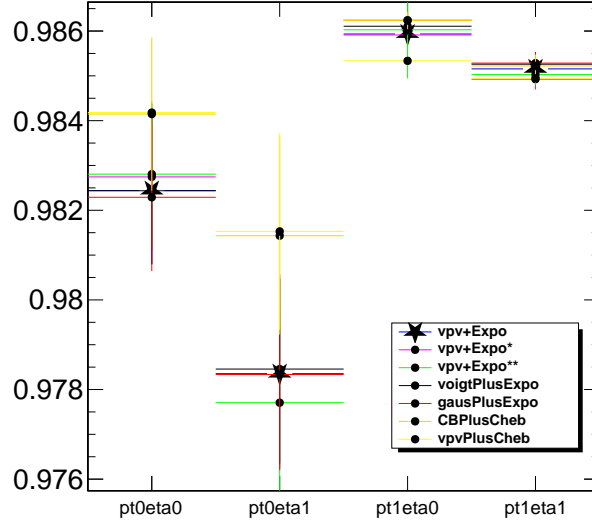

Figure 3.10: Example systematic spread from various fit models and binnings for muons. Includes the four combinations of regions either low or high pt and central and forward eta.

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	$0 \leq |\eta| < 0.8$	$0.8 \leq |\eta| < 1.479$	$|\eta| \geq 1.479$
$p_\mathrm{T} < 20$ [GeV]	0.003	0.001	0.005
$p_\mathrm{T} \geq 20$ [GeV]	0.001	0.001	0.002
Iso | ID			
$p_\mathrm{T} < 20$ [GeV]	0.002	0.003	0.003
$p_\mathrm{T} \geq 20$ [GeV]	0.001	0.001	0.002
SIP | Iso $\cap$ ID			
$p_\mathrm{T} < 20$ [GeV]	0.006	0.004	0.007
$p_\mathrm{T} \geq 20$ [GeV]	0.002	0.002	0.0006
VeryLoose			
$p_\mathrm{T} < 20$ [GeV]	0.002	0.007	0.03
$p_\mathrm{T} \geq 20$ [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	$|\eta| < 1.2$	$|\eta| \geq 1.2$
$p_\mathrm{T} < 20$ [GeV](J)	0.001	0.001
$p_\mathrm{T} \geq 20$ [GeV](Z)	0.001	0.0003
Iso | ID		
$p_\mathrm{T} < 20$ [GeV]	0.007	0.004
$p_\mathrm{T} \geq 20$ [GeV]	0.007	0.002
SIP | Iso $\cap$ ID		
$p_\mathrm{T} < 20$ [GeV]	0.005	0.003
$p_\mathrm{T} \geq 20$ [GeV]	0.001	0.002
Very Loose		
$p_\mathrm{T} < 20$ [GeV]	0.001	0.0003
$p_\mathrm{T} \geq 20$ [GeV]	0.001	0.001