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 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.

1.1 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 [1]. 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 1.1. 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 [2]. 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:

$$N^{\text{Pass}} = N_{\text{Total}}(\varepsilon \cdot f_{\text{All}}^{\text{sig}}) + \varepsilon_{\text{bkg}} \cdot (1 - f_{\text{All}}^{\text{sig}})) \tag{1.1}$$

$$N^{\text{Fail}} = N_{\text{Total}}((1 - \varepsilon) \cdot f_{\text{All}}^{\text{sig}} + (1 - \varepsilon_{\text{bkg}}) \cdot (1 - f_{\text{All}}^{\text{sig}})) \tag{1.2}$$

$N^{\text{Pass/Fail}}$ 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.

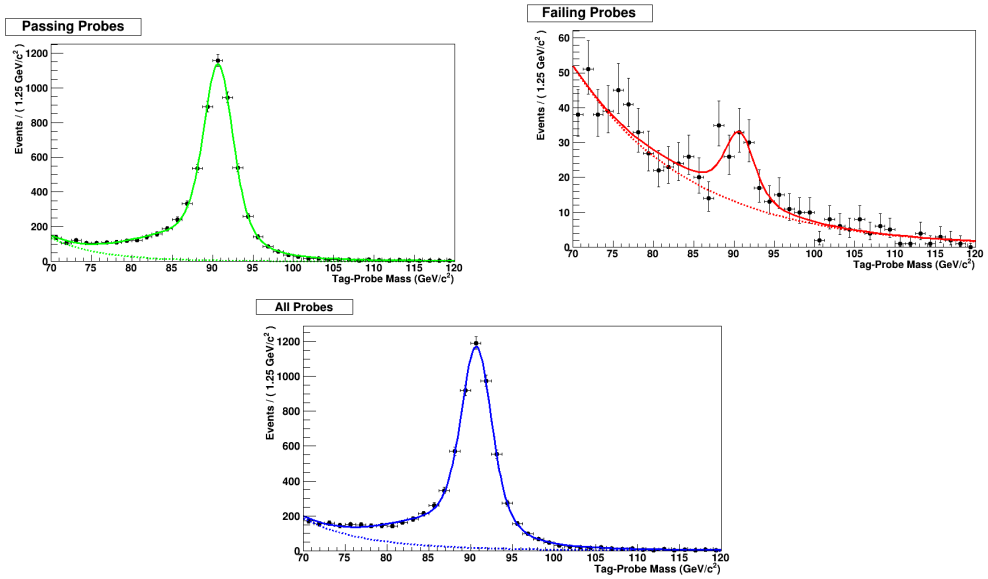

Figure 1.1: Example Tag-and-Probe Z di-muon fits for passing,failing, and all probes with the Medium Id, $|\eta| < 1.2$, and $p_T < 20$ GeV

1.2 Lepton Object Definitions

Leptons are selected according to the minimium requirement "VeryLoose" which depend kinematic and topological quantities which are shown in Table 1.1. 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 [3] and electrons use a more strict selection, due to their messy nature, with the Tight Id [4]. 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 [5] and MiniIso [6]. 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 1.3.

$$R_{\mathrm{miniIso}} = \begin{cases} 0.2 & p_{\mathrm{T}}^{\ell} < 50\mathrm{GeV} \\ \frac{10}{p_{\mathrm{T}}^{\ell}} & 50\mathrm{GeV} \leq p_{\mathrm{T}}^{\ell} \leq 200\mathrm{GeV} \\ 0.05 & p_{\mathrm{T}}^{\ell} > 200\mathrm{GeV} \end{cases} \tag{1.3}$$

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:

$$\begin{aligned} \epsilon_{\text{Gold}} &= \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated|ID}} \times \epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}} \\ \epsilon_{\text{Silver}} &= \epsilon_{\text{ID}} \times \epsilon_{\text{Isolated|ID}} \times (1 - \epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}}) \\ \epsilon_{\text{Bronze}} &= 1 - (\epsilon_{\text{ID}} \times \epsilon_{\text{Isolated|ID}}) \end{aligned} \tag{1.4}$$

The subscript for an efficiency, e.g. $\epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}}$, reads as the efficiency to pass the SIP3D requirement given the lepton passes the Id and Isolation requirements. From equation 1.4 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 1.1: 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
$\lvert\eta\rvert$	< 2.4	< 2.4
$\text{IP}_{3D}/\sigma_{\text{IP}_{3D}}$	< 8	< 8
$\lvert d_{xy}\rvert$	< 0.05 cm	< 0.05 cm
$\lvert d_z\rvert$	< 0.1 cm	< 0.1 cm
$\text{PFIso}_{\text{abs}}$	$< 20 + (300/p_{\text{T}})$ GeV	$< 20 + (300/p_{\text{T}})$ GeV
MVA VLooseFO ID	✓	–

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 1.2 and the overall efficiency for Gold, Silver, and Bronze on truth matched objects are shown in Figure 1.3. 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.

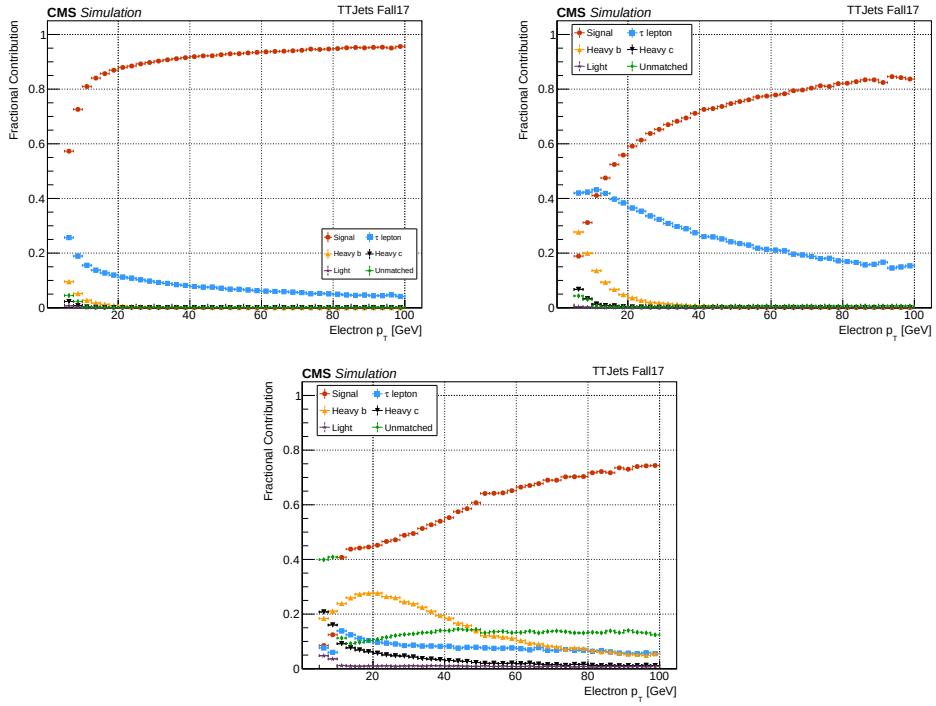

Figure 1.2: Gold (Top-Left), Silver (Top-Right) and Bronze (Bottom) MC truth matching in TTJets sample 2017. Signal is defined here as prompt electrons from a W decay.

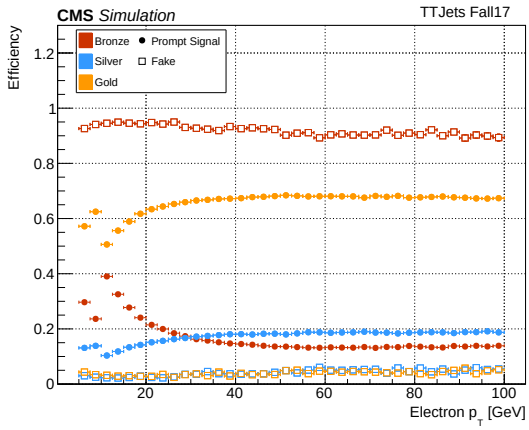

Figure 1.3: Gold, Silver, and Bronze efficiency on truth matched prompt electrons as signal and secondary electons as Fakes.

1.3 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` [7]. The samples used in the Ntuplizing stage are listed in Table 1.2. In the fitting stage, a random subset of of TnP pairs are sampled with `TnPTreeAnalyzer` [8]. 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 1.2: 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 1.3: selection

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

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 1.3. 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 1.4, are shown in Figure 1.4. 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 1.5. 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.

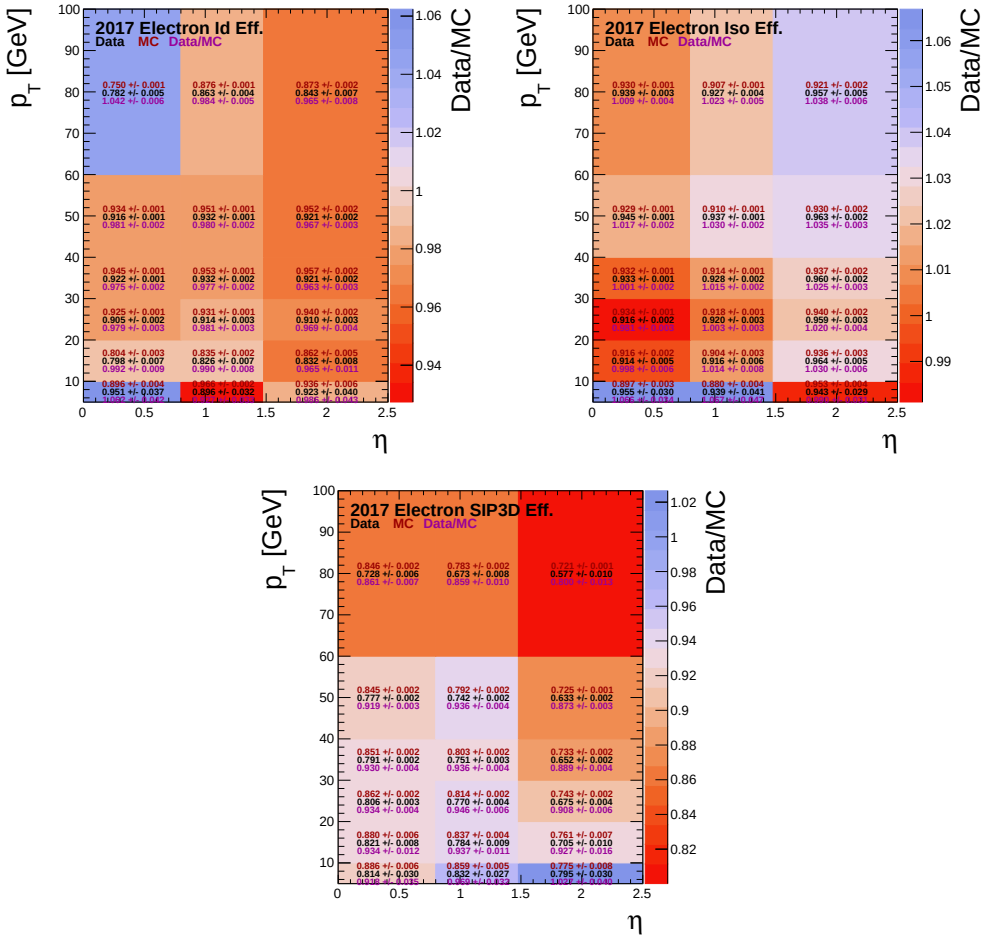

Figure 1.4: 2017 efficiencies

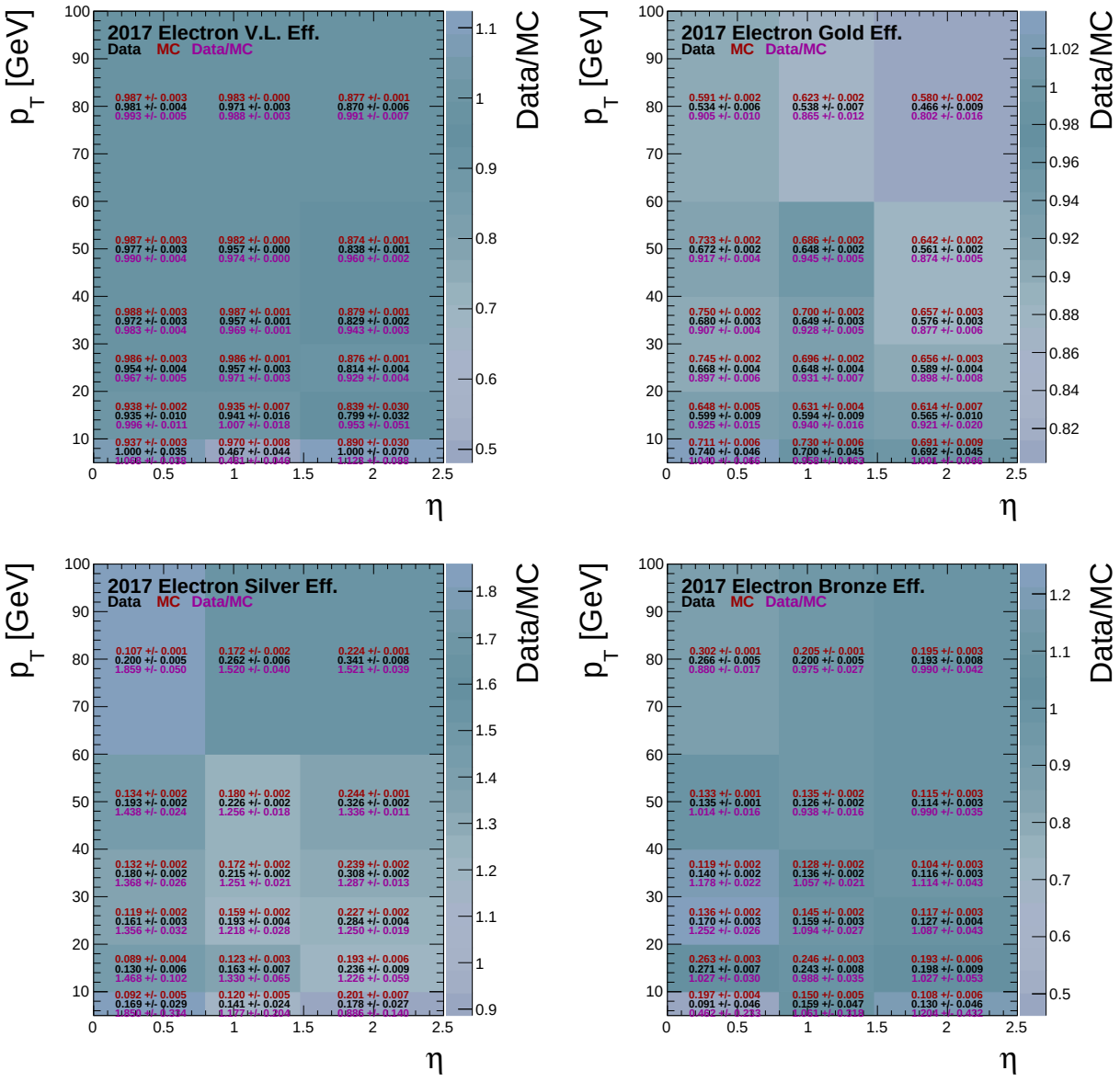

Figure 1.5: 2017 electron GSB efficiency and SF

1.4 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 [9] and fitting [10] stage. The code bases for muons and electrons are separate but functionally identical. The samples chosen for Z measurements are shown in Table 1.4. 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 [11]. 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 1.5.

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 [12]. Similary another 10% of all J/ψ are produced within b-jets and leading to non-prompt unisolated events [13].

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 1.4

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

Table 1.5: muon binning

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 $p_{\mathrm{T}} > 5$ GeV Matches hltIterL3MuonCandidates	Matches hltTracksIter OR Maches hltMuTrackJpsiEffCtfTrackCands	$2.8\text{GeV} < m_{\mu\mu} < 3.4\text{GeV}$ $\|z_{\mu_1} - z_{\mu_2}\| < 1$ cm
Z		
passes tightID $\sum p_{\mathrm{T}}^{ch}/p_{\mathrm{T}} < 0.2$ $p_{\mathrm{T}} > 15$ 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_{T} and reflect the high and low binning separations shown in Table 1.5. The low p_{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_{T} muons are composed of all the factors directly measured in Z_H.

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

$$\begin{aligned}
\epsilon_{\text{Gold}} &= \epsilon_{\text{ID}}^{J/\psi} \times \epsilon_{\text{Isolated|ID}}^{Z_L} \times \epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}}^{Z_L} \\
\epsilon_{\text{Silver}} &= \epsilon_{\text{ID}}^{J/\psi} \times \epsilon_{\text{Isolated|ID}}^{Z_L} \times (1 - \epsilon_{\text{Prompt|(ID}\cap\text{Isolated)}}^{Z_L}) \\
\epsilon_{\text{Bronze}} &= 1 - (\epsilon_{\text{ID}}^{J/\psi} \times \epsilon_{\text{Isolated|ID}}^{Z_L})
\end{aligned} \tag{1.5}$$

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

$$\begin{aligned}
\epsilon_{\mathrm{Gold}} &= \epsilon^{Z_H}_{\mathrm{ID}} \times \epsilon^{Z_H}_{\mathrm{Isolated|ID}} \times \epsilon^{Z_H}_{\mathrm{Prompt|(ID\cap Isolated)}} \\
\epsilon_{\mathrm{Silver}} &= \epsilon^{Z_H}_{\mathrm{ID}} \times \epsilon^{Z_H}_{\mathrm{Isolated|ID}} \times (1 - \epsilon^{Z_H}_{\mathrm{Prompt|(ID\cap Isolated)}}) \\
\epsilon_{\mathrm{Bronze}} &= 1 - (\epsilon^{Z_H}_{\mathrm{ID}} \times \epsilon^{Z_H}_{\mathrm{Isolated|ID}})
\end{aligned} \tag{1.6}$$

The 2017 Id efficiency with statistical errors for both data and MC are shown in Figure 1.6. 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 1.7. The errors for each bin are the combined statistical and systematic errors from Table 1.7 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 1.7 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 1.8. 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 $(5-10)\%$.

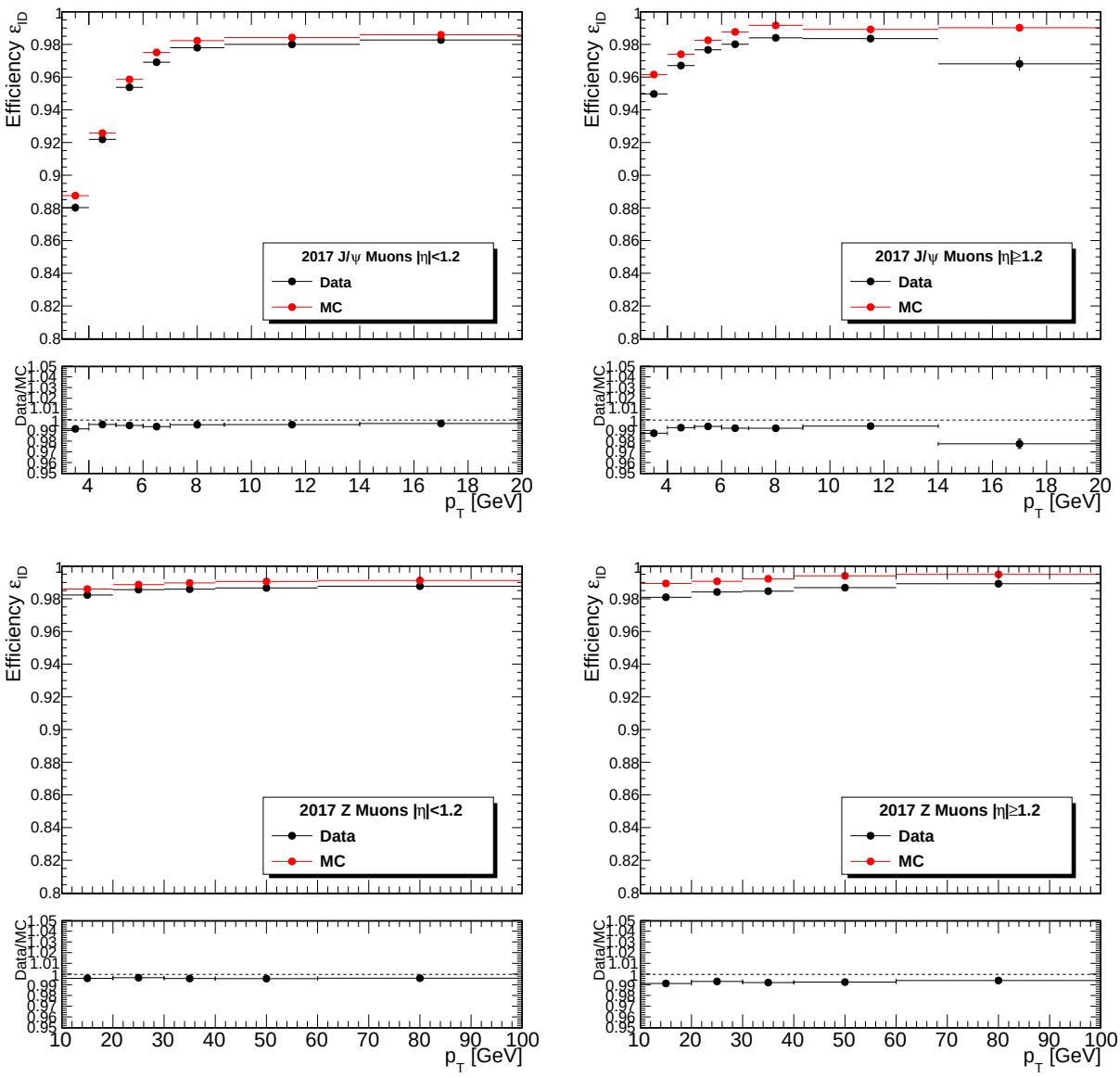

Figure 1.6: Tag-and-Probe efficiencies for the Medium Id in 2017. The left plots show the barrel while the right plots show the endcaps. The top fits use J/ψ resonance while the bottom use the Z resonance.

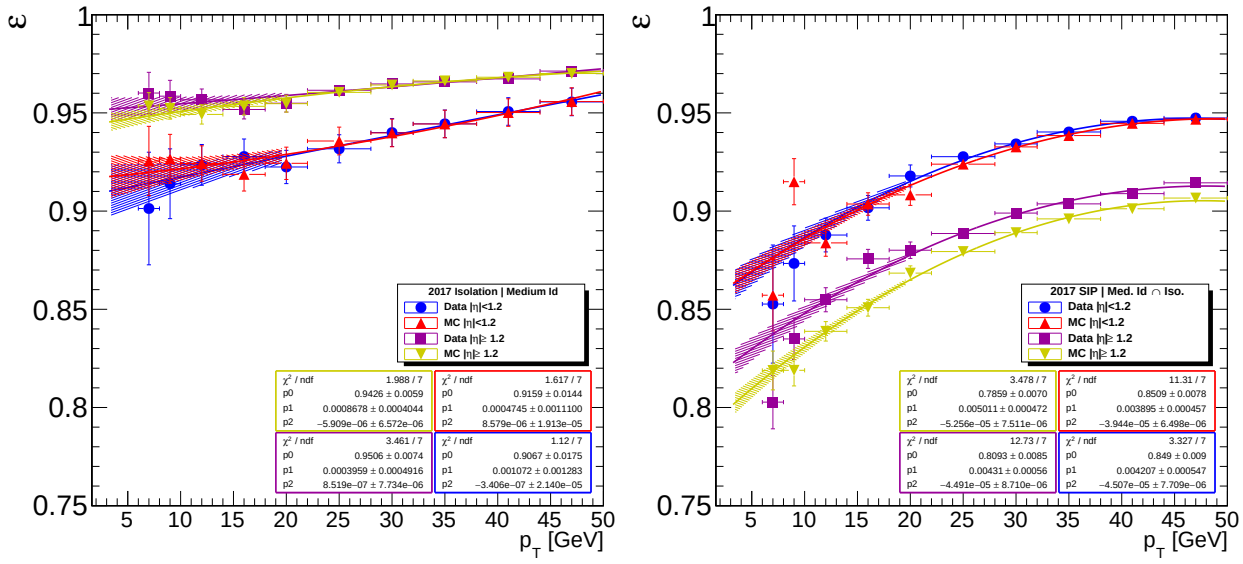

Figure 1.7: The fitted muon isolation and SIP3D efficiencies for 2017. Includes both data and MC which are separated between barrel and endcap.

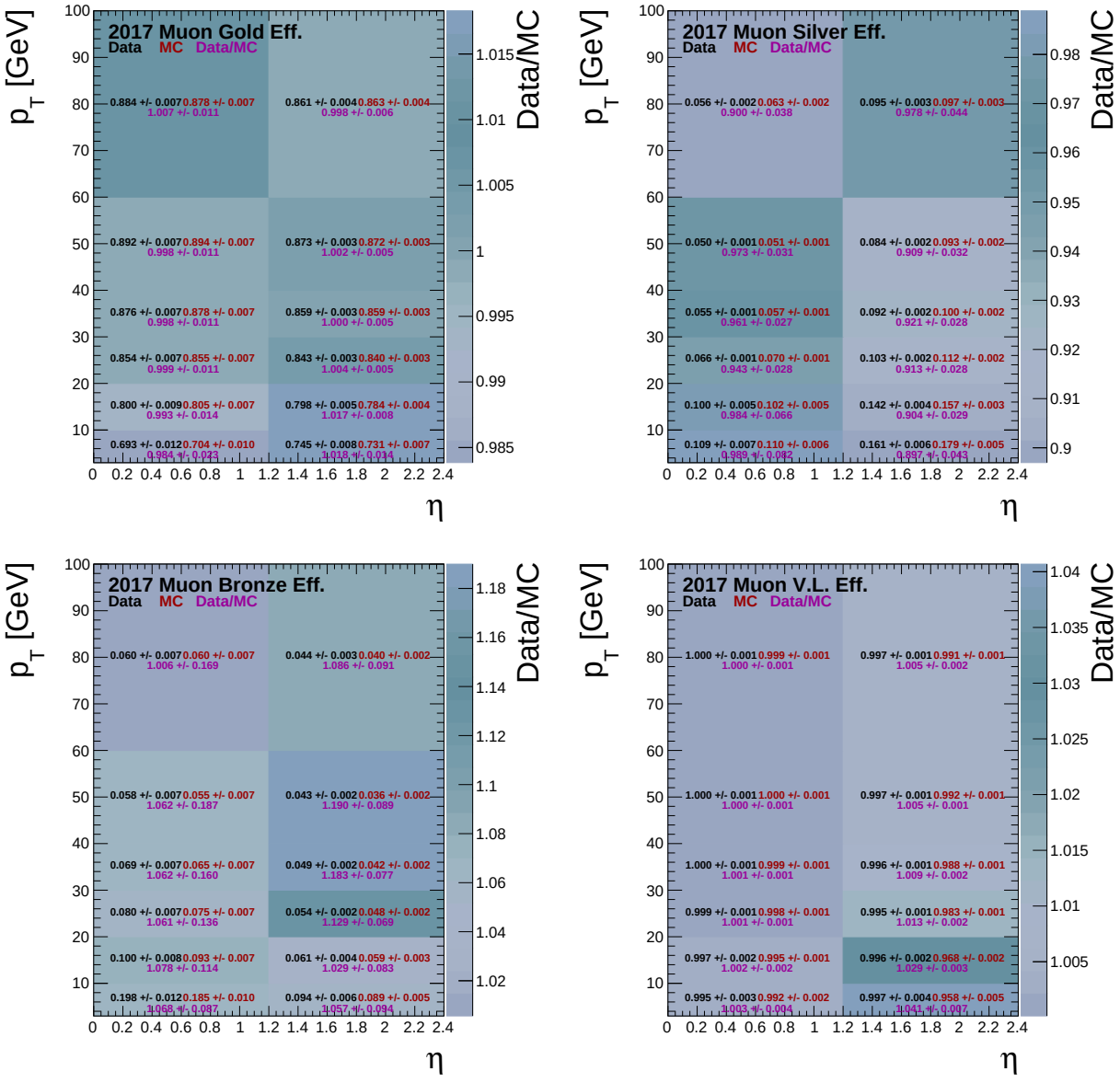

Figure 1.8: The combined efficiency components from equations 1.5 and 1.6 and Very Loose for 2017. The low-p_{T} region (< 20 GeV) includes the contributions from J/ψ as well as the isolation and SIP3D extrapolations. Propagated errors are treated as uncorrelated.

1.5 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 1.10. 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 1.9. 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.

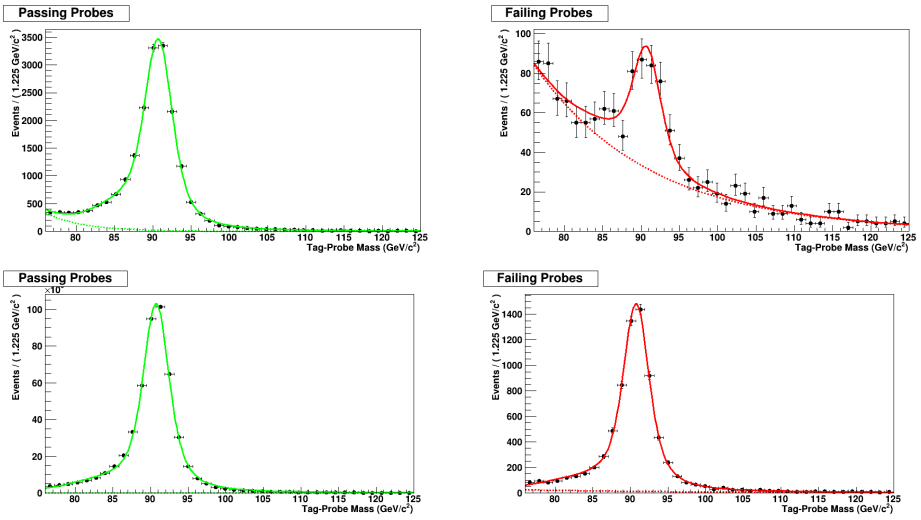

Figure 1.9: Tag-and-Probe di-muon mass distributions for both passing and failing probes. The top set of plots consist of probes below 20 GeV and the bottom set are about 20 GeV.

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 1.6 and Table 1.7.

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.

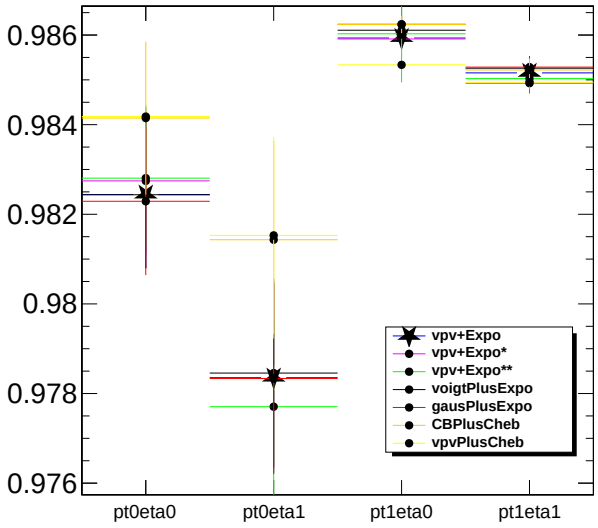

Figure 1.10: Example systematic spread from various fit models and binnings for muons. Includes the four combinations of regions either low or high pt and central and forward eta.

Table 1.6: The electron systematic error derived from the Tag-and-Probe for 2017 data and split into p_{T} and $|\eta|$ regions.

ID	$0 \leq \lvert\eta\rvert < 0.8$	$0.8 \leq \lvert\eta\rvert < 1.479$	$\lvert\eta\rvert \geq 1.479$
$p_{\mathrm{T}} < 20$ [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 1.7: The muon systematic error derived from the Tag-and-Probe data and split into p_{T} and $|\eta|$ regions.

ID	$\lvert\eta\rvert < 1.2$	$\lvert\eta\rvert \geq 1.2$
$p_{\mathrm{T}} < 20$ [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

References

[1] V. Halyo A. Hunt K. Mishra N. Adam, J. Berryhill. Generic tag and probe tool for measuring efficiency at cms with early data. June 2009.

[2] Jeffrey Berryhill, Georgios Daskalakis, Valerie Halyo, Jeremy Werner, and Si Xie. Electron efficiency measurements with 2.88 pb^{-1} of pp collision data at $\sqrt{s} = 7$ tev. Nov 2010.

[3] Proposal of a new "medium" muon id. `https://indico.cern.ch/event/357213/contributions/1769745/attachments/710701/975626/muonid-pog081214.pdf`. Presented 2014-08-12.

[4] Electron identification based on simple cuts. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/EgammaPublicData`. Late updated 2016-04-22.

[5] Baseline muon selections for run-ii. `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Particle_Flow_isolation`. Late updated 2022-01-19.

[6] Alternatives to standard isolation for 13 tev. `https://indico.cern.ch/event/368826/contributions/871862/attachments/733500/1006397/20150127_ald_susy_iso.pdf`. Presented 2015-01-27.

[7] Electron tag-and-probe producer git repository. `https://github.com/cms-egamma/EgammaAnalysis-TnPTreeProducer`. Last updated 2022-06-27.

[8] Electron tag-and-probe analyzer git repository. `https://github.com/cms-egamma/egm_tnp_analysis`. Last updated 2022-09-01.

[9] Muon tag-and-probe producer. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonTagAndProbe. Last updated 2012-02-11.

[10] Tag-and-probe analyzer. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideTagProbeFitTreeAnalyzer. Last updated 2010-06-03.

[11] Muon tag-and-probe instructions for run-ii. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonTagAndProbeTreesRun2. Late updated 2020-05-18.

[12] J. P. Lansberg. J/ψ, ψ ' and Υ production at hadron colliders: A Review. *Int. J. Mod. Phys. A*, 21:3857–3916, 2006.

[13] R Aaij et al. Production of J/psi and Upsilon mesons in pp collisions at sqrt(s) = 8 TeV. *JHEP*, 06:064, 2013.