

Questions:

1). I notice in Table 2 that FastSim and Pythia are used for the $t\bar{t}$, Z+jets and W+jets backgrounds. Since these are the backgrounds where the fake lepton effects dominate, I think they are the ones that most need to be studied with MadGraph and FullSim. Although general comparisons have been made between FastSim and FullSim, the fake leptons are dependent on the tails of distributions, which is where results will be most sensitive to simulation. I would expect that the current samples will underestimate the background from fakes, which will make the claimed sensitivity overly optimistic and could substantially change the result.

There are comparisons between FastSim and FullSim for relatively tight electron identification criteria (muon experts -- do you know anything about muons?) and these agreed well between each other (both efficiencies and fake rates). However, as we do not know how well data is going to agree with either of these simulation schemes, we added a sentence to PAG Monte Carlo simulation section to clarify the situation. Here is the sentence: "Fast simulation can result in a relatively simplified modeling of the misidentified leptons in the final state selection. However, as we plan to use data-driven methods to extract backgrounds and use Monte Carlo simulation as a cross-check, the dependence on simulation accuracy is not expected to affect the results."

2). How are the error bars determined for figures 3-6? I think they are not calculated correctly, e.g., because one high p_T bin in Fig. 3 has efficiency=1 with no visible uncertainty.

These are binomial errors which is the standard way to display efficiency numbers. However, once efficiency is close to 100% or 0%, the uncertainties cannot be trusted. It is not important as these Figures are used only to illustrate the performance of the selection criteria. The numbers come from tag and probe method.

3). Section 10 describes how one might measure the lepton identification efficiency, particularly as a function of p_T . But, it is not clear how you do that for this particular result. What efficiency is used to calculate the sensitivity, i.e., the limit and significance plots? What efficiency errors are used for that? From line 339, I think that a 1% uncertainty is used. But, Figures 3-6 and the discussion in section 10 lead me to expect a larger uncertainty.

Yes, Figures 3-6 are used for illustration purposes. The uncertainties in these plots correspond to the size of Z->mumu W->enu sample and as such, these uncertainties are quite big. As we use Z->ee for electron ID measurement, we plan to have 1-2% uncertainties with as early as 100 /pb data sample. These numbers have been used in a number of results and originally they came from estimations from the tag and probe developers (in collaboration with Z->ee and W->enu measurement teams).

4). The H/E cut in table 3 could be sensitive to pile-up because a cut of $H/E < 0.016$ corresponds to less than a GeV for low p_T electrons. How have you studied the effect of pile-up on the sensitivity?

We completely agree that pile-up could affect the sensitivity. Unfortunately, we do not have samples generated with the pile-up to estimate the effect, and the H/E definition has been redesigned starting with 31X release. We have modified the PAS to clarify the point. We have added the following sentence to the selection section: "Any potential difference in the selection that may arise due to pile-up and different detector conditions will be investigated in future versions of the analysis and eventually with collision data." We do think that if the pile-up activity renders H/E less effective, there are other variables that we can tighten up to achieve the desired level of performance (H/E is not a crucial ingredient of the electron ID anyways).

5). I don't see that WW+jets has been included in the background estimate. I'd imagine that it could be important since it could give a real e and mu and leave the second e only requiring loose cuts and the mass window. How have you calculated the WW background?

WW+jets has the same final state as Z+jets and $t\bar{t}$, and as such it is going to contribute with the same scale factor as the latter two processes. As cross section is very small compared to Z+jets and about 1/5 of the $t\bar{t}$

production we neglected this process in this analysis. We also plan to estimate backgrounds without genuine Z using the sideband subtraction method, and this would take $t\bar{t}$ and WW+jets (together with W+jets and QCD) as one source.

6). The conclusion of section 8 is that the QCD background is small. However, that relies on the efficiency for the WZ mass cut that is determined from the highest pT QCD MC samples. Those high pT bins will have high HT and therefore I expect them to reconstruct mostly to higher WZ masses. As such, there is likely to be a higher efficiency for the WZ mass cut in the lower pT QCD MC samples. Since that is where most of the, pre-mass cut, QCD background events are, I expect that the quoted QCD background could be a substantial underestimate. How can you predict the WZ mass cut efficiency for the lower pT QCD MC samples?

Thanks for catching this. It is just a poor wording of the text. The whole pT-hat range was considered, but the low-pT bins did not contribute to the signal region (mostly due to the fact that high pT jets become more narrow and can easily fake a matching of a track to EM cluster in case of electrons. Thus, the electron ID fake rate somewhat increases with the pT). The good thing is that the QCD process has a smooth WZ transverse mass distribution and thus an estimation of the background in the signal region can be approximated by a simple polynomial.

7). Line 239 states that the Zgamma background "can be determined from data once the FSR Z signal is measured at CMS". But, how do you determine this background for the current estimate?

We used the pythia Zgamma sample to estimate the background and found it to be negligible in the WZ transverse mass region of interest. We have modified the PAS to clarify this point.

8). Section 9 states that the $t\bar{t}$ background might be measured in data using the fraction of b-tagged events. I expect that such a method will have limited applicability in this analysis since the fake lepton in a trimuon signature from $t\bar{t}$ will contain b-quarks that fragment mostly into a single lepton. That is a region of the fragmentation function that is not well modeled in the b-tagging efficiency.

Agreed. We also plan to use the sideband subtraction method (a la standard model WZ production) together with the MC estimation and the b-tagging method as a cross check. We have added a line to this effect to the text.

9). The data-driven determination of isolation efficiency for fake leptons apparently does not include the WZ, WW, and $t\bar{t}$ samples. Those processes will contribute to actual use of this method in data. Why aren't they included here?

Completely agree with you. The text does not explain the method completely and we fixed it in the current version. We require the probe lepton to be of the same charge with the tag lepton. This takes care of WW and $t\bar{t}$. We also impose a Z boson veto if there is an additional lepton of the same flavor is found in the event.

10). I'm puzzled by Tables 10 and 11. First, I would expect that the electron fake rate would be considerably higher than the muon fake rate (despite the 5GeV difference in the pT cut). Why is that not so? Second, I don't understand why $B_{\{TT\}}$ is higher for muons than electrons (which is actually negative). Third, the errors quoted on $P_{\{fake\}}$ are for 1/fb. Scaling this down to 300/pb, which is relevant for the low mass pT limit, I expect a much larger error on $P_{\{fake\}}$; that could degrade the sensitivity. Are these errors propagated into the limit and sensitivity calculations?

$B_{\{TT\}}$ is negative??? Whoever calculated this -- please comment. It is either zero or positive. I can comment on why electron fake rate is similar to muons -- this is due to much improved isolation criteria. Also, can someone confirm uncertainties? (Jeff) Yes, the method I have been using fits a Breit-Wigner and a linear, then I integrate underneath each of those curves in a window around the peak. The linear function always ends up near zero. I will try to repeat this calculation now, with the newly processed Z+jets with 2-lepton preselection to see if I see any differences.

11). I think that the conclusion that "it is possible to discover ρ_T up to masses of about 300 GeV...using $\sim 500/\text{pb}$ " is not correct. That luminosity gives only a 3 sigma significance, which is not sufficient for a discovery. Furthermore, the 313/pb prediction for a 95% CL limit could vary up to 600/pb due to the stated cross-section uncertainty. Taken together, I think the general tone of the conclusion should be that there is only limited sensitivity with 10TeV, i.e., a much weaker conclusion than is stated.

We agree and we restated the conclusion in the PAS.

This topic: Sandbox > TechniWZQuestions

Topic revision: r4 - 2009-06-18 - TulikaBose



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

or Ideas, requests, problems regarding TWiki? use Discourse or Send feedback