1, AN section 4.1 - Soft muon requires at least 6 tracker layers, while the AN says that 6 tracker hits are required. That is not the same. Is it just a typo in the AN?

- This is a typo and we will fix it in the updated AN.

2, MC - some of the trigger efficiencies shown in the dropbox (line 129 in the AN, https://www.dropbox.com/s/47nwk9uu1tnspry/3DMthEffAndTrigEff.pdf?dl=0) have large difference for mu+ vs mu-, much larger than the error bars. Moreover, they jump up and down quite erratically. Do you know why? (In the dropbox, the biggest differences are around page 70,71, but in some other places as well)

- This should be related to muon detector acceptance in azimuthal direction (not 100% coverage). Because the bending curvatures of mu+ vs mu- are different, the efficiencies between mu+ and mu- are expected to be different in the edges of each muon sector (this analysis uses 60 fine phi bins). As shown in Figs. 2 and 3 of AN v4, there seems to be some acceptance holes (non-functional muon sector(s)?) in mid-rapidity range (-0.3 < eta < 0.3), which will amplify the efficiency difference between mu+ and mu- causey by acceptance, as shown in p69-p74 of 3-D single muon efficiencies.
    - Ota: Related to this: Phi efficiency dependence is based just on MC. I talked to Devin (muon POG convener) and we weren't sure how reliable the MC is in that regard. First of all, does it even make a difference? Can you please check how much the result changes if you use different binning for the efficiency (e.g. just integrated in phi?) If it doesn't, then I don't think we'd need to worry about it
        - We checked the consistency between MC and data, and found the MC describes the data reasonably well in phi direction, see p5-6 in UPC TnP Study.
            - Ota: the TnP shows (slide 5-6) 2 large dips in phi, around -2.5 and 0.5 (both in MC and data. It is entries, but the dips are presumably caused by inefficiency). If I compare it with the figure 2 (part b and d) in the AN, I can clearly see the dip around 0.5, but nothing around -2.5. Can you explain where the difference between the MCs comes from?
                - The single muon distributions in slides 5-6 are from reconstructed dimuon pairs with 8<M<60 GeV and alpha<0.006. If one muon of pair is missing, the other muon cannot be filled into the histograms in slides 5-6 (UPC TnP Study). The dimuon pairs from leading order photon-photon scattering is almost back to back, therefore, the dip around phi ~ -2.5 is purely caused by the inefficiency of phi ~ 0.5 (delta_phi ~ pi). This phenomena actually tells us the phi description in MC is good.

    - The reason we use phi dependent efficiencies is the dimuon pair from photon-photon scattering has strong correlation in phi direction (back to back). The phi integrated efficiencies (average effect) may not accurately account for this back to back feature.
        - Ota: That is fine. My question was if the phi differential efficiencies make an actual difference for the results. If it matters, then we should make sure that the phi description in MC is good (see the comment above)
            - I tried to use phi integrated 2D efficiency to correct the alpha and mass spectra. Compared to the results with 3D efficiency correction, the difference is <0.4% on <alpha> and <0.2% on <M>.
- Alice: Why do you say that the bending curvature is different? The magnetic field is totally uniform and the bending of mu+ and mu- are the same. It is the first I hear such an argument. Can you explain what you mean? I can understand that depending of the physics process that creates the muons and depending of the PDF, mu+ ad mu+ can be pushed in different region of the detector (like for W or Z) and if a phi region is less efficient than another one then it can indeed induce efficiency difference between mu+ and mu-. But the muon object itself is the same.
    - Sorry for the misleading. We intended to say the **bending direction** (NOT deflection curvature radius) between mu+ and mu- is different, thus, the muon related efficiency could be different in a small differential phi region, especially at the edge of the muon detector sector. As you mentioned, because of the uniform magnetic field, the overall efficiencies from mu+ and mu- should be the same. The detailed differential and overall efficiency comparisons between mu+ and mu- can be found in our responses to your other comments (ARC_Alice_comments)

3, Equation 1 says that it uses GEN kinematic variables for N_Reco. How do you do the matching?

- We do the matching between GEN and RECO using DeltaR variable (deltaR = genMom.DeltaR(recMom))

4, Figure 4, right - what causes the big jump in the resolution? Is it muons reaching the barrel?

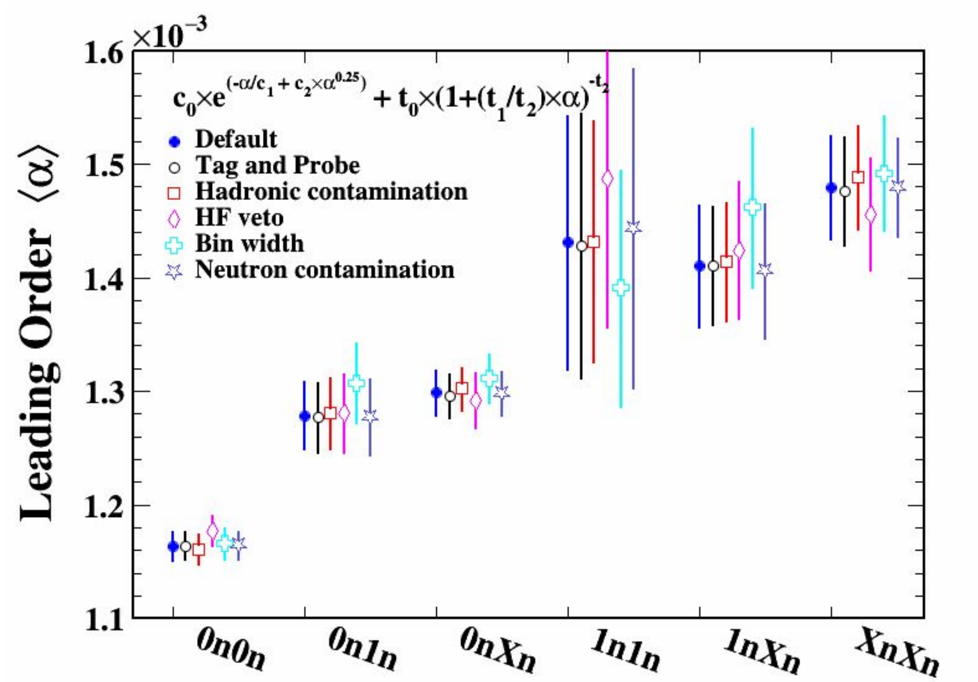
- The big jump should be caused by the end-cap muon chambers because All muon with pt<3 are from end-cap sectors. Anyway, in this analysis, we only use muon with pt>3.5 GeV

5, AN section 4 - 140-142. I'm not sure about the official T&P being the upper limit on the data/MC difference. The problem is that the official tag and probe is derived from the difference between data and MC PYTHIA8 with the CP5 underlying event tune. While your MC is

STARlight. So even if the muID and trigger were similar, the differences to MC might be rather different. Maybe André can comment on it, he knows more about the generators.

- The logic here is 1) The requirements of soft muon are looser than that of the hybrid muon while the tracking performance in UPC is at least no worse than in hadronic collisions; 2) The trigger requirements in UPC (L1) are looser than that in hadronic collisions (L2 or L3). Therefore, the tracking, muon identification and trigger efficiencies in this analysis are expected to be higher than that of hybrid muons in hadronic collisions and the corresponding TnP SFs are expected to be smaller and flatter than the ones of official TnP for hybrid muons [1]. This expectation is true through Marek's preliminary TnP study in UPC [2] (p7).
- [1]: Official TnP SFs for hybrid muon in hadronic collisions
- [2]: Marek's study for soft muon TnP in UPC
    - Ota: I understand the logic and it makes sense to me (in absence of more relevant TnP). However, taking the difference of results with/without the SFs as an uncertainty might actually underestimate the uncertainty from the TnP (if the scale factors are closer to 1 than the size of the error bars, e.g. 0.99+-0.03). Looking at figures 6,7 and 8 there are quite a few bins that are like this, with 1 covered by the uncertainty. So, to be really conservative, I think that you should calculate the standard uncertainty from the TnP (see how to: https://twiki.cern.ch/twiki/pub/CMS/HIMuonTagProbe/TnpHeaderFile.pdf). Then check that it is smaller than the difference (with/without the SF). If not, use that instead.
        - The standard uncertainty from TnP is still negligible compared to our current statistical and systematic uncertainties. This is because our analysis does not measure the absolute cross-section and only cares about the spectral (alpha or mass) shape, therefore the efficiency shape is more relevant.
        - Meanwhile, we derived our own soft muon ID and L1 trigger TnP efficiencies using the STARlight MC and single muon UPC triggered data samples (see details in UPC TnP Study). The difference between with and without hybrid muon SF is < 0.3%, between with and without new derived soft muon SF is < 0.24%, between with hybrid muon SF and with soft muon SF is < 0.8%. All these cross-checks demonstrate our results are not sensitive to the TnP SF. Therefore, we would like not to apply TnP SF in this analysis and purely rely on MC simulation.
    - Ota: The TnP has not yet been approved by the muon POG (Batoul will present it to the muon POG on the 2nd of March), so it is not quite official yet. Is there any chance that Marek will finish the UPC TnP in time for this paper?
        - I am not sure. As we said in the last response, we plan to purely use MC efficiency without TnP SF in this analysis because: 1) the effect caused by TnP is super tiny; 2) we do not have approved TnP correction factors in hands.

- Ota: NB: I think that I originally misunderstood what STARlight is used for. It is just a generator, while GEANT is used for the detector simulation + rest is the usual. Could you please modify AN line 74-76 to include mention of GEANT ("...Monte Carlo (MC) samples, generated by STARlight [21], were produced to study the detector effect, including resolution, tracking, and trigger efficiencies")
    - Done. They will show up in the updated AN.
- Alice: The hybrid soft ID starts from a global muon while the soft muon ID starts a tracker muon. It is two different reconstruction and a part of the phase space you use is not covered by the global reconstruction (since tracker muon is more efficient at low pT, in the endcaps and in the barrel crack); You can't use hybrid soft ID TnP to estimate your TnP correction. Instead I recommend you to use the soft ID TnP measured in pp collisions or to derive your own correction.
    - To make our cross-checks more convincing, we used the STARlight MC and single muon UPC triggered data samples to generate our own **soft muon** ID and L1 trigger TnP efficiencies (see details in [UPC TnP Study]). The tracking TnP SF seems to be flat in pp collisions (0.99? Please correct me if I am wrong), thus we still use the **unflat tracking SF** (Fig.6 in AN) from hybrid muon and new derived soft muon ID and L1 trigger SFs to conservatively estimate possible effects caused by TnP. We found the TnP effect on <alpha> is <0.24% (blue solid circle vs. black open circle), which is negligible compared to the statistical and other systematic uncertainties. Therefore, we would like to argue not to use TnP efficiencies in this analysis because 1) the effect caused by TnP is super tiny; 2) we do not have approved TnP correction factors in hands



    -

- Alice: We need a comparison between STARlight MC and the MC used for TnP to make sure that there is no additional systematics coming from the difference of generator
    - I don't fully understand your suggestion, I guess your concern is that CMS GEANT4 may have different performance for different generators? I do know how to compare STARlight MC and the MC (PYTHIA8 with the CP5 underlying event tune) used for hybrid muon TnP efficiencies.
    - Anyway, We switched to our own soft muon TnP and we checked the consistency between MC and data. The STARlight MC describes UPC data reasonably well (p4-6 in UPC TnP Study)

6, Questionnaire, answer 4 ("The muon reconstruction and trigger efficiencies are applied as a function of pT, η, and φ on a candidate-by-candidate basis, where candidate means muon pair"). Could you please clarify what is done exactly?

- We apply the efficiency correction for each muon pair (opposite-sign and same-sign) during pairing process and each pair is weighted by $(\varepsilon_{trig} \times \varepsilon_{reco})^{-1}$, where $\varepsilon_{trig} = 1 - (1 - \varepsilon^{\mu+}_{trig})(1 - \varepsilon^{\mu-}_{trig})$ and $\varepsilon_{reco} = \varepsilon^{\mu+}_{reco} \times \varepsilon^{\mu-}_{reco}$. The corresponding efficiencies are sampled as a function of pt_{mu}, η_{mu}, and φ_{mu}.