

- BLUE answered
- PINK for after unblinding data or after pre-approval
- RED to be answered

## FR comments

We started a new twiki to prepare the paper for FR.

## CWR comments

We started a new twiki to prepare the paper for CWR.

## Comments on Paper draft v1 by (June 29th, 2017)

For completeness, we list a set of style comments by Bill Gary and Joe Cole. We implemented most of the comments unless they conflicted, were not relevant because of other text changes. In a few case we answer questions or explain why a comments were not considered.

Comments by Bill Gary (reading up to around line 124)

Abstract: line 4: "...missing TRANSVERSE momentum..."

line 5: "... with A large rapidity gap and A large dijet mass."

Text: 1 "particle physics" [not plural "particles"]

10 "energy" -> "momentum"

10 "... with A large rapidity gap and A large dijet mass."

31 define pseudorapidity on first occurrence: "...barrel (pseudorapidity coverage  $|\eta| < 1.6$ ) and two endcap (coverage  $1.6 < |\eta| < 3.0$ ) sections. [which is a bit awkward so I suggest you may not list these eta ranges at all since they are not relevant for the paper and you do not list the eta range for HF in the following sentence.]

we removed the details of the barrel and endcap coverage.

31 Fix the three "<" symbols.

33 The recommended verbiage is now "Muons are detected in gas-ionization chambers embedded..."

45 "...and is treated as a systematic uncertainty."

69 "single out" -> "identify"

71 define transverse momentum on first occurrence: ...with transverse momentum  $p_T > 20$  (25)..."

71 add space before "(20)" and reduce it before "GeV": "25 (20) GeV"

72 "...and pseudorapidity  $|\eta| < 2.4$  (2.5)..." [assuming you do not mention pseudorapidity on line 31 and thus that the first occurrence would be here]

75 "ET" -> " $p_T$ "

76 Do you really mean "highest momentum", or rather "highest  $p_T$ " ?

77 "...have A large dijet mass .... A large pseudorapidity separation..."

78 "max" should be in roman.

78 You need to define the variables in the  $z_l^*$  definition, viz.,  $\eta_{\ell}$ ,  $\eta_{j1}$ ,  $\eta_{j2}$ ,  $\eta_{jj}$

85 "...an identified hadronically decaying tau lepton with  $p_T$ ..."

88 "nonnegligible" [CMS style]

91 "nonprompt" [no hyphen, here and in numerous other places]

92 remove comma before "are"

101 "charge misidentification" [no hyphen]

115 Use an en-dash "20--40%" [two dashes]

118 "... and, depending on the kinematic region, are found to be as large as 12% for the signal normalization and 20% for..."

121 "...and is treated as a systematic uncertainty of up to ..."

Comments by Jo Cole

---

Abstract:

L3: "... with the CMS detector at the LHC."

L4: should it be "moderate missing transverse momentum"? I note that you say "moderate missing transverse energy" in the introduction (L10), but then later define missing transverse momentum on L66 (which appears to be what you ultimately use?)

Main text:

L1: "... of particle physics"

L9: "Candidate events contain exactly two ..."

L11: "The selection of same-sign lepton events ..."

L14: "Figure 1 shows representative Feynman diagrams for electroweak and quantum chromodynamics (QCD) induced same-sign W boson pair production.

L37: "several" -> "several different"

L39: The definition of QCD needs to appear earlier than this - see L14 comment above.

L39-40: "... via diagrams with two or fewer QCD vertices and up to six electroweak vertices."

L43: "The background from QCD production is small ..." or "This background is small ..."

L44-46: "... and differs from the signal only in terms of kinematics. The interference between the EW and QCD processes is found to be at the level of a few percent in the signal region and is accounted for in the systematic uncertainty."

L48-49: "The WZ and ZZ production processes via qqbar annihilation, and the Wgamma process are all simulated at LO using MADGRAPH 5.2."

L50: "Z+jets" is CMS jargon and should be defined before it is used. Something like "Z production in association with one or more jets (denoted Z+jets)" should appear somewhere. I also note that this is referred to later on as "Drell Yan" (L70). A consistent naming convention should be adopted and used throughout.

L54: "... is simulated with the GEANT4 package [20], based on a detailed description of the"

L57: "... per event included in the"

L69: "The event selection aims to single out ..."

L75: "Events are required to contain at least two jets with ..."

L79, L80, L83 etc.: If you are trying to save space "bottom" could be replace with "b" in all of these lines.

L80-81: "... combine the information from displaced tracks with the information from secondary ..."

L84: "W+jets" - again this is jargon, but also this appears to be missing from the list of simulated samples earlier in the paper ...?

L97-98: "... lepton pair with an invariant mass consistent with that of the Z boson. The background ..."

L100: What do you mean by "Drell-Yan boson events"? remove the word boson. We mean DY events

L103: I don't understand why the sentence beginning "The signal efficiencies are estimated ..." is needed here. It does not seem to fit with the rest of the paragraph. The next sentence starting "In the analysis ..." is also rather awkward.

L107: "within" -> "less than"

L111-112: Presumably the "statistical uncertainty for each process" means MC stats. I think this should be made clear.

L112: "The uncertainty of 2.5% on the integrated"

L117-118: "... factor of two from their nominal value in each event, and are found to be up to 12% depending on the kinematic region used for the signal normalisation and up to 20% for"

L120: Delete "the" at the start of the line.

L121-122: "be small and is accounted for by a systematic uncertainty of up to 4.5%, estimated using the PHANTOM ..."

L122: "A PDF uncertainty on the signal efficiency"

L123-124: "The systematic uncertainties associated with the background normalizations are also taken into account."

L125: "The simulated signal and background yields, as well as the observed data yields ..."

L126: "... separately, along with their sum. The two dominant sources of background"

L131: "with a two-dimensional fit of the  $m_{jj}$  and  $m_{ll}$  variables."

L131-132: "The fit is performed simultaneously in the signal region and in the WZ control region, although in the latter only the  $m_{jj}$  distribution is used. All background contributions can vary" Alternatively, you could use some wording based on what you have used in L165-167 instead.

L135-136: "The best-fit signal strength, defined as the ratio of measured event yield to that expected from the SM, is  $0.90 \pm 0.22$ ."

Table 1 caption: Assuming I have understood correctly, you need to indicate which of these yields are from simulation and which are from data more clearly.

L137: "The cross section is extracted in a fiducial signal region."

L138: "... two same-sign leptons from the W boson decays with ..."

L141: "... the MADGRAPH 5.2 generator"

L146-147: "... while the fraction of events that are outside the fiducial region, but are nonetheless selected at reconstruction level is 20.6 ..."

L149: "the leptonic tau decays reduces the value to 4.9 ..."

L153: "... expected 95% CL limits for the nine"

L157: "Run-I" has not been defined up to this point. This is jargon that should not be used without defining what it means. I would recommend something like: "... reported by CMS using data collected at a centre-of-mass energy of 8 TeV."

L160: "in which the couplings depend on ..." [NB. The sentence that begins "Some of these scenarios on L158 is rather a long sentence. If it could be shortened or divided in two, that would probably be preferable.]

L162: "... VBF production of  $H^{+/-+/-}$  decaying to  $W^{+/-}W^{+/-}$  is directly ..."

Table 2 caption: See earlier comment about Run-I usage. If you wish to use it here, you need to define it where you used it before. The start of the last sentence should also read: "The last column summarises the ..."

L168: "are derived" -> "can be derived"

L173: "Results based on this model ..."

Figure 2 caption: "The normalization of the EW WW signal and background distributions corresponds to that determined in the fit. The hatched bands ..." "... to a cross section of 0.1 pb is shown in on the  $m_{jj}$  distribution (left). The  $m_{ll}$  distribution including aQGCs for the T0 operator is also shown (right). Other operators result in similar distributions when compared at the value that corresponds to the 95% CL limit. In both distributions, the histogram labelled "Others" includes the contributions from QCD WW, ..."

L175: "a first observation" does this mean ATLAS have already seen it? If not, then it should be "the first observation".

L178: See earlier comments about "missing transverse energy"

L179: "The two dominant sources..."

L180: "after the event selection requirements have been applied are nonprompt ..."

## Comments on Paper draft v1 by Francisco (June 28th, 2017)

Hi, here are our latest comments, only one more or less general and a few small details.

The general comment is about our request of quoting systematic errors, we meant it would be nice to have them on table 2. Quoting only MC stat errors seems useless. Any reason preventing to include them there? Maybe it is difficult for every background but you could at least have it for the total background, so the total error has some meaning.

We added the systematic uncertainty to the table for the total yields. The counts per flavor are not used in the analysis and only given as additional information.

the details

L88, remove "to reduce this background", to avoid repetition fixed

L131 my request was to say which is the free parameter, a two dimensional fit to  $m_{jj}$  and  $m_{ll}$  is not enough we added that the signal and WZ normalization are free parameter.

L146 a +- is missing fixed

L150-157 Should explicitly say what you do with the  $W+W+$  expectation in this interpretation we added that the  $ssWW$  signal is treated as background in the BSM search. The new text reads: "The EW production is treated as a background consistent with the SM expectation and can vary within the estimated uncertainties."

Caption figure 2, if BSM figures are kept, the sentences about aQGC need to be rewritten remove the BSM histograms from the plots.

Francisco

## Comments on Paper draft v1 by Francisco (June 26th, 2017)

hi, we will send detailed comments to the draft soon. For now I want to insist on one issue might need some discussion. I am a bit disturbed by figure 2, where you pretend to show at the same time the agreement of the fit for the WW and the apparent absence of  $H_{++}$  or aQGC. I first find confusing that you show higgs in one projection and aQGC in the other. I see that aQGC only use  $m_{ll}$ , but  $H_{++}$  uses both. Even more worrying, you show the WW after the fit, which is fine to justify that the post-fit looks good, but it is not correct (I guess ) for setting limits. I suppose you are setting WW to SM when you set limits, aren't you? I would recommend to remove the BSM from figure 2. I guess it is not mandatory for the limits and your paper (check the title 😊 is about WW's

we removed the BSM signals from the plots in Figure 2.

## Comments on Paper draft v0 (June 13th, 2017)

From Francisco and the ARC

Hi, we found the paper in quite good shape but still have a (longish?) list of comments. I tried to avoid some repetitions, but there might still be some after merging the comments of the 4 ARC. apologies in advance. Tell us if you have question or there is something to discuss.

Francisco

More or less general comments.

\* You need to clarify what you mean by constrain the WZ. This is still not clear neither on lines L142-144 nor on L175-178. You are fitting, rather than constraining. This is not statistically equivalent of doing the two fits separately, and the point has been raised a couple of times before already on the previous drafts.

We tried to be more clear. For explanation here only, the simultaneously fit using the background region constraints the WZ background in the signal region. The sentence did not specify whether or not this is done in sequence or simultaneously. We were stating a fact but did not describe the method. The new sentence reads as follows (~1142) The fit is performed simultaneously in the signal region and the mjj distribution of events in the WZ control region. This procedure constraints the WZ contribution in the signal region. All background contributions can vary within the estimated uncertainties. and (~;175) By using the (mjj, mll) two-dimensional distribution in the signal region and the mjj distribution in the WZ control region simultaneously to discriminate (late sentence removed) .

\* it is not clear which of the various uncertainty sources (if any) are propagated to the MET Removed the sentence 1122. The 1% MET systematics on the acceptance is irrelevant.

\*about renormalization and factorization scales: the (1/2, 2) and (2, 1/2) cases should not be used but the current text is unclear, confirmation that this is just a missing item from the text would be great the case is not used and the text clarifies this now.

Referencing to unpublished physics results as you do in table 2 is not permitted. We believe that [29] and [30] are now published as <http://cds.cern.ch/record/2146658> [http://dx.doi.org/10.1007/JHEP08\(2016\)119](http://dx.doi.org/10.1007/JHEP08(2016)119) <<http://dx.doi.org/10.1007/JHEP08%282016%29119>> <http://cds.cern.ch/record/2245557> <http://dx.doi.org/10.1016/j.physletb.2017.04.071> If not, you have to remove this column. Thanks for catching this. References were updated.

Can you include the full uncertainties quoted for the table of yields and the efficiencies in L155-158. We quote the full uncertainty for the efficiency.

\*don't understand what you refer to about the genuine pmiss and how it is assessed (L121) sentence has been removed. Meant was here events which have MET because a particle (neutrino) carried transverse momentum which was not detected.

\*Figure 2, I forgot the reason if any to show on projection for H++ and another for anomalous coupling, Looks strange, can we put everything together and still manage to see something in the plot? the binning for the double charged Higgs and the aQGC analysis are not the same. In Run-I we received the comment that we should not show the same data twice. We could, if that is the wish, add the aQGC and H++ signal in both plots but it makes it very busy.

Details

L14 and LARGE? Dijet mass? added "large"

L14 Requiring->the requirement of changed

L36 MC simulation tools->MC generators? done

L38-L41, shouldn't these lines go after L58? agreed

L60 the tau decays ... is confusing, I'd say that these e or mu, can be either directly produced from a W decay or from a  $W \rightarrow \tau \rightarrow \text{lepton}$  changed to "The electrons and muons can be directly produced from a  $W$  boson decay or from  $W$  boson with an intermediate  $W$  lepton decay."

L61-62 I suggest to revert the logic, saying that the trigger only rejects a 0.2% of the selected events We prefer to talk about the trigger efficiency here.

L69 top quark means  $t\bar{t}$  or includes also single top? Clarify we mean events with top-quarks including single top ( $tW$ ). More details are given in line 79 (of v0).

L75 leading in  $P_t \rightarrow$  highest momentum jets changed

L76 how do you define  $Z_l$  for two leptons? it's defined per lepton. The requirement is on the maximum value for both leptons.

L84 you mention a loose lepton without mentioning what is a tight lepton. I know you cannot enter into many details into a letter, but should avoid confusion We don't think this needs to be fixed.

L86-88 i think it would be clearer to reverse the order a bit: ... is measured incorrectly, which is only non-negligible for ee. To reduce this background... modified the order and added "The charge confusion in dimuon events is negligible, while this background is non-negligible for dielectron events."

L89-104 i know we asked for this, but can you maybe reduce it a bit and reference to another paper? We referenced the Run-I paper for the method.

L105 remove one with . Need to say why you need a WZ CR done. The name WZ control region explain why we need it - to control the WZ background. We explain later how this region is used in the analysis.

L107 this opposite sign wrong sign is confusing and partially was already mentioned above, can you just say that the background from charge misid is estimated from data, etc? yes, changed.

L113 statistical analysis at this stage sound strange, can just say analysis or statistical interpretation changed.

L113 shape and normalization uncertainties is jargon, I'd say that the uncertainties are calculate bin by bin, including global normalization effect or something similar. I'd include here the comment in L124 about mc stat modified the sentence. Just saying now "In the analysis, systematic uncertainties are incorporated by varying the predicted distribution of a given observable for each source of uncertainty.

L117 you mention several systematics an quote a number, it is the total contribution? this is specified per source, e.g. 2% per lepton .

L134 add a reference and mention this 5% is calculated, not assigned we were saying that 5% is included, not assigned. Changed to calculated and added a reference.

L138 and following, do you need to repeat everywhere  $WZ \rightarrow 3l$ ? I guess not changed

L142 should mention what you fit we do

L147-150 make clear this are cuts on generation added using MC generator quantities Slang?

L153 again it is not taken, it is calculated and much better if you can add a reference changed to calculated

L152-154, we are usually asked by pubcomm to reverse the argument, it is not that we are reproducing theory, but that theory calculations are in agreement with our measurement 😊 agree

Caption figure 2 the use of signal for WW is misleading when you show H<sup>++</sup>. Predicted... corresponding to a fit seems contradictory. I'd say the EW WW contribution is normalized to the fit result. changed

L180 the sentence about the WZ (that you have to change) is repeated twice left unchanged.

L181 for that the H<sup>++</sup>-> in which the H<sup>++</sup> ok

Figure 3, there is a new recommendation

<<https://twiki.cern.ch/twiki/bin/view/CMS/Internal/FigGuidelines>> about brazilian flag plots. I favour the first option, but it is up to you fixed

Title:

- shouldn't 'two jet' and 'two lepton' be plural? No, this is correct.

Fig 1:

- caption: guage -> gauge fixed

- left diagram: I believe the top and bottom quark lines should be q - q - q' and not q - q' - q' yes, yes, yes, fixed. Very good catch - thx.

Table 2 header: Run I limits -> Run I observed limits (to be consistent with the rest of the header) done

Fig 2:

- caption: "VVV processes" not defined since last rephrasing of the paragraph of L53... please state WWW, ZZZ, etc. explicitly instead changed to triboson

- right figure: while the  $\mu$  of 0.61 is indeed a very relevant case (expected limit), why show the same coupling for 0.42? Wouldn't it be more interesting to show another coupling for another value? eg  $\mu$ 1 also shows a factor 10 improvement wrt the previous result... the distributions for the other operators at the respective limits are very similar. We added a statement in the text.

L22: weak -> electroweak removed the weak

L28: transition a bit harsh, please consider some smoothing moved line 8-10 ( $v_0$ ) down.

L42 and 55: please give the exact MG5 version: I believe you mean versions 2.2.2.2 and 2.3.2.2 we added 5.2 as version number in line 54

L42, 54, 55, 151: MG5 and aMCatNLO are one now, they should be quoted as `\textsc{MadGraph5\_aMC@NLO}`, and you can then specify if the program has been used in LO or NLO mode see above. We also write 5.2 in line 151.

L56: I believe the pythia version is more recent than 8.205 if you use > Summer15 CMS samples, please double check double checked.

L58: please consider defining PDF for L134, or write PDF in plain text at L134 done

L60: the sentence is a bit weird... maybe "This includes tau lepton decays (...)" ? change the sentence (see

above)

L81: unclear which algorithm are you talking about, CSV, CMVA, deepCSV? added (CVS)

L82 - related: please quote the Run II BTV performance PAS the algorithm is defined in [27]. We do not see a reason to add a reference to a technical document.

L83: no spaces in " $\sqrt{s}$ +jets" done

L100-104: sentence too long, please consider splitting it sentence is not in the new version

L102: 'tight' is undefined... I believe there has been some rephrasing and you mean "standard-to-loose" ? this part was shortened

L105: with with -> with fixed

L110 and 116: two notations for (presumably) the same thing, and you have been introducing Drell-Yan before. Please consider introducing the DY abbreviation throughout the paper and use it here. Also, if you mean the method is the tag-and-probe, I would suggest naming it explicitly and citing the standard CMS reference. changed to Drell-Yan. Don't like DY

L116: please add the reference to the Run II Electron and Muon performance PAses We don't think it is good style to reference this PAses. Also, there is no benefit.

L120 or somewhere else in the paragraph: it is not clear which of the various uncertainty sources (if any) are propagated to ptmiss, I recall at least the JEC / JER was, but this is currently missing removed the reference to MET. This is not relevant to this paper.

L130 about scales: please specify that the (1/2, 2) and (2, 1/2) cases are not used (and also please confirm this is indeed done for the analysis) these cases are not used.

L134: see L58 defined PDF in line 58 now.

L142-144: please rephrase as the fit is done simultaneously in the signal and WZ regions done.

L152-153: insert a non-breaking space in between 0.21 and fb done

L153: unclear if the 5% should be added on top of the +- 0.21 which already corresponds to 5% fixed

L162: extract the results -> perform the stat analysis ok

L170 179 and 183: please switch to your otherwise consistent notations:  $m(H\mu)$  ->  $m_{\{PH^{\{\mu\}}\}}$  and  $m(H++)$  ->  $m_{\{PH^{\{++\}}\}}$  done

L171: unnecessary 'where  $m_W$  is the W mass' (or if you do consider it necessary, introduce it for every single occurrence for every mass quantity...) changed sentence

L172: unnecessary (vev): not used anywhere as far as I can see removed

L175-178: please rephrase as the fit is done simultaneously in the signal and WZ regions done

L186: please also specify the final state being studied in this summary We explain that we study same-sign Ws and use event with two leptons. I think this is clear.

L190: WZ **process** added process"

L191: same complaint has before: the fit is done in mjj, mll **AND** WZ control region, please rephrase removed the sentence from the summary

L198 onwards: incorrect formatting (copy-pasted line breaks?) fixed

References: the formatting of reference is done with the PRL style. CADI does not pick this up. We tried to integrate the following as well as possible,

[1] and [2]: add spaces in the journal name "Phys. Lett. B" done

[12]: remove the month of publication removed month

[13]: title: GEANT4 casing is incorrect (should be  $\textsc{Geant}$ ) this is the title of the paper and I am following the bibtex derived from the publication.

[13]: 'A' should be part of the journal name, not the journal volume fixed

[23] and [25]:  $k_t \rightarrow k_T$  (corresponding to  $k_T$ , in line with the text) this is the title of the paper. The text follows the CMS guidelines

[26]: 'D' should be part of the journal name, not the journal volume fixed

[33] title:  $\sqrt{s} \rightarrow \sqrt{s}$  fixed

[33]: please add the journal of submission note added

- caption to Fig 1 "gauge couplings **of** electroweak" fixed to of the "

L68 or so: please mention the non-prompt background so that we have all backgrounds introduced in this paragraph. we are listing top quark which is included in the non-prompt background.

L76, 77 " $z_{\ell}$ " instead of " $z_l$ " ? fixed

L105 remove a spurious "with" done

L110-111 perhaps the barrel and endcap should be defined earlier in pseudorapidity range so that this detail (0.01% in the barrel and 0.3% in the endcap) is appreciated better. Otherwise, it may be better to just say "found to be between 0.01% and 0.3%, increasing with  $\eta$ ." introduced in the detector section.

L120-122 I do not recognize this systematics from the past documentation. It does not appear in Table 14 of the ANv7. this is not relevant and removed from the document.

L124 drop "in the yield of each bin and" to get just "The statistical uncertainty for each process is also taken into account." The issue here is that binning is not yet defined. done

L130 and 134 mention uncertainties on signal normalization, which in fact apparently is supposed to mean the uncertainties in "acceptance\*efficiency". Please check and rephrase. In contrast, IIUC, e.g. 20% of the triboson background normalization corresponds to the total number of events. right, re-phrased

On L130 "found to be 12%" is different from the Table 14 in ANv7 which has 3 to 12 %. Is this 12% the final impact factor on the inclusive signal strength or should it be changed to say "up to 12% depending on kinematic region. fixed

Table 1: -(accidental spot-check): WZ total shows up as " $25.12 \pm 0.54$ " in the AN and " $25.1 \pm 0.6$ ". Is there a typo? different number of digits are given but the numbers are consistent.

L154 "measured to \*be\*" ok

Table 2 caption should be above the table. fixed

L198-216 Is there any particular reason that all text is flushed left, compared to the rest of the paper with the text fully justified? The first few lines starting from L198 look particularly odd on the right hand side. Note that this is not the formatting used for submission, but the issue is fixed.

Fig. 1 caption: "quartic gauge couplings electroweak" -> "quartic gauge couplings for electroweak" fixed by for the "

L42/54/55: Same comment as for the PAS already; you refer to the same software in different ways. How about just calling it "MG5\_aMC@NLO" everywhere? In L42 you might add that it was run in leading-order mode. We left this because the community still distinguishes the MG and MG aMC.

L105: "with with" fixed

L105ff: You write one sentence on the definition of a WZ control region and then immediately write about the charge misid background. I had to read it three times before realizing that they have nothing to do with each other. I would suggest to change the start of the second sentence (L107) to read something like: "The contribution of events with charge-misidentified leptons to the signal region...". That should make it clear that you're not talking about WZ anymore. changed this part

L108: "charged misidentification" -> "charge misidentification" (might become obsolete with the previous comment) yes

L131: Suggest "for the WZ background normalization", since you're talking about the physics process, not the event sample. Don't understand this comment. We are talking about the uncertainty of the triboson background normalization.

L132: I find "QCD background processes" a bit confusing. Maybe change the sentence to something like "The interference between the EW induced signal and the QCD induced same-sign W boson production, considered as background, is expected ..." ok

L144: "...using the **\*trilepton\*** control region." fixed.

L145: You haven't defined "signal strength". Maybe just add "..., defined as the ratio of measured and expected cross sections,". This is fairly important, since this is the only place from where a reader can derive the actual measured (non-fiducial) cross section. defined

Tab. 1: Summed totals are customarily at the bottom of a table, so maybe move those rows (and the data row) down. we think it is easier to read the table this way.

L152: make the space non-breaking between number ( $4.25 \pm 0.21$ ) and unit (fb). done

Tab. 2 caption: You haven't defined "BSM" removed BSM

L170: Shouldn't this be " $m(H_{\pm\pm})$ " instead of " $m(H_{\pm})$ "? we now consistently use the underscore notation.

Fig. 2: Indicate in the plot itself that the  $H_{\pm\pm}$  and aQGC signals are normalized to 0.1 pb. this would be wrong and we disagree respectfully

Fig. 2 caption: "with aQGC **\*(right)\*** are shown." Define "DPS" somewhere. added right and writing double parton scattering now

L181: "...parameter space **\*in which\*** the  $H_{\pm\pm}$  " fixed

## Comments on PAS draft v3 (May 9-11th, 2017)

From Matthias

During the approval for SMP-17-004 I noticed that you are still using the Feynman diagrams from the 8TeV analysis. Those are technically correct, but the one with the quartic vertex looks horrible with that displaced vertex / asymmetry. we fixed the quartic vertex diagram

From Slava

Specific comments follow: - abstract "35.9 fb" space missing fixed

Conclusion and abstract should probably provide the value of the measured cross section. we think the cross section value only makes sense with a definition of the fiducial region which seems to be to much information for the abstract.

L1 "of particle physics" fixed

L5-6 "Physics models" fixed

L42-51 may be better to count QCD vertices than  $\alpha_s$  powers. The  $\alpha_s$  power counting is also confusing because as done,  $t_i$  refers to amplitude squared, while a diagram is usually representing the matrix element. fixed

L47 "taken into the backgrounds" -> "considered to be a background" wrote: considered as background. (see above)

L53 "V" is not defined spelled out  $t\bar{t}W$  and  $t\bar{t}Z$

L75 drop "so-called" fixed

L76-80 needs to say "top-quark veto" here, to be able to refer to it in L119 we say used to veto instead of used to discriminate"

L105-108 needs to be rewritten somewhat: the first sentence is talking about charge misid scale factor, while the second about charge misid fraction. There is some disconnect here. %Blue% now writing: The contribution of opposite-sign (wrong-sign) lepton events to the signal region due to charged misidentification is estimated by applying data-to-simulation scale factors to simulated events. The charge-misidentification rates and the scale factors are estimated using  $\$cPZ\$$  boson events. The charge-misidentification rate is found to be between 0.1\% and 0.5\% for electrons, while it is negligible for muons."

L123-124 unclear why there is a new paragraph here. fixed

L141-142: I think we expect to see the expected and measured cross section values here, with syst and stat contributions separated. we only give fiducial cross sections which are give right after the definition of the fiducial region.

L150 vs 153 use inconsistent text for syst/stat (with/without a period). fixed

L151-155 should we add syst uncertainties to the eff\*acc ? we are actually taking about the efficiency here. Changed the sentence.

From Bill

I noticed that a version of the PAS was uploaded today

(May 9). Is that intended to serve as a draft for the paper ? In principle, it should (we don't want people to write a PAS for the sake of a PAS; they should be writing the paper). I see a few small style issues with this PAS (v3) o In the title it doesn't make much sense to state the c.m. energy without stating the process. I would stick with the usual phrase "in proton-proton collisions at 13 TeV" (or " $\sqrt{s}=13$  TeV") o "W boson pairs" and not a hyphen "W-boson pairs" o "doubly charged Higgs bosons" and not a hyphen "doubly-charged" ["doubly is an adverb] o line 13 "missing transverse momentum" [not "energy"] o line 24: "ATLAS and CMS Collaborations" o 65 "\ptmiss", not "\MET" addressed all above

From Francisco

think you should highlight the observation in the abstract, that is somewhat hidden by many details. You start "studying" SS, then claim observation, then describe the sample and finally give your significance. I would change the order to something: "from the study of ... a signal of SS WW is observed with a significance of 5.5 sigma" You don't need to quote the expected significance on the abstract. Then you can say that this result can also be used to establish limits in that or that. going back to a previous version of the abstract which starts with "The observation of electroweak production ..." Adding the expected significance carries the information that the result agrees with the SM expectation.

L20 the sentence about H++ is odd, do you mean of a new resonance such as a H++? now saying explicitly or the existence of a new resonance, such as a doubly charged Higgs boson."

L27 is expected-> WAS expected fixed

L47 I'd add to justify they are background that the contribution is small and the kinematics different added a sentence

L59 as it is the sentence of the tau is wrong. I'd say e or mu produced in the decay of a tau are accepted. fixed

L132 I'd drop "using log-normal distributions". As it does not add information and rather adds confusion fixed

L140 need to explain what you mean by is constrained in bins of mjj changed to as function of . Not sure if this improves the text.

L142 need to add wrt SM? added

Figure 2, the signal lines are barely visible if you could repeat the plots with thicker line much better updated

L11: "...to identify **same-sign** W-boson pair events..."? fixed

Fig. 1: For a paper you'll probably need to increase the size of the labels in these diagrams, or just make them all larger. for now I just adjusted the quartic W vertex diagram which had an extra vertex causing some asymmetry (comment by Matthias).

L20: "...existence of a new resonance, such as a doubly-charged Higgs boson." fixed

L24: "collaborations" (should be plural) fixed

L42/54: You're using the same reference ([14]) for both "MADGRAPH 5.2" and "MG5\_AMC@NLO 2.3". I had also understood you produced the signal samples with MG5\_AMC@NLO, so this is a bit confusing. Is one just the LO mode and the other the NLO mode? But it's the same package, right? right

L64: serial comma: "...photons, and leptons." fixed

L77: "...criteria that combine..." (plural) fixed

L79: "soft muon" (no hyphen) fixed

L80: "bottom quark" (no hyphen) fixed

L108: The charge mis-id probability is much smaller than 0.1% in the bulk of the distributions (e.g. according to Tab. 11 in the note). Either you mean something else by "charge-misidentification fraction", in which case I find it confusing, or you might want to change the numbers to something like "between about 0.01% in the barrel region and about 0.3% in the endcap regions for electrons,..." fixed

L141: Why "about" 5.5 sigma? In the abstract you just write "5.5 sigma". fixed

L141: Need to define signal strength? we added with respect to the SM expectation which should make clear what this is.

L170: Define here the "GM" acronym that you use in the caption of Fig. 3. Either that or spell it out in the caption. fixed in caption

Abstract:

non-breaking space missing between : 13 TeV 35.9 fb should be fine

14: 'experimental signature' repeated twice in two lines removed the first

30 lead tungsten -> lead tungstate (?) I think it both can be used. Tungsten is use in CMS's default detector descriptions. I'd rather leave it because I am sure there have been hours of discussion on this question, which I am not aware of, leading to the wording.

40 leading order (**LO**) (as you use this notation eg in line 52) fixed

paragraph on l52: - V is not defined, WWW and ZZZ backgrounds not mentioned, neither are Wgamma nor W+jets - please use PYTHIA fixed

70: id and isolation requirements should be mentioned as you do refer to these on l89 Added that isolation is required. We can discuss for the paper how much detail (cone size, relative isolation cut, etc) should be give.

103 WZ is normalized in data control region -> incorrect, see comment above I'd say that this is technically correct, but it is a bit confusing because it implies a sequence of normalizations and not a combined fit. In any case, we adjusted the text.

115 contribute 1% -> contribute about 1% (I guess) fixed

135 and the WZ **process** fixed

136 "In order to quantify the significance" -> I note only now that it's not mentioned that there is an excess with respect to the b-only hypothesis, here would be a good place to do so before talking about significance fixed

138-139: "2D fit ; WZ is further constrained" -> incorrect, see comment above fixed

141:  $\sigma$  -> standard deviations fixed

142 what is the expected uncertainty for the best fit signal strength ? 0.22

table 1,  $W\gamma$  in  $\mu\mu$ -: is it exactly 0.0, just a rounding or neglected ? replace by - / rounding

152 non-breaking spaces missing for 20.6 +- 0.3 fixed

154 non breaking spaces missing for 4.9 +- 0.1 fixed

174 WZ is further constrained -> still incorrect, see comment above fixed

185 'extracted using a 2D fit' -> WZ should be mentioned, see comment above fixed

188 what do you mean by 'exotic structure beyond the SM' ? I would suggest to rephrase we mean aQGCs fixed

## Comments from Approval (May 9th,, 2017)

Show that the ratio of the signal plus background yields with the data on every final state is roughly consistent with unity

Below we show those ratios. As one can see all values are consistent with 1 within the 20-30% statistical uncertainty on every individual channel.

	m+m+	e+e+	e+m+	m-m-	e-e-	e-m-	++	--	all
Data	40	14	63	26	10	48	117	84	201
sig+bkg.	44.08 +/- 3.39	18.90 +/- 1.89	67.33 +/- 3.80	23.90 +/- 2.83	11.74 +/- 1.72	38.81 +/- 3.29	130.31 +/- 5.43	74.45 +/- 4.67	204.76 +/- 7.16
Data/(sig+bkg)	0.91	0.74	0.94	1.09	0.85	1.24	0.90	1.13	0.98

Instead of using a 30% fake rate uncertainty, decrease that value to 20-25% and add a shape uncertainty related to the different fake rate measurements as a function of the recoiling jet

An alternative set of fake rate measurements was obtained by varying the recoiling jet pt, which modifies the actual values by up to 15-20%. In the original analysis, we took 20% as a conservative estimation of this uncertainty, while for this cross-check analysis the full propagation of the difference between them was taken as an uncertainty, which produces both normalization and shape dependence on the nonprompt background estimation. Given we are adding that new nuisance, the flat fake rate uncertainty is now taken as 25% (although anything between 20 and 25% would be valid). After repeating the analysis, the actual expected significance is slightly increased (from  $\sim 5.70$  to 5.75) while the uncertainty in the signal strength changes by less than 0.01 units. This is not a surprise since our approach was slightly conservative given that the statistical uncertainties completely dominate the region with high S/B, and given that we know that a variation on the fake rate up to 15-20% doesn't mean such variation at the final analysis level. In other words 25%+ shape uncertainty is more aggressive than 30%. It's possibly more correct, but also more aggressive; and given the statistical uncertainties, it's safer for now to be on the conservative side.

Add the pre-fit and post-fit mll distributions in the WZ region for the aQGC analysis in the AN, and expand the fit explanation.

Done

## Comments from Slava (May 7th,, 2017)

> *\* please provide details on how the efficiency of the 0 missing hit requirement was measured. Was it a tag-and-probe on a single-lepton triggered events? I want to make sure that the value is not biased by the trigger requirements.*

>

> *The efficiency is measured -on top of- the standard tight cut-based electron selection, so it doesn't matter which triggers are used since both leptons must pass the trigger paths anyway. In other words, we apply all previous corrections, and then we look at the triple charge+Nhits=0 requirements MC to data ratios. This is a much easier measurement than the standard tag-and-probe studies because the denominator is already very clean, and pretty much only real electrons are selected.*

I will need a bit more pedantic explanation, as it seems like your answer still suggests that there may be a problem. Here is an exaggerated case: Imagine the trigger requires that electrons have 0 missing pixel hits. In this case you measure higher (if not 100%) efficiency of the 0 missing hit. I note that the trigger efficiency measurement is done after full ID selection in a non-lepton triggered sample. For this exaggerated case you will measure that the trigger has 100% efficiency for the 0 missingpixel hits and we have a problem. The ID efficiency has to be measured unbiased from the trigger efficiency.

The trigger efficiency has been measured not only after the nominal full selection, but also after applying a looser selection, e.g. the standard tight cut-based Id and the HLT-safe Id, and in all cases the trigger efficiency is ~100%. In fact, not very significant differences were found as a function of pt/eta of the leptons in the kinematic regime of the analysis. Therefore, we are sure that it's irrelevant for this study to require or not the trigger selection.

Nevertheless, we have performed the same study, but using only single electron triggers. Although we lose about 10-15% of the di-electron events, the efficiency results are totally consistent w.r.t. our nominal analysis, as shown below:

Efficiency ratio between the new analysis using single lepton trigger paths only and the nominal analysis is shown here

on the response to the Q about prompt contributions to non-prompt: > *Below we provide the contribution from real leptons in the non-prompt estimation. One important point to mention is the fact that the 30% uncertainty considers the whole estimation, and we think the prediction, including the real lepton estimation, is well within that uncertainty.*

The table shows that the prompt part is ~20% of the final non-prompt estimate on average. Looking bin by bin, the value is around or above 5% for bins with electrons. Your argument that 30% systematics covers it does not fully work because you let the statistical analysis to constrain this nuisance by using incorrect assumptions about its correlation (lack of it) with the prompt contribution. Either the correlation should be put back into the stat model or the prompt-related nuisance taken out of profiling.

Our main argument is that what you are talking about is a very small effect. Below we show the nonprompt yields for the nominal implementation, or using the post-fit normalizations for the signal and backgrounds, or for the signal only. As you can see the changes in the yields are well below 1% in all cases. In addition, bear in mind the fit will also make some small corrections which will completely dilute these differences. The fact that the effect is so tiny is easy to understand due to the small changes on the process-by-process normalization (within 10%), and the fact that different processes may compensate each other. Therefore one expects these effects below 2-3% upfront, and they are actually below that.

Type (nonprompt muons)	m+m+	e+e+	e+m+	m-m-	e-e-	e-m-	all
Nominal							

SameSignWWjj13TeVReview < Sandbox < TWiki

	18.38 +/- 3.34	0.00 +/- 0.00	8.92 +/- 2.66	14.18 +/- 2.81	0.00 +/- 0.00	6.87 +/- 2.29	48.35 +/- 5.60
Post fit norm.	18.41 +/- 3.34	0.00 +/- 0.00	8.92 +/- 2.66	14.18 +/- 2.81	0.00 +/- 0.00	6.87 +/- 2.29	48.37 +/- 5.60
Post fit norm signal only	18.42 +/- 3.34	0.00 +/- 0.00	8.94 +/- 2.66	14.19 +/- 2.81	0.00 +/- 0.00	6.88 +/- 2.29	48.43 +/- 5.60

Type (nonprompt electrons)	m+m+	e+e+	e+m+	m-m-	e-e-	e-m-	all
Nominal	0.00 +/- 0.00	5.53 +/- 1.68	15.93 +/- 2.48	0.00 +/- 0.00	5.00 +/- 1.59	12.93 +/- 2.18	39.39 +/- 4.03
Post fit norm.	0.00 +/- 0.00	5.61 +/- 1.68	16.05 +/- 2.48	0.00 +/- 0.00	5.05 +/- 1.58	13.01 +/- 2.18	39.72 +/- 4.03
Post fit norm signal only	0.00 +/- 0.00	5.56 +/- 1.68	15.94 +/- 2.48	0.00 +/- 0.00	5.03 +/- 1.58	12.98 +/- 2.18	39.51 +/- 4.03

As a cross-checked, we have built a new model where we add a correlated uncertainty between a given process and the nonprompt component. The idea is to add a new nuisance per process to correlate the given process and the nonprompt background. For every process we know the overall uncertainty, we know how much the contribute to the nonprompt background estimation, we know the total amount of nonprompt background, and the yield for the given process. With that information in hands, one can compute the uncertainty. Let us put an example. Imagine one bin with 20 EWK events, 10 nonprompt estimated events with 1 event from the EWK component, and with a EWK uncertainty of 20%. Therefore, in that bin we have a  $\ln N$  uncertainty of  $1+0.2*1/10=1.02$  related to the nonprompt component, and  $1/(1+0.2*1/20)=0.99$  related to the EWK component. Why do I put 1/... for the EWK component? The reason is that it's a negative contribution to the nonprompt component, and therefore the signal and nonprompt estimations are anticorrelated, the larger nonprompt component the smaller EWK component, and viceversa. The effect on the performance is tiny, as we already anticipated by the previous tests shown above.

> \* WZ: please add yields with corresponding uncertainties to table 12 to be able to understand the ratios already provided.

>

> Data, WZ, and background yields have been added to Table 12.

I found the entries in the version in svn, but I did not understand the entries. E.g. in 500-800

"WZ scale factor" Data WZ Backgrounds 0.82 45 30.5 15.5 Are the last columns for expected WZ and backgrounds? If it were the case,  $(45-15.5)/30.5 \sim 0.97$  doesn't match 0.82.

We were showing the observed WZ post-fit normalization. We have now switched to the expected (pre-fit) WZ normalization, and we have added "Expected" on the table to make it more clear.

> \* there is a fairly significant dependence of WZ scale factor bin-to-bin in  $m_{jj}$ . ...

My concern was about the residual effect on the effective [same-sign]/[3 lepton] ratio that you are relying in each bin to constrain the SS yields by the 3L observations. The original question was not clear about it. The systematics table is somewhat opaque about it.

Yes, all systematic uncertainties are considered on the ratios, it's not just a single number being translated from the control to the signal region, but other systematic uncertainties are included in the fit.

> \* looking at Fig14, I see that non-prompt is significantly constrained. ...

Your response suggests that this can be ignored.

The fake rate systematics cannot be ignored when quoting the full significance. Nevertheless, the expected and observed signal strength remain unchanged. Furthermore, the nonprompt yields are just slightly modified well within 30%.

> \* *wrong sign and  $W_{\gamma}$  apparently have the same systematics for the lepton part (efficiency and trigger) as normal leptons. ...*

Your response for  $Z \rightarrow ee$  is correct for efficiencies (I have now written down the conditional efficiencies and scale factors used and see no grounds for the question on the topics of the efficiency scale factors). The point still remains for the momentum resolution, but this is now too insignificant to matter. I guess the remaining issue for  $W_{\gamma}$  is shuffled under the 30-90% statistical uncertainty. ~OK.

Regarding this point, we have studied the  $V_{\gamma}$  background in  $ll_{\gamma}$  events with a mass consistent with a Z boson, and we find a reasonable agreement. Indeed, the theoretical uncertainties are as large as 20% for this small background, and the statistical uncertainties are rather large, so any efficiency variation plays little role.

> *Limits on anomalous couplings \* please clarify how contribution of anomalous couplings on WZ is treated.*  
> *A: For the aQGC limits, we use the mll distribution, both in the signal region and in the WZ region.*

This did not come clear from the AN; it is more evident from the PAS draft. Please add some text in the AN and also provide a table similar to table 12. Similar questions on systematics of relating 3-lepton control region yields to the SS signal region apply here as well.

Yes, we will try to make it clear in the PAS,

## Comments from Francisco (May 3rd., 2017)

\* please quantify the efficiency of the electron triple charge requirement (I recall that it is 1-2%, which is not so small to ignore)

The efficiency is about 97%. Nevertheless, this is not neglected, and the MC to data scale factors take into account both the triple charge requirement and the 0 missing hit requirement. We mention this point in Section 4.3.

\* please provide details on how the efficiency of the 0 missing hit requirement was measured. Was it a tag-and-probe on a single-lepton triggered events? I want to make sure that the value is not biased by the trigger requirements.

The efficiency is measured -on top of- the standard tight cut-based electron selection, so it doesn't matter which triggers are used since both leptons must pass the trigger paths anyway. In other words, we apply all previous corrections, and then we look at the triple charge+Nhits=0 requirements MC to data ratios. This is a much easier measurement than the standard tag-and-probe studies because the denominator is already very clean, and pretty much only real electrons are selected.

\* was the bad muon filter applied in this analysis?

Yes, although the effect is pretty much null given the additional muon veto in the analysis

\* eq (4) and (5) apparently have a typo: the double-fake term is added to the total the way it's written

Yes, this was fixed in the latest AN version uploaded to iCMS.

\* please provide absolute values of the contributions to the terms in non-prompt prediction. My main target is to know the size of the MC-based contributions. If I recall correctly, you did not include common nuisances on these terms together with the contributions of this MC (including signal) to the statistical analysis. The only justification for this is that the MC contribution in non-prompt is really small.

Below we provide the contribution from real leptons in the non-prompt estimation. One important point to mention is the fact that the 30% uncertainty considers the whole estimation, and we think the prediction, including the real lepton estimation, is well within that uncertainty.

m+m+	e+e+	e+m+	m-m-	e-e-	e-m-	all
0.8 +/- 0.1	4.2 +/- 0.3	7.6 +/- 0.4	0.3 +/- 0.1	2.3 +/- 0.2	5.0 +/- 0.4	20.2 +/- 0.7

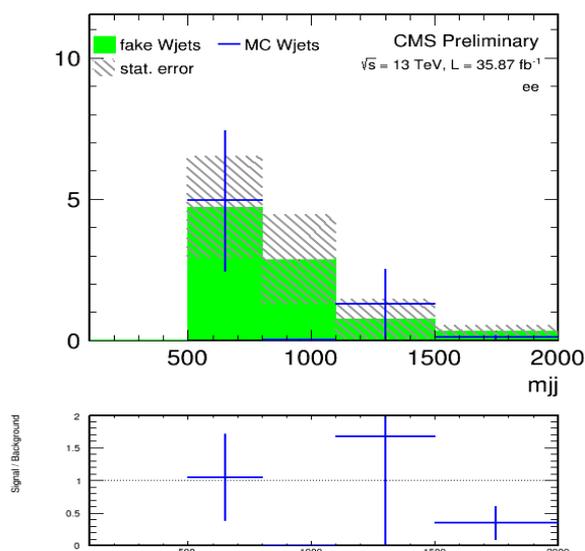
\* it would be nice to see the comparison between the two methods of estimating non-prompt split by channel. The variation of 4.5 on essentially the same inputs seems fairly large. Does it stand out in some specific channel?

Below we provide the yields for the six final states and the total for both methods. In any case, bear in mind the inputs and assumptions are rather different with the exception of the lepton fake rates. There are clearly pros and cons between the two methods, but we personally trust more the default method, we think the discussion about the two methods is a bit outside the scope of this single analysis. One should also take into account the relatively large statistical uncertainties.

Method	m+m+	e+e+	e+m+	m-m-	e-e-	e-m-	all
Alternative	18.37 +/- 3.56	7.43 +/- 2.14	25.51 +/- 4.36	14.30 +/- 2.99	5.45 +/- 2.08	20.67 +/- 3.79	91.74 +/- 7.99
Default	18.38 +/- 3.34	5.56 +/- 1.68	24.86 +/- 3.64	14.18 +/- 2.81	5.03 +/- 1.58	19.85 +/- 3.16	87.86 +/- 6.90

\* have you tested the non-prompt estimation closure on W+jets events? I recall from other analyses that it didn't close equally well for both. You have fairly good closure in ttbar.

- Below we provide a rough closure test on W+jets events. The available sample is rather small, but this is what we (quickly) get.



\* wrong-sign scale factor changes by 22% in going from 0.5-1 to 1-1.5 bin. Have you accounted for possible eta distribution systematics? It may be more appropriate to split up this eta bin in two.

We didn't, but we checked that the net effect was very small. The wrong sign efficiency is rather flat between 0 and 1, and it's again relatively flat between 1.5 and 2.5. Nevertheless, there is indeed a strong dependence between 1.0 and 1.5, which is not surprising since it's the barrel-endcap transition region. We tested in earlier versions of the analysis splitting the region 1.0-1.5 in two bins, and the difference with the nominal analysis was minimal. One point to consider is that the probability of having a selected electron in that region is rather small, and that's why a further splitting doesn't make a difference.

\* please clarify on the size of the  $Z \rightarrow ee$  sample used to extract the wrong-sign. The statistical uncertainties are very large. Is the full  $Z \rightarrow ee$  sample used or is this prescaled or is it a feature of the binned extraction?

We use the full dataset. The point is that the number of wrong sign events is rather small, both in data and MC and in particular in the central region, and that's why the overall uncertainty is relatively large. You can see that more than 1/2 of the uncertainty is actually from the available MC statistics, and we use the largest DY MC sample.

\* WZ: please add yields with corresponding uncertainties to table 12 to be able to understand the ratios already provided.

Data, WZ, and background yields have been added to Table 12.

On statistical analysis \* there is a fairly significant dependence of WZ scale factor bin-to-bin in  $m_{jj}$ . This suggests a jet or QCD scale dependence. If I understand correctly, you keep the WZ contributions in  $m_{jj}$  bins uncorrelated, which would miss to appropriately pass the dependence observed in Table 12. I think that one way out is to vary scales in MC used in estimation of WZ scale factor and add it as extra uncertainty.

First of all, the scale factors are all consistent with 1 within the statistical uncertainties, but the numerical values vary a bit. Indeed, we uncorrelate the bin because we don't want to add theoretical uncertainties. Another way to do the analysis, but it was decided several months ago not to do so, was to consider the theoretical uncertainties and making use of a single WZ scale factor. The net uncertainty will be similar. We don't think considering both theoretical and data approaches at the same time is correct, although both versions separately are valid in our opinion.

\* looking at Fig14, I see that non-prompt is significantly constrained. If I understand correctly, it's coming from bins with low  $m_{jj}$  and  $m_{ll}$  where the non-prompt is significant. As you mentioned in the description of the non-prompt estimates, the estimate depends on the underlying loose object jet spectrum. So, it's quite possible that non-prompt in low  $m_{jj}$  is systematically different from that at high  $m_{jj}$  where the signal is. This should be treated appropriately by introducing an independent nuisance per  $m_{jj}$  bin or having something similar for the non-prompt contribution.

Given that the statistical uncertainties are very large in the non-prompt estimation at high  $m_{jj}$ , we don't think this is necessary. We thought about it sometime ago, but given the available statistical precision, it won't add anything.

systematics: \* it would be good to show a table of systematics the way it contributes to the measured cross-section, instead of the way it contributes to each component.

The impact plots are exactly what you are asking for, we will have it in the approval talk.

\* wrong sign and  $W\gamma$  apparently have the same systematics for the lepton part (efficiency and trigger) as normal leptons. Have you validated this (at least in MC)? The wrong sign is more likely to be in showering or large brems electrons, which do not have the same efficiency as the regular electrons. Also, the  $\gamma \rightarrow e$  would appear from an asymmetric conversion, which also might not have the same object ID and reconstruction performance.

The trigger efficiency has explicitly checked that is consistent for all the processes. The point is that the trigger efficiency is considered after requiring two leptons in the event, and therefore it's quasi-irrelevant the underlying physics process. The wrong-sign efficiency ratio between data and simulation has been explicitly measured in  $Z \rightarrow ee$  events, and the associated uncertainty is part of the analysis. Indeed, there is an eta dependence for instance in wrong-sign events, but the  $Z \rightarrow ee$  studies were exactly made to study those effects. The gamma background has been checked by looking at  $Z \rightarrow \gamma\gamma$  events, although not an extensive study has been done since it's a rather small background after all with large statistical and theoretical uncertainties.

Cross section results: \* I think that we need a measured cross section value based on main analysis selections. Currently only the fiducial cross-section is spelled out.

We can quote  $\mu \cdot \sigma_{\text{theory}}$  without any fiducial definition, where  $\sigma_{\text{theory}}$  is the cross section from the full produced MC sample.

Limits on anomalous couplings \* please clarify how contribution of anomalous couplings on WZ is treated. In the last meeting someone mentioned that this contribution will self-correct via the WZ constraint from data. Since the WZ is constrained from data only in  $m_{jj}$  bins, the possible anomalous coupling signal will be diluted and WZ estimate will be dominated by the SM. If the sign of WZ contribution is positive, then the limits are weaker, which is probably OK. On the other hand, if the anomalous coupling drives WZ in opposite phase, the limits may need to be adjusted.

For the aQGC limits, we use the  $m_{ll}$  distribution, both in the signal region and in the WZ region. This means any possible signal in the WZ region will be completely diluted, as intended.

\* Why results are not compared to 8 TeV. It is true that it is a different  $\sqrt{s}$ , but could gain significance in terms of normalized signal strength

We can show the 8 TeV results in the approval, although the conditions are very different, it's not clear what we could learn. We have  $\times 2$  larger integrated luminosity and the cross section is more than  $\times 2.5$  larger, while the background is about  $\times 2$  larger. With those very rough numbers, one goes from  $\sim 3\sigma$  to  $\sim 5\sigma$  significance. The 2D fit gives another  $\sim 0.5$ - $1\sigma$  significance more.

the following are related to the PAS.

\* table 1, are you using SM or postfit cross-section? the numbers are pre-fit

\* how do you define the jets at GEN for your fiducial cuts? same anti-kt algorithm as in reconstruction

\* L145, where this 5% comes from exactly? standard variation of scales.

\* L157-158, how the effect of signal was accounted for? I remember you explained that but cannot find in the PAS nor AN we discussed this in the meeting. Since WZ is normalized from data, any effect from aQGC on the WZ yield and distribution is absorbed.

\* hadn't realized you had on table 2 some asymmetric limits, are you really affected by the sign of the  $f$ ? I would say you only are sensible to  $f^2$  and this difference is fake. BTW some of the references point to AN, aren't these yet published? will they be published by the time this paper is submitted? You know we cannot refer to unpublished physics results. there is a small non-zero interference between SM and aQGC which is sensitive to the sign. For consistency with the Run-I result we prefer to keep the table as is.

\* figure 2 left is very different from the equivalent in 8 TeV analysis (not surprisingly also fig2 left :-). Do we know why? The analysis is not that different. Don't understand your labeling signal region and aqgc signal region and some general comments on the current PAS this label should go. Will be fixed in later version.

\*As already said, by the time it is made public it is desirable you do not rely so much on the previous paper, you should rewrite many parts from scratch, rather than editing the previous text apologize again. We made some changes and will further work on this point towards the publication.

\* L19-22 need to be better explained changed the text a bit, but the comment is rather vague. Is the new version ok? "An excess of events could signal the presence of anomalous quartic gauge couplings (aQGC)~\cite{aqgc\_operators} or the existence of a new resonance. A new resonance could be a doubly-charged Higgs boson. These particles are predicted in Higgs sectors beyond the SM where weak isotriplet scalars are included~\cite{CE1, CE2}. They can be produced via weak vector-boson fusion (VBF) and decay to pairs of same-sign  $\mathcal{P}W$  bosons~\cite{1202.2014}."

\* L48 mention explicitly how interference is treated added more detail

\* L56-58 explain briefly which triggers are you using rather than speaking about a "suite" more detail as saying that we are using single and double lepton trigger does not seem to be necessary in a letter.

\* L110-111 need to better explain how WZ is constrained or refer to the fit we do this later in the document

Abstract and Summary miss a mention of what the main discriminant variables are. Maybe also a mention of the main backgrounds (à la line 105) The discriminant variable and the main backgrounds are now mentioned in the summary. We do not think that this level of detail is adequate for the abstract.

8 This Letter -> this document (for now it is a PAS) fixed

37-39 The average number (...) approx 20 (...) are included -> A number (...) approx 20 (...) are included changed to "Simultaneous proton-proton interactions overlapping with the event of interest, are included in the simulated samples. Approximately 20 additional  $\mathcal{P}p\mathcal{P}p$  interactions were observed in the 13 TeV data."

40-54 inconsistent naming of generators, please use  $\mathcal{MADGRAPH}$ , etc. replaced

56-57 A suite of triggers (...) These single and double lepton triggers -> please rephrase rephrased

72 citation for Zeppenfeld variable ? added reference to Phys. Rev. D54:6680-6689,1996

99 it sounds like WZ is normalized in data on one hand, and the fit is performed on the other hand, while the control region is included in the final fit, please rephrase the control region is indeed used in the fit. Not changed for know because I will need to rethink how to write this part coherently.

102 charge misidentification: somewhat a duplicate of line 81 ? line 81 motivates the cuts while 102 discussed how we estimate the background.

104 the expected signal & bkg **and observed data** yields fixed "The expected signal and background yields, as well as the observed data yields, are shown in "

108-111 again unclear that WZ is included in the fit, this part reads as if it's an independent control region added "The  $\mathcal{P}W\mathcal{C}PZ$  background contribution in the signal region is further constrained in bins of  $m_{jj}$  using the control region."

111 please consider adding a figure with the final discriminant ! We will add the 2d distribution in the section for additional material

118 and 121: the current phrasing is a bit confusing, some uncertainties are 'within 2%' or 'up to 7%', while lepton momentum scale and genuine MET seems to be exactly 1%. I believe this is taken as an estimate as well, and not the exact value yielded by the study ? I added about to 1%

125-127 non prompt sources seem somewhat different than on line 95, please rephrase removed the sources in 125-127 because of 95. Line 125-127 is using a different classification of the sources but the sources are the same.

127 pileup uncertainty is not mentioned, and a rough mention of the size of stat uncertainties could be interesting we are varying the minbias cross section by 5% as recommended by the lumi group. The impact on the analysis is minor.

132-133 I believe some additional explanation of why the interference is considered as uncertainty would be worthwhile, sth like the first paragraph of the AN section 9.7.1 could be enough (two sentences, so rather short) do we want to mention Phantom and give a reference?

137 the mjj **and** mll (...) are shown fixed

145 I believe the order of the calculation is LO, but it would probably be best to state it explicitly done

148 reconstruction analysis level -> pick one 😊 reconstruction it is

149-150 otherwise this value is reduced to -> I don't get this comment add Including the leptonic  $\tau$  decays the value is reduced to  $4.9 \pm 0.1 \sim (stat.)\%$ ."

166-167 2D distribution as discriminant variables: again unclear if WZ is in the fit or not added that WZ is constrained here too.

172 Results on this model 'have' also been reported fixed

table 1: - bkg processes contributing less than 1% -> these are accounted in the final yield it seems, please consider specifying it in the caption added

- charge mid-id for muons: if this is truly neglected I would prefer not showing 0.0 +- 0.0 in the table, but '< sth' or simply '-' using '-'

fig 2: - these are post-fit figures I presume ? please specify this in the caption added

- uncertainty is missing from the legend will be done

- signal cross section is also missing from the plot the aQGC normalization is fixed by the parameter choice. The H++ normalization is given in the caption. We think that is sufficient, but do not feel strong about this point.

fig 3 - is the y-axis sH or sH<sup>2</sup> ? if this is sH, is there a reason not to show (or at least mention in the caption) the negative side of the sine ? the plot shows sH

## Comments from Francisco (Apr. 19, 2017)

Three items to be explained in some slides:

- non-prompt background, how it is calculated and how the errors are estimated
- lepton ID, please summarize the additional checks performed
- how background normalizations are included in the fit, in particular the bin-dependent WZ

We will do it

table 16, I'm not sure I understand this test. Can you explain what you do and what you expect to get?

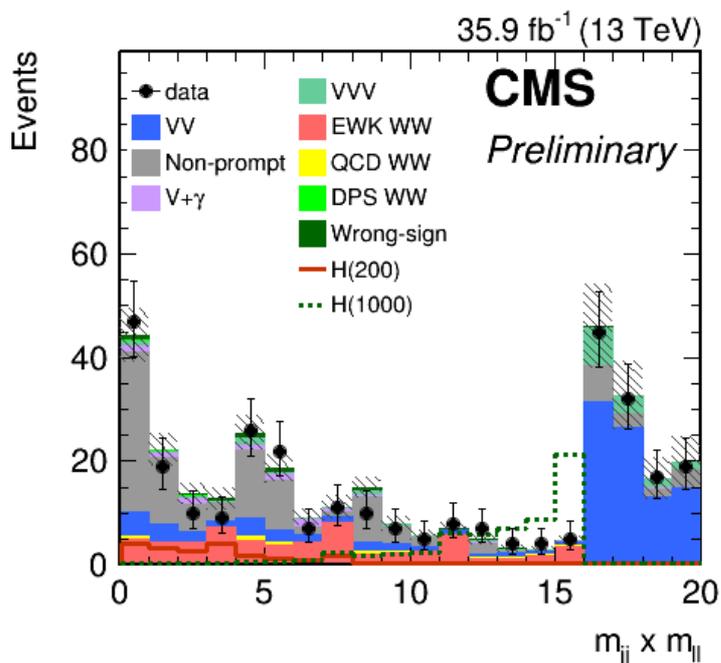
This is not a test, but these are the actual limits assuming no EWK process is present in the data. Indeed, it's obvious there is a significant excess of events (more than 5sigma!), but we wanted to show the effect on the upper limits. Those results won't be made public, they don't tell us anything once there is such a significant excess.

is fig 18, correct? Cannot understand the concentration of VV in the last 4 bins, does not seem to match fig 17. What  $m_{jj}$ - $m_{ll}$  range it corresponds to? If it is correct, you have an extremely good CR for VV!

No sorry, it's true that Figure 18 deserves some better explanation. That figure shows the actual bins used in the statistical analysis. In short, we have 16 bins in  $m_{jj} \times m_{ll}$ , and then there are 4 bins in  $m_{jj}$  in the WZ control region. Just for illustration purposes we put all together, maybe a vertical line between bins 16 and 17 would make it more clear. Regarding the matching with other figures, those last 4 bins should exactly match with Figure 7 top right. We have tried to improve the explanation in the AN.

not anything worrying, since you are withing  $\pm 1$  sigma, but i'm curious why you seem to point towards excess at high masses and defect at low masses. Do you know why? Table 18, doesn't help and it is hard to see from the scale in fig 18 (and the higher mass is 600). Can you produce (in addition to fig 18) another figure with more masses and with a scale we can more easily compare to data?

It's just what we see in the data, and you can see it all looking at Fig18. Bins 2-4 have a small deficit, and those are the driven bins for the low mass Higgs ( $\sim 200$ ). On the other hand, there are a few more events than expectations in bins 12 and 16 (high  $m_{ll}$  and  $m_{jj}$ ), which makes that little excess for high  $H^+$  masses. The plot below includes  $m_H=200$  and 1000 GeV, in both cases the signal is normalized to 1pb:



for the limits on  $H^{++}$  and anomalous couplings, how you introduce the  $WW^{++/--}$ ? You assume SM or your measurement? Not sure which is the most correct way, but you need to clarify and justify. Do you include additional errors?

In this case we are assuming that the EWK process exists, and it's just another background assuming the SM result. Since the signal strength measurement already includes the QCD scales and PDF uncertainties, no additional uncertainty is needed in the EWK process. There are obviously uncertainties related to the Higgs samples. We totally agree, this procedure needs to be spelled out clearly in the public documentation.

## Comments from Slava (Apr. 12, 2017)

- ANv3 table 8 has electron fake rates going down by a factor of 2 in  $\eta > 2.0$ . Please check that there is no trigger bias (most worrying would be if the  $\eta < 2.0$  is biased: if your FR estimates are pushed up by restricted trigger in the denominator and loose electrons in the signal sample come from single-lepton trigger, the bulk of fakes is overestimated).

The reason of the lower fake rate at high  $\eta$  is the effect of the tighter electron identification, and not the trigger efficiency. The trigger efficiency is very high, even for the fakeable object lepton selection, this is something we have explicitly measured. We have a decrease both of the electron selection efficiency and the electron fake rate at high  $\eta$  due to the specific tighter electron requirements (obviously, the electron efficiency decreases is smaller than the effect on the fakes). To give you an idea, below we quote the electron fake rates applying the tight cut-based Id -but- without including our tighter specific selection. You can see the clear effect of our selection in reducing fakes when comparing with our nominal results in the AN.

eta	20<pt<25	25<pt<30	30<pt<40	40<pt
0.0<eta<0.5	0.368	0.326	0.324	0.364
0.5<eta<1.0	0.366	0.345	0.327	0.378
1.0<eta<1.5	0.395	0.380	0.394	0.426
1.5<eta<2.0	0.376	0.351	0.353	0.346
2.0<eta<2.5	0.313	0.300	0.301	0.300

- table 14 vs table 20: why did fake ele estimate go up? Is the table after a fit or something?

In Table 14 we indicate that non-prompt(e) yields are ~40 events (nominal selection), while in Table 20 we indicate that non-prompt(e) yields are ~36 events (tighter muon selection). So, the electron fake rates are unchanged between both selections, but the muon selection efficiency goes down. Therefore, we lose events in the mu-electron category where the muon is a real muon and the electron is a fake.

I'm somewhat OK with the check with muon  $\text{reIso} < 0.1$ . The naive significance ( $S/\sqrt{B}$ ) is up by 5% in the last bin. It seems like The limits on NP will be affected a bit more.

As also described in that appendix, we tested a tighter selection for the aQGCs by requiring  $m_{jj} > 800 \text{ GeV}$ , and the expected limits were pretty much identical. For the  $H^{++}$  searches, using a tight muon isolation, the expected limits didn't change either (e.g. for  $m_H = 800 \text{ GeV}$  the expected cross section limit is 27fb in both cases)

## Comments from Benjamin (Apr. 12, 2017)

- For the coincidence of the EWK and QCD cross sections, I guess my question is what the source for these numbers is. Did you calculate them yourself with aMC@NLO with the processes you describe in 2.3.1 and 2.3.2? It's just that seeing them so close numerically makes me wonder if there is some significant overlap between the two processes that might have been missed somehow? This is somewhat important for your result (i.e. the normalization of your expected signal), so maybe you could provide a reference for what you use.

The described samples in sections 2.3 were officially all produced. The cross section values were obtained by several people using the GenXSecAnalyzer scripts, and they agree with the quoted numbers on McM. We are sure there is no overlap between both processes, and we know the cross section should be reasonable similar with the loose applied generator level selection. Technically speaking, we could send you the very simple scripts to reproduce the exact numbers, they can actually be used for any dataset available on DBS.

- For the non-prompt backgrounds, I think 30% plus the bin-by-bin statistical uncertainty is perfectly fine, but please document this in the note. It's your main systematic uncertainty.

Yes, we agree. We have tried to make it more clear in the AN.

- For the WZ scale factors, I'm still a bit confused that the scale factors from the fit are different for each bin. Does it mean you fit each bin separately? I somehow expected you to float the WZ normalization for all bins coherently in the fit. Or do I just misunderstand what you're doing?

It's not a different fit, but there are four different normalization factors for each mjj region. Using the "combine" language, there are four different rateParam nuisance, once for each mjj region. The reason to use different normalization factors is that in that way we avoid any theoretical uncertainty related to the mjj shape. In particular, we have WZ contributions from EWK and QCD processes. Since we want to avoid those issues, we simply measure in data each normalization factor. This approach was discussed in length in the SMP-VV group, and it was agreed it was the most systematic free way to do it.

## Comments from Benjamin (Apr. 11, 2017)

- Table 2: The cross sections for EWK and QCD ssWW reported here are almost exactly the same; 26 vs 27 fb. Is this just coincidence? Or something that can be easily understood?

The main reason why the cross sections are so similar is that the number of possible diagrams is very similar, but the actual ratios are a bit random. If we would apply looser requirements on the signal samples, we had different ratios. Notice that this is the only case where the initial cross sections are similar, for all other processes the QCD component is much larger. And this is one of the reasons why this is a golden mode to find the vector boson scattering. We invite you to look at Fig.8 in <https://arxiv.org/abs/1311.6738>. As you can see the looser the requirements are, the larger QCD component is. On the other hand, it's still true that the QCD and EWK components have similar sizes (for instance the QCD component is about 100 times larger with no selection in processes like Z or W, this is not the case for the same-sign WW).

- Signal selection (Sec. 5): Was there any attempt to optimize this selection? Or, in other words, where do these exact cuts come from?

The basic selection comes from the run-I analysis. Then, we try to optimize the final kinematic variables using ttbar MC events. Notice that we don't want to over-optimize the selection given the limited amount of (data-driven/MC) events at the final level for the main background processes. We should mention that the selection didn't change much after the optimization in any case since the run-I choices were already already very reasonable.

- WZ control region selection: Why do you use the tight CSV point for vetoing bjets in this selection? And why do you use a lower jet pt cut at 20 GeV here? I ask mostly because the two changes seem contradictory; one makes the selection tighter (vetoing with 20 GeV rather than 30 GeV), the other makes it looser (vetoing with tight tags rather than medium tags).

The main reason is that we want to reject a real b-jets, that's why a tight WP is the most suitable option. On the other hand, for a b-jet, a looser pt option makes sense, that's why we have chosen that approach. Notice that we follow the logic from SMP-16-002 (WZ inclusive cross section) where some of us were involved.

- Systematics on non-prompt backgrounds: You are being quite ambiguous about what systematic uncertainty is assigned to this background. The only statement I could find is in L386, in a sentence that must have been mangled a bit in editing. Can you clarify what exact uncertainties are being assigned and where they come from? This ends up being the nuisance parameter with by far the largest impact in your fit, so it would be good to be a bit more specific about it.

We have two components: (a) bin-by-bin uncertainty, (b) systematic uncertainty. Notice that (a) is rather important at this point, while for (b) we assign 30%, that's it. Indeed, this is coming from our previous knowledge together with the studies we showed in the AN. In principle, we think 30% -per lepton type- is

rather reasonable, but since we don't split in lepton flavors in the latest version of the analysis, we assign a total uncertainty of 30%, which is more conservative.

As a test, we have split the uncertainty in two terms, 1/2 for electrons and 1/2 for muons (i.e. two uncertainties with 21%), and the expected sensitivity is pretty much identical. Of course that approach would just make the impact of each one smaller, but no difference on the overall result. On the other hand, if we assign 30% uncertainty for muons and electrons separately, this means the uncertainties are about 15% each (since about 1/2 of the fakes are muons and 1/2 electrons). With this approach, the expected significance is increased by about 2%, while the signal strength uncertainty is reduced by 1%. In other words, the improvement is rather minor.

- Prompt leptons failing the lepton selection: There is a very short statement (L. 357) about a weight being subtracted from simulation. I didn't understand what you are doing here, can you expand a bit?

Yes, this was also asked by other ARC members. We have added the explicit formula in the current version of the AN in Section 6.5. That's the exact formula we use to take into account double fake and dileptons.

- Charge misID background: I'm a bit confused about what is being done here. What are the scale factors used for exactly? You write that you apply them to simulated opposite-sign ee events, but that does not give you a prediction for same-sign events. Is there a typo somewhere or do you do something else? Also, the probability is not the number of same-sign events divided by the number of opposite-sign events, it is (approximately) equal to the number of same-sign events divided by twice the number of opposite-sign events. (In the scale factor this would cancel out, so I'm not sure this would actually affect your prediction or not.) Please clarify.

Apologize, this was also spotted by other ARC members somehow. The explanation given in the AN was not right. Let us summarize here too for clarity. First of all, the charge mis-ID rate is measured in 5 eta bins taking into account all combinations, i.e. there are 25 different eta1-eta2 regions and 5 charge mis-ID rates to measure. A  $\chi^2$  fit is performed to obtain the charge mis-ID rates in data and simulation. With those numbers obtained in data and simulation, we get some scale factors. Then, those scale factors are applied to simulated events passing all of the selection requirements to obtain the final prediction in the signal region.

- WZ scale factors: Where do the scale factors come from? It looks like it's just data/MC from Fig. 5 top left, but shouldn't you subtract the non-WZ contributions before doing the ratio? Then, do you apply these scale factors bin-by-bin to WZ before floating it in the fit, to arrive at Fig. 5 top right?

The scale factors are coming from the full fit of the signal and WZ control region, although the WZ region has the largest weight to get them by far. Indeed, the fit is taking care of 'subtracting' the background for each region, this is something we have checked ourselves 'by hand'. What you see in Fig.5 right is the post-fit result. In this case the agreement between the data and the prediction is so good because the background level is so small that the WZ scale factors are obtained rather precisely. Notice that we don't apply those scale factors to the fit, but we are using them in the final fit. In other words, you are already seeing the post-fit distribution of the full fit for the WZ control region.

## Comments from Olivier (Apr. 10, 2017)

General comments: - you should show the final discriminant used for the observation analysis... you only show the two 1D distributions, the unfolded 2D  $m_{jj} \times m_{ll}$  distribution should be at the very least in the AN (but would also be worth adding in the PAS itself IMO)

We have added the unrolled distribution to the AN. We were already planning to do so, but for the unblinding.

- I'm a bit surprised you don't quote anywhere the SM Higgs background, not even to mention it is negligible... at the very least something should be said in the AN about VBF H(WW), which would be irreducible modulo a charge misID with the current selection

The real "irreducible" background is the EWK WW opposite-sign signal, which is already considered as part of the wrong-sign component. Obviously, the contribution is tiny due to the triple charge requirement. The VBF H(WW) component is actually very small due to the triple charge requirement, but also due to the high lepton pt requirements. Furthermore, the dilepton mass is used in the fit, and hence the H(WW) component gets completely de-weighted. In fact, there could be some contribution from VH and ttH processes, although it's also tiny due to the relatively tight VBS selection. We have added a small sentence about it in Section 2.

- the data/MC plots in general are almost unreadable in B&W... please improve the colors

We suggest to discuss it for the figures in the paper.

- some clarification on the systematics assigned to the determination of the non-prompt background is needed (see below)

OK.

Line by line: 148: could you detail how do you compute the trigger efficiencies ? Are you impacted by the inefficiency of closeby muons in the endcaps that affected part of the data taking ?

We have added a few sentences about it in the AN. The trigger efficiency is very high because there is (almost) always a lepton with  $pt > 50 \text{ GeV}$  in the central region. We have a  $pt1/\eta1$   $pt2/\eta2$  mapping, but the net data/MC scale factor is  $\sim 1$ . Given the presence of a very high pt lepton, it's not surprising that the efficiency is so high, and that's why the trigger efficiency plays a little role. Indeed, we find some discrepancies between data and the simulation, but they are totally irrelevant for our analysis.

233: "the requirements for [the looser] selection are shown in Table 7" --> please add them to table 7

Those sentences were not properly written, they are now fixed. We use the tight and loose definitions, and Table corresponds to the tight definition.

Fig 2: how are the uncertainties computed ?

The plots show statistical uncertainties only, there are no backgrounds in both denominator and numerator, and hence everything is within the overall 2% uncertainty.

288: why is the Z veto only applied to the ee channel ?

The fraction of events with  $Z \rightarrow \mu\mu$  with wrong charge is  $\sim 0$ , therefore there is no need to require the dimuon events to be outside the Z peak.

290: "event is not b-tagged" -> an event cannot be b-tagged... but a jet can. I guess you mean you ask here for 0 ( $< 1$ ) btag as with the CSVv2 WP ?

Correct, we have clarified the sentence in the AN: "event is not b-tagged, i.e. there are not jets in the event passing the medium CSV2 working point"

311: more details on WW DPS are supposed to be found in Section 2... but there is only the sample name and the xsection there, please do add more details

Sentence has been re-written.

344: I'm not sure I understood what is meant exactly by 'real lepton contamination / subtraction', could you please expand ?

In the tight-loose sample from data there are events with two real leptons where one of them didn't pass the full selection. Since we are evaluating the background from non-prompt events only, this contribution needs to be subtracted. We have added the full formula in the AN. Let us also mention this is the standard procedure in this method, we are not inventing anything.

370: the non-prompt rates (...) : these non-prompt scale factors that you use seem worth having a look at: could you please add a table with those ?

We have added a couple of plots. For technical reasons the numbers were quoted with not the very latest definitions, but we think this is good enough. Notice that those numbers aren't used anywhere in the actual analysis in any case.

384: I'm not sure I understood: your systematic uncertainty on the non-prompt estimation comes mainly from flavour composition and (fake) lepton kinematics ? if so, shouldn't this uncertainty be correlated somewhat with jet energy corrections ?

Not really, or not exactly. the point is that there is somehow a correlation between the lepton-jet pt distribution and the fake rate. A way to modify the lepton-jet pt distribution is to modify the recoiling jet requirement.

394: "these results are well covered by the syst uncertainties in the non-prompt estimation": you mention this several times, but it seems there is no actual number on this syst uncertainty due to the method (at least not in table 12, not in the list of syst considered in section 9.9) ?

Indeed, we don't take quantitative numbers from the MC study per se. We see that the net uncertainty is within or below 30%, which is the overall factor we consider.

432-433 : do these expected numbers already include the lepton fake rate and WZ SF ?

The non-prompt rates are truly from data without any scale factor and therefore they are included in all our quoted results. Nevertheless, the WZ scale factor(s) is not considered anywhere, except for the final result. Indeed, a posteriori, we see that all 4 scale factors are consistent with unity, as shown in Table 11.

494: please confirm that shape uncertainties are affecting mjj-WZ and mjj-mll distributions, not only mjj

Yes, we confirm. We have added a small sentence to emphasize it in the section about the systematic uncertainties.

section 8: - how was the binning of the final discriminant chosen ?

We have had a long debate with the SMP conveners about it. There was not full optimization, but we try to have a reasonable flat behavior for the signal on mjj, while having enough background events in the high sensitivity regions. For mll, we also tried to have a reasonable binning without large statistical uncertainties. Notice that we should not overoptimize the binning since some of the background processes have limited statistical precision, and hence we may bias the expectation too. We should also notice that the signal strength uncertainty shows a rather smooth behavior, and that's probably our most important result.

- please show the mjj x mll unfolded distribution

Done

511: could you add a figure (eg lepton kinematics) justifying your statement that the residual difference is covered by the trigger eff uncertainty ?

The lepton pt1 and pt2 are part of the distributions for the unblinded signal region. Since we are not sure if other plots fit well in the AN, we have added here the leading and trailing pt distributions for  $Z \rightarrow ee$  and  $Z \rightarrow mm$  events

- Leading pt in  $Z \rightarrow ee$
- trailing pt in  $Z \rightarrow ee$
- Leading pt in  $Z \rightarrow mm$
- trailing pt in  $Z \rightarrow mm$

section 9.4: here and after (section 9.9 / table 12) it's not clear if what you mean by 'MET uncertainty' is only the effect of JEC/JER ? If so please do clarify and remove 'MET uncertainty' from the list of unc. sources

You are correct, the terminology was not right (and just discussed with the MET convener too :-)).

section 9.9: - trigger efficiency uncertainties seem to be missing from the list

Done

References: AN numbers of 'technical reports' are not appearing

Thanks for noticing, we hope we fixed them all

769: I'm not sure to understand why this 14% difference in the signal yield is not significant: this seems to me larger than the various syst uncertainties on the signal ?

No, this is a choice and we feel strong about it. We are going to quote the cross section w.r.t. the Madgraph prediction, and therefore anything else shouldn't be taken as an uncertainty. In fact, we have doubts about the powheg sample since one expects a small NLO correction. Nevertheless, we will see once we unblind if the data is more similar to Madgraph (and VBFNLO) or powheg.

Some preliminary comments on the PAS (in no specific order): - the summary should mention what are the final discriminants - mention of the Zeppenfeld variable - list of background is incomplete (at the very least WWW and ZZZ) - 193: please clarify that other backgrounds are estimated from MC - hadronic tau veto not mentioned - background proportion not in sync with the AN (1304-306) - please consider adding the WZ control region that enters the final fit to the PAS - please consider adding the final discriminant (2D mjj-mll) to the PAS - all figures in the PAS should show post-fit uncertainties once unblinded - fig 2: the y axis label do not match the plot (which binning is not fixed) - 1132: the leading lepton pt is not shown in fig 2 - 1143: specify the order of the calculation

We will consider them, although the PAS needs a very serious re-writing.

## Comments from Francisco (Apr. 10, 2017)

- I am a bit confused about your signal definition. In sections 2.3.4 and 2.3.5 you seem to be including  $Wll$  and  $W\gamma$  as signal, is that true? Why?

We have clarified the text a bit in section 2.3. Those samples are not part of the signal definition, but the samples have been produced with requirements closer to the signal region definition to enhance the statistical precision of them. Since they are not standard samples used by everyone in CMS, we thought it was useful to add some information about them.

- section 3, I do not understand your logic to define the signal, since I cannot control the interference, I remove the QCD... Can you include everything and assign a systematic error due to the interference term? Looking to section 9.7.1, I guess you could estimate at GEN an approximate efficiency for the

interference and hence your systematics will be a fraction of the 6%. Is the efficiency and shape for EWK very different from QCD?

This is something which has also been discussed in the past. In principle, it's just a matter of definition. Do we want to observe the EWK process or the EWK+QCD process? We think it's probably better to look for the EWK process only, but we would be happy to agree with you in either choice. In any case, it's just a small additional percent. The shapes and efficiencies for the EWK and QCD processes are indeed very different. In short, the QCD component is  $\sim 0$  (or very small) at high dijet masses, and that's another reason why it's probably better to look for the EWK process only.

- How do you treat events with the decay chain  $W \rightarrow \tau \rightarrow \text{lepton}$ , signal or background

We have added an explanation in Section 3. The  $W \rightarrow \tau$  decays are treated as part of the signal component and selected in the analysis through the tau leptonic decays, with the exception of the fiducial cross section measurements where the  $W \rightarrow \tau$  decays are not part of the signal fiducial region. This is done to help external theory predictions to be compared with our signal fiducial cross section definition. This definition has been done in agreement with the SMP folks.

- section 3, it would be interesting to see in a plot or a table how your additional requirements on the electron work

We have added a new appendix summarizing the efficiency of additional requirements on signal and the rejection power on background. We have also added the expected (large) gain on the signal significance.

- figure 2 right, how are the error bars calculated? Unless there is an important systematic contribution they cannot be right (they are too large for the fluctuation observed). Maybe you are not considering binomial errors?

This is just assuming Gaussian statistics, but the number of events is so large, that the quoted results are good enough. On the other hand, we don't see any special fluctuation, the shape on the scale factors are reasonable to us/

- figure 2 right. Is this shape induced by you additional requirements or it is the same for the standard electron? Can we see the same plot at different stages of the definition? Do you understand why it is not symmetric around 0?

Yes, this means the electron selection is not symmetric in  $\eta$ . In fact, this is something that it has also been observed in looser electron definitions too. The plot just includes two stages: (a) triple charge requirement, (b) number of missing hits = 0. The efficiency of the first requirement is very high, and therefore all the inefficiency, and associated scale factor, comes from the missing hits requirements. It's indeed a tight definition, and it's not that surprising that data and simulation don't match perfect, nor they are symmetric in  $\eta$ .

- section 5. I am curious why you include a veto on leptonic taus, it is against WZ, with  $Z \rightarrow \tau\tau$ ? Have you considered cutting in the maximum number of jets?

Yes, this is mostly to reject WZ hadronic taus (we have added that point in Section 5). Requiring a maximum number of jets is not easy to be applied, and it won't help after the final selection. With our selection, both signal and backgrounds have a similar jet multiplicity at the final level.

- non-prompt backgrounds. In general I found a bit confusing that you mix the particle id with the physics channels. I mean, I think you have a background from  $t\bar{t}$ , where you selected a non-prompt lepton

Yes, the non-prompt background is mostly  $t\bar{b} \rightarrow l\nu b\bar{q}b$  where one of the jets is identified as a lepton, and the other two jets are identified as VBS jets.

- 6.3 do you know which of the WZ enter into your selection? I guess that most will be true leptons, one from the W and another from the Z, while the last one missing due to acceptance (or maybe a tau). How can you ensure your selection, which requires a reconstructed Z, represents this background?

Yes, this is something we studied to improve its rejection. Around 25% of the events have only two leptons (i.e. one lepton from W decay, one lepton from Z decay, and the other Z decay is a low pt hadronic tau), around 55% of the events contain 3 leptons, but only two of them are in the fiducial region (i.e. the other one has either low pt or large eta), and then finally about 20% of the events have three 'fiducial' leptons (and hence one of them is not identified as a lepton). Then, the ratio between the WZ events in the signal region and in the WZ control region comes from kinematic and leptonic branching ratios, which is something well know. In addition, we have also defined a WZ control region inverting the Z selection (see Section 7.2), which further checks that background.

- 6.5 beyond the calculation of the fake rate (BTW a name I don't particularly like :-), it is not clear to me how you build your data estimated samples. "events with fakable objects in MC", which MC?

For this, maybe it would be better to explain it in a meeting, but let's us try here. The data estimated samples come from data where we select events with one tight lepton and one loose lepton, or two loose leptons. Events with two loose leptons in data are used to be subtracted from the one tight and one loose lepton region to avoid double counting the QCD contribution. In addition, we use simulated events with one tight and one loose lepton -with two generator level lepton- to subtract the real lepton contamination. In short, neglecting the QCD contribution, we can say that

$N_{\text{fakes}} = (N_{\text{data}}(\text{tight-loose}) - N_{\text{MC}}(\text{dilepton}(\text{tight-loose})) * \text{fake\_rate}) / (1 - \text{fake\_rate})$ . We have added the full formula in Section 6.5.

- 6.6 First you have some confusion between the charge misID and the scale factor (the difference between this in MC and data). Then I think there is a factor two missing. If you use di-lepton events to check the charge, you have 2x the charge ID per lepton. You should divide by two. Of course if you apply later to ee, you recover the 2, but that is not the case for emu. And you should also correct emu. Didn't you check the dependence with Pt?

Unfortunately the explanation was not correct, and it was referring to an earlier stage of the analysis, this has been fixed in the AN, and summarized here too. The charge mis-ID rate is measured in 5 eta bins taking into account all combinations, i.e. there are 25 different eta1-eta2 regions and 5 charge mis-ID rates to measure. A chi2 fit is performed to obtain the charge mis-ID rates in data and simulation. No significant pt scale factor dependence was found, as shown in the attached figure.

Wrong sign scale factor as a function of pt

- 7.2, it is not clear to me what you do with these comparisons. It is just a check and you are fine demonstrating there is sort of an agreement? You calculate a SF from here? Or maybe you add a systematics? the same for 7.3, 7.4, 7.5 etc

We have clarified the text a bit (see introduction of Section 7) The first region described below is included in the final fit to constrain the WZ background normalization, while the rest of them are just different checks on the different background processes.

- 7.6 what do you mean by "the low met region allows...", you run any quantitative test? You get a SF?  
...

It's not a quantitative statement, just a qualitative observation. We simply see that the agreement between data and prediction at low MET is good, which indicates a good prediction of the non-prompt background. Once again, we are just building confidence that the non-prompt normalization agrees with the observations within 30%, it's another check.

- L481-485, do you have numbers supporting your statement that the charge does not help? Looking to table 13 you have very different signal/background for + and -, nearly a factor 4 difference. Also depending on the flavour, but I understand you have too low stats for that yet. Have you envisaged to calculate separate xs for ++ and -- or maybe compute an asymmetry?

If we perform a fit on mjj splitting in ++ and -- (and mm/ee/em), then the expected signal significance is about 5.2sigma, to be compared with our value of ~5.6sigma. Indeed, we could gain by using mjj, mll, -and-splitting in + and -. The problem is that the statistical uncertainties become too large. This is what we have tried to explain in the AN: "Nevertheless, at this stage this produces large statistical uncertainties, and it does not increase the sensitivity in a significant manner; therefore it has not been considered." In other words, this is something which should add in future version of the analyses (at least splitting in + and -) once a slightly larger dataset becomes available.

- L515 you cannot "assume" a systematic, you have to calculate or justify it

Correct, sorry, wrong text.

- 9.5 Again, you confuse (in the text, i'm sure you are doing right) what is the correction and what is the systematics (the propagation of the error on the correction)

You are correct again, the text has been improved.

- theoretical uncertainties. It is not totally true that your DD are not affected by theoretical errors, when you take the normalization from data and the shape from MC. Theoretical uncertainties can induce shape effect

Yes, this is true. We mention that we consider shape uncertainties for the theoretical effects too.

- table 12 all errors are identical for the different samples, how can that be?

Numbers are just indicative, this is something we discussed in length with the SMP conveners too. We think the impact plot is more useful. We can add even more precise numbers, although differences of a few percent here and there won't make a difference.

- fig 12. Apologies in advance, because I'm still not sure I fully understand this type of plot 😊 Some things that strike me. You seem to allow independent variation of the scale of WZ, shouldn't you have a single nuisance rather than for three bins? Then, shouldn't these be different from 1? Essentially all nuisances are unconstrained, except CMS\_fakeM. Do I take correctly that this is related to the fake normalization? Do you leave it free and get it from your fit? I mean, for example in fig 14 left, you have the first bin dominated by ttbar and can use it to constrain it. I'm also wondering whether you should introduce a different nuisance for prompt for each of the channels, since the contribution in size and in source are rather different

This will maybe easier to be explained in person, but let's try here. We consider a different WZ scale factor per mjj bin. We do this to allow for a possible shape systematic (effectively speaking in mjj for WZ, e.g. for different EWK and QCD components). By the way, you don't see bin2, just because it's much less important, so it's below 30 in the ranking. The values in data will be different from 1, but these are the expected values, that's why they are all centered at 1 (unless there is a bias in the fit). Once we unblind the data, we will show the unblinded impact plot. Indeed, CMS\_FakeM is the non-prompt normalization. It's not completely free in

the fit, but it's constrained within 30%. The statistical precision would be too poor to include more normalization factors, and we believe our normalization is correct within 30%. On the other hand the bin-by-bin uncertainty is considered per bin, and it's indeed sizable for the non-prompt rates.

- figure 13, middle left, where this nasty peak at 0.5 comes from? Can you add  $m_{ll}$  and  $\#jets$ ?

This is something real and well-known. The CSV2 output has a peak at around 0.5. We will add  $m_{ll}$  and  $N_{jets}$  in the AN.

- figure 14, since you do a 2D fit, it is interesting to see also  $m_{jj}$  vs  $m_{ll}$

Actually, we were planning to use the full 2D unroll distribution for the unblinding, but we will show it now.

- Section 12, wouldn't it be simpler to interpret if you give the result in terms of the signal strength

Well, not really because we are expecting not to see an excess of events over the SM prediction, that's why the limits make sense. We think it's reasonable to show the signal strength for the EWk analysis, and CLs limits for aQGCs and  $H^{++}$  searches... unless we see something significant. In that case, we will need to show it too.

## Comments from pre-approval (Mar. 27, 2017)

document in the an-note MC closure tests for the fake rate method (process dependence, control->signal region extrapolation)

Done

investigate systematic uncertainty on the extracted EW signal due the QCD-EW interference

Added in the analysis

isolation on the muons might be too loose, check if tightening can give better sensitivity, separate e-fake and mu-fake contributions in the yield table

Done, added an appendix with this study

## Comments from Senka and Ming (Mar. 24, 2017)

Table 8 : the electron fake rate is much smaller than the muon fake rate in the high eta region. Is it because the electron ID in high eta is much higher than muon ID ?

The reason is just because the extrapolation factor from the control region to the signal region is much larger for electrons. In other words, all muons regardless the eta region behave in a similar way, while this is not the case for electrons.

Section 7 : do you fit two control regions (top-tagged and WZ) simultaneously ? or you fit top-tagged CR first ? It seems to me the non-prompt contribution does not change before and after fit in Fig. 4.

All regions are fitted simultaneously. The WZ control regions is rather pure on WZ events, and furthermore there are four independent rate parameters, one per dijet mass bin. Nevertheless, there is a single normalization factor for the fakes, plus the btagged region is not so pure in fakes, there are contributions from wrong-sign and signal events. Therefore, the only noticeable change for the fake component is an overall factor (up/down) in the yields. With a larger dataset it will be possible to perform fits adding more parameters.

Fig. 9 : (plotting style) right now the aQGC histograms cannot be seen easily. usually we plot the signal on top of background and non-zero aQGC is stacked. I think we also want to include this figure to PAS.

There was a ROOT plotting issue with that plot, this is already fixed. We will show the proper plot during the pre-approval, we will be happy in adding it to the PAS/paper.

Fig 3: please also add prefit distributions the same as for WZ CR

This was already done, the left (right) figure is the pre-fit (post-fit) distribution.

table 10: please change indicative numbers to either range per bin (for example 3-5%) (preferable) or to the mean value of uncertainty per bin.

Some values have been added to the AN. We still write 'indicative' since there are several caveats to those numbers.

please report the postfit signal strengths for non-prompt and WZ in the AN (not the PAS)

Values have been added to the AN.

## Comments from Senka and Ming (Mar. 16, 2017)

- W+gamma\*

V+gamma\* is mislabeled in all plots, it should be Wgamma. The actual W+gamma\* is taken into account in what we called WZ category now

- MC statistics in table 10 and table 11, what do they mean

They are actually meant to show the same thing. Numbers in table 10 is indicative, exact numbers are in table 11.

- Fig.3, in order to show the results of validation of fake rate method, data points are needed

Done, although we disagreed in doing so.

- Fig.4 add both pre- and post-fit plots in AN, and only post-fit plots in PAS, also clarify the details of the options in fit

Added to the documentation

- anomalous coupling

A plot was added in the AN showing the effect on the SM signal of a non-zero coefficient for the T0 operator.

- clarify in cross section measurements, what leptons are used

bared leptons, added to the documentation

- add plots like Fig. 8 to Charged Higgs section

Figure 9 includes the aQGC effects

**Comments from Senka (Mar. 15, 2017)**

<https://paper.dropbox.com/doc/WWss-documentation-comments-March-15-smakxYmgRw3TCu40dwqKn>

**Physics:**

QED-QCD interference. Ok, this is now documented in AN section 3. I agree that neglecting the interference is ok since it is small. If we would want to include it it would mean to include it as the part of the signal definition and keep WW QCD only as background. So for the cross section measurement of the signal (=EWK+interference) the effect of the interference would be via acceptanceXefficiecny, and the effect on the signal shape in 2D mjj\_mll. The fiducial PS definition now is close to what you do at reco level, so the effect on acceptanceXefficiecny will be small. The effect from signal shape is also going to be small since measurement is stat dominated. Also the effect on both will be small in general anyway simply because the interference is <6%.

Agree.

WZ/gamma\*

- I understand that the normalization is taken from data CR and the shape from MC. Correct?

It's estimated from simulation, while the 4 normalization factors corresponding to 4 mjj bins are calculated from data control region. In other words, any shape dependence in mjj is also taken care by the fit.

- Please add the data to MC comparison from this CR: table or plot

Based on the estimation method from the answer to last question, Fig.4 is the comparison you want to look at.

table 7. I am a bit confused with these uncertainties. It says these are stat only.. For example WZ has unc of 2%.. Is this just MC uncertainty without stat uncertainty from CR? DPS WW has 9% and table 10 says 20%. Please check or clarify.

Table 7? We assume you mean Table 10. The WZ uncertainty is about 20%, which comes from the statistical power in the control region. Table 7 has nothing to do with systematic uncertainties as far as we can tell. This is something that can be trivially discussed in a meeting.

**Editorial:**

- VBS vs EWK production. As far as I understand your signal is EWK and not only the VBS. It is a bit confusing in few places in the documentation. Expecially in L860 (today's doc version) where you say significance of the observation of the production via VBS .

Fixed, changed it to "...of the EWK production of same-sign W boson pair"

- L79, L801: make the mean PU value consistent between AN and PAS

Fixed.

- L748: 36.2 35.9 fb<sup>-1</sup>

Fixed.

- L89: please update WZ and ZZ MC Powheg MG

Fixed.

- table 2: please add names of used official EWK and QCD WW samples

Added.

- perhaps I simply missed it but I did not see in the documentation described that QCD WW is taken from simulation..

Added.

- L403: requirement. requirement.

Fixed.

- figure 3 is missing data points in today's doc version

We prefer having the blinded distributions in the documentation at this point.

- figure 4, WZ control region. The uncertainty (shaded area) on the estimation from MC+non-prompt is much larger than it was in the past. If I compare to last SMP-VV report. Uncertainty should be dominated by WZ MC stat uncertainty.. Please check.

The plots are post-fit now.

- L496: WZ is not estimated from simulation

It's estimated from simulation, while the 4 normalization factors corresponding to 4 mjj bins are calculated from data control region. We can discuss the working at the CWR.

- table 10: missing WW QCD and Wgamma and corresponding systematics

Added two columns for WW QCD and Wgamma. Please, please, remember these are indicate numbers.

- L515: table 17 11

Fixed.

- L517: please update the signal definition!

Fixed.

- L520: fit is in mjj\_mll, not mjj

Fixed.

- L582: using a mjj binned template

Fixed.

- L754: Cross section measurements for  $W^\pm W^\pm$  and WZ processes      Cross section measurements for EWK  $W^\pm W^\pm$  process

Fixed.

Editorial:

- L791: please update signal definition!

Fixed.

- L796: powheg MG

Fixed.

- L895 and L529: jet eta 5.0 4.7, even though there are no events in between.

Fixed.

- Wgamma sample and bkg estimation is not mentioned in the PAS.. or did I miss it?

It can be done, we can discuss it at later stages.

- L857: please update the signal definition and interference!

Fixed.

- L858: The mjj and leading-lepton pT The mjj and mll

Fixed.

- L861: four bins in mjj with two bins in the lepton charge please update!

Fixed.

- L882: found to be 5% for the signal normalization please update

Fixed.

- L883: there is no PDF uncertainty for WZ since data-driven

Fixed.

- L889: please update signal definition!

Fixed.

- L891: Using the mjj distribution please update

Fixed.

## Comments from Senka (Mar. 07, 2017)

<https://paper.dropbox.com/doc/WWss-PASAN-v1-commentsquestions-Llkrhlls71Rd1HTJsdVp2>

### Text:

- please check that jet dr cone size, luminosity, signal and background definition for WW is correct everywhere in the documentation

We anit-kt jets with a distance parameter of 0.4. Fixed the paper draft.

- 8TeV 13 TeV in the documentation

The paper draft mentions 8 TeV twice. The first is the reference to the Run-I result and the 2nd a statement about the unitarity limit.

- please update: table 2 (WZ), table 6

Table 6 is updated. WZ samples are updated in Table 2. A few other processes have also been updated.

- L211: The muons used to reconstruct a Z candidate are selected inside the fiducial region of the muon spectrometer I assume this is from another analysis

Fixed. Should be W in this analysis.

- please document more details on the fit. Which regions are used, that the fit is simultaneous, what are the input yields and nuisances, and which parameters are fitted with which constraints. Once unblinded please report the (prefit and) postfit parameter values and uncertainties. So this would mean fake lepton normalization and WZ normalization.

There is a section about it in the AN.

- figures 2 and 3: once you unblind please include the postfit normalized plots

Will do

- 7.3 loose ll+jj control region. Please motivate this control region. I think it is good to have this as extra information but some motivation in the text is needed.

Added motivation for having this control region. Please find it in Section 7.3.

- L387: you do not fit MET here..

Fixed. Should be mjj.

- L432-435: please update, WZ is data-driven

Fixed. Removed WZ. Replaced with VVV, which is also from simulation.

- L442: DM sample

Fixed. Should be same sign WW samples.

- table 11: what is the definition of SM expectation ? WW QCD+EWK?

Fixed. Should be WW QCD only.

- Please include two different fiducial regions for cross section measurement. One region consistent to region used in 8TeV analysis and another tighter region that is close to selection cuts.

As discussed during the meeting (SMP-VV, Mar. 7th), we will work on cross section measurement results after pre-approval.

## Physics:

- please document QCD-QED interference studies in AN and add few sentences with conclusion to PAS

A reference was added to a study using the VBFNLO framework and cross checked with Sherpa using the Comix matrix element generator that shows the interference to be smaller than 6% for  $m_{jj} > 300$  GeV. (<https://arxiv.org/abs/1311.6738>) It was also checked using phantom for the High luminosity studies and confirms the interference is negligible using a soft VBS selection: [https://twiki.cern.ch/twiki/bin/view/Main/VbsForTpQandA#follow\\_up\\_with\\_Matt](https://twiki.cern.ch/twiki/bin/view/Main/VbsForTpQandA#follow_up_with_Matt) In our Madgraph samples the interference is not reliable because we didn't set a scale choice and the default will choose different type of scale for the same PS point depending on if you are generating [total]/[QED]/[QCD]/[INT].

- L236: do you use jets up to eta 5?

Up to 4.7, although it makes no difference. There are no reconstructed jets between 4.7 and 5.0.

- figure 5: the impact of QCD scale is ~10%. I would expect this value to be the same as QCD scale on signal in table 9 of 5%. Why are these so different? I guess I am missing something here..

5% in Table 9 was an indicative number. Please find the real number in Fig.6 in current version of documentation. It's about ~10%

Fake-rate related:

- please include the exact event selection used to derive the fake rate

Done.

- include the plots showing fake rate as a function of pt and eta

Done, although there is a table instead of plots.

- please make it clear in the documentation that you use both TL and LL regions to apply the fake rate to

The text is updated with clear statements saying that we extrapolate from both "Tight+Loose" and "Loose+Loose" samples. Please refer to Section 6.4.

- please motivate the 30% uncertainty on the fake rate

The 30% comes from the data/MC agreement in b-tagged control regions and from previous studies.

## Comments from Ming (Mar. 07, 2017)

<https://paper.dropbox.com/doc/WWss-PASAN-v1-commentsquestions-Llkrhlls71Rd1HTJsdVp2>

Text:

- Section 2 : Can you please document which MinBias cross section is used for PU reweighing into AN ?

Done.

Physics:

- table 1 : update to re-miniAOD

Updated.

- table 2 : the WWW is missing from the table. The dataset name for ttZ(qq) is incorrect. Can you please document the k-factor listed in the table in AN as well ?

WWW sample is added. ttZ(qq) sample is corrected.

- Section 4.1 : Can you please document the exact MET filters used in this analysis ?

We use all the filters recommended by the jetMET group, we think it's better for the reader to go to the twiki instead of writing them in a non-expert manner.

- Section 4.4 and 4.5 : please document the SFs you use. If it is measured by yourself, please document how it is done. Do you apply any electron energy regression and scale correction/smearing and muon momentum correction ?

We have added the appropriate links. In addition, we have added the studies regarding the Nmissing hit requirement scale factor. We apply the electron energy regression and scale correction/smearing and muon momentum correction. Either way, bear in mind this makes no difference in the analysis, we don't need/care about a high precision mass measurement.

- Section 7 : I have to say this section needs to include much more details for review. In addition, it is odd in the current AN when the loose dilepton and dijet CR is mentioned because it does not mention what it is for.

Texts are added saying the the purpose and motivation is not to include loose dilepton and dijet control region in the final fit, but add another region which is close to signal region, check data and simulation agreement. Please find it in Section 7.3. In reality, there is not much to say. The samples are selected in similar manner as the signal region, just with a few different requirements. In the first two cases, they are actually used in the final fit to extract some of the background normalizations directly from the fit.

- Section 9 : What does it mean to fit to the MET shape in L386-387 ?

Fixed. Should be mjj.

- L442 : What is DM in this analysis ?

Fixed. Should be same sign WW samples.

### Physics:

- table 3 : does the primary vertex selection given still hold for Run-2 ?

You are correct, we use the nominal selection. We have also removed any explanation in the text, there is no point. Indeed, at this point almost everyone uses the run-II approach, which is based not only on the sum  $pt^2$  of the tracks in the PV, but also in other quantities like sumET or MET.

- Section 4.6 : Do you plan to update to newest JEC ? I am just curious whether we can use jets between 4.7 and 5.0 in Run-2. Please also add the jet ID you use.

We use jets with  $l_{\text{eta}} < 4.7$  in this analysis. We are using the latest and greatest JEC recommendations.

Text:

- Section 9.4 : How do you smear jets when there is no matched gen jet ? Do you smear your jets in MC in nominal analysis ?

If there is no matching, then no smearing is applied. This is not done for the nominal analysis since it makes no difference given that we don't care about the mass resolution in a significant manner.

- Figure 6 : What are the shaded bands ? Does it represent the statistical uncertainty only ?

Yes, it's statistical only. For the final results, we will have the statistical and systematic uncertainties for the main distributions.

---

This topic: Sandbox > SameSignWWjj13TeVReview

Topic revision: r46 - 2017-09-04 - MarkusKlute



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