Dear Anadi,

Thank you for your comments. Most of them have been implemented or noted (see revised version of the paper) and detailed answers are added below.

General comments:

- the paper is overall very pleasant to read and complete. It will be a valuable document for the future.
- However Sec 1 to 3 (and section 6.3) would benefit from some consolidation. The same piece of information is repeated in various places, leading to redundancy and possible confusion. Table with specs and layout details would be beneficial.
- Also the language is at times too colloquial and not appropriate for a paper. It would good if you could review carefully these sections and ensure that statements are always quantitative and/or supported by references.
- Figures tend to lack units and the format is quite difference from fig to fig. The labels can be fixed (even in the pdf file), it would be good if plots were remade with a common format. I know that the latter is challenging.
- Please review the tense used in the paper (sometime it is past, sometime it is present tense). I expect the CLE to make this remark as well, but it would be quicker if you fixed this while you implement the reviewers' comments.

Line by line comments:

- L4 I find this statement a bit vague. I suggest you state the requirements or remove the statement → has been rephrased according to Ulich's suggestion.

- L5 Likewise (it is of course a matter of taste) I find the words "key"," indispensable" not appropriate for a paper. I would be factual and state how the pixel info is used. → rephrased

- L8 (and other cases) I don't think the interaction point is a reference in the CMS system, so you would not be able to define a distance with respect to that. → this is the standard term in the description of the position of the pixel layers, but can of course be discussed.
-
- I suggest you add "instantaneous" in front of luminosity → ok

General comments for Sec 2. Having a high level introduction is very useful, however the risk (as in every paper) is that we use terms that not defined (for example ladder on L51). I would streamline this section. For example I would build the section starting from sensor, then modules, ladder, large structure. → ok, this has been improved. In particular repetition in Section 1 and 2 has been removed. Also definition of ladders is removed as it is not needed at this point.

- L44, 45 are the hybrid modules the same as the modules? the sentence may be mis-leading to non expert readers → restructured sentence to avoid confusion.

- L46 the bonding is not electrical per se, I suggest you remove "electrically" → ok

- L65 While I understand what you mean, I think the correlation between "fitting in the same envelope" and "need for larger bandwidth" is not clear. I suggest you reword it. → ok

- L78-92 is a more precise description of the scope of the upgrade, with respect to L22-26. As I mentioned above, I think making sure that the statements in the introduction are more quantitive (for example the requirements in L78-92 could be moved to the introduction) and Sec 2 is consolidated (we don't add qualitative statements when you have a proper rigorous statements in the following section) will help. → Agreed. Instead of moving things to the introduction, however, the approach was to keep the introduction simple and have the details in Section 2.

Given that the ROC will differ depending on the layers, it would help if you added a table with the requirements for BPIX (for each layer) and FPIX: dose, hit rate, instantaneous luminosity, etc. → ok, added table with hit rate/dose& fluence in Section 2, and table with design requirements for ROCs in corresponding section.

- L84 spurious Fig? → removed.

- Fig 3 could you make the Fig with the same x axis? If there is a reason for having different x-axis, can you explain? → yes, can use same x axis. Noted. To be done.

- L85-ff Likewise above, I think it is better to introduce definition and descriptions only once. Here for example you mentioned the pixel size that was already mentioned in the introduction, and you

quote the size of the sensor which is then presented again on L104. I strongly encourage you to go through the paper and make sure there is no redundancy (otherwise the paper becomes hard to read) - it would help make the paper more compact. → agreed and redundancy removed.

Another example of redundancy that could be resolved is on L100 and L122. You state twice that the designs are different but you don't provide the details till L124/143. Please try to consolidate. → rephrased.

- L106-107 the sentence does not flow nicely grammatically → has been rephrased

- L117 given that you don't provide cost and schedule for the project, I don't think we have to add this piece of information → ok, has been removed

Title of 3.1.4 I find 'height' to be a bit jargon. I would simply state "signal". I find this section a bit too pedagogical for a peer reviewed publication. I suggest you remove the explanation of why the signal degrades (I would actually remove the description about the benefit of n-n silicon for the same reason) and give the technical requirements, specifications, description of the CMS specific sensors. Here for example what do you mean by 'small signal', can you provide the spec of the chip? → section is removed

- Fig 6 is a bit hard to read. There are more lines that documented in the legend. Can you explain why (adding the explanation to the caption) or fix the plot? → Figure is removed

- L167 This is the first time you write about the actual bias voltage. It would be useful to quote the planned range of bias voltages in the silicon sensor section to place this discussion into a context. → section removed
- L169 what do you mean by "probably"? I suggest you are quantitative. What is the impact of having on average 6k signals on hit finding? → section removed
- L171 is 3000e the threshold for Phase1? The number is different on L228. What ROC is used here? → section removed

- L172 This piece of information is important but should be introduced ahead, when you discuss the expected range of bias V (in the sensor section) → section removed, content moved.

- L173 how did you estimate 10k from Fig 6? → section removed
- L181 Another example of redundancy, you have already stated what your requirement is (600MH/cm2). and the need for a new ROC. Please consider consolidating, you could for example remove L179-185. → ok, removed L179-185.
- L185 citation for the ROCs? → have been added.
-
- L185: please add here where each chip is used. → section headers removed, so this now comes in the next sentence.
- L227-239 by the time the reader reaches this section, they should know what the requirements are, to place the test into a context → table with ROC design parameters has been added.
- L239 It is not clear why you chose 600MHz/cm2 as the boundary? From Fig 7 one would infer that PROC600 would be needed already for ~150MHz/cm2. → yes, but what is said is only that we need something that operates up to 600 MHz/cm2. What I think is missing is to specify the 120MHz/cm2 maximum rate in L2. Has been added.
- L241 please add the requirements → the requirements are high efficiency at 600 MHz/cm2. A table with requirements has been added.
- L251 As I wrote above, it would be helpful to have a table of requirements on L71 to help understand with what margin the specs are met → agreed.
- L253 can you be quantitative? What is the degradation? → rephrased slightly and moved ref [11], where this is discussed, closer
- L316-317 A table summarizing the various options and the number of modules per layer would help guide the reader → ok, added
- L321-322 Are you going to describe how the modules differ? If so please reference the section. If not, I suggest you remove the sentence (or add a detailed explanation) otherwise the reader does not learn much → it is explained in L348-352

I don’t fully appreciate why the discussion about the production schedule is presented in the paper. I would remove L323-325 and L328. If however you would like to keep it, please add the details of the L1 vs L2/3/4 schedules, what was the schedule driver, etc. Our objective is to have the reader learn about past experiences with detector construction. → ok, metion of different production centers now motivated by use of different bump-bonding technique. Schedule discussion moved to explanation of Figure where the yield of module production is shown.

- L323 The “test” itself did not contact the ROC, please reword this sentence (for example the “goodness of the bump bonding was verified using …”) → has been rephrased
-
- L337 how many centers? Please be specific. Also what made a given center capable of that operation? Why not sending all faulty modules to the centers capable of reworking the assembly? → new sentence is more specific: "The replacement of single ROCs that failed in an otherwise good module was used by four of the five centers and 10-20% were reworked with good yield." comment: Re-working modules from one center in another center was not possible due to the different bump bonding processes. It was, in principle possible everywhere, but PSI/Dectris decided not to do it. Often there was not enough indium left on the sensor after pulling off a previously bump-bonded ROC leading to a bad rework quality unless specially prepared ROCs with indium on the ROC side were used.

- L349 Do you have an exploded view of the module to help understand the mounting procedure better? Fig 4 is too compact. → added module cross section (Fig 8 in new draft).

- L359 made of —> instrumented with → ok

- L373 Was any of the issues reported in FPIX encountered in BPIX? If so, please specify in the previous section → not really, issues have been removed from FPIX part

- L382 please add the specs of the thermal testing→ ok, added

- L385 - 388 This test is described together with the others in the following part of the section. Why did you single it out here? I would move whatever relevant information you have here to the 7. IV Test to avoid redundancy and possibly confusion → has been removed

- L389 do you have a reference for the SW? → added ref to Simon Spannagel's thesis.

- Fig 11 If the z axis is the efficiency why doesn't reach 100% Please explain the figure in the text. Missing units on axis → figure has been removed.
-
- L402 and caption of Fig 11 please use subscript where appropriate (thr, Comp, Del) → ok, but kept CalDel.

- Fig 12 would it be possible to make these plot more inform? For example the x axis labels? Also for BPIX, what is he difference between the dark and light color? Please add the legend. Missing units → yes, noted. To be done.

- Fig 14 please fix the label Hitrate [MHz/cm2] —> Hit rate [MHz/cm^2] . Caption: please explain why the 50 o 120 range is used to determine the efficiency. → ok, caption added. It is the hit rate range expected in L2-L4. Labels to be improved.

- L487 please specify what you mean by 'defective pixel' (i.e. what test the pixel failed) → according to pixel alive.

- L488 is 600 the threshold for both chips? yes

- Fig 16 please explain what 'detector grade' means in the caption. Bottom plot: too small, hard to read. The black line (received) is not visible, 'received' is not defined. Please add. → plots to be remade. Noted.

- Fig 18 you have a note "update picture". Is the picture up to date and the text needs updating or the other way around? → The figure is still not updated. Drawing with latest version of BPIX supply tube needs to be made. Noted.

Both Fig 18 and 19 would be more useful if you added a description of the relevant components to the caption. → added labels to figure (to be done for fig 18 once updated)

- L590 any lesson learnt from this test that is worth documenting? → discussed in DAQ paper. Added reference.

- Sec 4.2.1 given the issue with the DC-DC converters, I think it would be very useful if you went into the details of what QC tests were made during the R&D, preproduction production phases. It is important to document what tests were carried out which - unfortunately - did not spot the issue that would emerge later on during data taking. →

  Since the problem is related to irradiation, it could not be found during QC testing, which is of course done on unirradiated DC-DC converter modules. We would thus prefer not to enroll the QC procedures in the paper. The actual problem was that in the irradiation testing by the chip designer no power-cycling was done. The relation to irradiation has been made more explicit in 4.2.4, in reaction to other comments. It is also now added how this was found, which implicitly tells that these tests have not been made before.

-

- Fig 22 Too small, please move the sub fig on the right to the bottom → ok

- Fig 25 the Fig is a bit hard to read. Please add a detailed explanation to the caption → figure has been removed.

- L807 what was the conclusion of this test? Again, it would be useful if you documented whether or not the test was successful and reliable for future references. → added statement that we lowered flow during operation based on the results of these tests

- L829 I am not sure you refer to Fig 18, as Fig 18 is not labelled → fixed and labels added.

- L890 was the yield 100%? Presumably not, but we don't have a number.

- L902 please add the dimension of the disks → rephrased, information about sensor radial coverage added.

- L962 is the sag within spec? → to be followed up.

- L972 As mentioned above it would be important to expand on QC and explain what (if) issues were encountered or all parts met the specs →to be followed up.

- Fig 34 is too small, please expand it. Please explain the Fig in the caption. What is ref 1/2/3? If those are preirradiated samples, why is the variation so large in some cases? For the irradiated samples, the araldite depends strongly on irradiation, please explain. → figure has been removed.

- Sec 5.2.7 Was any material used by FPIX tested for rad hardness? → section has been removed

- L1009 a reader may wonder why the DC-DC converter issue did not manifest itself in the pilot system. Please add a short explanation about it →

  In line 1011 added: "Unfortunately the problem with the DC-DC converter chips described in Sect. 4.2.4 was not found, though, as the critical cycles of disabling and enabling were not regularly performed. Also the number of installed DC-DC converters was very small (six devices), while the fraction of failed DC-DC converters in 2017 was 5%. "

General comment to Sec 6.3: the language is frequently too colloquial; technical details ought to be added to actually provide valuable information to the reader. Please have a look at this section again. I also give my line-by-line comments below. → language is improved.

- L1110 what does "the wire bonds were pulled" mean? → now explained in BPIX section: „The hub address for each module is set before mounting by removing the corresponding wire-bonds on the HDI.". Sentence removed here.

- Sec 6.3.1 Please specify the temperature and duration of the baking → added 30min at 55C.

- L1122, 1123 'heavy', 'bulky' too colloquial. Please report the details of the pigtails (size, etc) → rephrased.

- L1136 'inadvertent' colloquial as well. Please write that cables did not meet specs. Also explain why the issue was not identified ahead of time. Add if any other issue originate from the cables being shorter than needed. → This is proposed rephrasing by Will: "The half cylinder cabling procedure was modified when it was found that some flex cables did not reach the port card in the original cable routing schema. Subsequently we determined that it

was a miscommunication that led to a shorter length flex cable being ordered than originally planned.". Prefer to keep the original sentence.

- L1137 “went more smoothly” again too colloquial → sentence removed.

- L1138 in this case as well it would be important to be specific: what flex cables, how many broke, etc. → added info.

- L1141 please specify what 2% refers to → rephrased: „In all, 9 of 768 modules of the detector were not working after all practicable recovery procedures were used. Of these 9 non-working modules, 7 were due to flex cable issues, 1 had an issue with a broken TBM, and one had broken wire bonds due to improper handling.

- L1147 ‘abrupt’ too colloquial. Please specify DeltaTem/DeltaTime and explain in detail what issues were observed (including % of bump bonds lost). What was the new procedure? What was the improvement? → rephrased: „For the first half cylinder tested, 0.153\% of pixels were lost in patterns at the corners of modules that were thought to be due to cable strain, uneven pressure when screwing in a module support, and/or issues due the the number and nature of temperature cycles done: CO$_2$ gas was allowed to enter the detector until the temperature of the half cylinder under test plateaued to within about 10-15C of the set point of the coolant at wich point the CO$_2$ liquid was allowed to flow into the detector. In order to reduce possible temperature effects for the second half cylinder tested, we limited the number of temperature cycles by leaving the detector on and cooled, and we modified the CO$_2$ system to provide coolant at an adjustable temperature in 5C steps. For this second half cylinder, we found 0.024\% of pixels were lost in patterns at the corners of modules. For operations at CERN, the CO2 system was more adjustable and we limited the cooldown rate to 1C/min or less. We did not notice an increase in bump bond loss after reassembly and testing at CERN.“

- L1152-54 I would have this statement only if it can be made fo FPIX and BPIX. Also, in order to be informative, it should be more quantitative: how many people (students RA senior engineers tech), how many shifts etc? → statement is removed.

- Fig 43 please explain the spread among channels. Add units. → ok. Units are explained in the caption

- L1280 please quantify, indicate how much the thresholds were changing till we collected 10/fb → changed following Andrea's comments

- Fig 44 still missing FPIX? → yes. Plot needs to be produced.

- Fig 45 unfortunately the colors are really hard to distinguish → improved.

- L1317 how long? → rephrased

- L1320 Previously it was stated that only 0.4% of the BPIX channel only 2% of the overall FPIX were lost. Could you please explain how you get to 98.4 and 96.1%? → in FPIX there was one sector with a clock issue (this is added in the text). In BPIX there are few individual modules that did not work after installation, mainly because of issues at high trigger rate.

- L1343 what reduced the number from 100% to 96.7% after the DC-DC fix? This is due to the detector parts that were bad from the beginning (compare to 96.1% working fraction at the beginning of 2017). Few channels could be fixed during the YETS, but not all.

- L1376-1384 The language is too colloquial (mistake, unpleasant, periodically,…). Please review this part. Provide additional details about what transistor was actually causing the problem. Add the length of the power cycle and the corresponding down time. → language improved, we do not think that it helps the reader to provide details about the transistor, but added the effect. Statement about power cycles improved following also Andreas comments.

- L1384 I assume that a study of efficiency was made, in comparing the two approaches: reset during fill and in between fills. Can you please support the decision taken in a quantitative manner? → the reason for going back to power cycling during fill was made due to the issue with DCDC converters. Explanation added tot he text.

- L1390 It would be useful to have a table with the threshold (spec, during data taking) for the 2 chips. On L228 we mention 1800e for the PSI46dig, on L171 we have 3000e (not clear for which one of the two chips) and here we state that instead of 2000e we lower the threshold to 1300e for PROC600. I assume the final numbers are provided on L1405-6. A Tab like Tab 2 for HV would be useful → will add table of specs. Threshold value of 3000e in L171 is indeed

confusing and is removed since section is removed. Number on L228 should be below 1500 instead of 1800.

- L1382 why is Fig 51 here? Some issues with the numbering of the Fig (it should be 47) → yes there was an issue. Reference is removed.
- L1418 please reword “sudden jump’ → changed to ‚steps’.
- Fig 47 please align the plots → plots need to be remade and merged.
- Fig 48 please explain the shape of L1 → due to cluster breaking, added to caption.
- Fig 50/52 is this the latest plot? → no, plot needs to be made for 2017 and 2018. Noted.
- Fig 55 is vey had to read. Please produce it again with a different format or include a description of the markers in the caption. Why is data/sim better in 2018? → noted. To be done. Comment about precision of temperature measurements added to paper.
- Fig 56 likewise for the start of 2017. Please enlarge this fig (it uses anyway the full page) → noted. To be done.