1) General comments:
* Please remove the list of contributors before the tracker wide review. → ok
* There is a mixture of present and past tense. It may be tedious but I recommend to go through the paper and fix it before the tracker wide review. Just as an example L520, 527, 538. → done to the best of my knowledge
* Different formats are used for the plots. This is understandable given that they are most likely produced by different colleagues and we cannot obtained data/macro at this stage. However you should decide whether you add CMS or not to each plot (also in this case you have a mixed approach). Also please remove "Preliminary" where applicable. → ok, added "CMS" for plots with CMS data or simulation. No label for other plots.
* Especially in the initial sections (introduction, section2,…), the authors tend to write "as described in the next sections" without referring to the sections themselves. This is not very useful and I suggest to either drop the sentence or add the proper references. → ok, done

2) Line by line comments
* End of L138 —> change line given that the improvements refer to all upgraded components and not only to the new ROCs → ok
* L169 I still find this sentence a bit strange: it indicates that latest studies confirmed the goodness of the sensors, and not those done during the R&D phase → but this is what it means. There has been an extensive R&D program to design the Phase-0 sensors. For Phase-1, the Phase-0 sensors where used, after checking that they also survive under the increased radiation conditions.
* L233 define the Lorentz Angle → ok
* L236 do you have a range of variation you could add to the paper? It's a general statement about the dependence of the LA. Prefer to keep as is.
* L239 missing closing bracket → ok
* L243 why 120/fb? → because 2017+2018 is 117.7/fb: https://cmslumi.web.cern.ch/cmslumi/publicplots/int_lumi_cumulative_pp_2.png would you put the exact number?
* Tab 3 I would add what layer uses which ROC → ok, added.
* L330 citation for PROC600? → there is no publication about the PROC. Only the irradiation studies that we cite in [14]

* L350 the cross-talk is discussed a few times in the paper. However it is refereed to as high, low, lower so it is hard to assess the impact for the various chips. Could you add quantitative statements and analyze the effect? → No this is not possible. Cross-talk is a very complex phenomenon and can be measured only in special structures where additional spy pads are included. In the final readout chip we have no access to any direct cross-talk measurement. We can only observe it indirectly by looking at high-rate inefficiencies, which are presented in the paper.
* L392 please explain why → it is said already: „to allow additional status information to be transmitted"
* L410 should this statement be updated with the actual status? → removed the sentence about the testing, since it is obvious and we do not have any statements to be made about the results.
* L422 Can you add a few words to explain the 53% yield? → Added footnote: A follow on batch of TBMs produced for the new layer 1 of BPIX were found to have a yield close to 90% after a new wafer cleaning method was used before probing.

* The paragraphs describing the assembly of BPIX are not that clear. Could you consider adding a table indicating what steps have been done at which institute? For example on L428 you mention "combination of an in house flip chip step": is this in combination with RTI? In some cases you mention the bumps material, in some you don't. Likewise for the UBM. → The idea is that more details are given for the non-commercial processes.

Added information about bumps material for all centers. The process used in CH is the same as used for Phase 0 and described in Ref. [27]. The "combination" refers to the combination of the flip chipping and the UBM, which is described in this sentence.

* L450 the sentence is grammatically incorrect? → fixed

* Sec 3.6 is extremely useful but it would benefit from adding a table listing the various tests and what they probe → We had this more structured approach in a previous version of the paper, where we introduced paragraphs and names for each test. This was changed based on the reviewers request to shorten this section. Prefer to not extend it more and keep as is.

* FIg 11 explain the various populations (why BPIX and FPIX are so different) → As explained in the text the BPIX and FPIX use different sensors which influences their noise and threshold distributions. The precise explanation of each feature seen in the plots is not possible. Some tail can be attributed to the exiostance of large (edge/corner) pixels. Add a sentence at the end of the paragraph.

     caption: we don't mention grade A and B anymore, please update the text → fixed

* Fig 12 motivate the choice of range for the fit --> Well, it only makes sense to do the linera fit in the linear part of the curve. I have added a sentence aboiut this.

* L575 shouldn't the uncertainty on the offset be truncated? → yes, fixed

* L576 roc-to-roc —> ROC-to-ROC →ok

* Fig 15 just to be sure: all BPIX modules where assembled by week 20 and then, out of those, a fraction was tagged as detector grade? I.e. the y is not cut? → y was cut. Fixed now.

* L729 Do we have this value for BPIX? Add it to the paper for completeness? → it is similar for BPIX. But since this statement is not so important to repeat it twice, prefer to remove the sentence for consistency, rather than adding it.

* L730 Was this test done for BPIX? Add it to the paper for completeness → no, this has not been done for BPIX

* L763 How did we know it was necessary before assembly? → -**This is mentioned in the text.** A prototype was loaded with the expected weight. Maybe I'm not understanding the question?

* Was the high pressure stress test done for FPIX? If so, add it to the paper for completeness → -**Add as sentence to end of paragraph**: Extensive pressure testing, along with temperature cycling, was done after the laser welding before the pipes were integrated into the detector structures. Pressure tests were also perfomed after assembly steps and prior to installation certification.

* I would add the fraction of good channels after installation to this section as well (you have it later but the reader may wonder at this point in the paper) → not sure what this refers to.

* L827 Does FPIX have the same sensors? If so, add it to the paper for completeness → ok, added in L848.

* L948 I am surprised that BPIX only was tested the full chain. Is this statement accurate? This statement is correct. It is in the power section and refers to the testing that has been done during the development of the power chain. For this the BPIX chain was used. Of course there has also been testing of the full chain for FPIX which is the system test mentioned in the paragraph starting in L850

* Section 7.5 seems to belong to 4.1 (like L731 and ff belong to 4.2) → in some sense, yes. The sections are closely linked. The reason we have put it here is that sections 4.1 and 4.2 refer to the design and construction of the cooling lines and mention tests that have been done for all loops for Q&A. This goal of the tests in this section on the other

hand was to find the optimal parameters to operate the cooling plant and thus is mentioned in the cooling section.

* Fig 31 please explain better what you mean by "this set of points". not clear what you refer to → yes, fixed by mentioning ladder number.

* L1125 what modules were used? please clarify → added, "FPIX" modules

* Fig 32 there is a spurious text box on the right photo → yes, fixed.

* L1173 not clear what this means, please clarify → the wire-bonds for the hub address are introduced in the TBM section (L374). Added here "between the TBM and the HDI" to make it more clear.

* L1412 See my comment above about the cross talk → answered above

* L1425 the reference is wrong (it does not point to the Sec but to a figure) → fixed

* L1423 Would it be possible to add the range of variations? → The future pixel performance paper will include the detailed historical evolution of the thresholds. I have modified the text to mention the changes.

* L1472 please use the same number of significant digits (I would use 91% instead of 90.9% unless you know the value at the level of precision) → changed to 85.0%. We do know the value at that level of precision as it is given by counting bad ROCs.

* Fig 43 shouldn't you explain why L2 differs from L3 and 4 as well? → Well, it is not very relevant and quite arbitrary. L2 is on a completely different clock distribution than L3/4. So various settings are set differently. Also it was shifted to a non-optimal position in order to have an overlap with L1.

* Fig 44 Left is this really BPIX? why 3 rings? → yes, legend was wrong. Fixed.

* Fig 45 needs updating? Yes. Still waiting for updated plot.

* Fig 46 explain the range of fit → We just fit the central, linear, region in order to measure the deflection, so the LA in the center of the sensor. I have added a sentence.

* Fig 47 explain why efficiency of BPIX is systematically lower than FPIX →

Well, we make a big deal of the fact that dynamic inefficiencies are rate dependent. The data rates in L1 and L2 are, on the average, higher than in the fpix.The rate dependency is mentioned in the text. Eventually the radiation damage will also kick in affecting L1 and L2 more the average fpix, except if fpix is radially split.

Bibliography
* Reference 40: I doubt a journal will accept a private file as reference, Could that be stored somewhere else? → moved to CMS docdb.
* Also webpages (other than home pages companies) should have a date for traceability → ok, added.
* Reference 60 does not have a link → ok, added.