Georg comments

Dear authors,
congrats to this very nice paper with beautiful results from the early phase of the pixel R&D.
The language is fine for the most part, but some improvements are needed.
Please find my comments below.
Next time, please add line numbers to a draft!
Content:
Fig 5/6 upper:
I have not understood why both the data points have uncertainties (colored bands), as well as the horizontal lines of the expected signal!
You right! It's confusing. BTW, even the expected MIP MPV have uncertainties, ~ 1%

Fig 6 (lower plot) in general and
discussion of Fig. 6 in last paragraph of page 3:
To me it looks like "charge collection efficiency" and "particle detection efficiency" are confused/ mixed up here! At least it is unclear what you plot, how it is defined, etc.
Please explain how the plotted efficiency is defined!
Is this now the charge collection efficiency (as you say on p 3), or the particle detection efficiency (as you say on page 4)? These are two very different things!
Will do my best to make it clearer.
Nevertheless, Charge Collection Efficiency is obviously the ratio between the collected charge and the number of e-h pairs generated by the ionizing particle, therefore $MPV_{measured}$ / $MPV_{expected}$
Detection Efficiency is the usual particle detection efficiency.
The bottom plot of Fig 6 is for detection efficiency as specified in the caption.

Fig 8:
Do you understand the rather large inefficiency at 0 fluence, 150 V (middle plot) at the lower and upper right corner of the pixel? You say it's due to the bias grid, but don't say why it is so large at higher V.
This is what we measure! I have an explanation, but cannot rigorously prove it. Nonetheless, it is evident that the loss is due to the bias grid. Just compare the det eff on the side with bias grid with that on the opposite side.

Grammar/clarity
Abstract:
fluencies-àfluences
remove "To this extent"
the sentence …which allows for planar thickness 100 and 130 and 3D of 130 ….sounds funny.
Almost sounds like a restriction of Si-Si: Why is 3D in 100 mum not possible? Àrephrase
I agree! Funny!

Intro:
collected dataàdelivered lumi
ten times such a fluence à ten times this fluence
remove “in fact”
saturated velocity àsaturation velocity (2 times)
footnote 1: remove “indeed”
page 2:
underlying p-type layer à underlying p+ layer
fig 2 caption: between couples of àbetween two

“In addition, p spray implants can be Low, Medium and High, while p-stop rings can be Open or Close.”
àwhat does this mean? Explain!
I can add a footnote saying that “The actual p-spray concentrations are subjected to nondisclosure agreements”

I assume Low refers to low p-spray concentration. Then also mention the concentration if you can. Explain what Open or Close*d* means!
Hope that everybody knows what open and close p-stop ring means!

It is worth noting…àremove
number of junction/ Ohmic ànumbers of junction/ Ohmic
page 3:
Figure 5 shows the trend of the MIP-MPV as a function of bias voltage
àremove “the trend of” (and also in other places)
Remove “By contrast”
Sentence starting with “the MIP-MPV trends” rephrase to sth like àThe MPV for MIPs is shown in Fig 6. (top) as a function of bias voltage for three different fluences.

P4 first paragraph: the thick sensor àa thicker sensor
You say, that at high fluence for a thicker sensor a much higher Vbias is needed to get the same signal. While this statement is eventually correct, your plot does not really support this: The signal for 130 mum is always above the one for 100 mum. The problem is that you don’t compare the two thicknesses for the same fluence….
I agree! It’s obvious to me, but probably such to require more explanations.
I drop it.

àRephrase
Maybe rephrase:
This underlines the fact that at high fluences as expected at the HL-LHC thinner sensors can be operated at lower bias voltages.
, and, namely àremove “, and”
Fig 7: put a label on each plot with the thickness
P4, left column:

For what follows àFor the following discussion, it is important to keep in mind that the sensors have punch through…
P5 before the irradiation àbefore irradiation
The observed loss of efficiency is practically independent from either the active layer thickness, or the p-spray and p-stop isolation structure.
àThe observed loss in efficiency is practically independent of the active layer thickness and the p-spray and p-stop isolation.
I would remove "structure" since there is no p-spray structure
Later:
The complete uniformity of the efficiency can be recovered->Complete uniformity of the efficiency can be obtained by tilting…
Fig 8 caption:
Rephrase last sentence: "The efficiency clearly results affected by the bias structure…"
Conclusion:
The last part of the conclusion is unclear:
First you make rather general statements about the sensors studies for this paper up to 5E15. It looks like your statements ("performance is excellent") apply to planar and 3D.
àtry to make a more specific/ quantitative statement about the efficiencies than just saying they are excellent(?)
The last part of the conclusion is unclear:
In the outlook, you make a statement that full detection efficiency should be possible for 25x100 and 50x50 *3D* sensors at the target fluence for voltages below 300 V.
àwhat is the target fluence? I assume 50% layer 1, I.e. 1E16? Or something else? Where does the 300 V come from??? Is 100 % efficiency realistic?
And: I think you need a similar outlook statement for planar sensors to complete the picture

You completely right! I should rephrase it!

Katja comments

Comments and questions to content:

-In Section 3 you discuss a fiducial area cut for the larger (100x150) pixels, and you say that no such cut is used for the smaller pixels. I found this not very clear. Can you motivate this? Does this introduce a bias/penalty for the larger pixels? Can you comment on its relevance/quantify this? I think 1-2 additional sentences would help.
OK

-In Section 3 the concept of „threshold" is mentioned for the first time (top of right text column). This should be introduced much earlier, and explicitly. In Section 2, when you

explain that the PSI46dig was used, you should introduce its readout scheme, and say that it digitizes the charge for hits that exceed a setable threshold.
I don't understand. Threshold is quite well known concept. And all the info about the employed readout chip is properly documented in the cited reference.

-In the last paragraph of Section 3 you explain that you were limited in fluence by the radiation hardness of the ROC. This could/should come much earlier, maybe in the last paragraph of Section 1, because the reader wonders why you did not go higher in fluence.
OK

-The concept of annealing is never mentioned. In particular, how were the sensors stored after irradiation? Did they receive a controlled annealing? Or is it irrelavant for some reason? Isn't annealing the reason why the ROC must be bump-bonded to the sensor before irradiation, which is at the source of your troubles? This should also be explained early.
We followed the standard procedure: once irradiated the detectors were stored in a refrigerator at -20 C. I will add a sentence near the end of Sec III

-The loss of detection efficiency after irradiation seems to be a combined effect of 1) less charge collection in the sensor and 2) higher threshold needed in the ROC used for testing. I think this should be spelled out explictly and be explained and, if possible, quantified.
I think it's clearly explained and quantified whenever possible. Explicitly, the efficiency loss due to the punch-through bias is 6.4% = 96.4%-90%.

It also means that your efficiencies are conservative, which could also be brought out more prominently.
This is obviously clear. Anyhow, in the conclusions we stress the fact that if we could rely on a better chip with 1000 e thresholds we could largely improve the detection efficiency.

-You should say at the beginning what bias voltage will be used in ATLAS and CMS, so that people can have a fealing for how this compares with your measurement range and by how much in can be turned up later.
The final bias voltage will only be limited by the maximum value the PS system can provide (as long as the sensors can handle them). We will constantly increase the bias voltage to follow the damage of the sensors.

-For Fig. 6, why did you stop at different bias voltages for different sensors, or what did determine your bias voltage range in the first place? Why did you not go up further to see where you end with the MPV and efficiency? Add 1-2 sentences.
You absolutely right! This is what we were able to do. Nevertheless, as you can notice, our measured performance is pretty close to saturation, i.e. further increase of the bias voltage wouldn't have changed the picture.

-For the discussion of Fig. 6 it is not optimal that you show sensors that differ not only in received fluence, but also in thickness and presence of bias punch-through structures. It would be much better to compare the same type of sensor. I suppose the reason is that you do not have more dat
You know how things are working: a set of sensors are sent for irradiation; another set for tests on beam. When the irradiated sensors are back, only a fraction of them are still working (in our case, we had several failures due to the ROC!). Therefore, you to live with them.

Comments to phrasing, style, etc.
OK, will implements them!