*Dear Editor,*

*We would like to thank the referees for reviewing our updated paper and producing second reports.*

*We have considered all comments, and we have applied several changes to the original version of the paper to address the issues raised. Detailed responses to all the comments can be found below. For ease of review, the original question is written in black font, the author response is written in blue font, and any modification to the paper draft is written in bolded green font.*

*We are at your disposal for any further clarifications and/or additional information.*

*Sincerely,*

*Artur, Si, Cristian, …etc*

Reviewer #2: Please consider reviwing the following remarks:

P(age)#L(ine)#

Abstract:  
The efficiency and timing resolution before irradiation were found to be 100% and 30-40 ps, respectively. A "no-response" area between pads ………..  
The sensors are divided in 'efficient' part, and no response area. This leads to 100% efficiency but it is misleading if the concept is not clarified (as the geometrical efficiency is not this value).

To distinguish efficiency from the no response area, we will use the term “sensitive” area. We have changed the text to:

**“The signal detection efficiency and timing resolution in the sensitive areas”**

P2L58 … pile-up contamination ….  
I would consider a different adjective…………..

We have changed the text to :

**“in order to overcome the event reconstruction challenges posed by  the high rate of concurrent collisions per beam crossing.”**

P3L67 …. We also compare the uniformity of 50 and 80 m LGAD sensors, as well as the HPK and CNM sensors irradiated to an equivalent neutron flux of 6 1014 n/cm2. Detailed measurements of irradiated HPK sensors were presented in Ref.  
of 50 and 80 m thick LGAD sensors ……  
…. As well as…   
Here you put in the same sentence two different comparison (thickness and manufacturer) which I found confusing.

We have separated the sentences as follows:

**“We  also compare the uniformity of 50 and 80 μm LGAD sensors. Uniformity and and time  resolution of the HPK and CNM sensors irradiated to an equivalent neutron fluence of  6 × 1014 n/cm2 are also presented.”**

. an equivalent neutron flux of 6 1014 n/cm2…   
The unit of flux are n/cm2/s. This is integrated flux.

We have changed “flux” to “fluence”:

**“…equivalent neutron fluence…”**

P2L102 The primary beam (bunched at 53 MHz) consists 102 of high energy protons (120 GeV) at variable intensities between 1 and 300 kHz.  
Incomprehensible. Thought Hz is frequency, not intensity.

We have changed the text to:

**“The beam is resonantly extracted in a slow spill for each Main Injector cycle delivering a single 4.2 sec long spill per minute. The primary beam (bunched at 53 MHz) consists of 120~GeV protons. All measurements presented in this paper were taken with the primary beam particles.”**

Figure 4: Photographs of the HPK 50D-PIX 22 array sensor (top left), the CNM-W9HG11 22 array sensor (top right), the 50D-GR single sensor (bottom left), and the CNM W11LGA35 single sensor (bottom right) are shown.  
….. , the 50D-GR single sensor…….                     
……., the HPK 50D-GR single sensor …………

We have changed the text to:

**“Photographs of the HPK 50D-PIX 2 × 2 array sensor (top left), the CNM W9HG11 2 × 2 array sensor (top right), the HPK 50D-GR single sensor (bottom left), and the CNM W11LGA35 single sensor (bottom right) are shown.  ”**

P6L129 ……. with a different p+ dose of the gain layer to study the optimal parameters of the charge multiplication mechanism.  
I do not think you are studying the optimal parameters for charge multiplication (could e.g. maximum gain….), but rather the parameters functional to time resolution.

We have modified the sentence to reflect the optimization for fast timing detectors:

**“Four gain splits, identified with the letters A (lowest gain) to D (highest gain), were produced identical in the mask design but with a different $p^+$ dose of the gain layer to study the optimal parameters for fast timing detectors.”**

P7L161 …. about 130 W per channel.   
I hope it is much less!

We have corrected the typo to 130 mW per channel.

P8L184 … The above choices of timestamp reconstruction algorithms used are motivated by the result of a dedicated study of different timestamp reconstruction algorithms for each of the three readout boards.  
The choice of the above timestamp reconstruction algorithms was motivated by the result of a dedicated study of various algorithms for each of the three readout boards.

Corrected according to reviewer suggestion.

P9L209 … timing resolution …  
… time resolution …..

Corrected according to reviewer suggestion.

P9L212 …. is the same for all runs and requires that  the signal amplitude is between 160 mV and 320 mV….  
This information is redundant as the mip signal, or a signal calibration, is not provided.

We have changed the sentence to say:

**“The signal selection in the Photek MCP-PMT is the same for all runs and requires that the signal is consistent with a MIP corresponding to amplitude values in the range between 160 mV and 320 mV.”**

P9L213 … The signal selection for LGAD boards was optimized for each board individually, by selecting the MIP signal peak fitted with a Landau function.   
What does this mean? The signal selection here is not distinguishing real events from noise?

We selected signal events by requiring that the pulse amplitude is above noise. The noise for each board was slightly different and therefore the amplitude cut was adjusted for that. We have changed the text as follows:

**“Signal events in LGAD sensors  are selected such that they are above the noise levels listed for each board in Sec. 4.  ”**

P10L237  
Fig. 5 Can explain the error bars? Why they are bigger for the CNM sensor, and bigger in the y-coordinate, and increasing with it?

The dataset accumulated with the HPK 50D-PIX sensor was the largest among various sensors we tested in these measurements, hence the plots with this sensor have the smallest statistical uncertainties. The measurements presented in the upper and lower panels of this figure were obtained from different datasets and therefore do not have the same statistical uncertainties. The reason that the error bars grow with increasing Y-coordinate is due to the fact that beam is more centered at low Y-coordinate, and hence there are more events at lower values of the Y-coordinate. We have added the following sentence in the end of last paragraph in Section 6, before Section 6.1:

**“Measurements presented for various sensors were obtained from different datasets and therefore the statistical precision is not always the same. The reason that in some measurements the error bars are not the same across either X- or Y-coordinate is due to the fact that the beam does not uniformly illuminate the whole sensor area, and hence the number of events is not the same across sensor surface.”**

P10L239 … impacts its timing ….  
… impacts on its timing ….

Corrected according to reviewer suggestion.

P10L241 …. Is fit to a Landau …..  
…. Is fitted to a Landau …..

Corrected according to reviewer suggestion.

P11 Figure caption.  
Figure 7: t measurement across the X- and Y-axes of the HPK 50D-PIX sensor mounted on the FNAL board. The scans of pixels 1 and 2 along the X-axis, and pixels 1 and 3 along the Y-axis are shown. The pixel numberng scheme is defined in Fig. 4.  
Why the caption does not mentions the CNM sensor?

We have added the CNM sensor to the caption as follows:

**“Δt measurement across the X- and Y-axes of the HPK 50D-PIX sensor mounted on the FNAL board, and the CNM W9HG11 sensor mounted on the UCSC board. “**

P10L249 … A possible explanation for this effect is a small difference in the rise time of the pulses that originate from the passage of particles in the metallized and non-metallized areas……  
Too much hand-waving. Is the rise time difference well identified? I would no try to give explanation of the effect before verification.

The rise time did not show a statistically significant difference to explain this effect. We have modified the text following your suggestion:

**“This effect cannot be attributed to the algorithm used to time-stamp the events, since the same behavior is observed with the CFD and CDT algorithms. Furthermore, the same behavior is observed on all HPK sensor varieties mounted on KU board, as presented in Sec. 6.3. Further studies are needed to understand the effect.”**

P11L254 The results for the CNM sensor also show hints of this behavior, though the larger statistical uncertainties make it less significant.  
They seem to go in the other direction (concerning the effect of metal), or not?

After further discussion among the authors, we have confirmed that there is no metalized area in the CNM sensor, and we believe that the feature alluded to in the original text is unconvincing. Therefore we have changed the sentence to:

**“The CNM sensor does not contain metalized areas on its surface and we do not observe the same effect.”**

P11L270 … We define the width of the \no-response" area as the distance between the half-maxima of the two  fitted S-curves, as shown in Fig. 9.   
This definition contradicts the statement before that the sensors are 100% efficient (excluding the interdetector region). The area between the 100% to 50% efficient is obviously not fully efficient. But with the current definition, it is part of the active sensor area. This must be corrected.  
Besides, if a definition of the inter-detector region should be chosen, the 0.5 value is not good, as it should be chosen a value high enough to be used in experimental condition (90%, 95%, ???).

Indeed, at 50% efficiency, the measurement would not be very powerful. We have changed all relevant plots to show the distance between the two sensors at the 90% efficiency level.

Fig. 9 Some of the not efficiency points (0%) for both ch1 and 2 are significantly higher than 0. What is the noise hit occupancy for these sensors?

We know that in a small fraction of events (<1%) the reconstructed position of the track is wrong. The points outside the sensor area in these plots actually had hit the sensor. These events are not a result of noise in the LGAD sensor.

We have added the following sentence in the text to clarify this point:

**“Data points outside the sensor area in Figs. 9, 10 actually  had hit the sensor active area, but the coordinate of the track is incorrectly assigned, due to a small  probability (< 1%) to misreconstruct the position of the track.”**

Fig. 10: same comments as above.

Please see the answer above.

P18L342 …. The distribution of the  t between the reference timestamp and the timestamp from the HPK 50D-PIX sensor is shown in Fig. 16. We observe no significant changes in the behavior of the  t as the temperature varies.  
It does look like there is a shift in the x-direction of about 1mm from high to low temperatures? Is it possible?

**Admit in the text that there is a shift.**

**CRISTIAN/SI: plots need to be remade such that the boundaries are the same on Figs 15-17.**

Figure 16 and 17: please use the same markers in the two figures for corresponding temperatures.

**CRISTIAN/SI: Colors and markers need to be fixed in one of the plots so that all of them are the same in Figs 15-17.**

P21L367 .. The distribution on the right of Fig. 18 shows that the amplitude under the aluminum (periphery) is about 2.5 times larger than that without aluminum (center).  
How does it compare with the pre-irradiation spectra?

Unfortunately, we didn’t have unirradiated sensors at the FNAL test beam campaign. However, measurements were presented on page 12 of Nicolo Cartiglia’s talk at RD50 meeting in Krakow:

<https://indico.cern.ch/event/637212/contributions/2608659/attachments/1471224/2276633/Cartiglia_BeamTest_FNAL.pdf>

As can be seen in the presentation above, a similar effect, but of much lesser magnitude (~10%) is observed in the pre-irradiated device.   
  
  
  
  
  
  
  
Reviewer #3: Dear authors, dear editor,  
The manuscript addresses the timing resolution of LGAD sensors measured in beam tests. These detectors are foreseen for both CMS and ATLAS after phase II upgrade for precise timing measurements and are currently very intensively studied. The presented work certainly importantly adds to understanding of their operation and their performance (different manufacturers, electronics, geometries analysed). The amount of data and the work done is impressive.   
However, I have several issues that need to be addressed before the paper can be recommended for publication. They are listed below.   
  
With best regards,  
  
Your reviewer  
  
MA - major remark that needs to be addressed  
MI - minor remark that needs to be addressed  
OP - optional remark or comment from reviewer  
  
Abstract:    
"…factor of 2.5 when" ->  at lower voltages should be less so "… of up to a factor of 2.5 …"

DONE

Experimental setup  
L88: MI - "electronic time resolution"  it is not clear to me what you mean by that. Please make it clearer.

The definition of the electronic time resolution is added to Sec 2, third paragraph:

**“One of the main parameters of  DAQ system for precise time measurements is the “electronic time resolution”, defined as  the measured time jitter between two signals that are split from the same source. These  two signals are used as “start” and “stop” signals to electronic system measuring the  time interval between them. The electronic time resolution of the CAEN V1742 digitizer  was measured to be less than 4 ps, and thus, its impact on the timing measurements  presented in these studies can be neglected.”**

L102: MI - "at variable intensities between 1 and 300 kHz" - it is not clear what the intensity refers to.

This sentence has been changed, also at the request of the other reviewer. Now we have:

**“The beam is resonantly extracted in a slow spill for each Main Injector cycle delivering a single 4.2 sec long spill per minute. The primary beam (bunched at 53 MHz) consists of 120~GeV protons. All measurements presented in this paper were taken with the primary beam particles.”**

L109-L114 : MI - the  detailed description of trigger board is in a way to detailed as not all terms can be explained in enough details such as "front end board". I would prefer a simpler statement. It is not crucial for the remainder of the paper. In this context I think Fig.3 isn't needed, particularly for an already lengthy paper.

While we do see your point, the custom board was indeed critical to establish the synchronized readout of two independent DAQ systems, those of pixels and V1742. Only due to the achieved synchronous readout the precision tracking was possible, hence we think the board deserves to be mentioned in some detail. A similar readout board could have implemented in other facilities, to achieve similar goals. There is a similar telescope being built now at CERN, where they intend to use a similar system.

LGAD Sensor Properties  
MA: Table 1 is not referenced in the text. It is vital to list also the information about the gain of the detectors at given voltage, current, and break down voltage. It would help the reader to understand manuscript better. There has been no mentioning of the difference between the dose of W9HG11 and W11LGA35 from CNM. Are these two different only in geometry or also in dose?

Reference to Tab. 1 has been added:

**“A brief summary of the sensors dimensions and capacitances is presented in Tab. 1”**

We added a table summarizing the bias voltages and corresponding gains that the sensors were operated at during the test beam campaign. These are listed in Tab. 2 of the new paper draft, and added a the following text in Sec. 3

**“The list of sensors studied in this article, as well as the temperature and the sensor bias voltage used during their operation are listed in Tab 2.”**  
  
L131 : MI - Are they also identical in thermal budget ?

We are not sure what you are asking. If the question is whether they behave in the same way with respect of changes in temperature, then the answer is yes. If you are asking whether the generate the same amount of leakage current at the same gain, the answer is also a yes.

Readout Electronics  
L161 : MI - "power consumption of the board is about 130 W …" . is  130 W a typo? This is hard to believe.

Typo is fixed now.

Timestamp Reconstruction  
L176 : MI -  "decay time" is not an appropriate wording, maybe "fall time"…

Decay time is a commonly used term for describing the time taken for the amplitude of a pulse to decrease from e.g. 90% to 10%, and is used throughout the literature. We prefer to keep this terminology.

L194 : MI - "… by a linear function…" -> "… by linear dependence of timestamp on pulse height"

Done

Sensor Studies and Analysis  
L215 : MA  - The Landau function doesn't accurately describe the data. The spectra are broadened by noise and also by fluctuations in multiplication/gain. Therefore convolution of Gaus+Landau would be more accurate. Can you comment on uncertainty arising from this in your analysis?

**XXXX**

Cristian/Si: We can fit it to the gaus(x)landau and check the difference…

L227 : MA - What are the gains at these voltages?

This information has been added in Tab. 2 of the new paper draft.

L229-233 : MA -  What was the cut on LGAD signal to be considered as a signal (can be expressed in charge or gain fraction)?  If the cut is very low you may end up with 100% efficiency, but with very high noise occupancy, which would be unacceptable at real experiment. A value of noise occupancy would be also welcome.

The noise was measured using dedicated runs with no particles, and data collected using random triggers. The RMS of the amplitude of noise was measured to be 1.2 mV on FNAL boards. The cut used in the efficiency plots is at 20 mV, well above the noise level.

Added the following text in Sec. 6.1:

“Noise  values for different boards used in the experiments are listed in Sec. 4, and were measured  using dedicated runs with no particles, and data collected using random triggers.”

Figure 5: MA - How did you determine the error bars? Why are they larger for W9HG11 (bottom right plot)?

Error bars are evaluated as Clopper–Pearson intervals for calculating binomial confidence intervals. They are larger for the CNM sensor since there was much more data collected for the 50D-PIX sensor.

Added sentence in Section 6.1:

“Error bars  in all efficiency measurements are evaluated as Clopper-Pearson intervals for calculating  binomial confidence intervals.”

L241 : MA - Can you show the fit of Landau to the data? This goes hand in hand with the previous comment in L215.

**XXXX**

CRISTIAN/SI: Yes, we should show one of these fit examples.

Figure 6 : MA - Why is uncertainty of HPK smaller than for CNM? Can you convert MPV[V] into gain/charge? This conversion would be welcome already earlier in the text.

The gain vs voltage has been added in Tab. 2 of the new paper draft.

Added sentence in Section 6, last paragraph before Section 6.1:

**“Measurements presented  for various sensors were obtained from different datasets and therefore the statistical precision is not always the same. The reason that in some measurements the error bars  are not the same across either X- or Y-coordinate is due to the fact that the beam does  not uniformly illuminate the whole sensor area, and hence the number of events is not  the same across sensor surface.”**

L247 : OP - I think the reason for this difference in timestamp can be different signal formation in the  electrode for hit positions. The resistivity of n++ layers in not the same as that of Al electrodes, hence the difference in rise time. This holds for the section from L303-308.

The rise time did not show a statistically significant difference to explain this effect. We have modified the text following your suggestion:

**“This effect cannot be attributed to the algorithm used to time-stamp the events, since the same behavior is observed with the CFD and CDT algorithms. Furthermore, the same behavior is observed on all HPK sensor varieties mounted on KU board, as presented in Sec. 6.3. Further studies are needed to understand the effect.”**

L256/Figure 8 : MA - It is a absolutely necessary to present the reader with the gain values for the data in Fig. 8 in order to mislead the reader into believing that HPK in "intrinsically better"- see a comment L227. The difference is already in average electric fields achieved for both sensors, gain and also in capacitances.

This information has been added in Tab. 2 of the new paper draft.

L268 : MI - This looks like a ROOT code - please write erf full form. "floated in the fit" is colloquial - "were free parameters of the fit" is better

DONE

Figure 12 - OP:  Is it clear why time stamp is negative for 50D-PIX device in the gap region?

To clarify: this is not a plot of a time-stamp of a single detector, but rather the Δt=t1-t0 between the reference timestamp  in Photek 240 MCP-PMT (t0) and the timestamp of the LGAD sensors (t1). Due to different cable lengths used during different device testing, the absolute value of Δt is not relevant. Therefore, all measurements are calibrated such that the central, flat part of these plots aligns with the Δt=0. As a result, some of the points will appear negative.

Figure 14 / L318-322 - MA: What is Landau fluctuations limit for the 80 um sensor timing resolution? It seems that more charge for 80 um detector (although at lower gain) compensates for longer rise time and Landau fluctuations.  Can you show also "constant function" fit to data, so that it would be more evident?

The Landau fluctuation are worse for thicker sensors, for 80 micron is ~ 40 ps, e.g. see H. Sadrozinski, A. Seiden and N. Cartiglia, “4-Dimensional Tracking with Ultra-Fast Silicon Detectors”, 2018 Rep. Prog. Phys. **81** 026101.

Figure 15 - section 6.5 : MA -  There is one significant piece of information missing for the reader. It is true that lowering the temperature increases the gain (impact ionization depends strongly on temperature, also mobility), but also the breakdown voltage. Vbd shifts to higher values for larger temperatures hence higher gain at lower temperatures can be compensated by larger applied bias at higher temperatures. Hence, similar or only slightly worse resolutions can be obtained at 20C for non-irradiated sensors.

Yes, you are correct. The main reason we want to operate at lower temperatures is to avoid large currents for irradiated devices. This is reflected in the first sentence of Section 6.5 **“In order to maintain their optimal performance at the highest fluences envisioned at the HL-LHC, the LGAD sensors will be cooled to temperatures below -20C degrees. Operation at such low temperatures will allow to significantly reduce the leakage current, and additionally improve the timing characteristics of the sensors.”**

L355 : MA - There is no mentioning (even in a single sentence) of what radiation does to the LGADs nor citations to previous works on LGAD radiation hardness.

We added references 15 and 16, and the following sentence in Sec. 6.6:

**“Effects of neutron irradiation on LGAD sensors is documented in [6], [16] and [17]”.**

L358-359 : MA - missing citation for spectrum and flux.

Added citation [15] in Sec. 6.6

Figure 18 : MA -  What is the gain? What is the bias voltage of the W11LGA35 device?

This information has been added in Tab. 2 of the new paper draft.

L368 : MA - 2.5 times at what voltage?

This sentence describes the difference in signal size at the same bias voltage. Changed the sentence to:

**“The distribution on the right of Fig. 18 shows that at the same bias voltage the amplitude under the**

**aluminum (periphery) is about 2.5 times larger than that without aluminum (center).”**

L372 : MA - Looking at Fig. 4, I would say that the whole active area of HPK50D is without metallization.

**ASKED HARTMUT AND ABE TO CLARIFY**

L374 : MA - Please give the gain values - not only voltages. How close are the applied voltages to breakdown voltages?

This information has been added in Tab. 2 of the new paper draft. The breakdown voltage of HPK 50D sensor is about 650 V, and that of CNM is **XXX**.

Figure 20 : OP - Is the data set size for CNM detector smaller, hence larger uncertainty?

Yes. We added a sentence in Section 6, last paragraph before Section 6.1

**“Measurements presented  for various sensors were obtained from different datasets and therefore the statistical precision is not always the same. The reason that in some measurements the error bars  are not the same across either X- or Y-coordinate is due to the fact that the beam does  not uniformly illuminate the whole sensor area, and hence the number of events is not  the same across sensor surface.”**

L390 : OP - Do you have any idea how is this possible?

We think this is due to the same effect observed on non-irradiated HPK sensors with difference in time of arrival from metalized vs non-metalized areas of the sensor.

L391 : MI - The decrease of timing resolution for CNM at higher voltages can only be due to increase of noise, hence jitter, as Fig. 21 shows the increase of MPV. Please mention this or even better give the relative increase of noise.

**CRISTIAN/SI: I DO NOT SEE DIFFERENCE IN NOISE. MAYBE THIS IS DUE TO NON-OPTIMAL VALUE OF THE CFD THAT WE USED? CAN YOU LOOK AT SOME OTHER VALUE OF CFD, IS THE CONCLUSION THE SAME?**