Detailed Response to Reviewers

Reviewers' comments:

Reviewer #1:

Recommendation: major revision

This paper presents an architecture for high hit-rates based on the well-known column drain architecture. However, in contrast to other pixel readout chips with these architectures (e.g. FE-I3), the time stamp information are not stored per pixel and are only available per double column. This makes the system inherently paralyzable, and the claimed detection efficiency of > 99% with the stated hit occupancy questionable.

I might have missed important information and would like to ask the authors to answer the section "main question" below, where the issue is explained more in detail.

All the other points are minor. The paper is generally well written, has an appropriate length, is of interest to the community, and fits the journal.

Main question:

The "FASTOR" is a signal indicating if there are non-zero pixels in a DCOL. It is derived by an OR operation of all the pixel states. During read-out of the hits in a DCOL, no additional hits in this DCOL can get a time stamp.

This dead-time is about 150 ns (cluster size 3, 50 ns per hit).

However, the probability to get another hit in a DCOL during this 150 ns is roughly:

150 ns \* 24 / 25 ns / 512 = 28% (with 24 hits per 25 ns per chip [9]; 512 DCOLs)

I wonder, how the detection efficiency with such a paralyzable system can be > 99%?

Answer: Thank you very much for your comments. In response to your main question, we have checked the hit density listed in the Table 1 of [9]. The hit densities are 7.87 hits/bunch/chip, 7.54 hits/bunch/chip and 0.82 hits/bunch/chip, respectively for the Higgs, *W*, and *Z* operational modes. The bunch spacing are 680 ns, 210 ns and 25 ns, respectively. Based on these parameters, the probability to get another hit in a DCOL during 150 ns is roughly:

150 ns \* 7.87/ 680 ns / 512 = 0.34 % (Higgs mode)

150 ns \* 7.54/ 210 ns / 512 = 1.05 % (*W* mode)

150 ns \* 0.82/ 25 ns / 512 = 0.96 % (*Z* mode)

Therefore, the probability to get another hit in a DCOL during 150 ns is less than 1.1%. If there are two hits in the same DCOL during 150 ns, they will be read out to the matrix periphery sharing the same timestamp. Only one of them may affect the track reconstruction, thus the effect of this dead time of 150 ns on the detection efficiency could be minor.

Line-by-line:

l. 24: It is very much of interest to the reader to mention the CMOS fab providing the 180 nm CMOS process.

Answer: This work uses a TowerJazz 180 nm CMOS Image Sensor process. This project has a requirement on the TID radiation tolerance of 1 Mrad as listed in table 1, we could not specify the name of the foundry in the publication due to our confidentiality agreement.

l. 41: "CPS": Please use the more common term "MAPS", as your community also did in a previous publication about TaichuPix.

Answer: Changed “CPS” to “MAPS”.

Table 1.: Better mention the pixel pitch instead of the spatial resolution of the complete vertex detector (see comment to line 241).

Answer: Changed “spatial resolution” to “pixel size”.

l. 66 registers -> register?

Answer: Changed “registers” to “register”.

l. 67: While for TaichuPix the "column-drain" architecture might be "new proposed" it is with over 15 years not new. Therefore, I would propose to write: "column-drain, a proven architecture for high hit rate operations" or similar.

Answer: We have rephrased the sentence “The readout of the pixel array is built based on the “column-drain” scheme, a proven architecture for high hit rate operations.”

Fig. 1 "FIOF" -> "FIFO"

Answer: Changed.

l. 105: A citation for the FE-I3 could be added here, e.g.: https://doi.org/10.1016/j.nima.2006.05.032

Answer: We have added a citation [13] for the FE-I3 in line 106.

l. 117 "are implemented in pixel matrix" -> "are implemented in the pixel matrix"

Answer: Changed.

l. 118 "The matrix digital readout circuit" -> "The matrix's digital readout circuit"

Answer: Changed.

l. 119 "architecture of proposed" -> "architecture as proposed"?

Answer: Changed to “architecture as proposed”

l. 137 "ALPIDE-like version of in-pixel readout logic" -> "ALPIDE-like version of the in-pixel readout logic"

Answer: Changed.

l. 165 "performed 1000 times, and then analyzed the mean delay time" --> "performed 1000 times to calculate the mean delay time"

Answer: Changed.

l. 167 I guess the injection capacitance is extracted from post layout simulation? Maybe you can mention this and give an error?

Answer: Indeed, the injection capacitance is extracted from the post layout simulation. We have rephrased the sentence “we assume the capacity of the charge injection system to amount precisely its layout-extracted value of 172 aF”. The injection capacitance is implemented by the parasitic capacitance between two different metal layers. Unfortunately, the parasitic capacitance variation is unknown.

Fig. 5: I guess the deviation from a 1/x curve around 500e comes from the measurement discretization error (25 ns binning)? Maybe you can state this somewhere?

Answer: We’ve added a sentence in line 175 “The bumps in the red and blue curves in figure 5 are likely attributed to the measurement discretization error (25 ns binning).”

l. 194 This is unfortunate. I recommend making all important biases accessible and overwritable for future prototype chips.

Answer: Thanks for the suggestion, and we’ll adopt it in the future design.

l. 197 It would help the reader to state what the minimum threshold design value is.

Answer: We have added a sentence in Line 203 “The simulated charge threshold with the nominal ITHR setting is about 150 e-.”

l. 203 I have never heard about significant noise introduced by the charge injection circuitry. What noise do you mean here and how large is it?

Answer: The injection capacitance is a parasitic capacitance, and its variation affects the threshold. But the parasitic capacitance variation is unknown.

We have rephrased the sentence “The threshold dispersion measured in this work is the sum of both the fluctuations on the injected signal and the fluctuations on the threshold itself.”

l. 204 "which denotes the dispersion on thresholds of" -> "which denotes the dispersion of thresholds of"

Answer: Changed.

l. 207 I would expect in first order that different gain settings also lead to different noise level in first order. Is there an explanation, why this cannot be observed here?

Answer: The simulation result indicates that a higher gain design (i.e. S3 & S4) shows a lower noise level, but the measurement result does not. With preliminary analysis, we suppose the inconsistency between the simulation and measurement is related to the simulation model of the PMOS transistor, i.e. M6 in our case. The capacitance contributed by the bulk n-well of the PMOS transistor cannot be extracted by the simulation tool. Therefore, the capacitive load represented by M6 is not accurately extracted in simulations. Moreover, a normal layout for M6 is used in S1&S2, while an enclosed layout and different transistor dimensions for M6 are used in S3&S4. The effect of the enclosed layout is not included in the simulation model, which could lead to a deviation from the simulation result. To verify the analysis, we plan to include two different front-end designs with the same transistor dimensions but the only different layout for M6 in the future design.

l. 210 What is the conclusion about S4 that was designed for lower threshold dispersion, but does not show this?

Answer: An enclosed layout for M6 is used. The effect of the enclosed layout is not included in the simulation model, which could lead to a deviation from the simulation result.

l. 220 What electrical test result is meant here? Can you add a reference to simulated design expectations?

Answer: “the electrical test” refers to the test by applying a negative voltage step through the injection capacitance. We find that it’s improper to compare the test with a Sr-90 source to the test with injected charges. Therefore, we removed “This feature coincides with the result from the electrical test, as expected from the design.”

l. 230 One must be careful to draw conclusions about charge sharing with Sr-90 electrons, since they are of too low energy and heavily scattered within the silicon.

To draw conclusions about the cluster size from charge sharing and not by Sr-90 electron scattering the ratio thickness-over-pixel pitch must be small. What is the thickness of the sensor contributing to charge collection? Also, electronic crosstalk can lead to larger cluster sizes and this is not beneficial (no position information) and cannot be excluded here.

Charge sharing can be better estimated from more point- like charge depositions from laser/x-ray/high energetic beam measurements.

Better phrase the conclusion more carefully. Maybe something like "The measured cluster size is larger and is likely attributed to charge sharing, that is beneficial for the spatial resolution."

Answer: Thanks for pointing out the issue. We have rephrased the sentence “The measured cluster size is larger than one and is likely attributed to charge sharing, that is beneficial for the spatial resolution.”

l. 241 To my understanding, the complete vertex detector will have to fulfill the requirement to have 5 µm spatial resolution. This is affected by the intrinsic resolution of the pixel chip (pixel pitch, charge sharing), but also by the overall geometry (number of layers, distances) and the overall material budget. The chip itself can only fulfill thickness and intrinsic resolution requirements. I recommend rephrasing.

Answer: We have rephrased the sentence “The detector prototype needs to fulfill the requirement of 5 µm spatial resolution. The TaichuPix chip requires to achieve a 25 µm pixel pitch, 50 µm thickness and a hit rate of up to 36 MHz/cm2.”

l. 253 "The measured average cluster size in the range of 1.6 to 2.2 will advantage the position resolution". Again, I would be more careful. Maybe write "First indications of average cluster sizes in the range of 1.6 to 2.2 could lead to advantages in position resolution".

Answer: We have rephrased the sentence “First indications of average cluster sizes in the range of 1.6 to 2.2 could lead to benefit the position resolution.”

l. 255 "beam in a close future. " -> "beam in the near future. "

Answer: Changed.

Reviewer #2: The paper is well written and gives a good description of this development for CEPC. I have a few minor remarks mentioned below. If possible it would be good to give more explanation on the measurements shown in figure 5.

Answer: We have added more explanation in Line 171-175 and Line 183.

Some other detailed comments:

Line 41: Sentence better in plural: CMOS monolithic pixel sensors have become

Answer: Considering the suggestion from reviwer#1, we have changed the sentence to “The Monolithic Active Pixel Sensor (MAPS) technology has become extremely attractive for future high-performance tracking detectors.”

Line 56: TaichuPix, a dedicated CPS chip, …

Answer: According to the change in Line 41, line 56 has been replaced by “TaichuPix, a dedicated MAPS chip, is being developed for the first 6-layer vertex detector prototype for CEPC based on the baseline design”

Table I: include pixel size and time resolution.

Answer: We have included pixel size in the table1. Since this project has no requirement on time resolution, we have included the bunch spacing for CEPC in the Higgs, *W*, and *Z* operational modes.

Line 66 logic singular

Answer: Changed to “logic”.

Line 69: while timestamps are recorded at the periphery

Answer: Changed.

Line 71: a fast-or busy signal is delivered to the EOC

Answer: Changed.

Line 76: logic singular

Answer: Changed to “logic”.

Line 82: data multiplexer instead of data multiplier

Answer: Changed to “data multiplexer”.

Line 88: Can you expand on how many data bits per hit are transferred and the encoding (8/10 bit encoding or other) for the data transmission ?

Answer: As we state in Table 1, the maximum hit rate from the pixel matrix is 36 MHz/cm2. It translates to a hit frequency of ~120 MHz per chip (considering the full-scale TaichuPix chip). Since each hit pixel is recorded with a 32-bit word, the output data rate of the serializer achieves 3.84 Gbps. To keep some design margin, we plan to read the hit at a higher frequency than 120 MHz (e.g.140-160 MHz). In the TaichuPix-2, the readout frequency is designed to be 140 MHz, and thus the output data rate will be 4.48 Gbps.

We have rewritten Line 87-92 to make it clear “In the trigger-less mode, all the data in the 512 column level FIFOs are designed to be read out to the data interface at a frequency of 140 MHz. Each hit data includes 32 bits. A high-speed serial data transmission unit is designed to satisfy the largest data capacity up to 4.48 Gbps. An 8b10b encoder is available in the second prototype named TaichuPix2. When enable the 8b10b encoding, the data compression function [9] should be activated to fit the data transmission speed of the serializer.”

Line 102: n-well to p-well spacing is 2 or 3 um, refer to table 2.

Answer: Changed.

Line 105-106: give reference for FEI-3 and ALPIDE.

Answer: We have added references [13] and [14] in lines 106-107.

Line 112: both schemes

Answer: Changed.

Line 129: first instead of firstly

Answer: Changed.

Line 159: 3.2 Pixel properties

Answer: Changed.

Line 162: fix sentence: The timestamp of hits was recorded …

Answer: It has been changed to “The timestamp of hits was recorded with a resolution of 25 ns by the EOC circuit, and then was sent out to the DAQ system generating a FASTOR timestamp.”

Line 165: was then analyzed to extract the mean delay time.

Answer: It has been changed to “For each injected signal value, the test was performed 1000 times to calculate the mean delay time.”

Line 171, 179: it would be useful to also mention the power consumption per front end, and not only quote power densities.

Answer: We have added the bias current of the front-end in Line 177 and 188.

Line 177-179: “The time walk (variation of the delay as a function of the injected charge) varies with the bias current of the front-end (480 nA, 327 nA, 168 nA), which leads to different analog power consumption.”

Line186-188: “the pixels can be set with an analog power of ~100 mW/cm2 (with the bias current of ~327 nA for the front-end) to be compatible with the default trigger window.”

Figure 5: is there any idea why there is the bump in the red and blue curve in figure 5? It would be good to comment in the text.

Answer: We’ve added a sentence in line 175 “The bumps in the red and blue curves in figure 5 are likely attributed to the measurement discretization error (25 ns binning).”

Line 253: will benefit the position resolution

Answer: We have rephrased the sentence “First indications of average cluster sizes in the range of 1.6 to 2.2 could lead to benefit the position resolution.”

Line 255: in the near future.

Answer: Changed.