Dear Reviewers

Thank you for the very careful reading of the manuscript and the kind words, comments and questions.

Here below the replies/answers to questions and remarks and actions in blue that were taken.

See you Peter Kluit

Reviewer #1: This paper deals with the construction and the performance evaluation of a Time Projection Chamber (TPC) module featuring 32 GridPix detectors. Tested at DESY using electron beams of 5 and 6 GeV/c with magnetic fields 1 Tesla and without magnetic field, the TPC achieved impressive particle identification (PID) resolutions with two different analysis methods: (1) the dE/dx Truncation Method and (2) the Template Fit Method achieving respectively 3.6% of 2.9% of resolutions for the same conditions (1 T data for electrons).

Both methods demonstrate the GridPix detector's effective capabilities in particle identification.

The single-electron efficiency remained largely stable even at high hit rates, highlighting the GridPix's capabilities for effective tracking and PID, particularly in distinguishing pions from kaons at various momenta in proposed future experiments.

This paper is written is written in a well understandable English.

A minor revision has to be done and comments have to be taken into account before publication.

Suggested revisions:

* In Abstract:

- l. 30: Add a space: "... using the B = 1 T test beam..."

-> Done

* In § 1. Introduction

- l. 39: "... was constructed. A GridPix has a very fine ..."

-> Done

- l. 48: Precise the 'small amount of oxygen and water vapour'. (i.e. < 400 ppm).

-> Changed text "amount of oxygen ($<$ 620 ppm) and water vapour ($<$ 7000 ppm). "

- l. 56: "The time-over-threshold (ToT) is related ..."

-> Done

* In § 2.

- l. 66: Add a comma in "In a GridPix detector, one can measure both ..."

-> There is now a more elaborated text (see answer to Reviewer #2)

- l. 80: Need to say how is made the calibration (add a reference...).

-> Answer: per chip a constant weight is applied.

- Changed text: leave out "calibrated" that suggests a (more) complicated procedure.

- l. 99: Add a comma in "At distances above approximately 10 pixels, the

distribution ..."

-> Done

* In § 3.

- l. 167: remove one "runs"

-> Done

- l. 168: "Runs with high rate and low rate were taken for beam momentum of 6 GeV.

* In the Table 2. the column 'E_e' should be added for more clarity (instead of run numbers).

- l. 179: "... two high rate runs (121.7 and 122.5 Hz) taken at beam momentum of 5 GeV/c ..."

-> Answer

1) Added a column to Table 2 with the electron beam momentum (P).

2) replaced the sentence with "The analysed runs with a trigger rate above 90 Hz were taken at a beam momentum of 5 GeV/c. The runs that were taken at a beam momentum of 6 GeV/c have a trigger rate below 3 Hz."

- l. 187: add the comma after "... (lower half). The rate ..."

-> Added a dot "."

* In § 5.

- l. 228: "... has been measured in the run of B = 1 T data ..."

-> Done

- l. 246: "... and in z = 1 mm".

-> Replaced by "z equal to 1 mm"

- l. 254: "... radius of 155 pixels (8.5 mm) ...

-> Done

- l. 259: Remove the Em dash "-" in the sentence.

-> Done

- l. 277: For better readability, it could be better to replace track length 'tlength_0' of 1441 mm by explicit term like 'track_{\rm length}^{0}' and track length 'tlength' by 'track_{\rm length}^{0}'.

-> Indeed, we replaced "tlength" by $t_{\rm length}$" etc.

Comments:

- Avoid wordiness "In order to ...", please replace it in l. 74, l. 165 and l.234 by "To .."

-> Done

- The runs numbers are not relevant. Everything can be replaced by electron beam momentum (E_e) at 5 and 6 Gev/c.

-> The run numbers are relevant in case one wants to reproduce the results. We just mention them for this purpose and list them in the Table. The beam momentum is added.

Reviewer #2: The manuscript describes measurements with a GridPix-based gaseous

detector. A measurement campaign at an accelerator site has been conducted and the results

from the analysis are presented.

The topic of the paper is very important for the future of particle detectors.

The concept has the potential to significantly improve key parameters of TPCs, e.g. the energy resolution and therefore the particle identification capabilities.

On the one hand the manuscript is well written, I also had the impression that the measurements were carefully conducted and analysed.

On the other hand, I found important aspects of the paper not well explained if at all.

For example the concept of dN/dx is not well explained, which is one of the core topics of the paper.

I assume that the authors could improve the paper by adding a few sentences of explanation for the mentioned paragraphs.

Therefore, I ask the authors to provide a revision of the paper.

Major comments:

line 64ff:

The difference between dE/dx and dN/dx is one of the main topics of the paper.

In my opinion, the dE/dx is a concept that is widely used in the community.

The dN/dx concept is a rather new approach and one of the highlights of this

paper.

Therefore I suggest to give a brief introduction to dN/dx and to point out why

it is better (compared to dE/dx) and why GridPix are capable of measuring it.

->We added to the text:

" In a classical TPC with pad read out, the charge is measured and used to estimate dE/dx using a truncation method that reduces the Landau tail. In a Pixel TPC based on a GridPix detector, the number of ionisation electrons is measured with high granularity as a function of the distance along the track.

The number of ionisation electrons is proportional to the energy loss of the particle in the gas and has a Landau-like distribution. However, if the charge is used to estimate the energy loss - as in the pad read out scheme - the large Polya fluctuations from the gas-amplication process will contribute to the measurement. Due to the digital read out of the GridPix, these fluctuations do not contribute. Due to the fine granularity, a GridPix is sensitive to the primary clusters that are described by a Poissonian distribution. The best PID performance is expected to be reached by counting the clusters and measuring dN/dx.

A Pixel TPC is sensitive to dE/dx by counting the number of electrons and to dN/dx by exploiting the distance along the track."

Line 76f:

Why do you combine tracks to get a new 1 m long track?

This seems to be an important choice but it is not motivated at all.

-> Changed text to:

"The chosen length of 1 m is typical for a TPC and allows for extrapolations to different detector size."

One could also use a shorter or longer track length The disadvantage of a shorter distance e.g. 30 cm is that one gets slightly more sensitive to the Landau tail. Extrapolation to other track length values can be done with Eq (4) in the paper.

Line 115ff

It is not clear to me how you can "define" the response of a MIP by dropping 30%of the hits. Can you explain this to me? And how did you decide which hits to drop?

Was the percentage of dropped hits varied, e.g. between 20% and 40% or was it fix?

-> Answer

The procedure is as follows: We draw for each hit a flat random number[0-1] and if this is below 0.3, the hit gets dropped. So we use a fixed/constant number of 0.3.

In the text we use the word "define" because a real MIP does not have on average exactly 70% of the Eloss of an electron of 5/6 GeV. One could also say 70% is a convention for a MIP.

-> Changed text to:

"By dropping - randomly chosen - 30\% of the hits associated... ."

It is not clear to me Line 127ff:Why could one argue that the results from the template fit method will move more towards the results of the dE/dx truncation method when more diffusion occurs?

-> Answer

The dE/dx truncation method is insensitive to the diffusion (the fine granularity of the distance distribution is not really exploited). The slope method is expected to be sensitive to the transverse diffusion because the multi-electron peak at low distance -diffusion will spread more to higher distance values and will reduce (a bit) the sensitivity of the fitted slope. So the spread on the slope will increase.

This effect was studied and confirmed on a small toy MC, where additional smearing in the transverse plane was added to the hits.

->Changed text to add reason:

"One might argue that with more diffusion the results from the template fit method will move more towards the results of the dE/dx truncation method,

because the statistical error on the slope will increase."

Line 151:

Single-electron efficiency is not well defined. I would suggest calling it the "Single-electron detection efficiency".

Furthermore, I would clarify that you are referring to ionization electrons and not electrons from your beam.

-> Done

We changed the text to "Single-electron detection efficiency"

We start now the section by: “The detection efficiency of single ionisation electrons in…”

Line 172ff:

Why do you separate the results into upper and lower half? Was there a difference in the chips, e.g. in the resistive layer?

-> Answer

One would not expect a difference between the two halfs of the module, so one expects the same numbers and behaviour for the two regions. The chips are the same but they will vary (a bit) e.g. in mean <ToT>.

Line 189ff:

What is your explanation that the single electron efficiency rises by 1% if the magnetic field is set to 1T? Could it be that the measurement uncertainty on the value is greater and not negligible (as stated in line 182)?

-> Answer

Indeed, the 1% change is within the expected uncertainty (the temperature and pressure for the 0 and 1 T runs were also different)

Line 226:

To me, it is not clear what a "single-electron resolution" is.

-> Added to text: “The resolution is measured using the single-electron residuals to the track”.

More details about the single electron resolution can be found in our part I paper (ref [3], doi:10.1016/j.nima.2025.170397).

Line 273ff:

Here, you argue that the template fit yields a significantly better energy resolution compared to the truncation method. Earlier (in lines 127ff), you argued that with increasing diffusion, you would expect that these two methods yield the same result.

In the ILD, the maximum drift length is 2350 mm, therefore diffusion dominates. Why do the two methods still differ?

-> Answer

Indeed, we argued that increasing the diffusion would bring the performance of the slope method closer to the truncation method.

The diffusion in xy is always dominant (both in the module and ILD). Note that ILD will run at 3.5 T and that will bring down significantly the transverse diffusion to 25 microns/sqrt(cm). What matters is the product of Dxy sqrt(L). And in that respect the ILD and test beam conditions are not that different.

One can say for sure is that the performance will lie in between the two methods. Where exactly, one can only answer by using a model based on these key performance numbers. E.g. for angles with a shorter drift length the performance will be closer to the slope method results.

Why do the two methods still differ?

-> Answer

Because we extrapolate the test beam results (so the two resolutions) that represent the best and the worst case scenarios. The real situation will be somewhere between.

Not mentioned:

The quality of your gas was rather poor (according to your previous paper, you had up to 620 ppm oxygen and up to 7000 ppm water in your detector gas). How does this influence your results especially with respect to the energy resolution? I would assume that especially due to the high oxygen content, you suffer from electron attachment? Has this somehow been taken into account?

-> Answer

That is a very correct observation. For a large TPC gas tightness is very important. The drift distance in the module was rather short (max 35 mm) so the impact of electron attachment is rather small. No corrections were made for this effect.

The results on the PID performance include all possible single-electron detection efficiency losses (due to limited acceptance, different chips operating points, electron attachment etc.).

Minor:

Line 76f:

How do you define a hit?

-> Answer

A signal in a pixel over threshold

What is an event?

-> Answer

Hits in the detector in a time slice (around the trigger time)

Line 147ff:

You state that these measurements show the best dE/dx and dN/dx resolution of

TPCs at atmospheric pressures. Could you confirm this by citing the corresponding values for some TPCs, e.g. ALICE or sPHENIX?

-> Answer

We added to the text a reference to the performance of the ALICE experiment:

"E.g. the ALICE experiment published a \$dE/dx\$ resolution of 5\% for a track length of 1.65 m \cite{YU201355}." https://doi.org/10.1016/j.nima.2012.05.022}

Line 228:

Phi is not defined.

-> Added to the text:

" defined as the azimuthal angle in the \$xy\$ plane,"

Line 290ff:

How do these numbers compare to other TPCs?

-> Answer

1) We added in section 7 more context: a reference for ILD dEdx performance to line:

"It is clear from the above that a GridPix Pixel TPC in ILD will provide powerful particle identification".

Added:

"Studies have shown that for a pad readout in ILD, the PID resolution is expected to be 5\% \cite{einhaus2019}." (https://arxiv.org/abs/1902.05519)

2) For other TPCs we now quote the ALICE dEdx performance in section 2.

Table 2:

Include beam momentum in the table.

-> Done