# Table of Contents

FPGA-based tracking for the CMS L1 trigger using the tracklet algorithm

- General information ................................................................. 1
- First version of the draft ............................................................ 1
- Second version of the draft ........................................................ 1
- Third version of the draft ............................................................ 1
- Fourth version of the draft ........................................................ 1
- Comments to the first version of the draft .................................. 1
  - Comments by Rainer, 01.07.2019 ........................................... 1
    - General comments (Type B): .............................................. 2
    - Detail comments (Type B): ............................................... 6
    - Typos (Type A): .............................................................. 6
  - Comments by Viktor, 03.07.2019 ........................................... 6
    - General comments: ........................................................... 6
    - Detail comments (Type A): ................................................. 7
    - Detail comments (Type B): ............................................... 10
  - Comments by Malte, 03.07.2019 ........................................... 11
    - General comments: ........................................................... 12
    - Detailed comments: ........................................................ 12
- Comments to the second version of the draft ............................. 17
  - Comments by Katja, 26.9.2019: pdf pptx .............................. 17
    - Type B comments: ............................................................ 17
  - Comments by Rainer, 02.10.2019: ........................................ 17
    - General comments (Type B): .............................................. 17
    - Detailed Type B comments: ............................................. 18
    - Type A comments: ............................................................ 18
  - Comments by Malte, 04.10.2019: ........................................... 19
    - Type B: ................................................................. 19
    - Type A: ................................................................. 19
  - Comments by Viktor, 04.10.2019: ........................................... 20
    - Type A: ................................................................. 20
- Comments to the third version of the draft ................................ 21
  - Comments by Katja, 18.10.2019: pptx .................................... 21
FPGA-based tracking for the CMS L1 trigger using the tracklet algorithm

General information

- Contact person: Peter Wittich
- Reviewers: Malte Backhaus, Rainer Mankel (chair), Viktor Veszpremi, Katja Klein (ex-officio)
- Published as CMS Note-2019/005
- Published as 2020 JINST 15 P06024

First version of the draft

- First version of draft, as received on 7th of June 2019: pdf

Second version of the draft

- main_v03.pdf: second draft to committee

Third version of the draft

- main_v04.pdf: third draft to the committee

Fourth version of the draft

- main_05.pdf: fourth draft to the committee

Comments to the first version of the draft

- Comments by Katja, 30.6.2019: pdf

These comments were incorporated in the text.

Comments by Rainer, 01.07.2019

Dear authors,

thank you very much for preparing this highly interesting manuscript. I find the paper generally very well written, and it makes a very good read. Consequently, I have hardly any comments to the wording, writing style etc., and only list some obvious typos at the end.

General comments (Type B):

- What about the hybrid track finder... is it an evolution or a competitor of the tracklet algorithm? If the former, is this manuscript already slightly outdated in this respect?
The hybrid algorithm includes the pattern recognition part of this paper with a Kalman Filter fit. This manuscript is intended to document the demonstrator and the standalone algorithm.

- Does the method work with a realistic detector geometry, where each module can have alignment-related individual shifts or tilts, or does it assume some symmetry?

Yes, the algorithm works in global coordinates which allows us to incorporate shifts and tilts.

- In the whole paper one misses an assessment of the ghost tracks, i.e. wrongly reconstructed tracks or fake tracks that the method produces. A related information is the track purity, i.e. the fraction of "correct" hits. How much do ghosts or fake tracks add to the trigger rate?

We have added figures of track rates to the paper. The "fake rate" was always a big discussion in relation to the tracker TDR as it depends on the MC truth matching definition. As a consequence we never include fake rate plots in the TDR. We have added some discussion to the paper, but the key points are: overall fake tracks have a small contribution to the overall track rate, which is dominated by the rate of genuine low pt tracks. However, fake high-pt tracks (i.e. rare w.r.t. total number of tracks, but sizeable fake rate at high pt) can completely spoil e.g. track MET and so depending on the use-case, applying stringent track quality cuts are important.

**Detail comments (Type B):**

- L 61: "L1 trigger ... is applied to decide whether ... should be saved for further analysis":

   This is misleading, since the full trigger chain including HLT decides whether the event will be stored. Or do you mean the HLT with "further analysis"? Then the word "saved" is misleading... perhaps "passed to the High Level trigger for further analysis". Yes, "saved" here as opposed to dropped. You are assuming "saved" means "saved to tape"? Changed to 'considered'

- L 65-66: What about the ATLAS HTT?

   ATLAS plans are still very unclear, and they do not plan to do full detector, 40 MHz readout, regardless.

- Fig. 1 and L 73: does "L1 Tracking" refer to exactly the method that is reported in this paper, or is there any difference? It probably corresponds to what was assumed in [3], but is it comparable?

Yes, what is described here is an evolution of the ideas in [3].

- L 99: "while only the PS modules...": this means the combination of two PS modules, right?

   no, each PS module gives a precise z measurement with its pixel side.

- L 104: "remaining": this seems (strangely) not to include the stubs of interest?

   "Remaining" here is precisely the stubs of interest; i.e., it is those not filtered.

- L 107: You could indicate what development makes it possible now... FPGA development?

   yes, this is what is described in the following paragraphs.

- L 132: do you mean r \phi ? Otherwise the unit is wrong.

   added r
- L 136: the required computing power

added

- L 174-176: There seems to be less redundancy in the barrel/end cap transition region... is this intended? Seeding with pairs of 2S layers is not very effective, so this is what our detector design gives us.

- L 184: Is this the fit by Ricardo et al? In this case, reference would be in order.

No, this is a different fitter that was developed simultaneously and independently.

- L 188: You already discuss duplicates but have not yet mentioned how duplicates arise, e.g. from the duplication discussed in the first paragraph if Section 2.2.

Added a brief explanation about redundant seeding layers.

- L 190: What is the minimal number of stubs required?

At least 3 independent stubs are required, else the track is called a duplicate.

How is z handled in a 5+6 seed?

The beam spot is used as a constraint and wide projection windows are assumed.

- L 211-213: It would be interesting to see a histogram of the processing time, or equivalent the number of clock cycles.

The time is deterministic to the clock tick. The time to the appearance of the first track is deterministic. Table1 contains this information. Maybe that is sufficient?

Is there a full clock-level simulation?

There is a model that is built up that matched measurement to within 0.4% (see again Table 1).

- L 226: In this context it would be helpful to indicate the \phi range traversed by a typical track.

A "typical" track would be a track with 2-3 GeV in pt, since for pileup events the pt spectrum is a steeply falling spectrum. A 2 GeV track spans maximally 2/28 of the detector in phi, so a 2 GeV track spans a phi range of \pi/28.

- L 236: What difference does it make whether the z bins are "internal" or external to the VM?

- L 242: Introduce Verilog

Done

- Fig. 4: Where does the duplicate removal happen?

after the track fit. I presume you are asking why it's not included here in this figure. Katja also asked. we should discuss. The figure just gets too messy...

- L 255: Introduce HDL

Done earlier now
- L 260ff: The roles of "sector" and "virtual module" are not always clear.

TODO

- L 308: "a few clock cycles": this is to be compared with the 108 cycles that are available, right? (From 25ns * 240 MHz * 18) Perhaps point this out.

No this affects the latency, not the amount of time available for processing.

Can you give a histogram of the number of clock cycles that are needed? This is only implicitly reflected in Fig. 7. We refer again to Table 1? Maybe it needs to come earlier?.

- L 341: It is not clear what goes into this simulation. Is this G4, at which PU, which event type, is detector noise included?

Full CMSSW. This is mentioned on line 362.

- Fig. 7: Why does the blue curve fall off more steeply? Could you plot this for different PU scenarios?

This figure has been updated and clarified. What is shown is the number of stub pairs that are considered vs are accepted.

- L 362: Since you mention Geant simulation here, was it something else in the previous section?

No it is always CMSSW.

- L 370-378: Is the math connecting track parameters to hits not significantly different for barrel and end cap regions? And thus different in both regions considered here?

yes, the math is different as is the way the hits are organized. Not sure if you are suggesting a change here?

- Fig. 11: Why compare histograms... would it not be more informative to plot the differences of track parameters between C++ and HDL, track-by-track?

we do that too, but those histograms are boring to look at (delta function at zero). The histograms here are intended as a visualization. The text in lines 385ff is intended to explain this.

- Table 1: The bottom line numbers like 1500 and 3345 ns sound very large. How would they translate to the design system?

the specification was 4000 ns as described in the abstract.

- L 443: Do you really mean azimuth here (which is the \phi direction)?

no we mean polar angle fixed

- L 505: How does the maximum of 210 stubs compare to the average?

TODO

- L 523: "single object events": do you mean just one particle? Is this kind of event used for the first time now, or has it been used in earlier sections?

Yes, just one particle. Re first time, the performance plots really start here so most of the previous figures...
don't use MC data. Figure 11 is ttbar + PU=200. In figure 7 we don't specify the data set; here we are just
discussing the qualitative behavior of one part of the algorithm. Do you want us to make a change?

- Fig. 15: It is hard to distinguish the graphs. Make clear whether these are just seeds, or full L1 tracks.

added "final" to describe full tracking efficiency. The point of this figure is to show that there is redundancy in
the seeding, hence it being hard to distinguish the graphs is kind of the point :-). I have added a sentence to the
caption to this effect.

- Fig. 16: Not clear what \sigma means. The plots look like these are parameter residuals, reconstructed minus
true (gen level?), and in this case the nomenclature is plain misleading. You would get the resolution by a fit
to these residual distributions, and then it would be good to quote the numbers resulting from such fits.

These are the residuals indeed. The figure text has been updated.

"Events without pileup": does it mean ttbar events without pileup, or even single object events? One notes that
there is no "background" resulting from ghost tracks etc in these distributions, and the latter would explain it.

(I presume you are talking about the caption to Figure 16 here.) These events use a particle gun to generate a
single muon. This is the 'best case' scenario. We also use samples with single mu + PU, single e, single pi, and
ttbar. We show distributions from ttbar + PU in other places, e.g. figure 19. Do you want us to change
anything here?

- Fig. 17: Do you have an expectation for this number of tracks, based on seeding scheme etc?

since this is single muon w/o PU we expect exactly one track at the end. The peak at 3 is presumably from the
barrel region with its triple redundancy seeding. I have added a comment that we expect one track after
duplicate removal to the text.

Why \( \geq 1 \) in abscissa title?

to draw attention to the fact that we suppressed the zero bin.

- Fig. 18: What kind of events are used here?

Updated caption to say 'muons or electrons ... in single particle events ... ' rather than just 'muons or electrons.'

Please add an explanation why the turn-on of the efficiencies is like this

I have added a sentence explaining the relative difference between e's and mu's.

- L 563ff: Has the effect of hot channels ("bad components") already been considered?

Numerous performance studies were performed as part of the "stress test" in 2018. For example, we
considered looked at the impact of stuck bits, as well as dead modules. For the "stuck bit" case, we used ttbar
+ PU200 events and added 5% (25%) of stubs from the 1st event to every subsequent event. The impact for
5% (25%) stuck was: (a) efficiency drops by 0.06% (0.4%) and (b) the track rate increases by +0.4% (4%).
That is, the system is able to handle such scenarios, and we need to go to rather extreme levels of stuck bits to
have significant impact from it.

- Fig. 19: The small PU dependence is really surprising. It looks as if the system could easily handle PU500
and more. Does it mean the system is over-designed? Where does it break down? Or are important effects not
taken into account?
For the 'stress test' we looked into this and found that the readout electronics will start killing stubs between 275-300. For the track trigger it's not as much pileup (diffuse low pt jets everywhere) as localized higher density (inside jets) that gets you in trouble. One of the tracklet 1.0 -> 2.0 changes was to re-shape the VM such that the load is balanced more evenly (i.e., make them long and skinny so they don't look like a jet.)

Ultimately we want to do dynamic load balancing to even out the processing (mostly we provision for the tails of the occupancy) and then use the extra processing power for things like displaced tracking and/or extra tracking for electrons...

- L 597: Why does adding the -z side increase the occupancy? Is it because both sides are in the same "long" virtual module? And if yes, why is this increase of the occupancy harmless, given that both sides interfere now with each other? In general, tracking performance can degrade even non-linearly with occupancy.

TODO

Typos (Type A):

- L 118: were considered DONE
- L 128: are presented DONE
- L 186: radius DONE
- L 208: the time DONE
- L 485: Figure number missing TODO
- L 493: space after "parameter" DONE
- L 495: 9 -> nine DONE
- L 530: Figure 16? no, but fixed
- L 589: implementation DONE

Comments by Viktor, 03.07.2019

Dear authors, I hope you will find my comments below mostly useful.

General comments:

- the definitions of the sector are somehow a bit confusing, because it is usually dictated by the read-out electronics, but in this case it is also a variable parameter of the L1 trigger system

Since the track trigger is downstream of the readout electronics, there is some freedom to re-arrange sectors if it is beneficial. We explored this freedom.

- punctuation inside or outside quotation mark could be systematically checked, e.g. L96 vs L162

Based on Katja's comments we tried to unify around having punctuation outside quotes. I will double-check to make sure we are consistent.

- single and double quotation marks seem to be used inconsistently; also, sector in Figure 2 is italic while following the logic of the text it should be in quotes fixed

Typos (Type A):
- beyond effects of misalignment and fake tracks mentioned by Rainer, the effects of stub inefficiency would be interesting to consider

In as much as the CMSSW simulation includes a simulation of the FE this is included, though it's not really complete now. CMSSW has a hard time with the boxcar average in the FE electronics... For the 'stress test' we looked into this too.

**Detail comments (Type A):**

- L17 the higher lumi poses major challenges not the upgrades, perhaps could write something like "higher inst lumi in the upgraded LHC", also makes "under these conditions" in L19 refer to something updated

- L19 "charged particle trajectories" --> "charged particle tracking" We prefer not to use 'tracking' in the abstract as it is jargon.

- L20 "This paper presents... FPGA solution" --> sentence may be deleted, it is repeated in L25 rewritten

- L22 "Tbps": Tbit/s Based on Katja's comments this is Tb/s.

- L23-24 could implicit definition and use of "stubs" be omitted? E.g. "pattern recognition on these measurement points" We use stubs a lot and we have an explicit definition here -- we prefer to keep it

- L24 "producing the list of trajectory parameters within 4 us" sounds clearer? DONE

- L26 add "in this paper," after "is described" rewritten following earlier comment

- L27 move "based on Xilinx Virtex-7 FPGAs" to the end of previous sentence to make it sound more like the "demonstrator system" is what "meets timing ... performance" rewritten

- L40,41 "Demonstrator" could be deleted, since L39 already has it DONE

- L46 "performance" replaced with more precise word choice, E.g. "resolution"? Performance is already used in L44 and L45 rewritten

- L47 "performance" --> "effectiveness"? rewritten

- L57 "collide proton bunches every 25 ns: once upgraded, each of these" --> "collide proton bunches in every 25 ns. Each of these" rewritten

**Figure 1:** PhaseII but Run 1; should use Phase 2 or Phase-2 as later in the text (Not sure about the convention.) these figures are from the CMS Phase-2 Technical Proposal, we cannot change them here (but in general it should be Phase-2 yes)

**Figure 1:** has not been defined Now pseudorapidity, and PubCom doesn't want us to define this any more ...

- L65 sentence repeats much of the previous, perhaps something along: "this will be the first time a solid state detector is used in this way at the LHC"? Also avoids the use of "include/integrate in trigger" in three consecutive sentences removed the previous instance instead

- L73 word "incorporating" could perhaps be deleted DONE

**Figure 2:** move "(left)" to after "detector" and "(right)" to the end of the corresponding sentence as is done for all the other captions in the text DONE

**General comments:**
Figure 2: "One quarter": actually, this seems like the entire Outer Tracker on the Z+ side, since ‘r’ is plotted on the ‘y’ axis. Maybe call it "Positive side"? well it the sense that you need to mirror along the z axis to get the full +z it's only ¼...

Figure 2: "In the central barrel" --> "In the central, barrel, region up to z=130 cm," fixed

- L85 suggest to rephrase the rest of the sentence: "travel outwards in a 3.8T uniform magnetic field with field lines parallel to the z axis" partly rewrote (left out "field lines")

- L86 "A charged particle ... is bent .. that its trajectory forms a helix." --> "The trajectory of a charged particle ... is bent ... that it forms a helix." DONE

- L87 suggest "forms a circle and the radius of this circle is" --> "forms a circle. The radius of the circle is" DONE

- L91 "and the inner ... disk" --> can be deleted since better explained again in L98 DONE

- L92 "regions: The" --> "the" DONE

- L94 "and the outer half ... instead" --> could be deleted, better explained in L98 DONE

- L97-98 Numbers 5 and 7 should be spelled out (and perhaps even 10?) DONE

- L100 "z0" has not been defined DONE

- L111 "Tbps" --> "Tbit/s" See above

- L118 "was" --> "were" DONE

- L128 comma after "demonstrator" DONE

- L193 Figure number is missing fixed

- L200 9 --> nine; Figure 2 shows 28. DONE

- L227 and in other later occurrence "Tracklet" so far been lowercase DONE

- L236 "covers" --> "cover" DONE

- L236 "but are internally to the VM the data" --> perhaps "but, internally to the VM, the data"? fixed

- L242 "version" --> "versions" DONE

- L236 "covers" --> "cover" DONE

- L236 "but are internally to the VM" --> perhaps "but, internally to the VM," DONE

- L240 "This configuration..." sentence seems to repeat the first sentence in the paragraph removed

- L248 "neighbor and matches" --> "neighbor, matches" DONE

- L283 "drives" --> "drive" DONE

Figure 5: is on page 10 while first reference is only on L323 on page 11 fixed
Figure 7: not only quality, but missing labels on axes. Also, for me it looks like red and blue legends overlap fixed

Figure 10: could mention in caption that this is a simplified geometry

How simplified? This is the TDR geometry (before the tilts.)

- L388 "pileup" has not been defined DONE
- L412 "provide" --> "provides", "improvement" --> "improvements" DONE
- L413 "the the" --> "the" DONE
- L414 "removes" --> "remove" DONE
- L415 "provides" --> "provide" DONE
- L462 "inner half of the disks" --> "inner seven rings" DONE
- L463 "outer half of the disks" --> "outer five rings" DONE
- L470 and further on "Gbps" --> "Gbit/s" DONE, Gb/s based on comments from Katja
- L485 Figure number missing, also the figure appears to be missing fixed
- L490 and 491 "with" should be "within"? fixed

Figure 14: explanation of content could be improved this figure has been removed

- L493 space missing between "parameter" and r_c DONE
- L529 the very same sentence is written in L520 this paragraph was intended to be removed, fixed now
- L530 figure number missing, should probably be Figure 16, but then the order of Figure 15 and 16 are swapped same comment as above%
- L537 "is" --> "are" This section has been removed based on comments from Katja.

Figure 17: >= and 1 seems to be touching each other in the axis label looks fine on my screen...

- L571 for the same purpose, in L556 Arabic numerals were used (not sure if numbers are needed at all), same goes for L578

I do not understand this comment, what it is referring to or what the suggestion is?

- L579 "due to" is superfluous fixed
- L580 suggest to add "that with" before "2S modules" rewritten
- L608 subject missing

quoi?

- L610 Projection Router was introduced as one word, also Tables use two words fixed
Figure 1: "with pT>20 GeV" -- suggest to append "muon selection threshold" here added "threshold" (didn't think "muon" was neccessary as it just said "muon" earlier in the sentence)

- L70 "where the efficiency" -- presumably "where the event selection efficiency", as opposed to for example matching generated muon with L1 trigger muon candidate. Also, what events are simulated, only with one or more muons, or inclusive (may include fakes)

Rewrote this a bit. Also, as mentioned earlier, this figure is from the CMS Phase-2 Technical Proposal, which is still the best we have got since the CMS Phase-2 L1 trigger TDR is not out. The TP does not specify exactly which sample, nor does the old corresponding detector note, but I (Louise) am almost certain that it used single muon gun + PU140. This figure is just to illustrate how L1 tracking can help.

- L77 a hit is defined as energy deposit, but is used more like position measurement in L79 -- could add a sentence before "In the upgraded..." something like: the points where the traversing particles intersect the mid-plane of the sensor ('hits') are calculated from these energy deposits.

we don't think this is neccesary... we say "the relative position of the hits" etc.

- L79 "spacing of 1-4 mm" sounds like somethings being arranged at varying intervals, but seems to refer to the hits. Does 1-4 mm refer to the distance between sensors, or the relative distance between hits? Also, in what direction? In the transverse plane or in 3D?

didn't mention sensor spacing, it's really the distance between the sensors, so the direction depends on the module orientation (flat barrel vs tilted barrel vs disks)

Figure 2: "Collisions take place at (0,0)" -- perhaps "LHC beams are crossed at (0,0)" would be more true? fixed

Figure 2: "slices in phi that cover the entire azimuthal length" -- should it be "entire length of the detector along z"? fixed

- L85 "(0,0) in the figure" is superfluous, since footnote in the previous sentence and caption of Figure 1 both already convey this information removed

- L89-90 add "six" for the parallel and "five" for the perpendicular layers fixed

- L90 "large z" -- "large |z|" fixed

- L93 "higher z" --> "higher |z|" fixed

- L98-100 sentence "Both PS and 2S..." hard to interpret. Perhaps wants to say that stubs in both PS and 2S provide precise position and momentum measurements in the transverse plane, but do so only in PS along global z? As I understand, z0 is a property of the track not of the stubs? Also, it has not been defined. added definition

- L126 trigger latency is defined here, but if I understand correctly, the 4 us mentioned in the abstract and in L115 is the nominal trigger latency. Perhaps definition could be moved to L115 and used here? moved up latency definition
- L201 is a sector processor a single FPGA as stated in L167? A sector processor is an ATCA blade, as stated in line 201. I removed the reference in 167 to have it all in one place. While currently we envision a single FPGA per sector per TMUX slice, we might end up with more than one.

- L205 do I understand corretly that there are n number of FPGAs processing a sector, or the n number of sector processors live in the same FPGA fabric? Yes, as described in this section the system is replicated n times, and each system is "entirely independent." If this is not clear can you suggest an alternative formulation?

Figure 12 is not referenced anywhere in the text, also in caption "Virtex 7" --> "Virtex-7" TODO

- L443 "entire azimuth": do you mean mean length of the detector? Azimuthal angle was defined to be in the x-y plane on page 2 FIXED

- L460 should 12 be not 18? yes, fixed.

- L464 "nonants": is it the usual term in Phase 2 for read-out sectors? Not clear if it is introduced because the L1 TT is segmented to 9 sectors or it is an implicit definition for a read-out sector. clarified in the text, this is intended for nine-fold symmetry, which is proposed readout segmentation of the OT.

- L542 would it be interesting to comment why the L5+L6 seeding is significantly lower? Presumably, those trajectories are found by the other combinations, unless triggering on some secondary or long lived particles are of interest. TODO

- L566 what events are simulated? Since there is strong pT dependence (electron), the simulated pT spectrum affects the average of the vs eta plot. In that sense, perhaps, a ttbar plot would be more informative this is for a flat pt spectrum, added that comment. the corresponding figures for ttbar events is in the next plot

- L578 in i), should it be due to the modules not being tilted, or these results are already with the updated geometry from Figure 2 and not Figure 10 (in which case, this should be mentioned somewhere) it's tracklet 2.0. We do say earlier in the section that "This section shows the performance of the tracklet 2.0 configuration."

**Comments by Malte, 03.07.2019**

Dear Authors,

thank you a lot for distributing this very well understandable draft. The text is generally written such that the technically very complex topic is understandable and presented interesting also to non-FPGA programming experts - which I found very enjoyable. I find the structure very clear and well thought. Also the performance qualification is sound and well motivated. I personally would find a number of further studies very interesting, as for example the efficiency and resolution evaluation at higher pile-up than 200 (trigger efficiency vs. pile-up, for example, to see the pile-up where the system becomes inefficient), but I understand that this is a major effort (probably new event generation) that is out of scope of this paper. Despite the overall good language of the paper, I have a few general remarks, of which I believe that their implementation would result in a significant increase of the quality of the paper. I am aware that my comments are sometimes picky, please understand them as suggestions and my try to improve the quality of the paper as much as possible. Please find also a detailed list of Type A comments below.

Some of the comments are rather nit-picky indeed and make the text harder to read for no obvious benefit, we should discuss.
General comments:

- In several places the paper gives vague statements such as "good", "high", etc. Giving the according number would increase the quality of the paper. I will give examples in the detailed comments below.

- The paper concentrates (as appropriate for the content) on the CMS OT only. In some statements I miss either the mentioning of the CMS Inner Tracker, or the explanation that for this information exceptionally in the Track Trigger only the Outer Tracker is used. An example is the vertex identification. On that topic, I think that the location of the pixel detector should be very briefly described without too many details. added in a few places clarifications about inner pixel vs outer tracker

- In few places there are statements expressed too general, such that they are only correct in the focus of this paper, but not as they are written in the text. An example is, that with the CMS Track Trigger for the first time a solid state detector contributes to the L1 trigger, which is not generally correct, as ECAL is also a solid state detector. I list the places I found in detail below and propose a revision.

We feel that we have a good mixture of general statements in introductory parts of the text with details given in later parts. Having a wash of numbers at the very start leads to a confusing text without any additional clarity.

- I believe that introducing that the CMS Tracker will be build out of silicon pixel and silicon strip detectors would be beneficial, as well as very briefly introducing the PS/SS modules early on in the text. added better introduction of the tracker components

- pileup should be changed to pile-up throughout the paper CMS pubcom wants pileup without a hyphen. see the PubGuidelines twiki.

- numbers below 13 should be written out consistently throughout the paper. CMS Style says 10, see PubGuidelines

- acronyms should all be introduced, even common ones like FPGA, HDL, etc. DONE

- c++ or C++? I always find "C++", so I would suggest to change all appearances to C++ We now use \textsc{C++} everywhere, believe it or not...

Detailed comments:

- Introduction: the argumentation starts very abruptly without any motivation for the upgrade, and does then give the result very briefly. I think that a short motivation (2-3 sentences similar to "The sensitivity for discovery of very rare events as well as high precision measurements of known processes increase significantly with larger data sets recorded by the experiments. While a high instantaneous luminosity of the accelerator enables a high data recording rate, it poses significant experimental challenges to the detector systems. The instantaneous luminosity of the Large Hadron Collider (LHC) [1] at CERN is gradually increased during its operational time...") of the HL-LHC upgrade would be beneficial.

The text in lines 55ff describes the motivation for the upgrade as enabling searches for undiscovered rare particles and detailed measurements of the LHC. Those to me are a succinct summary of the motivation for the LHC. This paper is not about the upgrade generally.

- line 59: only a small fraction...: avoid vague statement and give the fraction (anyway given later in a footnote)

The exact fraction is irrelevant here and as you say is given later in the paper.

General comments: 12
- line 61: "should be saved" is imprecise. Events after L1 still have to pass the HLT before storage. I propose to change to "should be processed further". that is also imprecise, even after HLT they are 'processed further' offline. They are 'saved' for further processing... We have changed the text to 'be considered for further analysis'.

- line 62: "a large increase": give number updated

- line 63-64: "new handles are required." --> a reference or a plot to confirm this statement would be beneficial.

We are developing this concept here and the answer comes in Figure 1, referenced six lines later ...

- line 65-66: The sentence is not correct. I would change "solid state detector" to "solid state tracking detector", "silicon stracker", or similar. changed to solid state tracking detector

- line 67-68: "...improve lepton identification and momentum measurements": Isn't this statement only true for the identification / measurement before the L1 trigger? Then it should be clarified. Otherwise, I do not understand why the change in trigger information improves the final identification/measurement, so it should be explained. Also: isn't it the lepton identification efficiency? (in line 70-71 I think this is clarified, but I would use a clear statement already in line 67-68.

yes of course we are talking about integrating this information into the trigger...

- line 76-77: I would propose to move these two sentences before line 67 and additionally add one sentence that afterwards a track is reconstructed off-detector from this space point information (it is clear for our community, but the paper is interesting also for a large online data processing focussed community, to which this might not be clear).

this paragraph has been modified following earlier review comments. we explicitely already say that " the stubs can be linked together to reconstruct the trajectories of the charged particles."

- line 91: here the modules are intriduced. A very generic sketch would be useful and an explanation how modules are arranged. This could for example be in the caption of Figure 2, where the individual modules are indicated.

we already have figure 2 that shows precisely how the modules are arranged. not clear what else is useful to add here.

- line 103: I propose to give also the information that is available per stub, like: timestamp, coordinate, pT?, or coordinate of the two hits per stub?

we specify now in the previous paragraph, following other comments, the relevant piece of information we get about stubs from PS/2S modules.

- line 104: "many stubs" is vague, although trivial here for good style I would replace "many" with "90%". changed to "the majority of"

- line 105: "first time...data from tracking detector...included in L1". I think that this statement is not correct or written too general: The muon system is a tracking detector which contributes to the L1 (in ATLAS for sure, I believe also in CMS).

While it is true that the muon system detects muon tracks, its primary purpose is identification of muons. I think you are technically correct but the common usage of the term 'tracking chamber' or 'tracking detector'...
does not refer to muon systems.

- page 4: Maybe somewhere it should be mentioned, that the FPGAs are of course off-detector in a service cavern, and the rough distance could be useful. This is likely not usual information for non ATLAS/CMS communities.

Done

- line 161: allows precise determination of collision vertices": this is the major task of the pixel detector, why do we need pixels then? I think that "before full event reconstruction with information of all sub-detectors" or similar is meant here. I would propose to write it more precisely.

What do you mean why do we need pixels here? The L1TF's ability to assign jets e.g. to vertices is a major reason why it's worth doing... Clearly we are talking about the trigger here, that is obvious from the context of the entire paper.

- line 186: radius --> radius DONE

- line 193: Figure ?? DONE

- line 216: what is "N"? updated ENDCOLOR%

- line 241-242: two version --> two versions fixed

- line 242: VERILOG modules: I think that VERILOG needs to be be explained in a footnote or referenced, and the term "module" briefly introduced as it is slang.

Verilog and HDL are now defined. A module is a central concept in the language and in programming and is not slang .... See e.g. this wikipedia page for the language: https://en.wikipedia.org/wiki/Verilog

- line 255: Introduce HDL

DONE

- line 321: I would specify in the section title that these are HDL modules and not hardware detector modules or similar

I don't see how this confusion could happen, as the text has been replete with discussion of verilog modules in the preceding sections and no real hardware modules have ever been mentioned (except the 'virtual modules').

- line 324: every 150ns? Previously you motivated very clearly, that a new event is processed every 450ns. Where do these 150ns come from? They appear more often later, so that should be clarified.

In the new section 2.2 we now explain the provenance of 450 and 150 ns precisely. These numbers consist of two different configurations, one with 28 sectors, one with 9 sectors, and hence with different TMUX factors (6 vs 18).

- Figure 7: Revise the figure style significantly. The legnd is overlapping, the statistics box is automatically named... This figure must be processed to fit in style quality to the other figures.

fixed

-line 346: "inefficiency of the tracking algorithm": vague, I propose to give the inefficiency number as
estimated from this study.

There is a forward reference to section 6 where this is discussed. We intentionally collect all performance results in one section.

- line 347: "observed to be small" is again vague. Replace **small** with the actual number or below X.YZ %, or below the inefficiency introduced by, or similar ...

again there is a forward reference here...

- line 356: out of reach for Virtex7...: This reads like we don't have FPGAs that can fulfill the requirement, although later you mention a possible solution. I would give the reason why for the demonstrator Virtex7 is used and/or mention the envisaged model for the final system already here. we now make it clear that Virtex-7 are an older generation of FPGAs.

- line 359: "blades": What is a blade in this context? It needs introduction. Blade is the standard name for a piece of ATCA electronics.

- Figure 11: No difference is visible. I propose strongly to use a better discriminating plot, such as the "ratio of emulation over simulation" vs. pT or the difference vs. pT. Also: How can tracks have negative transverse momentum? This plot needs revision.

The fact that you cannot see a difference is the point. This is a very discriminating plot when there are differences. Re signed pT, as I'm sure you know, the sign here is just the charge of the particle, i.e., it is q times pt.

- line 410: What is "HLS"? Needs introduction. DONE

- line 412: "Tracklet 2.0 provideS several improvementS

DONE

- line 413: "the the"

DONE

- line 414: new sector definitions --> new sector definition (or remove the "s" in the verbs afterwards) DONE

- section 5.1: Information is given redundantly. I am aware of the immense difficulty of summarizing such information, but I believe that this section could benefit significantly from streamlining the given information.

???

- Figure 14: The caption needs significant expansion. All items shown in the figure should be explained in the caption. Also: here a sketch of a module appears without any explanation. I would provide this earlier including some explanation as mentioned before. based on a comment from Katja this figure has been removed. Its focus is not the TT and as such it's irrelevant here.

- line 493: "parameterr_c" lacks a space.

DONE

- line 493: "is a tunable" --> "is tunable"?
phi should be written as greek symbol to be consistent with the rest of the paper.

DONE

- line 530: Figure ?? DONE

- line 544: "...is good": vague. Give number.

the point is the comparison of the tracklet parameter resolutions w.r.t. to the full track fit, sentence updated

- line 546: "use narrow" is again vague. I propose to give the angle of the cone, the search area on the next layer, or similar.

this is already further discussed in the next paragraph

- line 552-554: The sentence is unclear, I suggest to rephrase it.

sentence rewritten

- equation 6.1: This is not an equation, so I would either include it in the text, or remove the equation number.

DONE

- line 561-562: "one track per event is expected, which is what is observed...": No, there are some percent of two tracks per event visible in the plot... Explain or rephrase the statement.

DONE

- 5666-567: The sentence "Electrons have a lower efficiency compared to muons due to 567 their additional interactions with the detector material." is wrong as it is written in the text! There are no additional interactions of electrons wrt. muons. What is meant is probably that due to the lower mass of the electrons they are deflected stronger by multiple coulomb scattering. I strongly suggest to write this precisely/correctly in a paper.

Yes, of course this is a shorthand for a full description of what is going on. For the readership of the paper this is obvious. None the less we have added a phrase that reminds people that the interaction rate with material depends on the momentum and species.

- line 580: "..., which are not finely segmented in z direction." I can not follow this argumentation. For eta > 2.2, there are only PS-modules, and they are additionally arranged in discs. So the z-position is known quite precisely. I would rather expect that the large lever arm for the extrapolation to the interaction point is responsible for the degraded z_0 resolution.

rephrased to say barrel PS module, disk PS modules given their orientation do not provide z coordinate measures, but radial measurement

- line 611: modules --> module

DONE
Comments to the second version of the draft

Comments by Katja, 26.9.2019: pdf pptx

Thanks for your comments and suggestions, all type A comments have been implemented except for the following:

- Fig 2: The left one is updated to what is shown in the TDR, the right one is from TkLayout, and we have no corresponding (i.e. x-y plane) version in the TDR to update to.

- Fig 14: This figure was approved a few years ago to be shown at conferences, so had the "preliminary" label then. We will try to remake this figure and then also fix the label.

- Fig 12 has been replaced by a table instead.

- Table 3 has now also been fixed.

Type B comments:

- section 6.1.2: Agreed, these subsection titles weren't necessary. They have now been removed. Also added "The residuals are tuned to be nearly 100\% efficient for trajectories from prompt particles originating from the primary interaction vertex." for the paragraph that was Section 6.1.2.

- Section 6.2 / fig 20: The curve with quality cuts is not above the curve without quality cuts -- the one with cuts is dashed red, while the one without is solid blue. Maybe you were comparing with the truth curve (black), which the one with cuts follow closely (=good)?

Comments by Rainer, 02.10.2019:

Dear authors,

thanks a lot for this thoroughly reworked second version of the paper draft. I have only few remaining comments which I summarize below, some of them are follow-ups to your responses.

Rainer

General comments (Type B):

- ghost tracks (follow-up): indeed the added Fig. 20 is informative in this respect, thanks! It it safe to assume that the "quality cuts" that appear to bring the ghost rate down do not strongly affect the tracking efficiencies (Fig. 17+18)?

This depends on the type of tracks you are looking at, and what the use-case in question is. For inclusive ttbar, i.e. what that plot was made using, the efficiency loss is about 5\%. If highest efficiency possible is most important, one should not place such stringent quality cuts, but for some scenarios where eliminating fake tracks is most important, e.g. in defining track MET, this shows that it is possible to set up such criteria. So these type of cuts can be tuned to define different working points for the downstream L1 trigger.

- make sure that wherever you write about simulated events it is unambiguously clear whether the 200 pileup interactions are added or not

should be OK now.
Detailed Type B comments:

- L 368-371: Suggest to write here very clearly in which cases the 200 pileup interactions are added/ not added. This is important e.g. for L 553.

the definition of pileup has been moved earlier following comments from Katja, and for all plots it is specified whether they are with pu=200 or without pileup.

- L 105 (follow-up): "while only the PS modules give a precise measurement of z0":

While I agree with your response that "each PS module gives a precise z measurement with its pixel side", I do not see how such single z measurement by a single PS module can give a measurement of 'z0', which you define as the distance of closest approach from the collision point in z direction. For this measurement you need to know the track slope in the rz projection, which needs two z measurements to compute. (This is probably what you mean, but it is not clear from the text.)

changed to "of the z coordinate"

- Fig. 4 caption (follow-up): Suggest to add: "The removal of duplicates is performed after the track fit (not shown)"

changed as suggested

- Fig. 7 caption: spell out L1, L2 to avoid confusion with L1=level 1 etc, like: "seeding in the two innermost layers"

added

- Fig. 15: If you mean single muons, call it single muons and not muons in events without pileup (which could still contain a lot of other tracks, like muons within ttbar events).

changed

- Fig. 16: The reader could wonder whether you purposefully hide the number of events that have zero tracks. Can one motivate the suppression of the zero bin somehow?

this is illustrating the performance of the duplicate removal, which by construction only considers events with at least one tracks, and by construction will not result, for events with >= track, in events with zero tracks. i.e. the efficiency to identify tracks in the first place is shown in other figures. added to avoid confusion "Only events with at least one track prior to duplicate removal are shown."

Type A comments:

- L 82: "based on that charged...": something is missing here

there is no grammatical problem with this expression but changed to "is based on the concept that"

- L 369: fully

fixed
Dear authors,

thanks you a lot for the new draft. As I tried to clarify in the first round, I did not expect all my nit-picky comments to be implemented. I agree that if the authors prefer a formulation that this formulation is kept. I find the new draft clearer in many aspects and really nice to read! I have only very few comments:

**Type B:**

- Figure 17 and lines 577-579: I am afraid I am missing something obvious, but also discussion with experienced colleagues did not clarify my doubt. The tracking efficiency is lowered by 20% for electrons wrt. muons (at fixed momentum). You state that this is due to their higher interaction rate. But what exactly creates this significant efficiency difference? My naive explanation is that it must be multiple scattering resulting in inefficiency to find matching stubs on the roads of the tracklets. If so, I don't understand why electrons are more affected than muons. The multiple scattering of muons should be stronger than of electrons (at the same momentum), as the multiple scattering is proportional to $1/(\beta p)$. If with "higher interaction rate" you mean bremsstrahlung, then I don't understand how this impacts the tracking efficiency significantly. Could you please explain what I miss (I mean here, not in the paper. I just would like to understand)

the problem for electrons is bremsstrahlung, whereby the electrons loose energy as they traverse the detector material. consequently, an electron track that undergoes interactions does not look like a muon track, for example. this is why CMS and others use gaussian sum filters to recover the efficiency loss for electrons, see e.g. https://arxiv.org/abs/physics/0306087 however, that is not easily (or at all?) implementable at L1, resulting in that the efficiency loss for electrons vs muons at L1 is more significant than you are perhaps used to for offline tracking. we can recover some of these tracks by using wider projection windows in phi (i.e. to capture the electrons that interact), but this is at a tradeoff of increased combinatorics, since you then are more likely to pick up random stub matches from pileup tracks.

**Type A:**

- L77: "In order for the tracks to be used in the trigger decision, they must be delivered within 4us." I am no native speaker, but wouldn't the formulation "The tracks must be delivered within 4us in order to be used in the trigger decision." be more understandable? I stumbled over the sentence a few times before I got it.

changed as suggested

- L369: "fully" --> "fully"

fixed

- L 402: "wass" --> "was"

fixed

- L 439: ".." --> "."

fixed

- L632: "a number of ... improvement" --> "a number of ... improvements"

fixed

Cheers,
Malte

Comments by Viktor, 04.10.2019:

Dear Authors,

thank you for the new draft. I only have a few more comments:

**Type A:**

- L19 the word trajectory is also ok with me; meant only to suggest to write that the procedure of trajectory reconstruction will be implemented in the trigger system not the result of that process (the trajectory). E.g. if your replace "inclusion" with "reconstruction"?

changed to "... the reconstruction of charged particle trajectories and their inclusion in ..." (since technically, tracks are included in the trigger, and we deal with the reconstruction of these, which happens before the CMS L1 trigger system

- L26 I think "problem, specifically..." --> "problem; specifically..."

I don't think that is appropriate here?

- L91 sentence sounds like the particles travel along z (even though that would be grammatically incorrect) as opposed to the magnetic field being aligned with z

changed to "... magnetic field, which is parallel ..." to be clear

-L98 "The" --> "the"

fixed

-L178 is it "r-phi" or "rphi" as in L140 or L567

changed to r-phi everywhere

-L264 "and as" --> "and, as"

fixed

Figure 7: in axis title "L1L2" -->"L1+L2"

fixed

-L438 "Fig. 12" --> "Figure 12"

fixed

-L583 and L590: as commented before: in L567, Arabic numerals are used in listing, while here Roman

changed

Best, Viktor
Comments to the third version of the draft

Comments by Katja, 18.10.2019: pptx

Section 1: L94ff: macro-pixel modules etc.: I have two problems with this paragraph. First of all, this is not correct, the PS stand for pixel-strip. Secondly, you do not explain what it means. So in my opinion in this paragraph you should do two things: 1) introduce correctly the acronyms, and 2) explain how the modules are made, namely from two strip sensors for 2S and one macro-pixel sensor and one strip sensor for PS. fixed

Section 5:

- Data are plural, at least according to CMS guidelines, and this is also how the word is used up to this point. In this section you use it as singular, though: L444, L447, L452, L453, L469 (I may not have caught all occurrences, please search)
- In Table 4, capitalize Unit for consistency... fixed

Section 5.1: L490: just a detail: th in n^th should better not be italic fixed

Section 6.1: L562: you use (i) at other places instead of i) I think (i) leads to better readability fixed

Section 6.2: L581: and the relative resolution... fixed

L584: seems like you do not use the latex small space for 1 mm fixed

Section 6.2: L630: seems like you do not use the latex small space for 1 mm and 5 mm fixed, now use \unit{1}{mm} here too. Well spotted, it looks weird to me now, but I guess this is correct.

Fig. 13, Fig. 15: how do we proceed with the Preliminary labels? fixed, changed to CMS Phase-2 Simulation

KatjaKlein - 2020-06-24