Sofia and the ACAT reviewing team,

Thank you for your insightful comments and help with layout and grammar. We've attended to each comment and written a response inline below.

Q1: Could you please give more details about the SCnet architecture and how is it different from JishNet? We have added an expanded description in section 2 on page 2 to describe the architecture and their differences.

Q2: What is the difference between SCnet and Facebook SparseConvNet in terms of architecture? Despite the misleading name, SparseConvNet is not a network architecture, it is a codebase which defines the SubmanifoldConvolution layer, which can replace typical dense convolutions. We have changed the language at the end of page 3 to make this clear.

Q3: in page 3 you mention training Facebook's SparseConvNet. I can't find any discussion about the performance you obtained on SparseConvNet. See the answer to Q2.

Q4. In order to speed up the layout review could you please make sure the paper follows the templates (e.g. in terms of institute address and citations) We have updated the citations to use the iopartnum bibstyle from CTAN and updated the institute address to match http://cms.iopscience.iop.org/alfresco/d/d/workspace/SpacesStore/ac424614-d325-11e0-adda-5d01ae4695ed/simpleinstructions.pdf. Further changes can be made if needed.

A few more points below:

Pg.1. l. 5 : "formats its simulated data" do you mean you use official MicroBoone simulated data or some "toy Monte Carlo resembling MicroBoone data"? maybe you could rephrase. The second is more correct, we've made appropriate changes to that sentence.

Pg. 3 l.7: "compression was needed... when loading" do you mean loading on CPU memory as in the subsequent sentence? Could you then rephrase and merge the two ? We've rephrased the sentence.

Pg.3 L.9 : "from memory and transferred to GPU memory" " from the host memory and transferred to the GPU memory" Thank you, We've made your suggested change.

Pg3. L. 21 "to parameters of Adam optimizer" "to the Adam optimizer parameters" Thank you, we've made your suggested change.

Pg3, L.23: "expected training" "the expected training" Thank you, we've made your suggested change.

Figure 3: TTL is explained in the text, but maybe you should explain the acronym also in the figure caption. We've added a short sentence to the caption of Figure 3a.

Pg4. L. 1: "Results show increasing" "Results show that increasing" Thank you, we've made your suggested change. *Pg.4 I did not understand the comparison in lines 13-15: "which is 60 % faster than SCNet." Could you please clarify?* We apologize, SCNet is 60% faster than JishNet, we've corrected the mistake.

Pg 4. L: "yet .. nevertheless." You need only one "yet" or "nevertheless". We've removed the word "nevertheless".

Table 1: do you understand where the differences in the two confusion matrices are coming from? Results for pi vs K are very different for SCnet and JishNet We have added a paragraph regarding those differences, which are likely due to the difficulty in discriminating between K and π and the probabilistic nature of stochastic gradient descent.

Pg. 5, L.9 "seeming inappropriateness of suggestions in the literature for training with SGD" pleae clarify this sentence The Horovod paper (Sergeev2018) provides recommendations on learning rates schemes with SGD for optimal training on distributed GPUs. We found these were did not perform well, and the Adam optimizer performed much better. We've removed the sentence as it was not discussed much in this work.

Thank you,

Alex Hagen