

## Round 2

### COMMENT 1

I thank the authors for carefully addressing my comments and answering my questions. However, I need to follow up on point 5 from my original review:

**Original point 5) Conclusions:** “10% of the improvement directly attributable to the SM interaction matrix.” - If I compare all background rejections of “int.” and “int. SM” in Table 6, I do not see in any of the numbers an improvement near 10%. The same is true for the background rejections in Table 4. Or do you mean that it is 10% of the overall 10-40% improvement, i.e. only 1-4%? If this is the case, please rephrase the conclusions and the abstract for more clarity.

**Authors’ answer:** The number 10 percent comes from the results of table 4 and 5. Our baseline was the signal efficiency of 70 percent. Overall (Table 5) the background is reduced e.g. for the 4-top case from 0.18 to 0.17 for overall (i.e. 6 percent absolute) and e.g. from 0.13 to 0.119 for the Z+jets sample (i.e. 9 percent absolute (Table 4). We changed the conclusion to “...approximately up to 10% of this improvement..”.

**Follow-up:** I am afraid I am not fully convinced by the answer. I acknowledge the more careful phrasing in Section 5.2 (“improvements of up to approximately 10%”), but the statements in the abstract and in the conclusions are still stronger (“with an additional gain of approximately 10% (absolute) due to the SM interaction matrix” and “with approximately 10% (absolute) of this improvement directly attributable to the inclusion of SM interaction matrices”). Moreover, I am afraid that I still find the statement about the improvements exaggerated and I apologize if I miss a point here. Following the authors’ argument to focus on improvements in background rejection at a signal efficiency of 70%, I calculate the following improvements from Tables 4 and 5:

Table 4 (comparing PN\_int to PN\_int.SM and ParT\_int to ParT\_int.SM for 70% efficiency):

PN (ttH): 2.8%  
PN (ttW): 6.2%  
PN (ttWW): 4.5%  
PN (ttZ): 7.4%  
ParT (ttH): 2.8%  
ParT (ttW): 1.1%  
ParT (ttWW): 2.4%  
ParT (ttZ): 8.5%

Table 5:

PN (4top): 4.7%  
ParT (4top): 3.5%

None of these improvements is as large as 10% and in total, there is an average improvement of 4.4%. I personally would think that a statement of “approximately 5% improvement” would better reflect the numbers in these tables. I ask the authors to consider this.

**Response:** We thank the referee for the careful follow-up and for recalculating the observed improvements in background rejection.

Our intention was to highlight the **maximum observed improvement**, which for some background processes (e.g.,  $t\bar{t}Z$  in Table 4) reaches  $8.5 \rightarrow \sim 9\%$ . We note that the “10%” figure previously quoted in Section 5.2 was meant as a rounded upper bound.

To better reflect the full set of results, we have revised the phrasing in both the abstract and the conclusions (as well as Section 5.2) to:

“...with an additional gain of up to 9% due to the SM interaction matrix.”

We hope this updated formulation accurately addresses the referee’s concern, while still highlighting the range of improvements observed.

## COMMENT 2

**In addition, I would like to point out two minor aspects that I trust the authors to take care of:**

**- Introduction, paragraph 2: Please double-check Refs. [6-10] and [11-15]. It seems that they have been mixed up, i.e. [6-10] seems to also include event-level references and [11-15] are only the jet references, while the text says otherwise.**

### Response:

We thank the referee for this careful observation. After re-checking the references, we agree with the comment: the original grouping in the text was not fully aligned with the scope of the cited works.

Specifically:

- Refs. [7, 10, 12–15] focus on jet-level classification tasks, including the Particle Transformer (Ref. [10]), which was originally developed for jet tagging.
- Refs. [6, 8, 9] contain studies that address broader or event-level ML applications in HEP.
- Ref. [11] is a duplicate of Ref. [7] (both referring to the Jet-Images paper) and has been removed.

We have revised the sentence in the second paragraph of the Introduction to reflect this change. The reference numbering has been updated accordingly in the bibliography.

## COMMENT 3

**- Conclusions, last sentence: “It is concluded that embedding SM interactions as physical information in network structures \*represents\* (?) an important avenue...” (typo?)**

**Response:** We thank the referee for pointing out this typo. The sentence was indeed missing a verb. We have corrected it to:

“It is concluded that embedding SM interactions as physical information in network structures **is** an important avenue in this field, which could lead to more accurate and efficient event classification in particle physics.”