Referee Report on "Normalizing Flows for High-Dimensional Detector Simulations"

The paper "Normalizing Flows for High-Dimensional Detector Simulations" investigates the performance of normalizing flows for fast calorimeter shower simulations with increasing phase space dimension. The authors present benchmarks for invertible networks applied to the CaloChallenge datasets and introduce a VAE+INN approach to address scaling issues for higher-dimensional phase spaces.

The paper is well-structured and provides a thorough analysis of normalizing flows for detector simulations. The benchmarks and comparisons are clearly presented, and the introduction of a VAE to address high-dimensional phase spaces is a valuable contribution.

The combination of VAEs with invertible neural networks to process high-dimensional data is innovative and represents a major challenge in this field. Furthermore, the authors provide a detailed evaluation of the proposed methods using high-level features and classifier weights. Finally, the results are presented clearly and concisely so that it is easy to track the performance of the different approaches. Making the code available in a GitHub repository promotes transparency and reproducibility, which is highly commendable.

Overall, the work makes a significant contribution to the field of high-dimensional detector simulations using normalizing flows. The methods are sound, and the results are promising. With minor revisions, I recommend the paper for publication in SciPost Physics after considering the comments below.

- The paper lacks a detailed discussion of the hyperparameter selection process in all models used in this work. The inclusion of this information would help to understand the tuning process and its impact on the results.

- The paper would benefit from a quantitative comparison with other existing methods using specific metrics. This would allow for clearer benchmarking and highlight the relative strengths and weaknesses of the proposed methods compared to other methods in the literature.

- The authors use of a very small beta value (e.g. 10^{-9}) in the loss function in addition to the binary cross entropy. It would be helpful if the authors could explain how this choice benefit the results and why is so.

- While the paper contains several references to normalizing flows for detector simulations, it could benefit from citing other recent work that has made important contributions in this area. Furthermore, there are no references to the β-VAE model, which is used extensively in the literature.