

Machine Learning for Particle Flow Reconstruction at CMS

Joosep Pata^{1,*}, Javier Duarte², Farouk Mokhtar², Eric Wulff³, Jieun Yoo⁴, Jean-Roch Vlimant⁵, Maurizio Pierini³, Maria Girone³

(on behalf of the CMS Collaboration)

¹NICPB, R vala pst 10, 10143 Tallinn, Estonia

²University of California San Diego, La Jolla, CA 92093, USA

³European Center for Nuclear Research (CERN), CH 1211, Geneva 23, Switzerland

⁴University of Illinois Chicago, Chicago, IL 60605, USA

⁵California Institute of Technology, Pasadena, CA 91125, USA

E-mail: *joosep.pata@cern.ch

Abstract. We provide details on the implementation of a machine-learning based particle flow algorithm for CMS. The standard particle flow algorithm reconstructs stable particles based on calorimeter clusters and tracks to provide a global event reconstruction that exploits the combined information of multiple detector subsystems, leading to strong improvements for quantities such as jets and missing transverse energy. We have studied a possible evolution of particle flow towards heterogeneous computing platforms such as GPUs using a graph neural network. The machine-learned PF model reconstructs particle candidates based on the full list of tracks and calorimeter clusters in the event. For validation, we determine the physics performance directly in the CMS software framework when the proposed algorithm is interfaced with the offline reconstruction of jets and missing transverse energy. We also report the computational performance of the algorithm, which scales approximately linearly in runtime and memory usage with the input size.

1. Introduction

Reconstruction algorithms at general-purpose high-energy particle detectors aim to provide a holistic, well-calibrated physics interpretation of the collision event. Variants of the particle-flow (PF) algorithm have been used at the CELLO [1], ALEPH [2], H1 [3], ZEUS [4, 5], DELPHI [6], CDF [7, 8, 9], D0 [10], CMS [11, 12] and ATLAS [13] experiments to reconstruct a particle-level interpretation of high-multiplicity hadron collision events, given individual detector elements such as tracks and calorimeter clusters from a multi-layered, heterogeneous, irregular-geometry detector. The PF algorithm generally associates tracks and calorimeter clusters from detector layers such as the electromagnetic calorimeter (ECAL), hadron calorimeter (HCAL) and others to reconstruct charged and neutral hadron candidates as well as photons, electrons, and muons with an optimized efficiency and resolution. Existing PF reconstruction implementations are tuned using simulation for each specific experiment because detailed detector characteristics and geometry are critical for the best possible physics performance.

Recently, there has been significant interest in the use of supervised machine learning (ML) to perform the reconstruction in order to improve the physics reach of the experiments as well



as reduce the computational requirements. ML-based reconstruction approaches using graph neural networks (GNNs) [14, 15, 16, 17, 18] have been proposed for various tasks in particle physics [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33], including PF reconstruction [34, 35, 36, 37]. In particular in Ref. [37], a ML-based PF algorithm is proposed to reconstruct particle candidates in events with a large number of simultaneous pileup (PU) collisions using a simplified benchmark dataset generated with PYTHIA [38, 39] and DELPHES [40].

In this work, we provide details on the implementation of a similar machine-learned particle-flow (MLPF) algorithm for CMS [41]. The goal in these proceedings is to demonstrate that MLPF can conceptually be applied to CMS reconstruction, as a preliminary step towards a future implementation that directly optimizes the desired reconstruction target and runs natively on heterogeneous accelerators. The CMS PF algorithm reconstructs stable particles based on calorimeter clusters and tracks [11]. An example reconstructed event, comparing PF and MLPF, is shown in fig. 1. We present a possible evolution of PF towards heterogeneous computing platforms such as GPUs using a GNN in the following proceeding, structured as follows: The training dataset is introduced in section 2. We show the details of the MLPF implementation and the computational performance of the algorithm in section 3. In section 4, we validate the physics performance directly in the CMS software framework (CMSSW) when the proposed algorithm is interfaced with the offline reconstruction of jets and missing transverse energy. Finally, we provide a summary in section 5.

2. Dataset

Simulated samples are used to train the MLPF model. For each event, we store the input detector signals and the target particles. The input features for the detector signals are chosen based on the PF reconstruction algorithm [11]:

- ECAL, HCAL, hadron forward (HF) calorimeter clusters: cluster energy, corrected energy, η , ϕ , x , y , z position; number of hits; layer; depth; cluster flags
- ECAL supercluster: cluster energy, η , ϕ , x , y , z position; number of hits
- Kalman filter (KF) tracks: p_T , η , ϕ , p_x , p_y , p_z , $|p|$ at the vertex; η , ϕ extrapolated to the ECAL shower max and HCAL entrance; number of hits; track charge; number of drift tube (DT), cathode strip chamber (CSC) hits
- Gaussian sum filter (GSF) tracks: p_T , η , ϕ , p_x , p_y , p_z , E at the inner point; η , ϕ at the outer point; track charge; number of hits; a flag to denote if the electron seed is ECAL or tracker driven.
- Bremsstrahlung (BREM) points: index of the trajectory point, ΔP , $\sigma(\Delta P)$

We choose the existing PF particles as the regression target in order to constrain the problem to a well-defined and well-understood detector response. It is clear that the MLPF model trained on the baseline PF as the target does not allow to exceed PF performance, but we expect to broadly reproduce the existing physics response both on particle level as well as for jets and missing transverse energy (MET). In addition, this allows to test the computational performance and integration with downstream reconstruction. Training on a generator-level set of target particles is left for a future study.

We use a mixture of simulated physics collisions such as $t\bar{t}$ with PU, as well as simulations of single particles shot from the interaction point. All samples are generated under identical Run 3 conditions. We split the simulated samples into a training (80%) and test set (20%) after randomizing the order. The sample sizes are reported in table 1.

3. Model and Optimization

The task in the MLPF setup is to predict the set of particles $y'_k \in Y'$ in the event, given the set of detector signals $x_i \in X$. The inputs comprise the features of ECAL, HCAL, HF

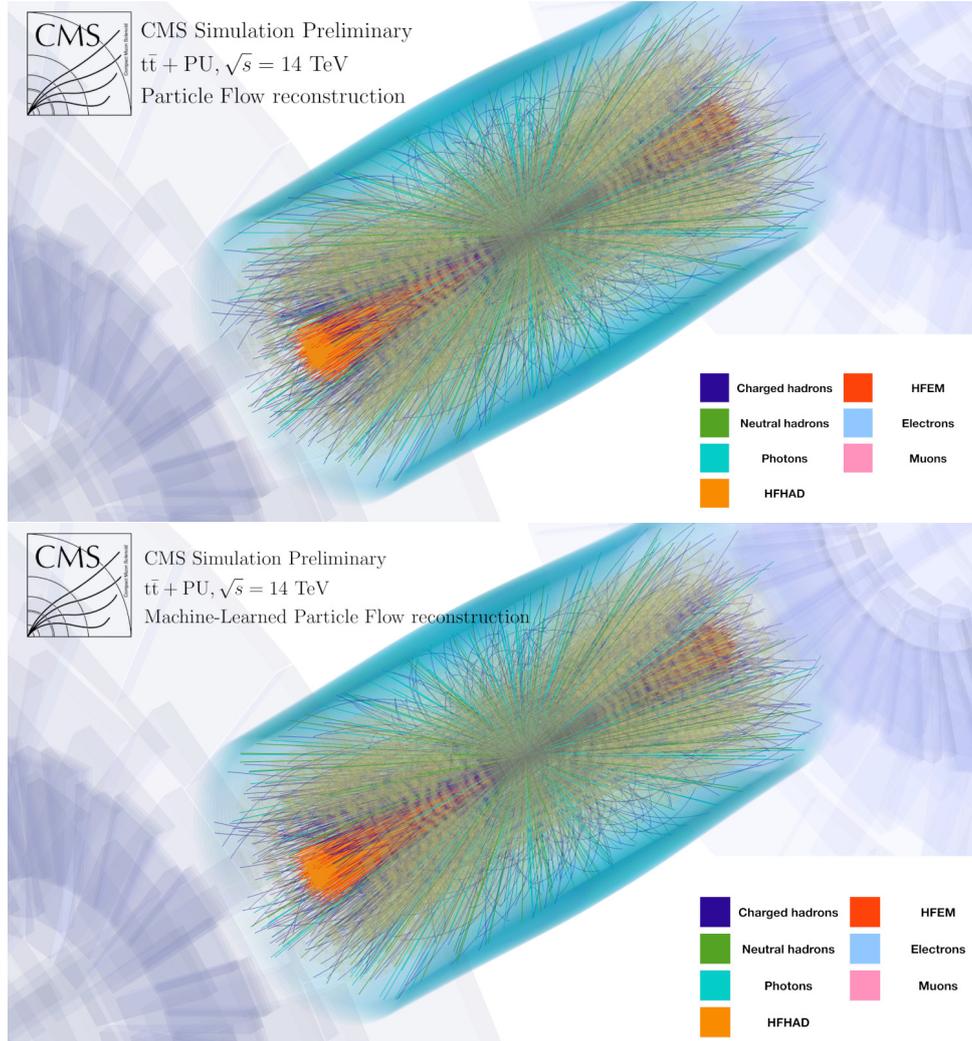


Figure 1: One simulated $t\bar{t}$ event with pileup under LHC Run 3 conditions, reconstructed with particle flow (top) and machine-learned particle flow (bottom). The trajectories correspond to the particle flow candidates extrapolated to the ECAL surface, with candidates of different type shown in different colors. We also show the ECAL detector surface (cyan) and the muon stations (blue).

Table 1: Simulation samples used for optimizing the MLPF model.

Sample fragment	PU Configuration	MC events
Top quark-antiquark pairs ($t\bar{t}$)	Flat 55–75	20 k
$Z \rightarrow \tau\tau$ all-hadronic	Flat 55–75	20 k
Single electron flat $p_T \in [1, 100] \text{ GeV}$	No PU	400 k
Single muon flat $p_T \in [0.7, 10] \text{ GeV}$	No PU	400 k
Single π^0 flat $p_T \in [0, 10] \text{ GeV}$	No PU	400 k
Single π flat $p_T \in [0.7, 10] \text{ GeV}$	No PU	400 k
Single τ flat $p_T \in [2, 150] \text{ GeV}$	No PU	400 k
Single γ flat $p_T \in [10, 100] \text{ GeV}$	No PU	400 k

calorimeter clusters, ECAL superclusters, KF tracks, GSF tracks, and BREM points described in section 2. The target particles y_i are described by a feature vector $y_k = [\text{ID}, p_T, \eta, \phi, E, q]$, where ID is a one-hot encoded vector representing the PF particle candidate type among eight options: charged hadron, neutral hadron, HF electromagnetic energy (HFEM), HF hadronic energy (HFHAD), photon, electron, muon, or none.

The model is optimized with respect to the set of true target particles $y_k \in Y$ in each event. In order to practically compute the loss function between two sets of arbitrary (and possibly different) size, we follow the object condensation approach first introduced in Ref. [34], and as implemented for particle flow reconstruction in Ref. [37]. Effectively, we first zero-pad the target set Y such that $|Y| = |X|$ and evaluate how well each predicted particle's type is classified and momentum is regressed. The association between the target and reconstructed particles follows the simulation association of the particles and input elements. We associate each target particle to a single unique primary element, i.e. each charged hadron is associated to a track, and each neutral hadron to the calorimeter cluster with the largest amount of energy with the particle, based on the links available in simulation. We find that this is always possible in this dataset and results in meaningful associations between the target particles and elements, as quantified by angular differences. Subsequent comparisons of particle-level and event-level distributions are not affected by this modelling simplification. This per-particle loss is a physics-based simplification of the generic set-to-set loss function. The total loss function for one event is then

$$L(Y, Y') = \sum_{k=1}^{|X|} \text{focal}(y_k, y'_k) + \text{Huber}(y_k, y'_k), \quad (1)$$

where we use the focal loss [42] with balancing factor $\alpha = 0.25$ and modulating factor $\gamma = 2.0$ for classification to deal with the highly imbalanced classes and the Huber loss [43] with shape parameter $\delta = 1.0$ for momentum regression to mitigate the effect of outliers.

The MLPF GNN model is implemented in TENSORFLOW [44] (the latest stable version, v2.6, at the time of writing) and can be exported to ONNX [45] for inference. Only standard matrix operations using dense arrays are used, thus it is expected to be highly portable across platforms. Particular care is taken to ensure scalability of the algorithm by using local context binning.

The local context binning is inspired by the existing PF block algorithm, but has been reformulated to be optimizable using ML and ensures the approximately linear scaling of the runtime and memory with the input size. In each local context bin, a graph is built dynamically using a Gaussian kernel. Once the graphs are built dynamically in the event, information can be propagated between the elements in a learnable fashion. For generality, several layers of graph building and message propagation can be stacked. A schematic overview of the network architecture and the scalable combined graph layer is presented in fig. 2. The code to build, train, and evaluate the model is publicly available [46].

The best performing MLPF hyperparameters were found after two separate hyperparameter searches. The first search used the Bayesian optimization hyperband (BOHB) algorithm [47] to tune parameters: the learning rate, learning rate schedule, and dropout percentage. The second search fixed the best hyperparameters found in the first search and used the asynchronous successive halving algorithm (ASHA) [48] to tune model-architecture related parameters: the number of graph layers, the sizes of input encoding and output decoding layers and the linearization bin size. ASHA allows for an efficient use of compute resources when performing distributed multi-worker hypertuning by aggressively early stopping trials that underperform relative to other trials. More details about the hyperparameter optimization can be found in [49].

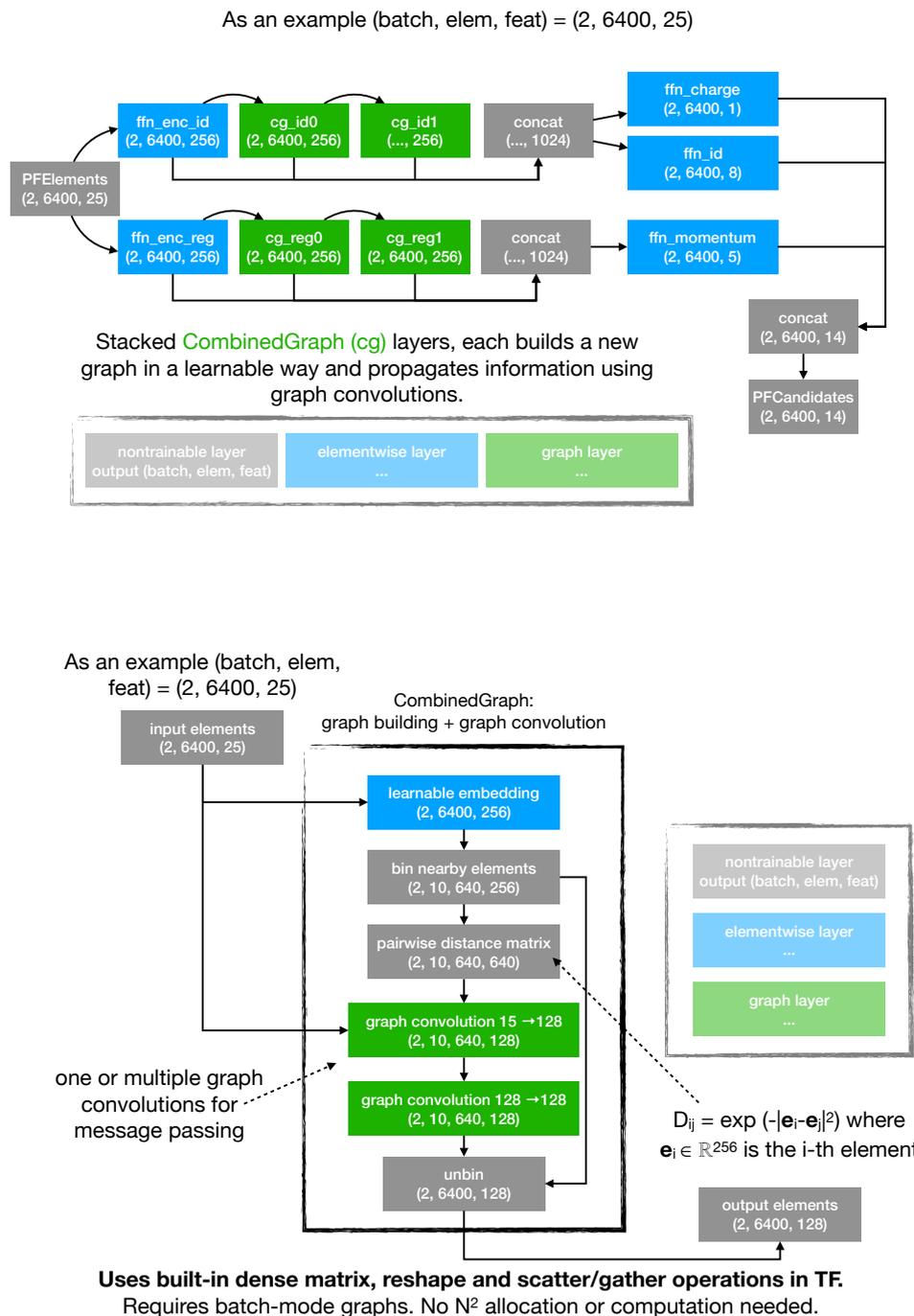


Figure 2: A schematic overview of the scalable, multi-layered graph neural network as implemented in TENSORFLOW (top) and the scalable combined graph layer (bottom). The model consists of a classification and regression branch. Both branches consist of multiple scalable graph building and convolution layers. Elementwise feedforward networks are used to decode the encoded inputs to a classification and regression prediction. The input elements are projected into a learnable embedding space. Nearby elements in the embedding space are binned to fixed-size bins. A fully-connected graph is built in each bin, which is used for one or multiple graph convolutions that are used to transform the input elements. Finally, the transformed elements are unbinned.

4. Results

The MLPF model is interfaced with offline reconstruction in CMSSW, though only the standard PF reconstruction is used for CMS reconstruction by default. For inference, the model is exported from TENSORFLOW to ONNX, such that the native ONNXRUNTIME that is already integrated with CMSSW can be used. The validation plots that follow are independent of the training setup as well as the training samples. This ensures that the model is tested under realistic conditions in actual physics reconstruction. We validate the model on both $t\bar{t}$ and QCD events, the latter in particular is to verify the generalization capabilities of the ML-based approach for particle reconstruction, as QCD events have a different momentum distribution compared to $t\bar{t}$ and were never seen during training.

Particle-level comparisons between PF and MLPF candidates are shown in fig. 3. In general, we observe a good correspondence between the p_T and η distributions for all particle types, within the available statistics, except for the overall electron multiplicity.

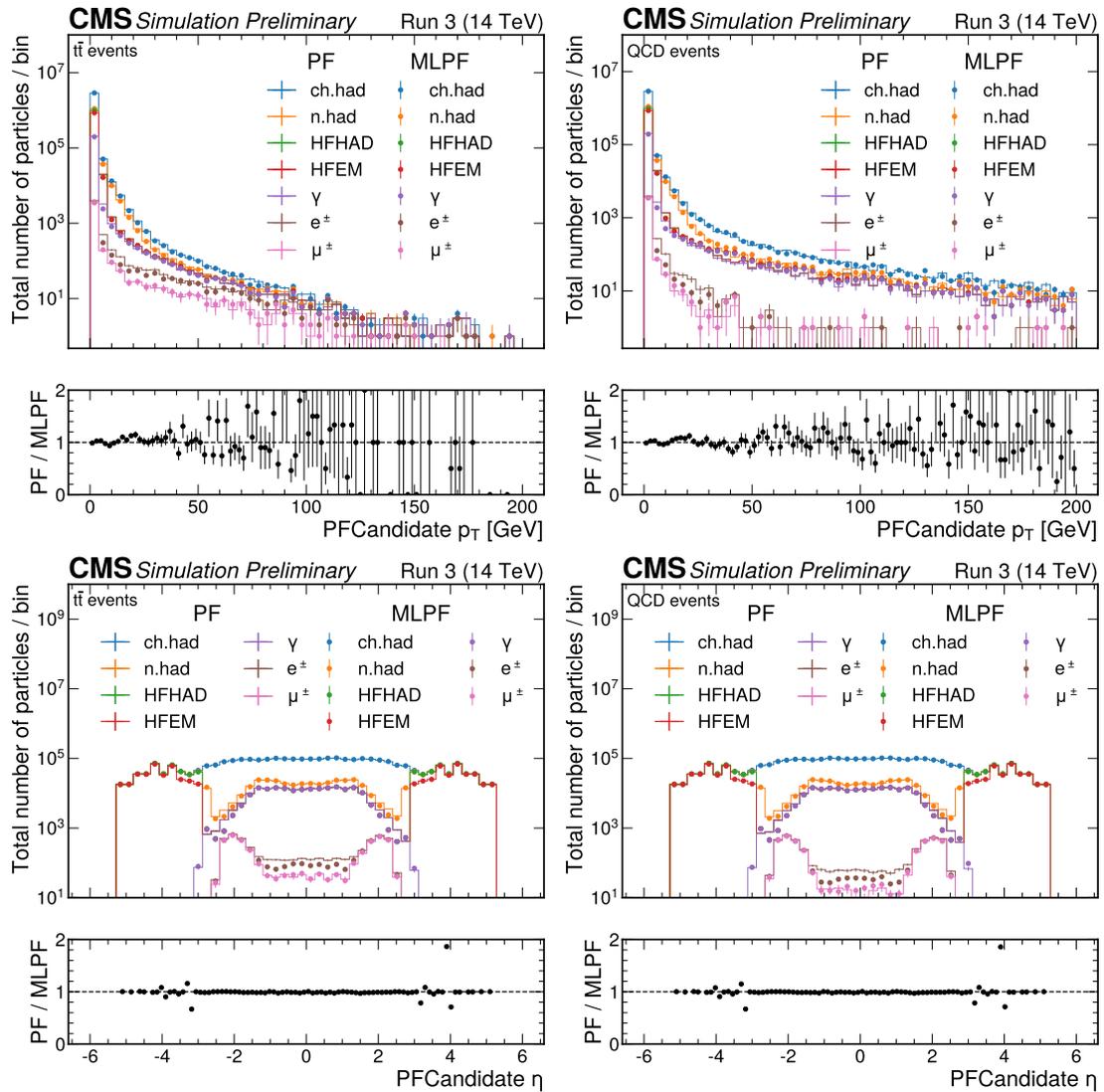


Figure 3: Reconstructed particle candidate p_T (top), η (bottom) distributions between PF (lines) and MLPF (dots) for $t\bar{t}$ (left) and QCD (right). Different reconstructed particle candidate types are shown using colors, the ratio is calculated by summing over all particle types.

Object-level comparisons (jets, MET) are shown in figs. 4 and 5. In general, we observe a good correspondence between the baseline PF and the proposed MLPF algorithm in the bulk of the distributions, within the available statistics. However, for MET, we observe a misreconstructed high-MET tail that is most prominent for the QCD sample not used in training. This could potentially be attributed to limited training statistics for high-energy neutral particles, which are correspondingly not well reconstructed by the MLPF algorithm in the current iteration and require further study.

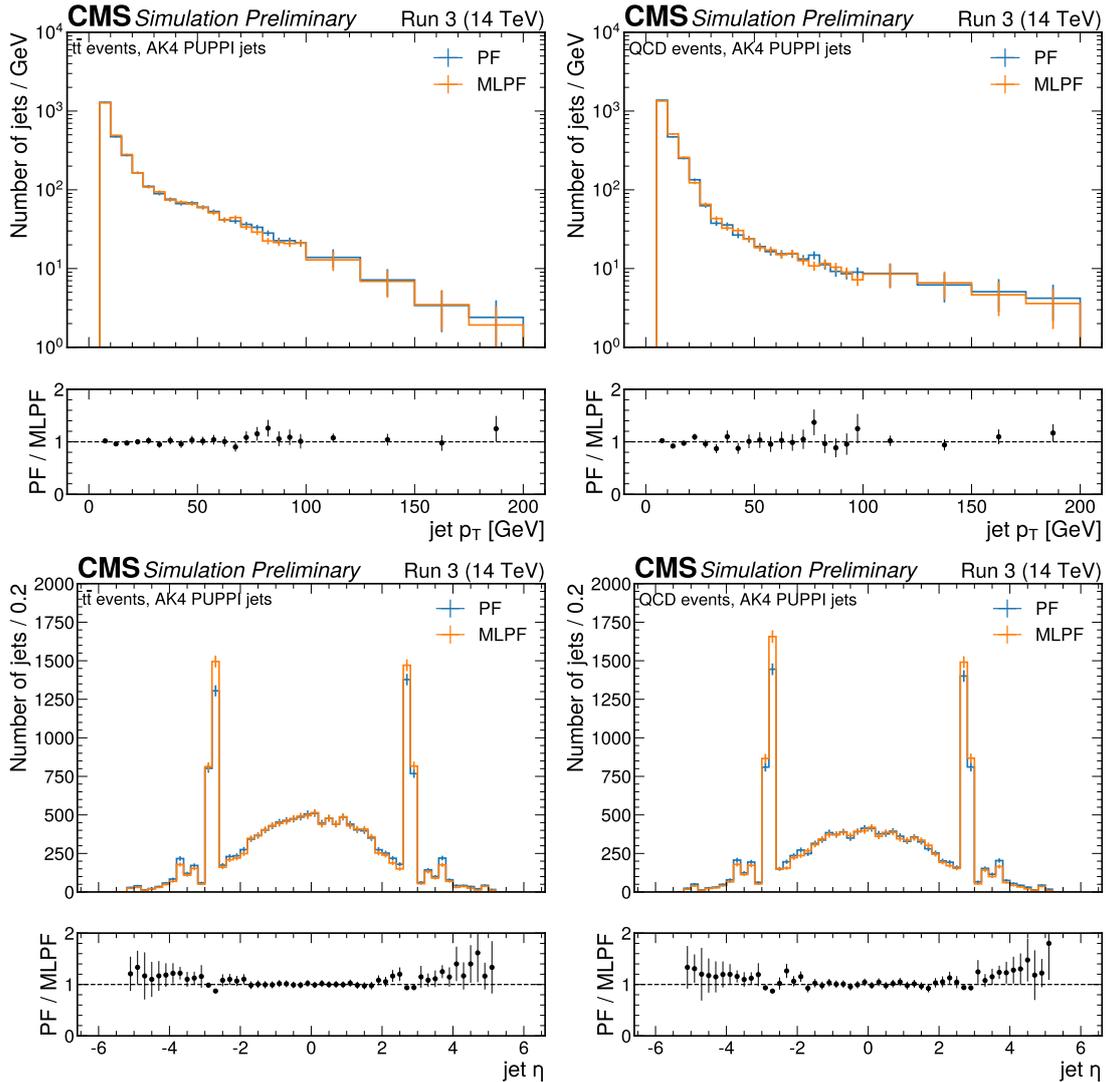


Figure 4: Transverse momentum distributions for AK4-PUPPI [50, 51] jets, with PF and MLPF, for $t\bar{t}$ (left column) and QCD (right column).

We report the computational performance of the model on a single stream on a single GPU in fig. 6. We observe an approximately linear dependence of runtime and memory consumption with increasing particle multiplicity, with a typical Run 3 event requiring around 10 ms of wall time and around 1 GB of RAM on a GPU¹.

¹ Due to the still-evolving support for GPU ML evaluation in CMSSW, this measurement was carried out in a standalone environment, outside the standard reconstruction software.

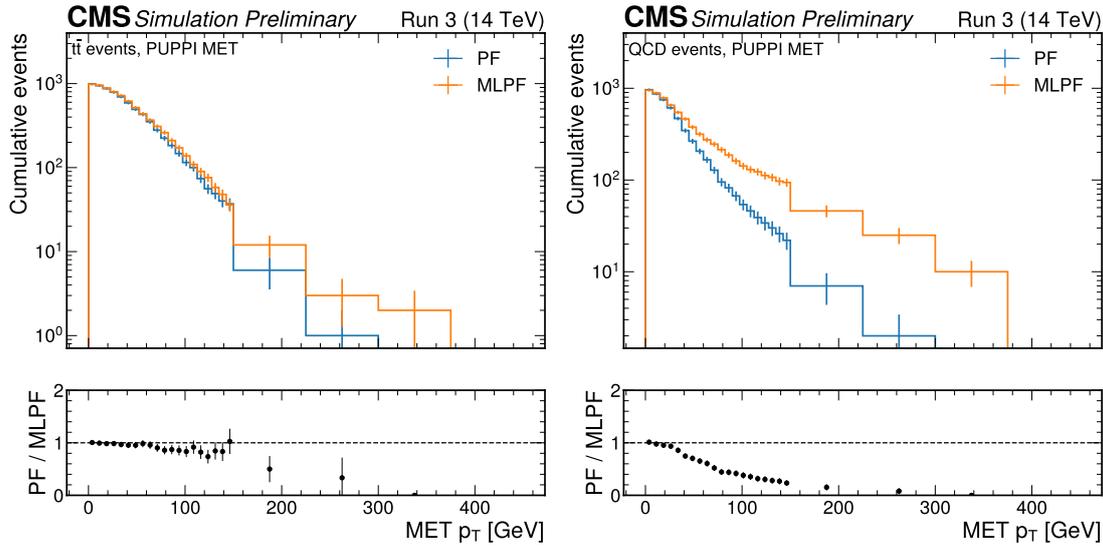


Figure 5: Missing transverse energy distribution, with PF and MLPF, for $t\bar{t}$ (left) and QCD (right).

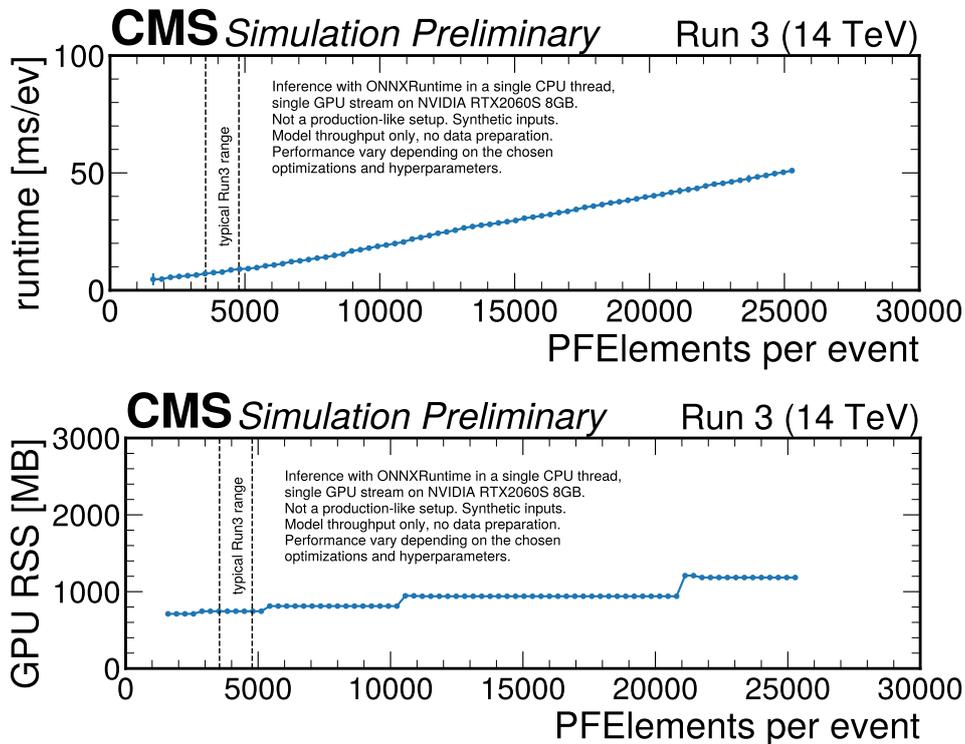


Figure 6: Runtime and memory use on an NVIDIA RTX2060S GPU with a varying event size. The 1σ range of a typical Run3 $t\bar{t}$ event with PU is shown with the dashed lines. Note that this is not a production-like setup, as a single GPU execution stream is used at a time. A realistic production-like test would involve the GPU running multiple different models/kernels from multiple execution streams in parallel.

5. Summary

We have developed a GPU-native algorithm for PF reconstruction at CMS and presented a first integration with offline reconstruction. The new MLPF algorithm is based on a supervised ML setup, where we optimize a scalable GNN model to reconstruct the output particles of the current PF algorithm, based on the input detector elements. By training on the current PF as the target, we demonstrate that it is possible to integrate the MLPF approach in CMS reconstruction, paving the way for a subsequent optimization on a generator-level target to explore potential physics improvements. In general, we observe a high degree of correspondence between the MLPF algorithm and the baseline PF both at the particle level, as well as at the object level in jets and MET, when the new model is interfaced with offline reconstruction in CMSSW. Some differences are observed in neutral hadron and electron performance, which we plan to address with additional training statistics and optimization. The MLPF model has an approximately linear scaling of runtime and memory with increasing particle multiplicity and runs natively on a GPU, while the standard rule-based PF algorithm is not straightforward to re-implement on a GPU in full detail.

Acknowledgments

We thank our colleagues in the CMS Collaboration, especially in the Particle Flow, Physics Performance and Dataset, Offline and Computing, and Machine Learning groups, in particular Kenichi Hatakeyama, Lindsey Gray, Jan Kieseler, Danilo Piparo, Gregor Kasieczka, Salvatore Rappoccio, Kaori Maeshima, Kenneth Long, and Juska Pekkanen for helpful feedback in the course of this work. JP was supported by the Mobilitas Plus Grant No. MOBTP187 of the Estonian Research Council. JD and FM were supported by DOE Award Nos. DE-SC0021187 and DE-SC0021396 (FAIR4HEP) and NSF Cooperative Agreement OAC-2117997 (A3D3). FM was also supported by a UCSD HDSI fellowship and an IRIS-HEP fellowship through NSF Cooperative Agreement OAC-1836650. EW was supported by the CoE RAISE Project which have received funding from the European Union's Horizon 2020 – Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733.

References

- [1] CELLO Collaboration. “An Analysis of the Charged and Neutral Energy Flow in e^+e^- Hadronic Annihilation at 34 GeV, and a Determination of the QCD Effective Coupling Constant”. In: *Phys. Lett. B* 113 (1982), p. 427. DOI: 10.1016/0370-2693(82)90778-X.
- [2] ALEPH Collaboration. “Performance of the ALEPH detector at LEP”. In: *Nucl. Instrum. Meth. A* 360 (1995), p. 481. DOI: 10.1016/0168-9002(95)00138-7.
- [3] H1 Collaboration. “Measurement of charged particle multiplicity distributions in DIS at HERA and its implication to entanglement entropy of partons”. In: *Eur. Phys. J. C* 81.3 (2021), p. 212. DOI: 10.1140/epjc/s10052-021-08896-1.
- [4] ZEUS Collaboration. “Measurement of the diffractive structure function $F_2(D(4))$ at HERA”. In: *Eur. Phys. J. C* 1 (1998), p. 81. DOI: 10.1007/s100520050063.
- [5] ZEUS Collaboration. “Measurement of the diffractive cross-section in deep inelastic scattering using ZEUS 1994 data”. In: *Eur. Phys. J. C* 6 (1999), p. 43. DOI: 10.1007/PL00021606.
- [6] DELPHI Collaboration. “Performance of the DELPHI detector”. In: *Nucl. Instrum. Meth. A* 378 (1996), p. 57. DOI: 10.1016/0168-9002(96)00463-9.
- [7] A. Bocci et al. “Study of jet energy resolution at CDF”. In: *Int. J. Mod. Phys. A* 16S1A (2001). Ed. by K. K. Gan and R. Kass, p. 255. DOI: 10.1142/S0217751X01006632.
- [8] A. L. Connolly. “A Search for Supersymmetric Higgs Bosons in the Di-tau Decay Mode in $p\bar{p}$ Collisions at 1.8 TeV”. PhD thesis. UC Berkeley, 2003. DOI: 10.2172/15017134.
- [9] CDF Collaboration. “Measurement of $\sigma(p\bar{p} \rightarrow Z)\text{Br}(Z \rightarrow 2\tau)$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. D* 75 (2007), p. 092004. DOI: 10.1103/PhysRevD.75.092004.
- [10] D0 Collaboration. “Measurement of $\sigma(p\bar{p} \rightarrow Z + X) \text{Br}(Z \rightarrow \tau^+\tau^-)$ at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Lett. B* 670 (2009), p. 292. DOI: 10.1016/j.physletb.2008.11.010.
- [11] CMS Collaboration. “Particle-flow reconstruction and global event description with the CMS detector”. In: *JINST* 12.10 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003.
- [12] CMS Collaboration. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.
- [13] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2.
- [14] F. Scarselli et al. “The graph neural network model”. In: *IEEE Trans. Neural Netw.* 20.1 (2009), p. 61. DOI: 10.1109/TNN.2008.2005605.
- [15] J. Gilmer et al. “Neural message passing for quantum chemistry”. In: *Proceedings of the 34th International Conference on Machine Learning*. Ed. by Doina Precup and Yee Whye Teh. Vol. 70. PMLR, 2017, p. 1263.
- [16] C.Q. Ruizhongtai et al. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation”. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. DOI: 10.1109/CVPR.2017.16.
- [17] P. W. Battaglia et al. “Interaction Networks for Learning about Objects, Relations and Physics”. In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee et al. Vol. 29. Curran Associates, Inc., 2016, p. 4502.
- [18] Y. Wang et al. “Dynamic Graph CNN for Learning on Point Clouds”. In: *ACM Trans. Graph.* 38 (2019). DOI: 10.1145/3326362.
- [19] J. Shlomi, P. Battaglia, and J.-R. Vlimant. “Graph Neural Networks in Particle Physics”. In: *Mach. Learn.: Sci. Technol.* 2 (2021), p. 021001. DOI: 10.1088/2632-2153/abbf9a.

- [20] S. Farrell et al. “Novel deep learning methods for track reconstruction”. In: *4th International Workshop Connecting the Dots*. 2018.
- [21] X. Ju et al. “Graph Neural Networks for Particle Reconstruction in High Energy Physics detectors”. In: *2nd Machine Learning and the Physical Sciences Workshop at the 33rd Conference on Neural Information Processing Systems*. Mar. 2020.
- [22] S. Amrouche et al. “The Tracking Machine Learning Challenge : Accuracy Phase”. In: *The NeurIPS '18 Competition*. 2020, p. 231. DOI: 10.1007/978-3-030-29135-8_9.
- [23] S. Amrouche et al. “Similarity hashing for charged particle tracking”. In: *IEEE International Conference on Big Data 2019*. 2019, p. 1595. DOI: 10.1109/BigData47090.2019.9006316.
- [24] N. Choma et al. “Track Seeding and Labelling with Embedded-space Graph Neural Networks”. In: *6th International Workshop Connecting the Dots*. 2020.
- [25] X. Ju and B. Nachman. “Supervised Jet Clustering with Graph Neural Networks for Lorentz Boosted Bosons”. In: *Phys. Rev. D* 102 (2020), p. 075014. DOI: 10.1103/PhysRevD.102.075014.
- [26] J. Li, T. Li, and F.-Z. Xu. “Reconstructing boosted Higgs jets from event image segmentation”. In: *JHEP* 04 (2021), p. 156. DOI: 10.1007/JHEP04(2021)156.
- [27] J. Guo, J. Li, and T. Li. “The Boosted Higgs Jet Reconstruction via Graph Neural Network”. In: *Phys. Rev. D* 103.11 (2021), p. 116025. DOI: 10.1103/PhysRevD.103.116025.
- [28] E. A. Moreno et al. “JEDI-net: a jet identification algorithm based on interaction networks”. In: *Eur. Phys. J. C* 80 (2020), p. 58. DOI: 10.1140/epjc/s10052-020-7608-4.
- [29] E. A. Moreno et al. “Interaction networks for the identification of boosted $H \rightarrow b\bar{b}$ decays”. In: *Phys. Rev. D* 102 (2020), p. 012010. DOI: 10.1103/PhysRevD.102.012010.
- [30] H. Qu and L. Gouskos. “ParticleNet: Jet Tagging via Particle Clouds”. In: *Phys. Rev. D* 101 (2020), p. 056019. DOI: 10.1103/PhysRevD.101.056019.
- [31] V. Mikuni and F. Canelli. “ABCNet: An attention-based method for particle tagging”. In: *Eur. Phys. J. Plus* 135.6 (2020), p. 463. DOI: 10.1140/epjp/s13360-020-00497-3.
- [32] S. R. Qasim et al. “Learning representations of irregular particle-detector geometry with distance-weighted graph networks”. In: *Eur. Phys. J. C* 79 (2019), p. 608. DOI: 10.1140/epjc/s10052-019-7113-9.
- [33] J. Arjona Martínez et al. “Pileup mitigation at the Large Hadron Collider with graph neural networks”. In: *Eur. Phys. J. Plus* 134 (2019), p. 333. DOI: 10.1140/epjp/i2019-12710-3.
- [34] J. Kieseler. “Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph and image data”. In: *Eur. Phys. J. C* 80.9 (2020), p. 886. DOI: 10.1140/epjc/s10052-020-08461-2.
- [35] F. A. Di Bello et al. “Towards a Computer Vision Particle Flow”. In: *Eur. Phys. J. C* 81.2 (2021), p. 107. DOI: 10.1140/epjc/s10052-021-08897-0.
- [36] Javier Duarte and Jean-Roch Vlimant. “Graph neural networks for particle tracking and reconstruction”. In: *Artificial Intelligence for High Energy Physics*. World Scientific, 2022, pp. 387–436. DOI: 10.1142/12200.
- [37] J. Pata et al. “MLPF: Efficient machine-learned particle-flow reconstruction using graph neural networks”. In: *Eur. Phys. J. C* 81.5 (2021), p. 381. DOI: 10.1140/epjc/s10052-021-09158-w.
- [38] T. Sjöstrand, S. Mrenna, and P. Z. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026.

- [39] T. Sjöstrand, S. Mrenna, and P. Z. Skands. “A Brief Introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), p. 852. DOI: 10.1016/j.cpc.2008.01.036.
- [40] J. de Favereau et al. “DELPHES 3, A modular framework for fast simulation of a generic collider experiment”. In: *JHEP* 02 (2014), p. 057. DOI: 10.1007/JHEP02(2014)057.
- [41] CMS Collaboration. *Machine Learning for Particle Flow Reconstruction at CMS*. CMS Detector Performance Note CMS-DP-2021-030. 2021. URL: <https://cds.cern.ch/record/2792320>.
- [42] T.-Y. Lin et al. “Focal Loss for Dense Object Detection”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017, p. 2999. DOI: 10.1109/ICCV.2017.324.
- [43] P. J. Huber. “Robust Estimation of a Location Parameter”. In: *Ann. Math. Stat.* 35.1 (1964), p. 73. DOI: 10.1214/aoms/1177703732.
- [44] M. Abadi et al. “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems”. 2015.
- [45] J. Bai, F. Lu, K. Zhang, et al. *Open Neural Network Exchange*. 2017. URL: <https://github.com/onnx/onnx>.
- [46] J. Pata et al. *jpata/particleflow: Baseline MLPF model for CMS*. Version v1.4. 2021. DOI: 10.5281/zenodo.5520559.
- [47] S. Falkner, A. Klein, and F. Hutter. “BOHB: Robust and Efficient Hyperparameter Optimization at Scale”. In: *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. Proceedings of Machine Learning Research. 2018, p. 1437.
- [48] L. Li et al. “A System for Massively Parallel Hyperparameter Tuning”. In: *Proceedings of Machine Learning and Systems*. Vol. 2. 2020, p. 230.
- [49] E. Wulff et al. “Hyperparameter optimization of data-driven AI models on HPC systems”. In: *J. Phys.: Conf. Series* these proceedings (2022).
- [50] M. Cacciari and G. P. Salam. “Dispelling the N^3 myth for the k_t jet-finder”. In: *Phys. Lett. B* 641 (2006), pp. 57–61. DOI: 10.1016/j.physletb.2006.08.037.
- [51] D. Bertolini et al. “Pileup Per Particle Identification”. In: *JHEP* 10 (2014), p. 059. DOI: 10.1007/JHEP10(2014)059.