Hyperparameter Tuning for Semi-Supervised Graph Neural Network for Pileup Mitigation

Tianchun Li, Shikun Liu, Yongbin Feng, Garyfallia Paspalaki, Nhan V. Tran, Miaoyuan Liu, Pan Li, **Jack Rodgers**, Yuji Li,
Purdue University, Fermi National Accelarator Laboratory, and Georgia Institute of Technology

Introduction

Often in particle accelerators such as the LHC, two proton beams running in the opposite direction collide at the center of the detector, producing many **charged and neutral particles**.

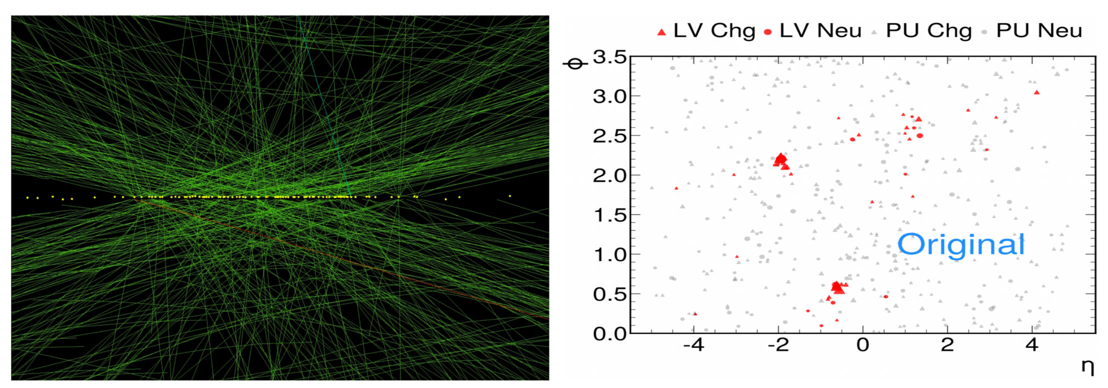

Figure 1: Computer representation of collision chamber in LHC

Two types of particles in collisions get overlaid:

- **Leading Vertex (LV):** Signal particles from the primary interaction
- **Pileup (PU):** Background particles from additional proton-proton interactions

Goal: Identify LV particles (Signal) from pileup particles (Background)
Challenge in Physics: Hard to identity neutral particles and physics reconstruction performance by the model must be able to generalize onto newer data. This generalization must be better than the current algorithm used, PUPPI.[1]

Experiment Results

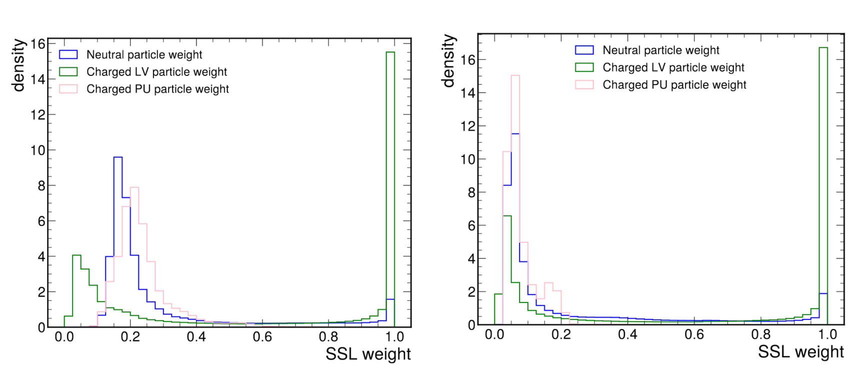

Figure 2: Weight distributions of the original baseline model (pictured left) compared to the tuned model using the validation metric $\frac{\sigma}{1-|\mu|}$ (pictured right).

- Weights from charged pileup significantly improve while neutral performance stays about the same.
- Tuned model introduces much more precision in correctly identifying pileup particles.

Problem Formulation

Approach: Perform hyperparameter tuning via Bayesian Optimization on a **Semi-supervised Graph Neural Network** that trains on charged particles (labeled) and infers on neutral particles (unlabeled).
Motivation of the hyperparameter tuning:

- SSL framework already undergone pretraining on CMS fast simulation samples and proven to be able to compete with existing algorithms like PUPPI[2].
- To extend this framework, CMS full simulation data was used which contains more complexity in geometry and replicates real world data and physics performance was prioritized.

Challenges:

- Graph structure itself is noisy and complex; neutral and charged particles don't have the same features.
- Many different packages were available to perform Bayesian Optimization. Choosing a package that was effective yet also provided more in-depth information about performance dependency from parameters was important.

Physics Performance

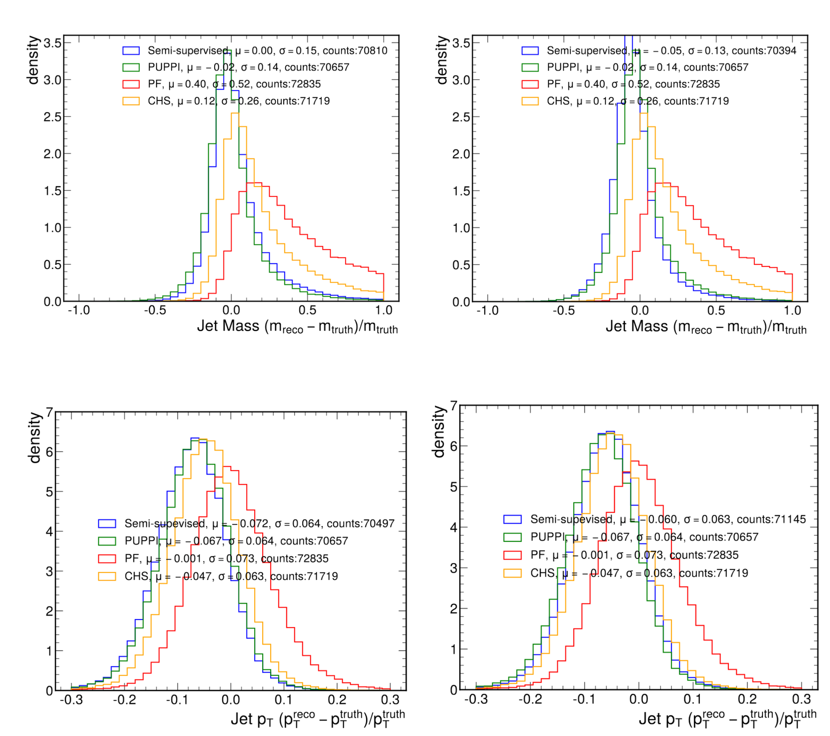

Figure 3: Both Mass and PT reconstruction pictured shows improvement over PUPPI before (left) and after (right) optimization.

Methodology

Configure parameters — Train Model — Acquire best parameters — Cluster on test events

- **Step a) Configure parameters:** The ratio of PU to LV particles that were masked was altered with dropout to attain the best performance in terms of neutral weights.
- **Step b) Train Model:** A separate model was trained each optimization step with the next parameters being decided by the Bayesian Optimization algorithm to be able to statistically find the best parameter set.
- **Step c) Acquire Best Parameters:** The parameters of the model that achieved the best physics reconstruction performance were found to be those that minimized a nonlinear physics metric from the validation dataset.
- **Step d) Cluster on test events:** With the model with the best parameters saved, the best model was then clustered on a much larger number of test events to see how well the model generalized.
- **Relevant Equations:**
  Bayesian Optimization algorithm: $X_k = argmax_X \; u(X|D_{1:k-1})$ where X_k is a point in the n-D parameter space at the kth optimization run and $D_{1:k-1}$ is the set of scores associated with the objective function. u represents the acquisition function. **Physics performance metric**: $\frac{\sigma}{1-|\mu|}$ where μ and σ are the validation mean and standard deviation of the mass reconstruction plots.

Conclusion

Future Direction:

- Experiment with internal model parameters to try to adapt the Gated model to this specific task.
- Adopt domain adaptation techniques to handle larger domain shifts situations.

Overall, using our aforementioned metric and focusing more on physics performance than the usual process of reducing the loss function, we have been able to show that this Semi-Supervised Graph Neural Network can perform better than PUPPI in PT and mass reconstruction simply by tuning for dropout and the ratio of which LV or PU particles are masked. Taking into account future optimizations for more internal parameters such as hidden dimensions, convolutional layers, and domain adaptation the future of Pileup Mitigation looks promising.

Main References

[1] Daniele Bertolini, Philip Harris, Matthew Low, and Nhan Tran. Pileup Per Particle Identification. *JHEP*, 10:059, 2014.

[2] Tianchun Li, Shikun Liu, Yongbin Feng, Garyfallia Paspalaki, Nhan V. Tran, Miaoyuan Liu, and Pan Li. Semi-supervised graph neural networks for pileup noise removal. *The European Physical Journal C*, 83(1), jan 2023.

[3] William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *NeurIPS*, 2017.

[4] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework, 2019.

[5] Vladimir Ivanchenko, Sunanda Banerjee, Gabrielle Hugo, Sergio Meo, Ianna Osborne, Kevin Pedro, Danilo Piparo, Dmitry Sorokin, Phat Srimanobhas, and Carl Vuosalo. Cms full simulation for run 3. *EPJ Web of Conferences*, 251:03016, 01 2021.

[6] Oac-2117997. A3D3 funding grant.