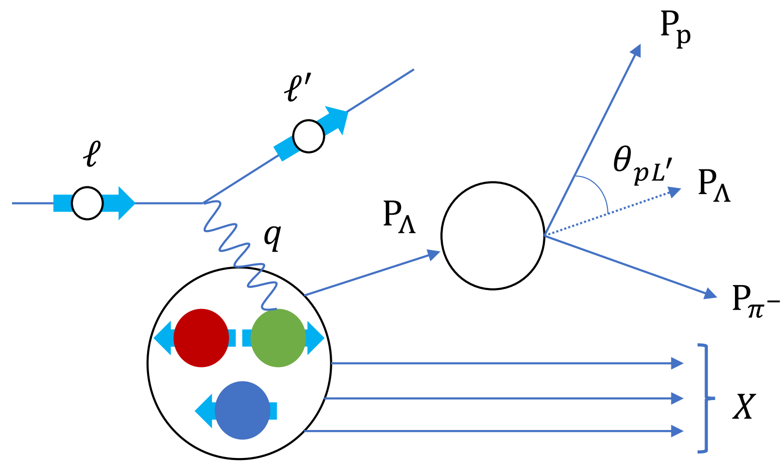
Domain-Adversarial GNNs for Λ Identification with CLAS12

12/Oct./22, Matthew McEneaney, Duke University

matthew.mceneaney@duke.edu

Research supported by:

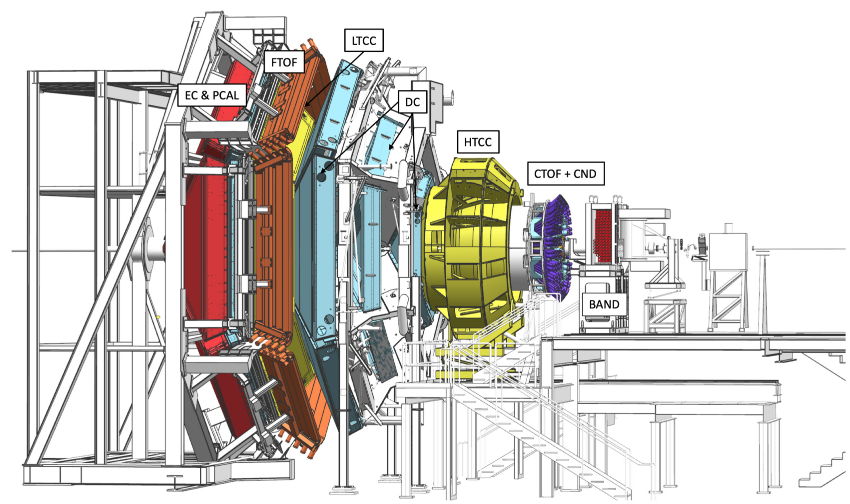
Λ Baryons at CLAS12



- Λ polarization is easily accessible from the $\Lambda \rightarrow p\pi^-$ channel:

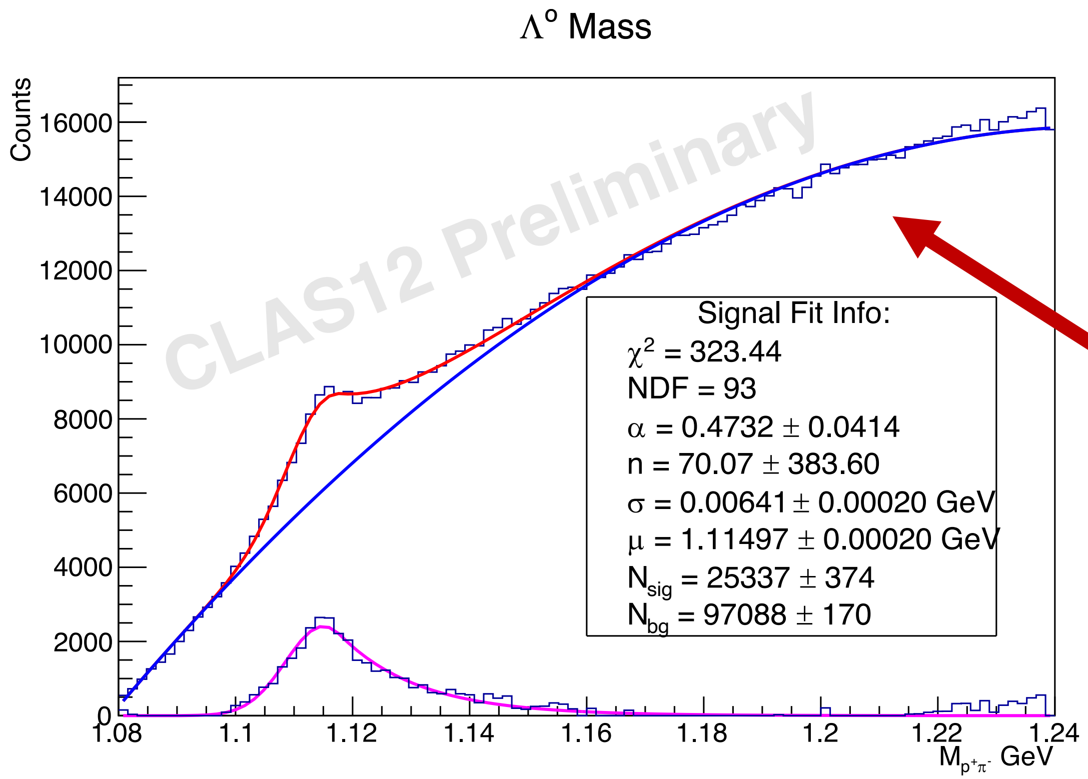
$$\frac{dN}{d\Omega_p} \propto 1 + \alpha P_b D(y) D^{\Lambda}_{LL'} \cos\theta_{pL'}$$

- CLAS12 has large angular coverage and good resolution/PID capabilities



V. Burkert et al., NIM A, January 2020

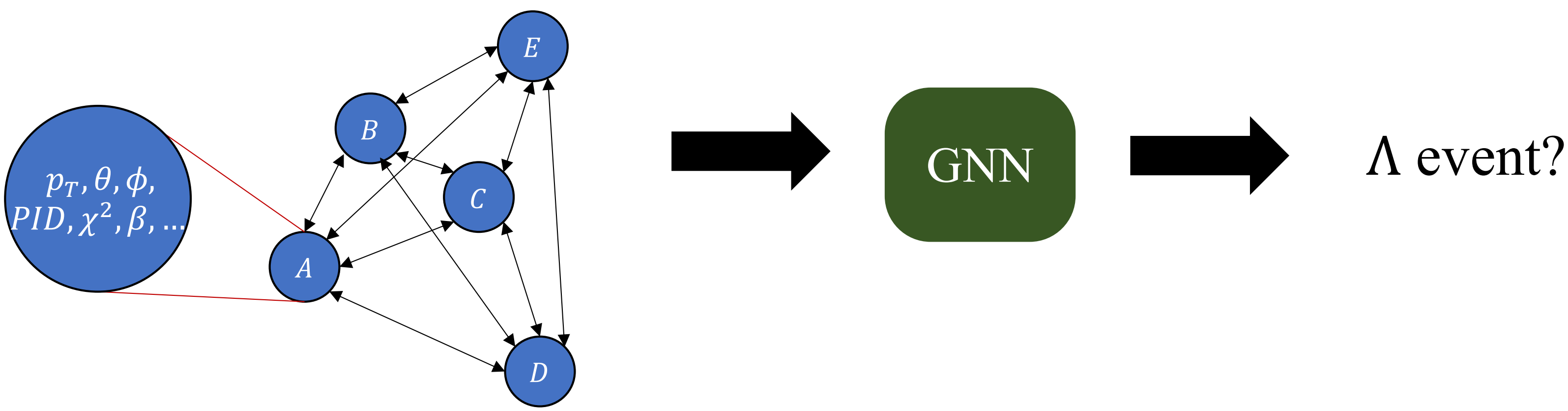
Invariant Mass Signal



We want to reduce this background

Graph Neural Networks (GNNs)

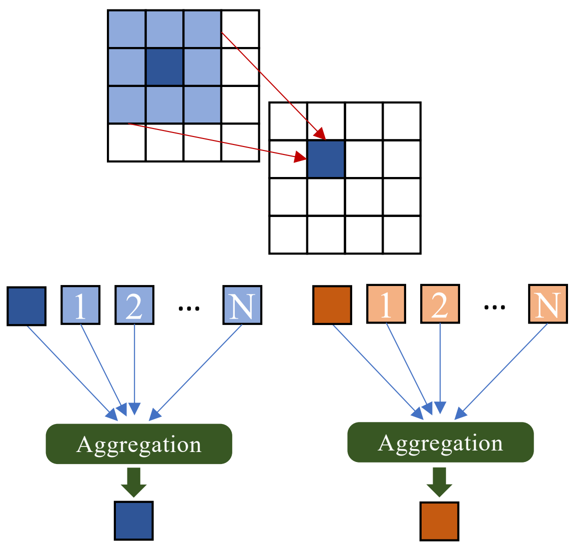
- **Idea**: use GNN to reduce background in invariant mass spectrum on event-by-event basis
- Pass each event as fully-connected, bidirectional graph
- Each particle is a node with its own data: p_T, θ, ϕ, etc.



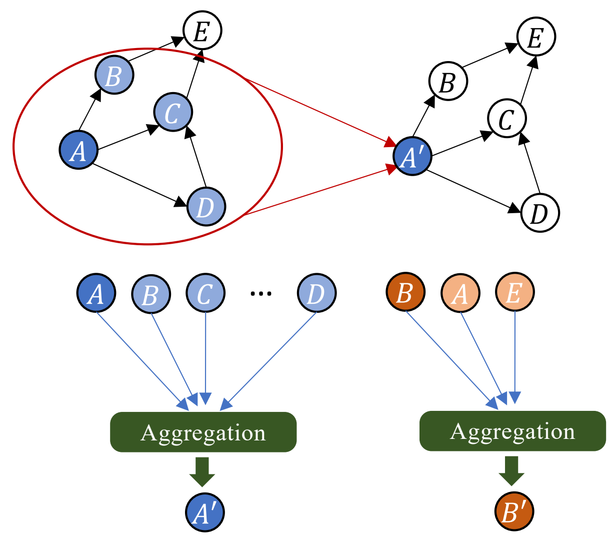
Graph Neural Networks (GNNs)

- At basic level, function as generalized form of CNNs

CNN



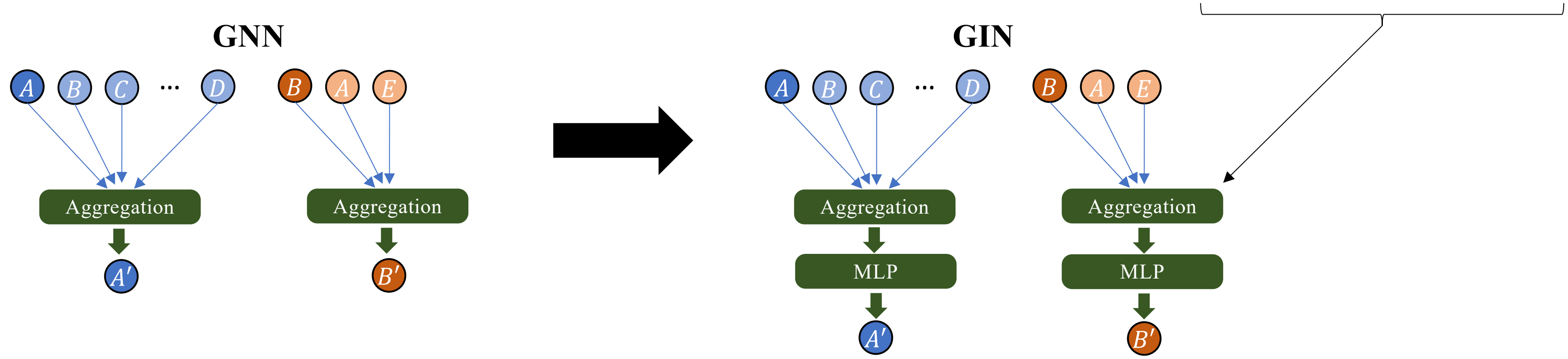
GNN



Graph Isomorphism Network (GIN)

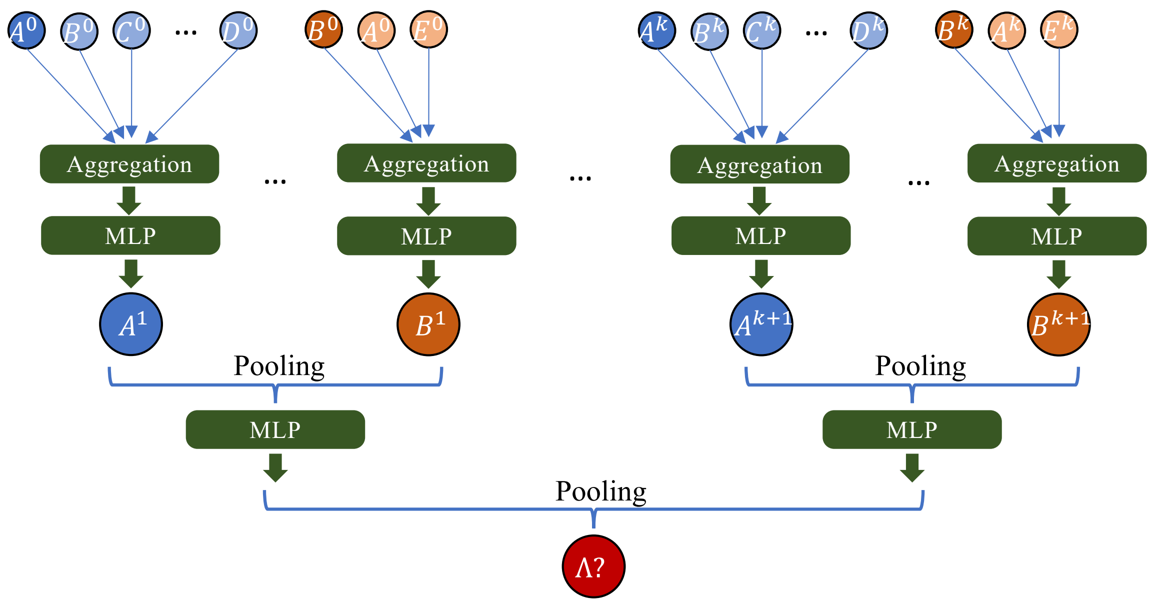
- GIN is in same class as algorithms testing graph isomorphism, and ensures aggregation is injective
- Compare with basic GNN convolution:

$$h_v^{(k)} = \mathrm{MLP}^{(k)}\left(\left(1+\epsilon^{(k)}\right)\cdot h_v^{(k-1)} + \sum\nolimits_{u\in\mathcal{N}(v)} h_u^{(k-1)}\right)$$



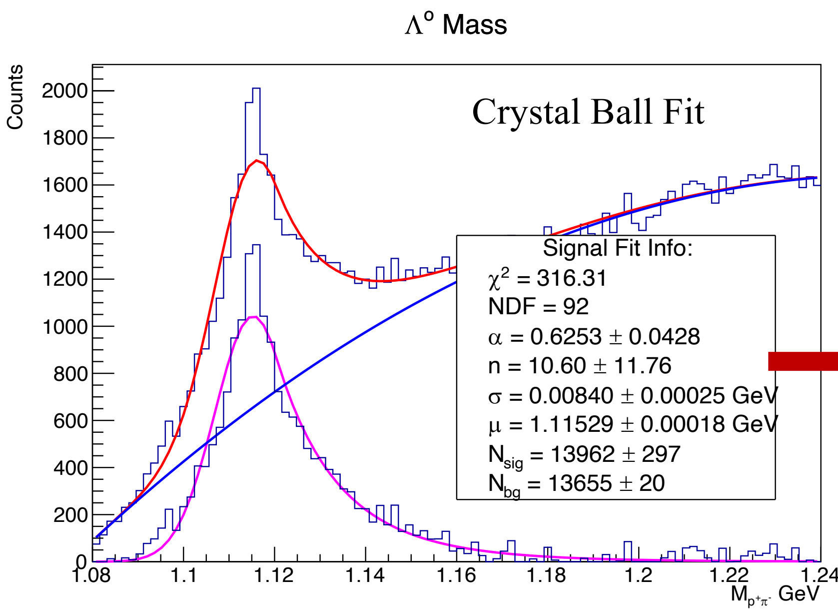
Graph Isomorphism Network (GIN)

- Aggregation in final layer is across all previous layers/iterations

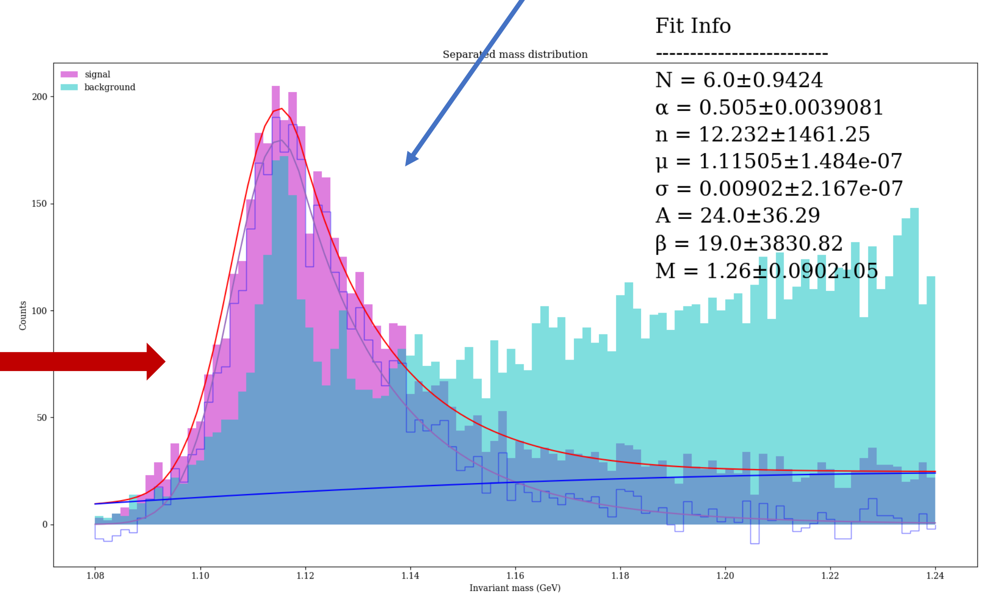


GIN Evaluation on MC

Signal shape is preserved



Before NN

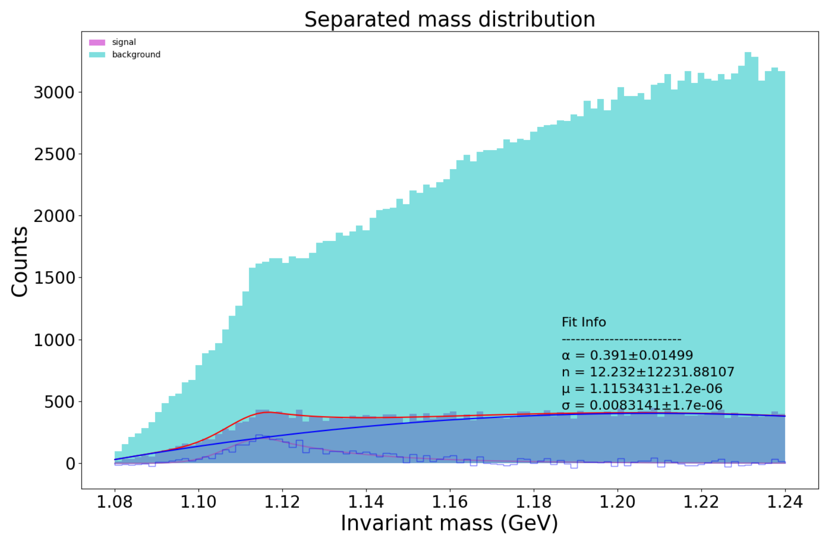


After NN

83.7% Test accuracy and background is significantly reduced!

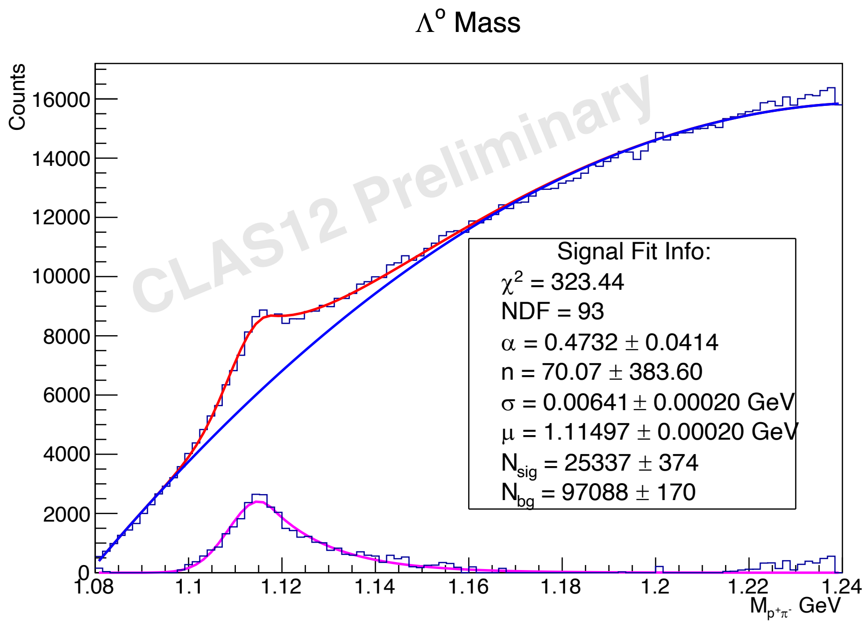
GIN Evaluation on Data

- Use ~240k events from Fall 2018 dataset
- FOM=$N_{sig}/\sqrt{N_{tot}}$ is ~33.53 is compared to ~36.92 without the GIN

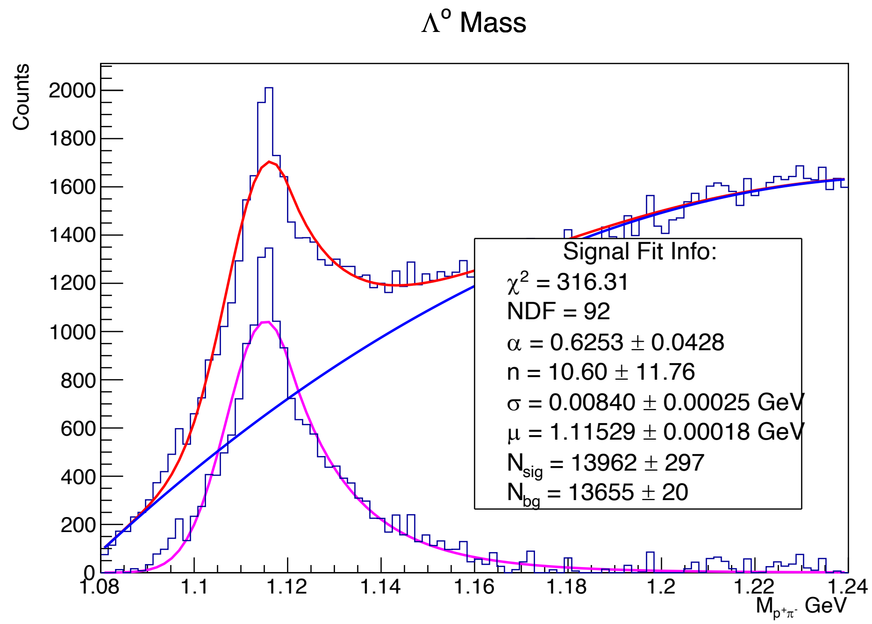


Domain Adaptation

- Problem: target domain does not match source domain



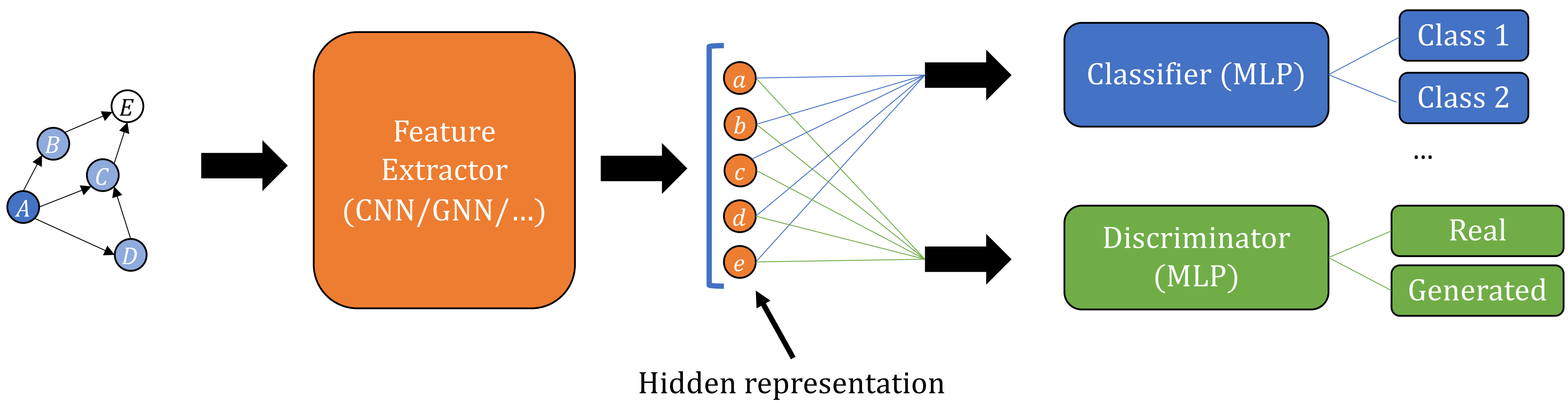
Data



MC

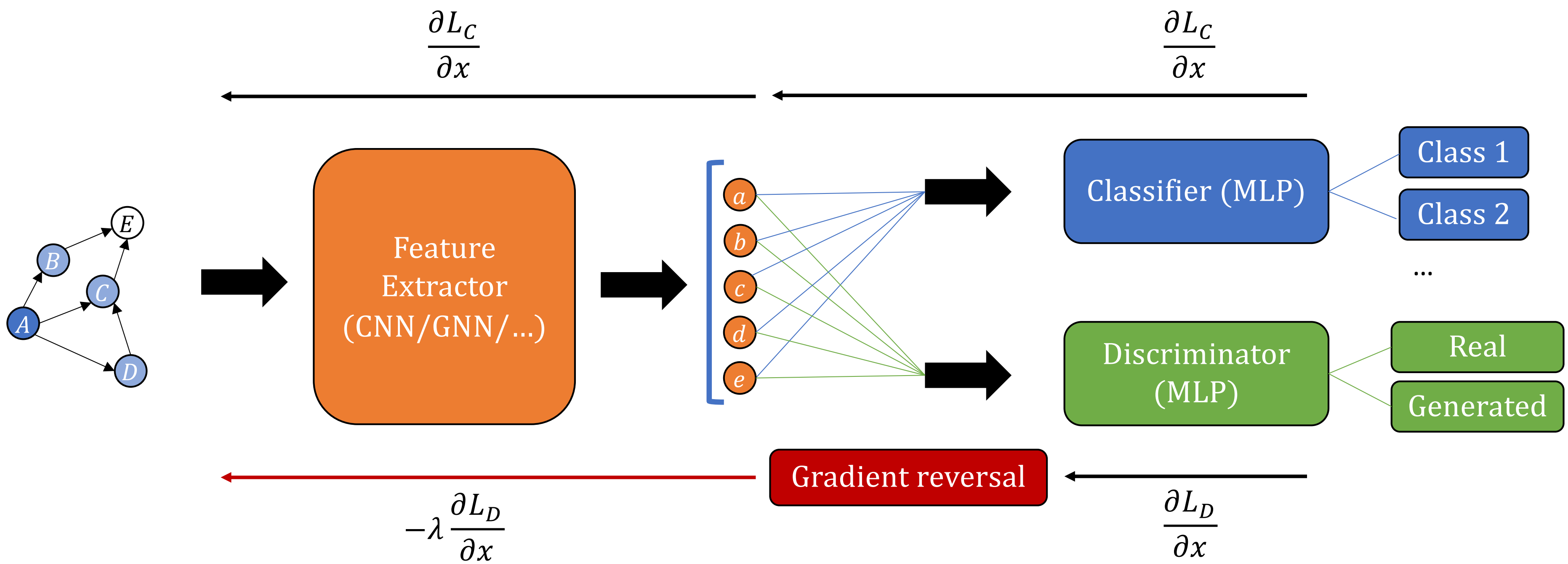
Domain Adversarial NNs

- Minimizes distinction between real and training data
- Two objectives: classification task and domain discrimination



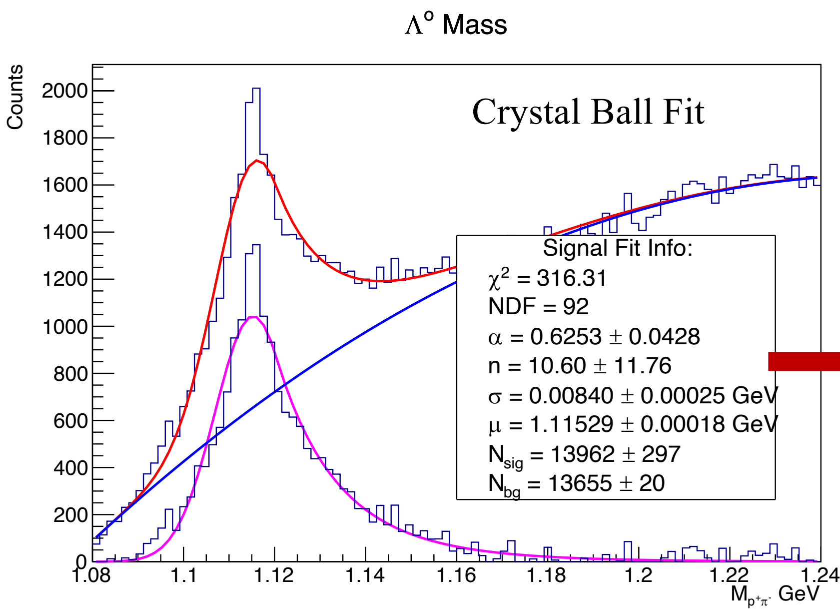
Domain Adversarial NNs

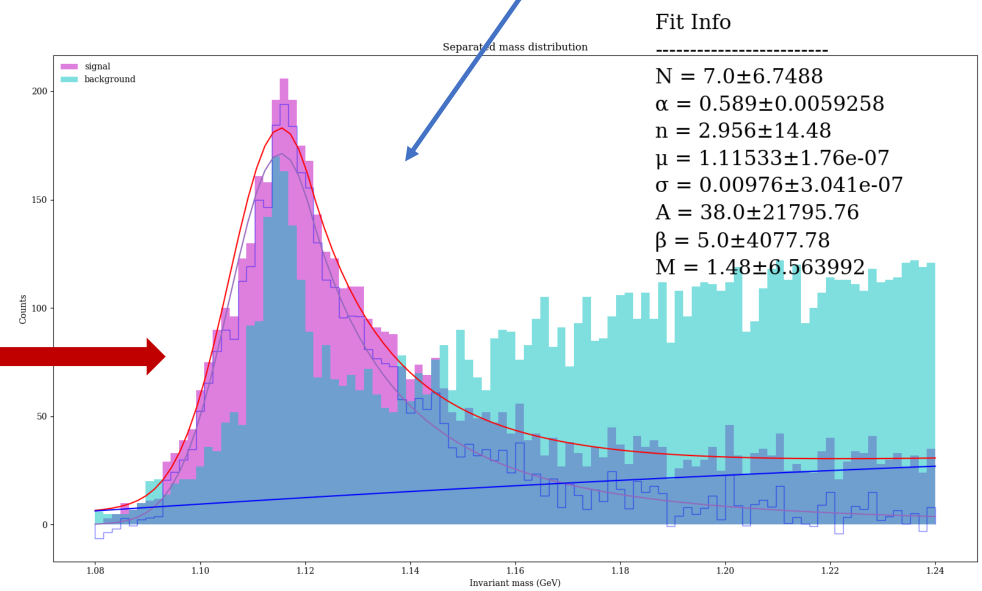
- Reverse gradient from discriminator loss during backpropagation



DAGIN Evaluation on MC

Signal shape is preserved





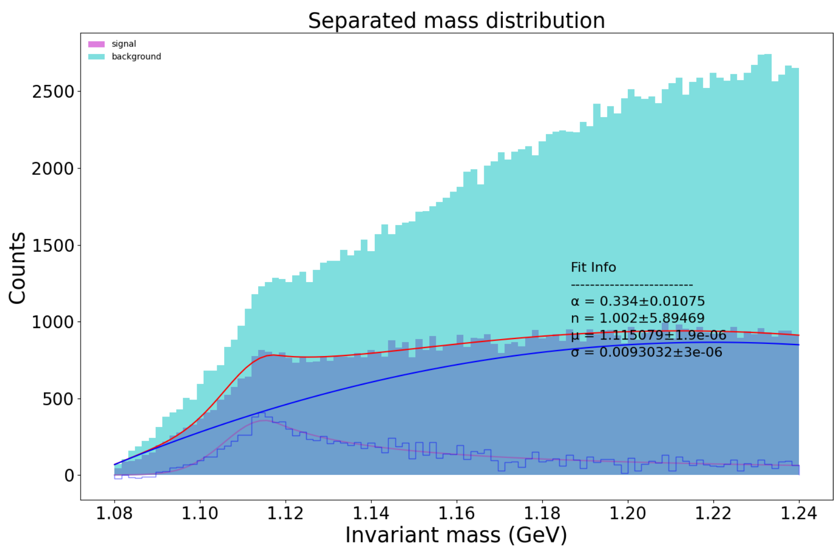
Before NN

After NN

82.9% Test accuracy and background is significantly reduced!

DAGIN Evaluation on Data

- Use ~240k events from Fall 2018 dataset
- FOM=$N_{sig}/\sqrt{N_{tot}}$ is ~47.47 is compared to ~36.92 without the GIN



Stability with Varied Background Fraction

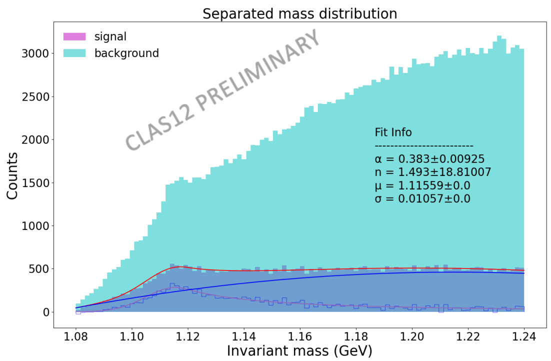
- Run the same hyperparameter optimization
- Train with 3 different background compositions: 80/20, 50/50, 20/80 Δ/Combinatorial background

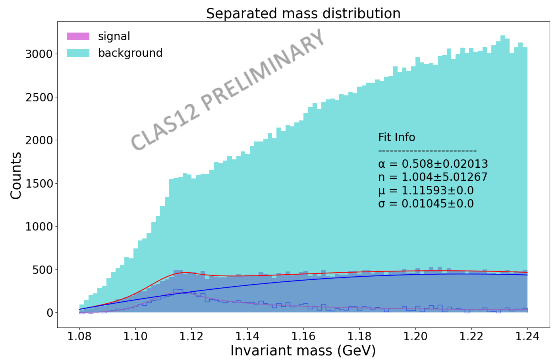
Δ Background Fraction	No GNN	GIN	DAGIN
0.80	36.92	48.27	57.11
0.50	36.92	41.33	43.73
0.20	36.92	46.40	58.21
0.09	36.92	33.53	47.47

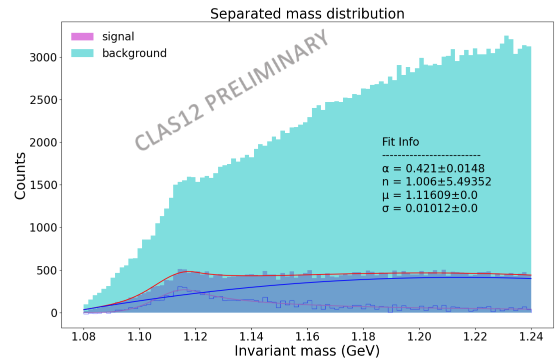
Evaluation on data

Dark blue/magenta is NN-identified signal
Light blue is NN-identified BG for comparison

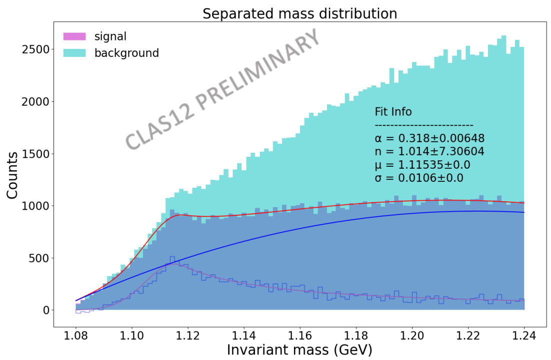
GIN

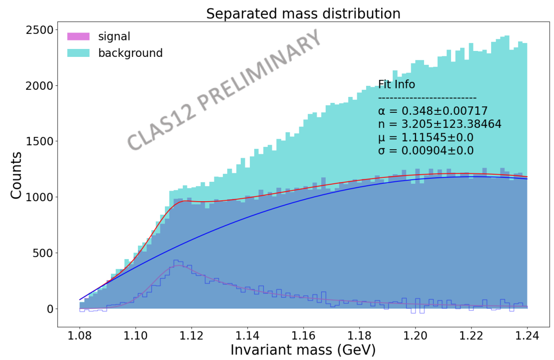


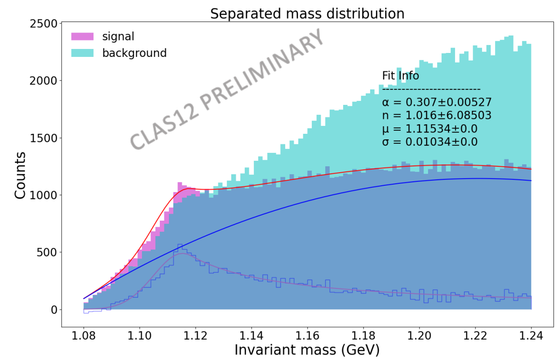




DAGIN







Conclusions and Outlook

- GIN network is a powerful method, other methods like ParticleFlow networks had little success
- **Domain-Adversarial method generally increases FOM**
- Potential improvement from adding detector data as inputs or using other networks, e.g., Subgraph GNNs, (arXiv:2206.11140)
- Performance limited by quality of MC simulation
- Similar method could be **useful for EIC Λ studies** where feed-down is expected to be more significant, (PRD 105, 094033, 2022)

Thank you!

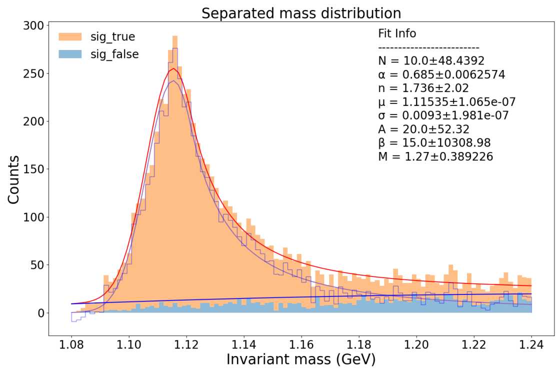
Research supported by:

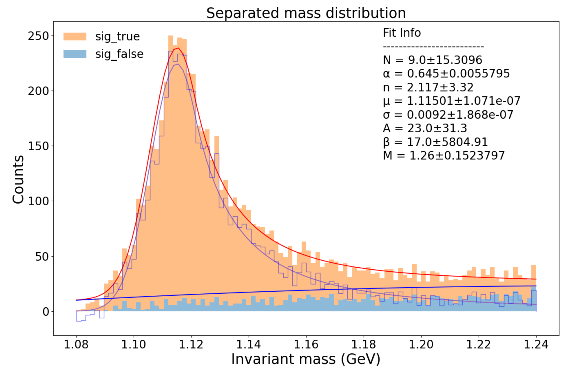
True NN-identified signal (yellow) over false signal events (blue)

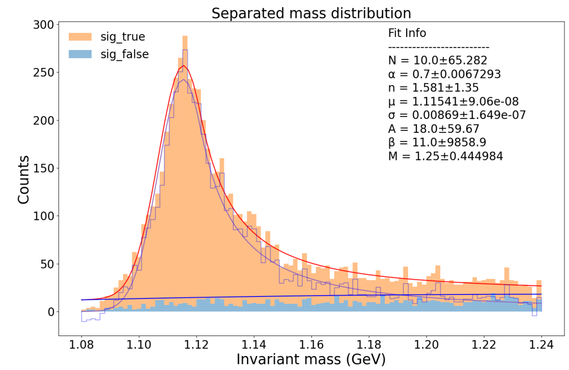
Comparison of MC Truth to Fit Results

80% Δ BG | 50% Δ BG | 20% Δ BG

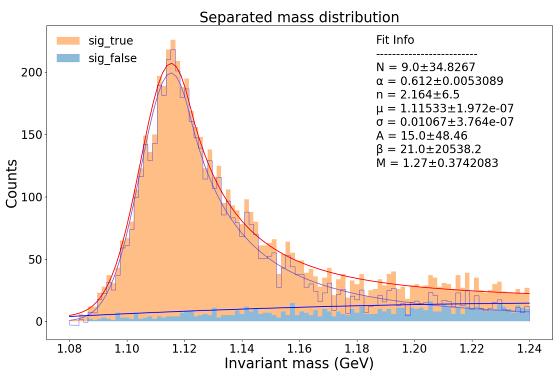
GIN

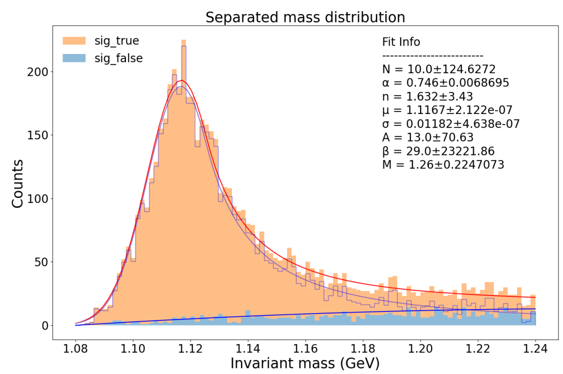


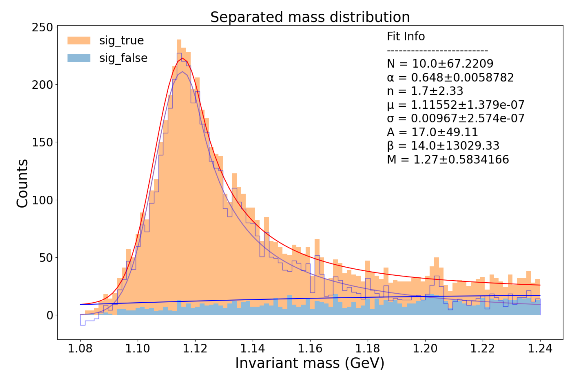




DAGIN



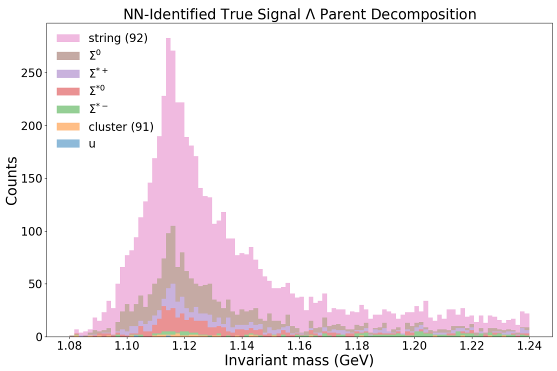


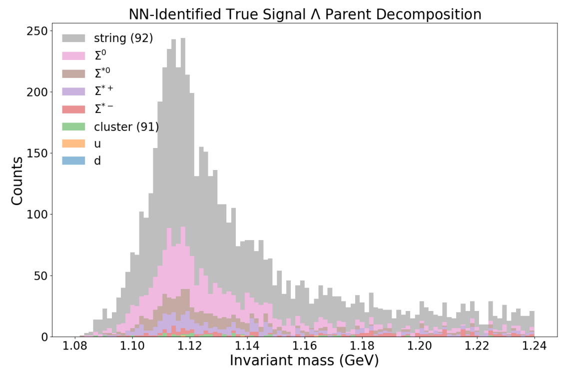


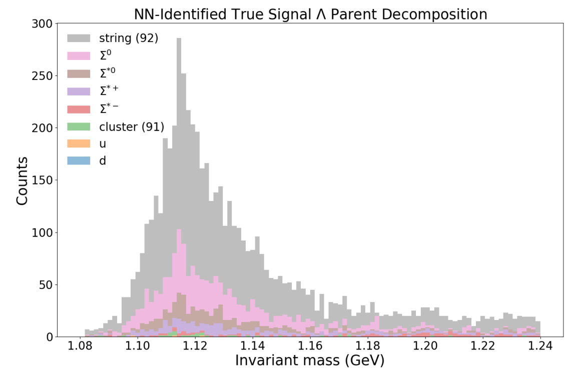
Λ Parent of True Signal/False BG GIN

80% Δ BG | 50% Δ BG | 20% Δ BG

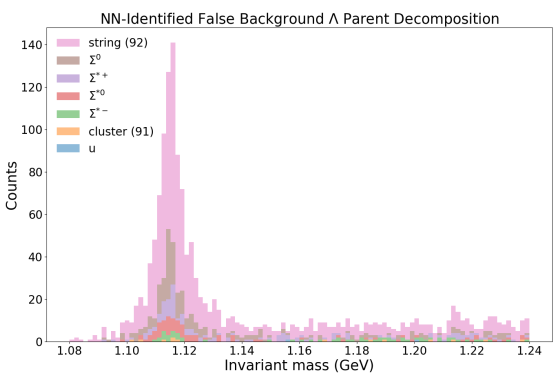
True signal

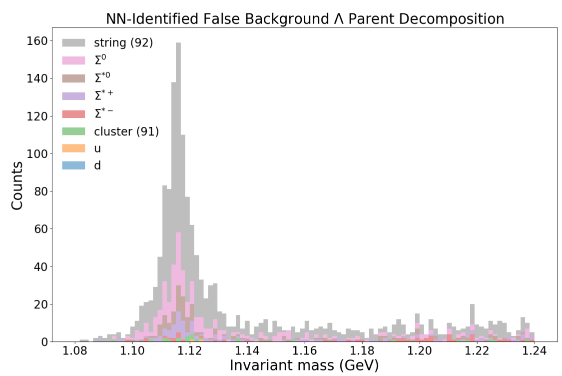


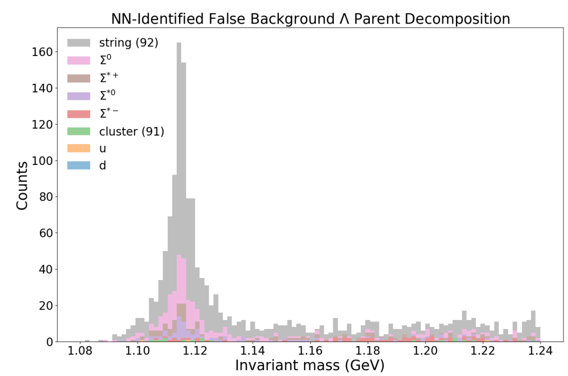




False BG



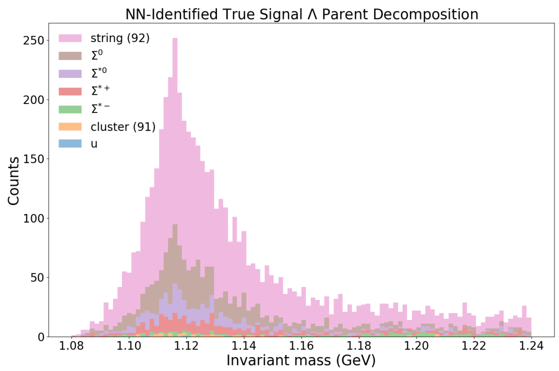


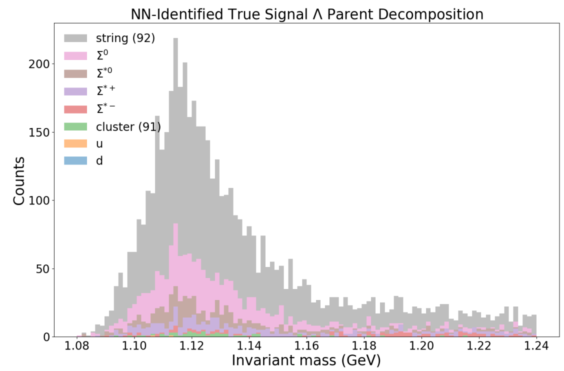


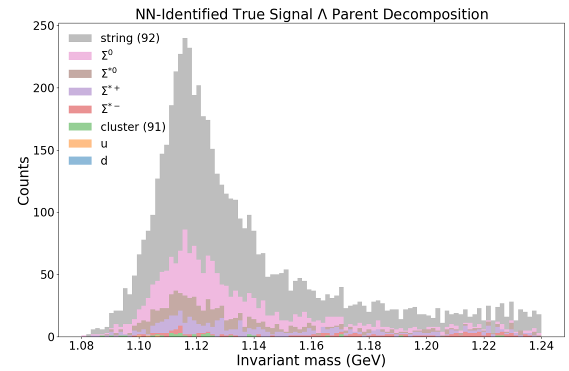
Λ Parent of True Signal/False BG DAGIN

80% Δ BG | 50% Δ BG | 20% Δ BG

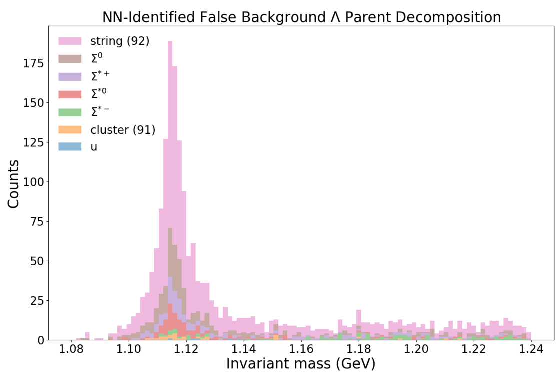
True signal

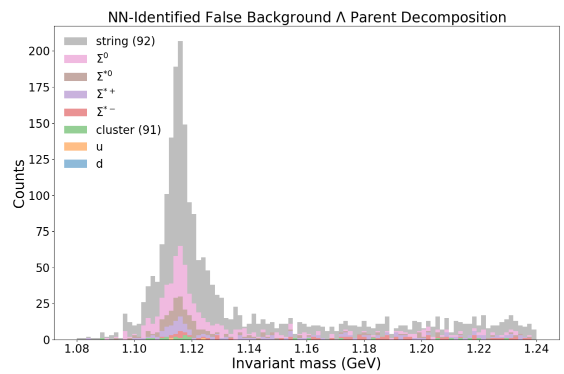


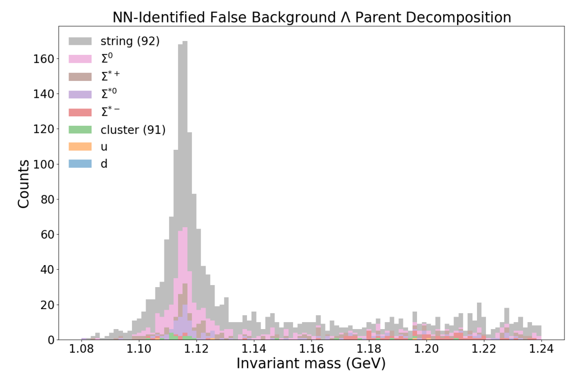




False BG

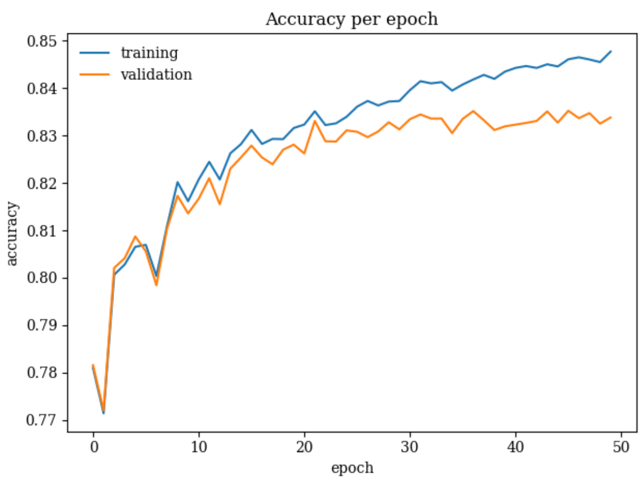


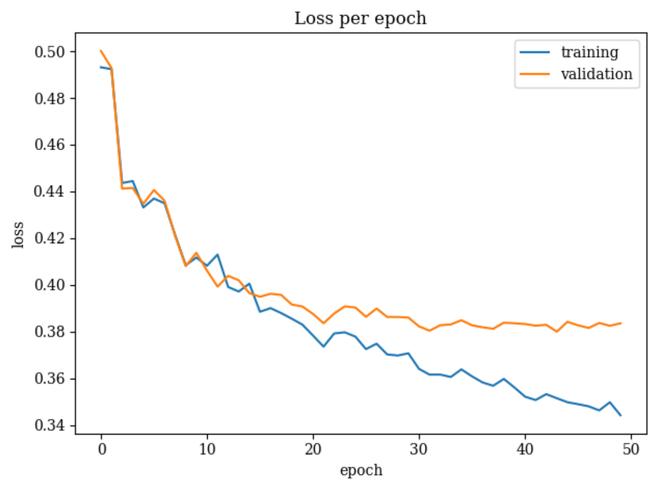


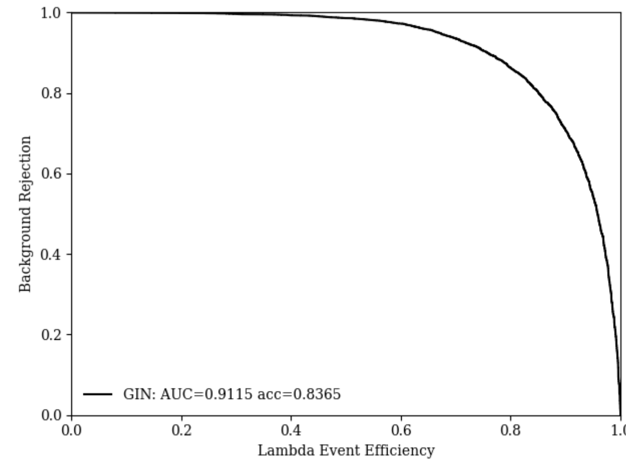


Training Results: GIN

- Implementation detailed in writeup
- Optimize hyperparameters with Optuna TPESampler
- Test accuracy is ~83.7% but still need to reduce overtraining







Training Results: DAGIN

- Optimize hyperparameters with Optuna TPESampler
- Test accuracy is ~82.9%