



東京工業大学
Tokyo Institute of Technology

GNN機械学習技術を用いた 電子 - 光子の分類アルゴリズムの開発

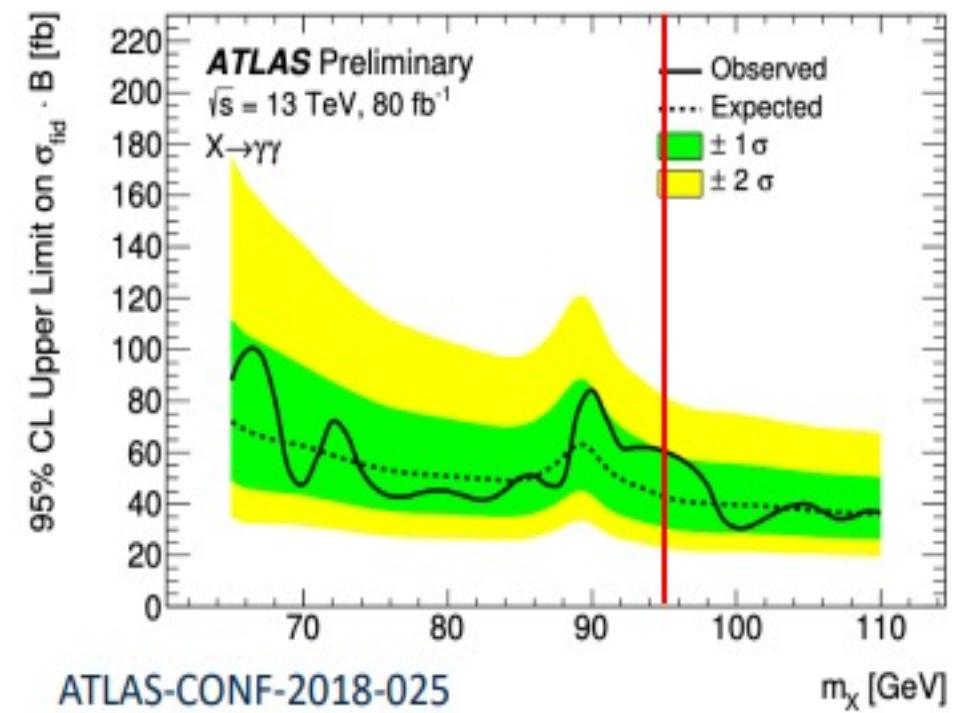
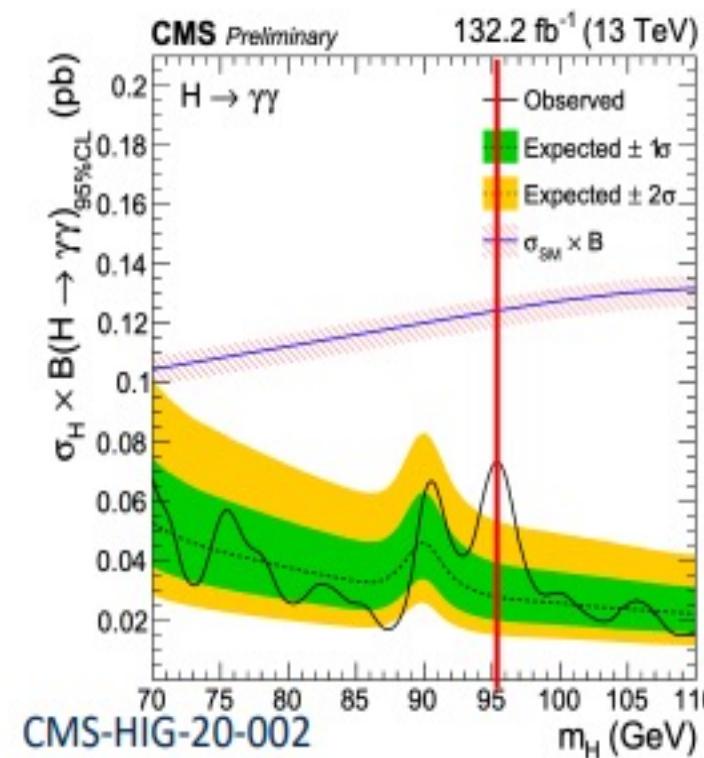
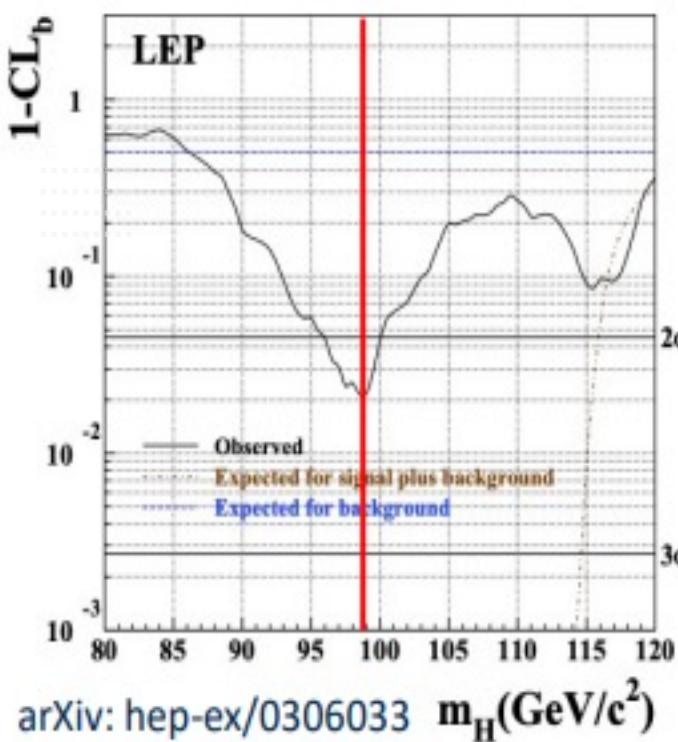
著者: Isai Roberto Sotarriva Alvarez, 陣内修, 留目和輝, 山口洋平, 増渕達也^A

所属: 東京工業大, 東京大^A

講演番号: 21pT2-10

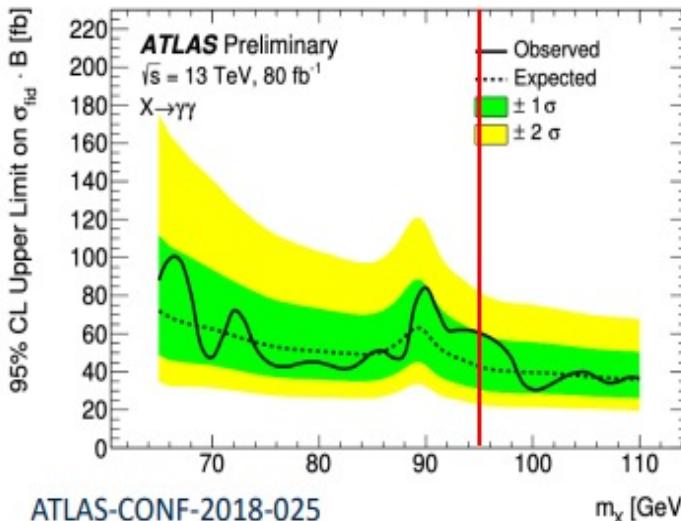
イントロダクション

- SUSYなどの理論は、標準模型のヒッグス・ボソンと同様のスピン0粒子の存在を示唆している。
- いくつかの実験では、96GeV付近のダイフォトン・スペクトルにわずかな超過が報告されている。
- (これは標準模型を超える新しいヒッグス粒子の兆候である可能性がある)

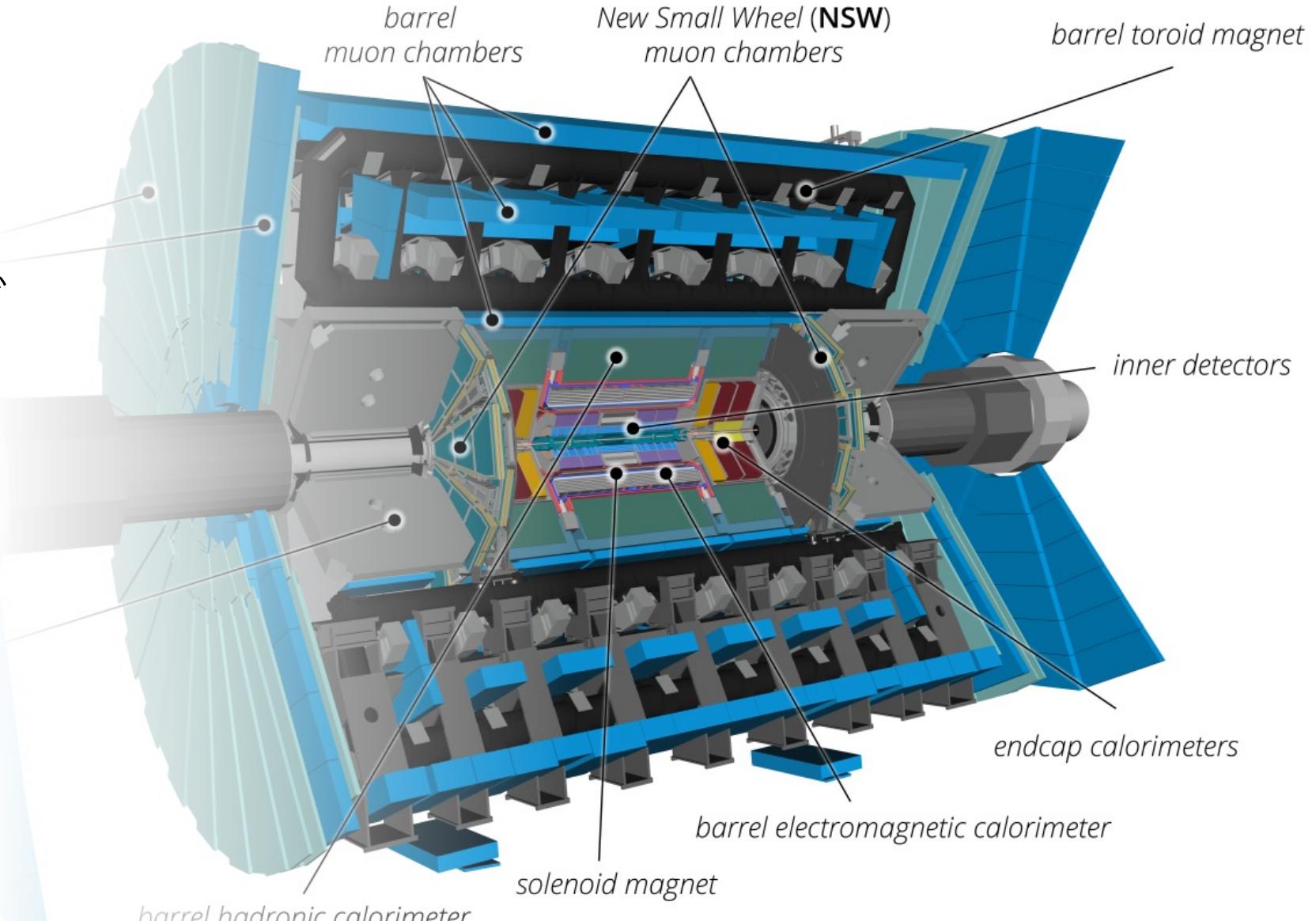


ATLAS

- ATLAS 実験は、宇宙についての理解を広げることを目的として、LHCにおける汎用大型検出器の一つである。

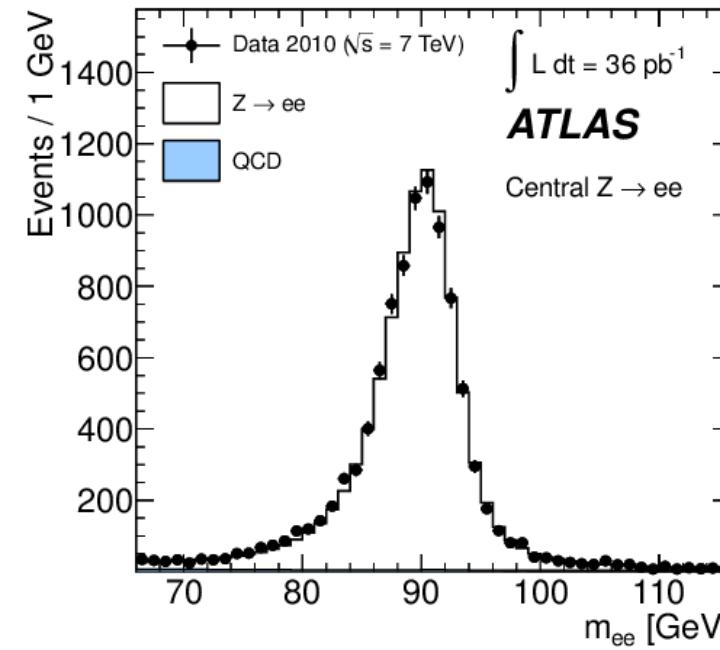
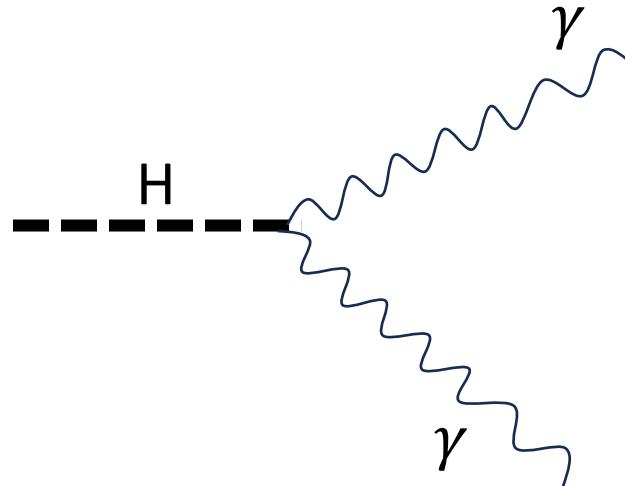


2024年03月21日



Bianchi, Riccardo Maria and ATLAS Collaboration,
ATLAS experiment schematic illustration(2022)

低質量[60 – 110] GeV $H \rightarrow \gamma\gamma$ 解析 シグナル バックグラウンド



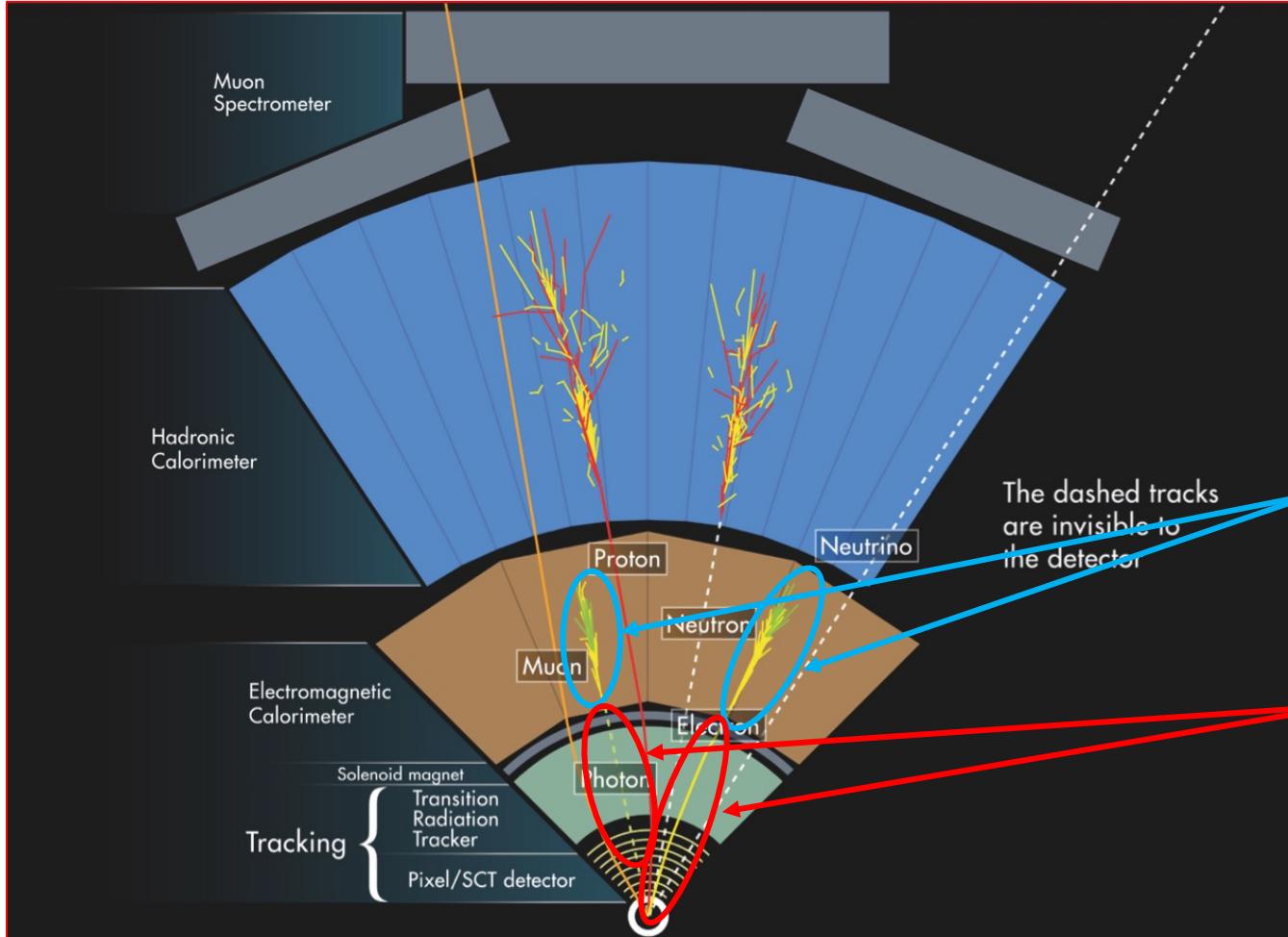
電子バックグラウンドの主な発生源の 1 つの例
[W/Z properties \(except mass\) form ATLAS and CMS](#)

この解析では、2 光子に崩壊する低質量ヒッグス粒子を探索する。

バックグラウンドは主に3つのコンポーネントで構成される。

- QCDプロセスから発生した光子。
- 光子と誤認されたジェット。
- 光子と誤認された電子。

光子検出



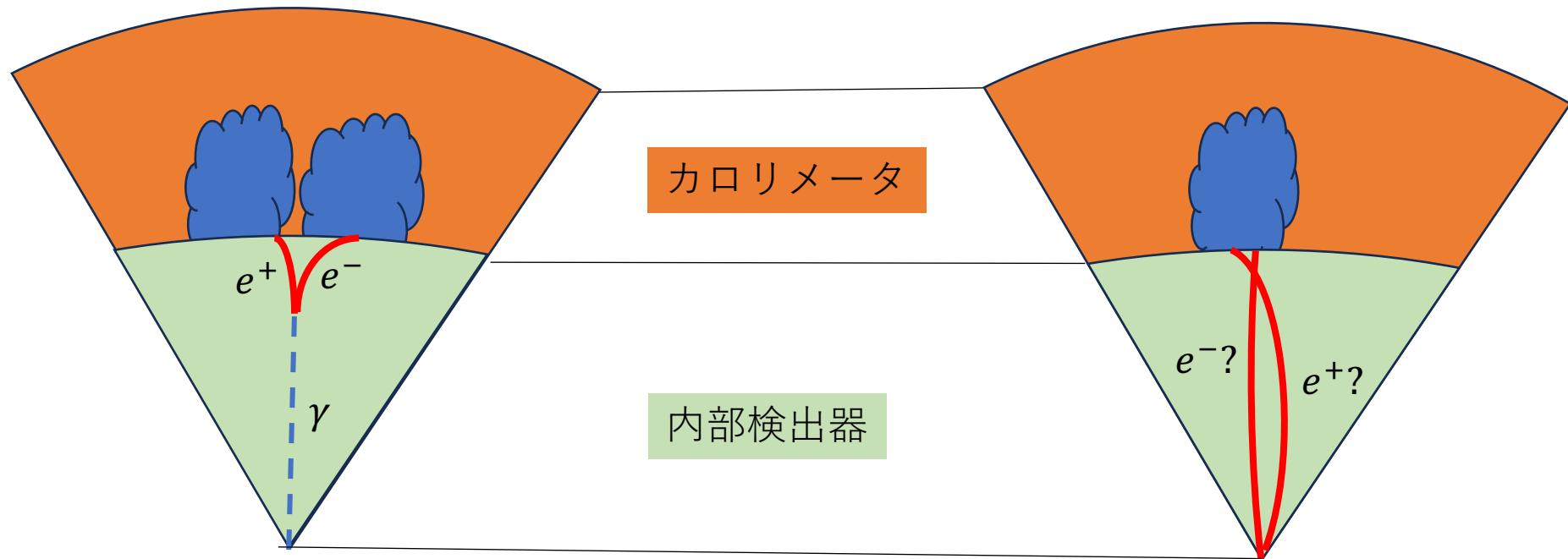
Vanden Broeck. Renilde, [ATLAS - Visualising particles. Visualiser les particules](#), 2019

ATLAS 実験は、粒子再構成のための特殊な検出器の層に依存している。

- 非常に似た信号を生じる電子と光子を識別するため、2つの検出器情報が主に用いられる。
 - カロリメータ: 電子または光子のエネルギーに関する情報を提供する。
 - 内部検出器: 荷電粒子(この場合は電子のみ)の飛跡に関する情報を提供する。

光子再構成

- ある信号が光子か電子か判断するのが難しい場合がある
- RUN2 では、電子/光子の識別に BDT が使用されたが、GNN を使用すると、将来の解析のパフォーマンスを向上できると考えている。



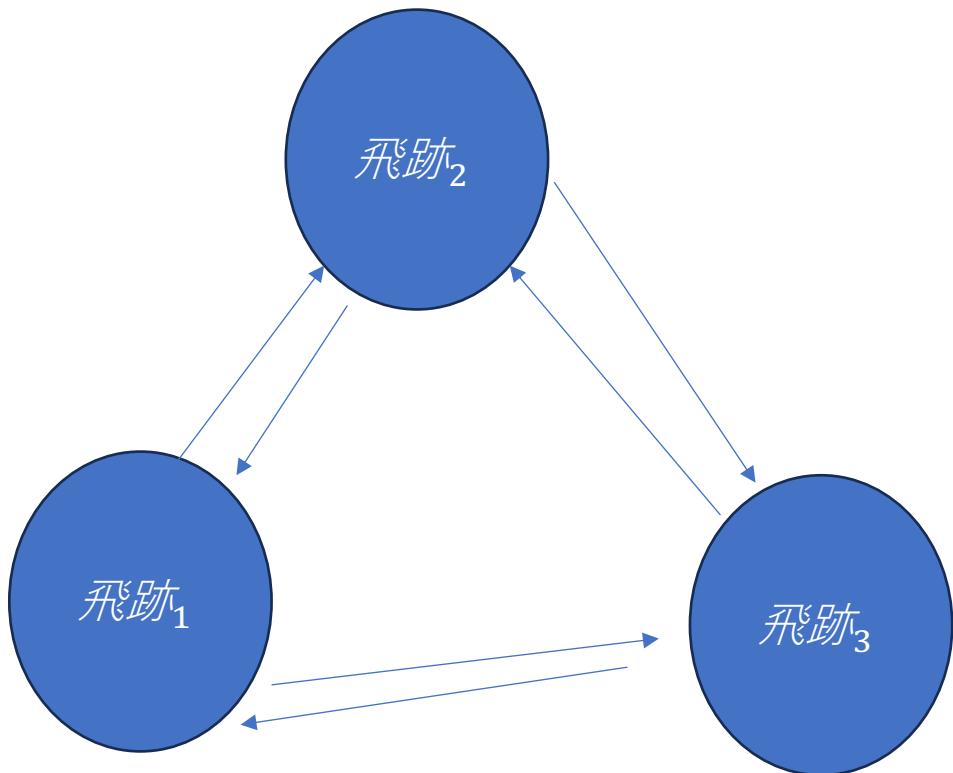
光子による電子・陽電子生成

光子は内部検出器上で電子陽電子対を生成し得る

近くの飛跡のどれが電磁カロリメータで観察されたクラスターに対応するものか不明瞭な場合がある

グラフニューラルネットワーク (GNN)

グラフはノードと呼ばれる個々の要素間の関係を表現する。

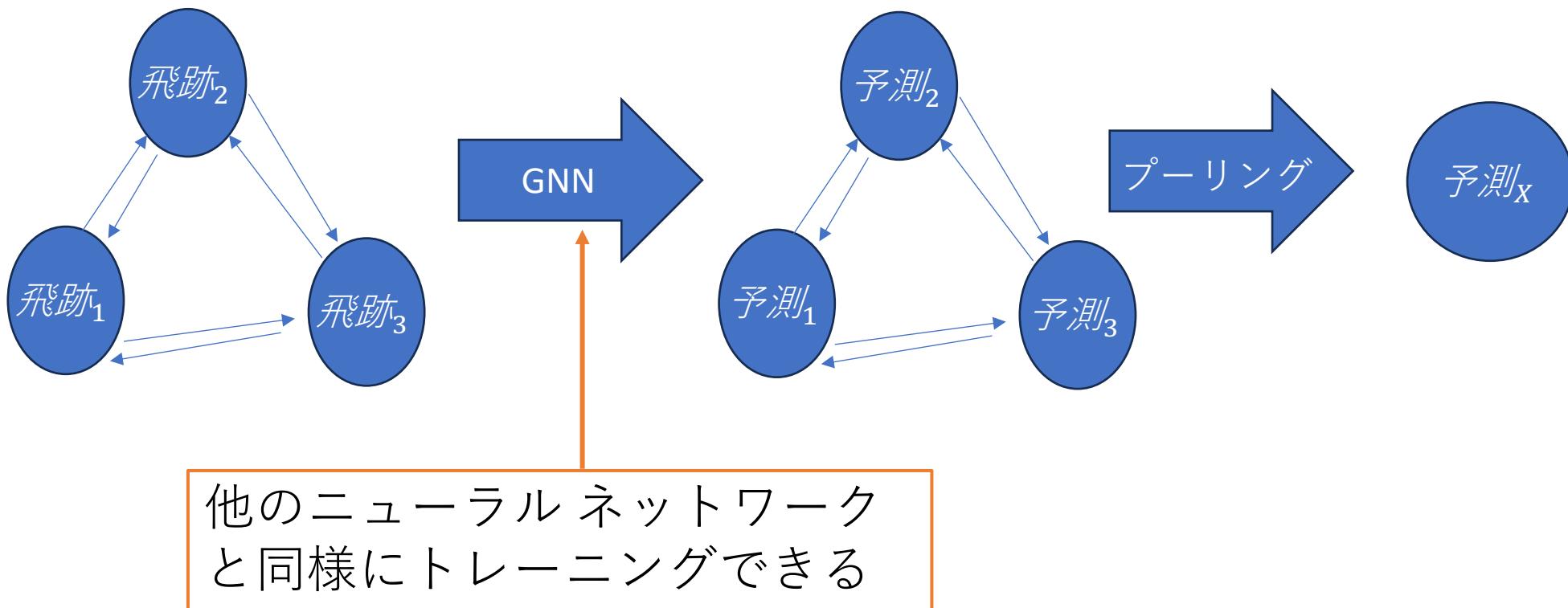


電子光子分類の場合、それぞれのノードは飛跡候補や崩壊点候補など。

再構築された粒子全体で共有され、特定のノードに属さない変数「特徴量」がいくつか存在する場合がある。これはグローバル特徴量と呼ばれる

グラフ・ポーリング

複数のノードの予測は、予測値の最大値、平均値、または最小値を計算することで集約できる。



Transformer Encoderモデルの表現

- グローバル特徴量
- エネルギー
- p_T
- etc.

G
N
N

グラフ・アテンション層

G
P
O
O
I

プーリング層

飛跡候補1+グローバル特徴量

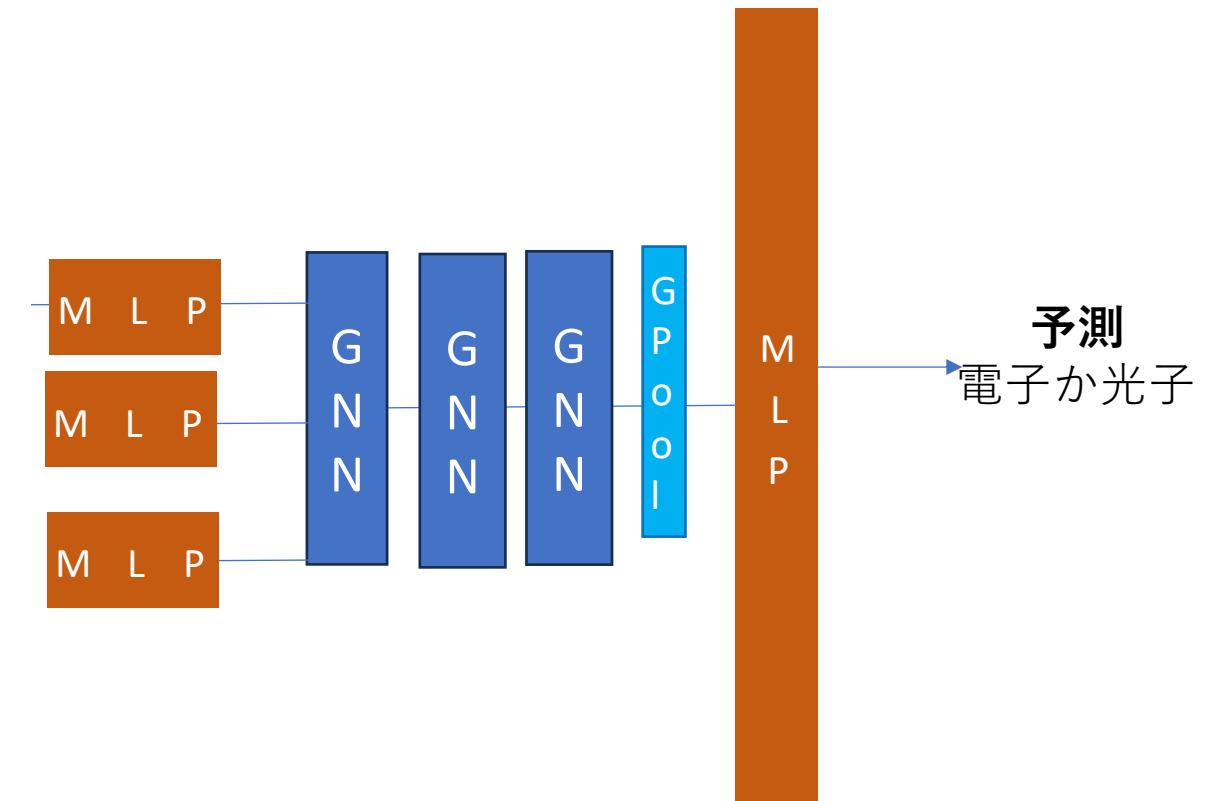
飛跡候補2+グローバル特徴量

:

飛跡候補N+グローバル特徴量

M
L
P

多層パーセプトロン



トレーニングに使用されるシミュレーション

- RUN2向けmc16シミュレーション@ $\sqrt{s} = 13TeV$

低質量 $H \rightarrow \gamma\gamma$ シグナル

Pythia低質量 $H \rightarrow \gamma\gamma$ @複数のエネルギー

- $m_{\gamma\gamma} = 40$
- $m_{\gamma\gamma} = 50$
- $m_{\gamma\gamma} = 60$
- $m_{\gamma\gamma} = 70$
- $m_{\gamma\gamma} = 80$
- $m_{\gamma\gamma} = 90$
- $m_{\gamma\gamma} = 100$
- $m_{\gamma\gamma} = 110$
- $m_{\gamma\gamma} = 120$

低質量 $H \rightarrow \gamma\gamma$ バックグラウンド

- Sherpaの $Z \rightarrow ee$ シミュレーション
- SherpaのQCDバックグラウンドシミュレーション

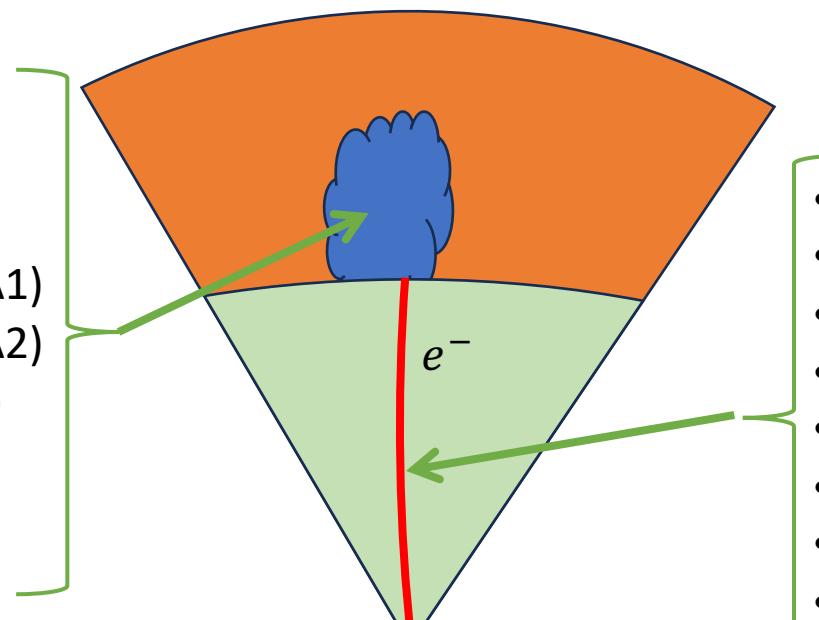
使った特徴量

グローバル特徴量

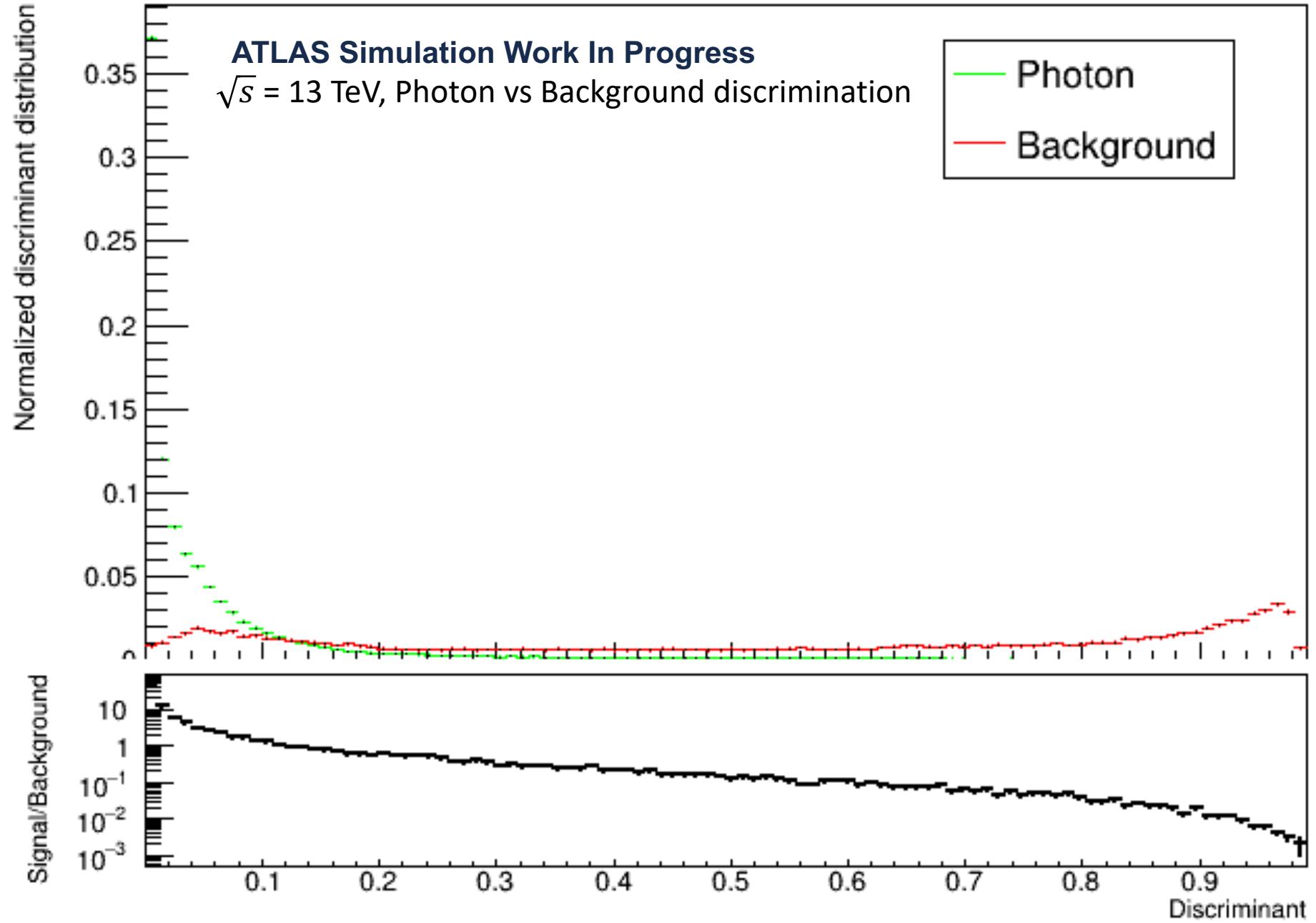
- クラスターのエネルギー
- シャワーシェイプ変数 (F1)
- シャワーシェイプ変数 (F3)
- シャワーシェイプ変数 (WETA1)
- シャワーシェイプ変数 (WETA2)
- シャワーシェイプ変数 (e277)
- p_T
- η
- ϕ
- ラピディティ
- 崩壊タイプ
- 崩壊点候補の数
- 飛跡候補の数

飛跡候補の特徴量

- $d\theta$
- $Z\theta$
- 飛跡の p_T
- 飛跡の η
- ピクセルレイヤーのヒット数
- SCTレイヤーのヒット数
- E/p
- 最内ピクセルレイヤーのヒット数



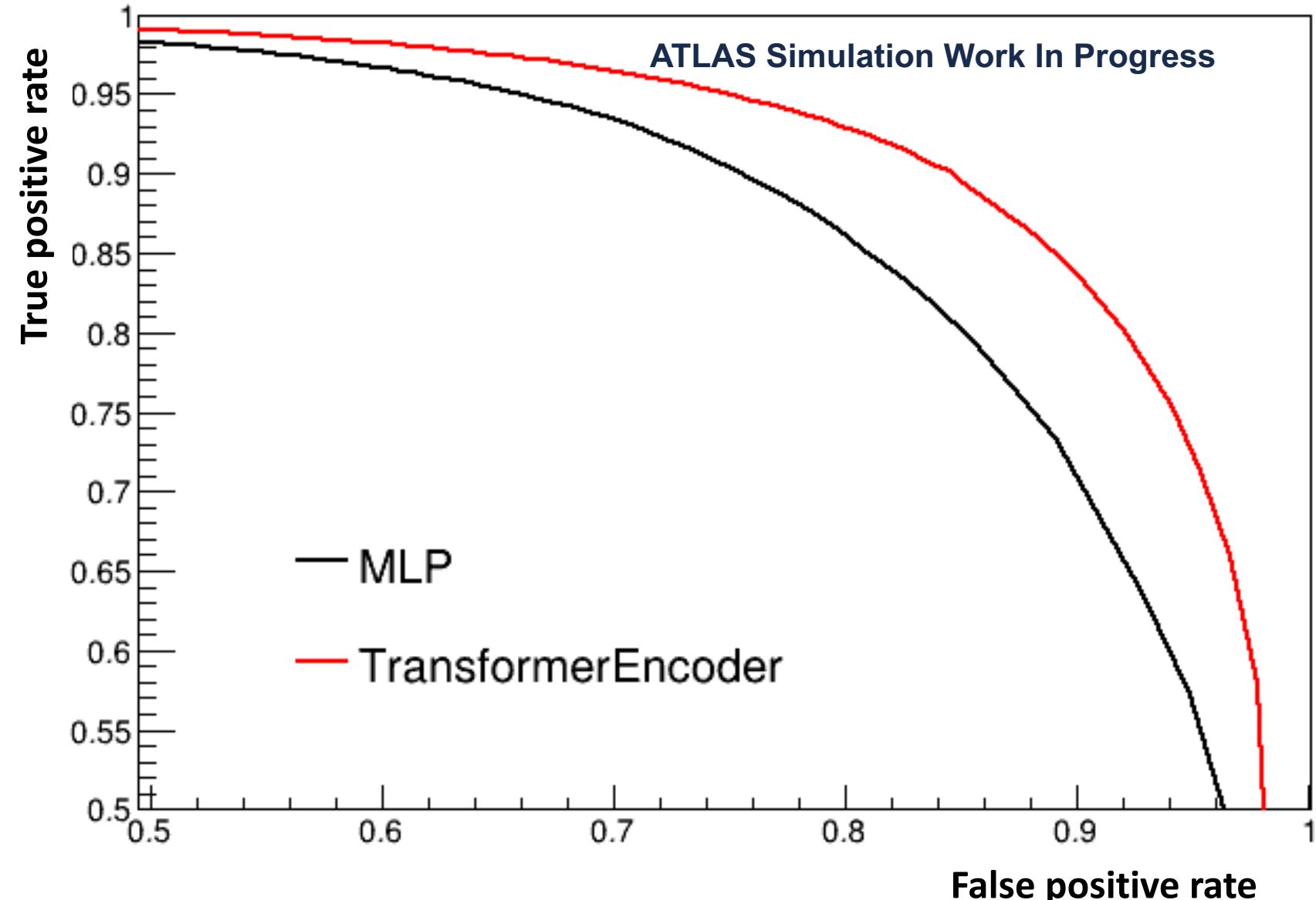
結果



結果

GNNベースのTransformer Encoderは、MLP単独よりも優れたパフォーマンスを示している。

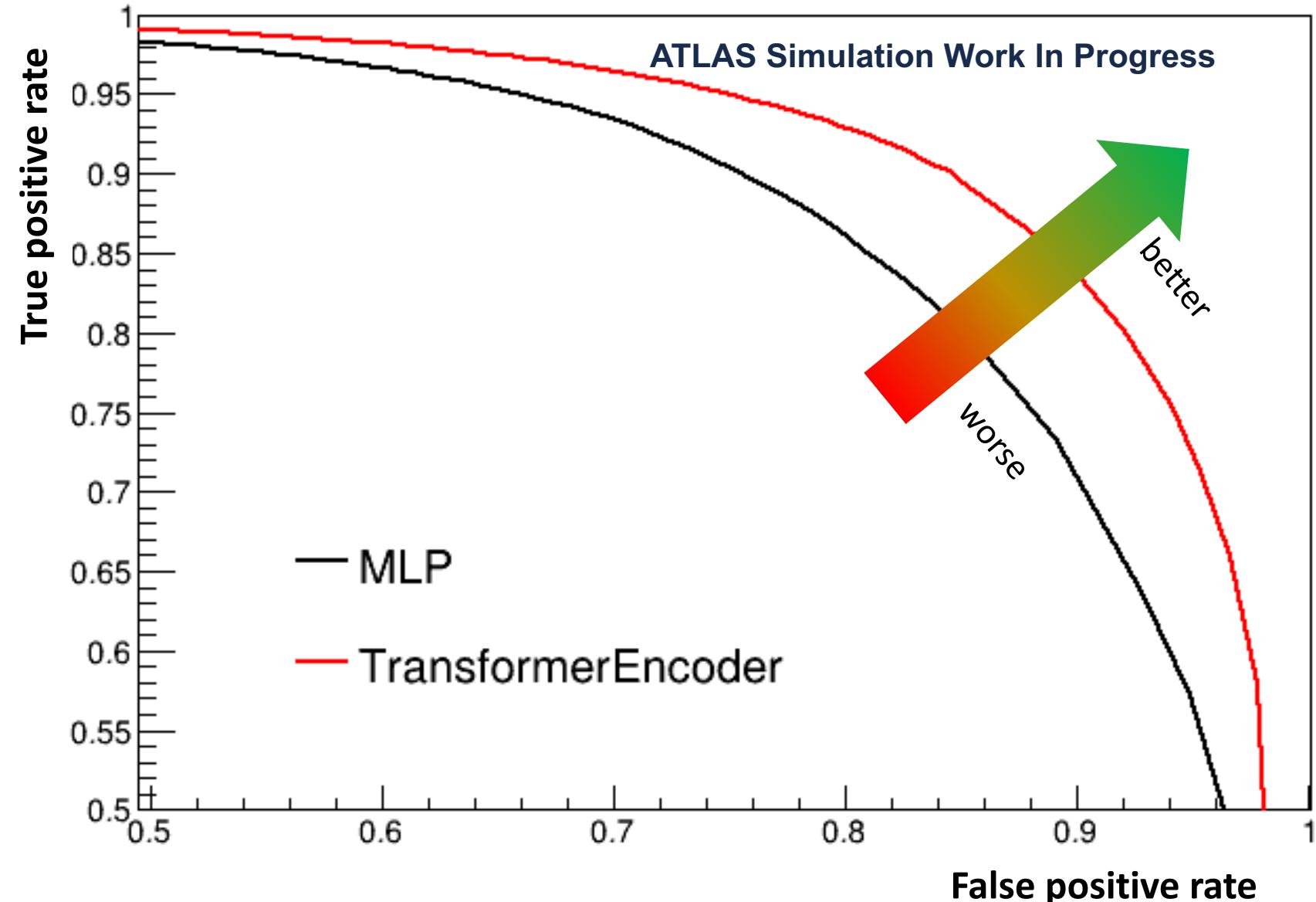
ROC curve



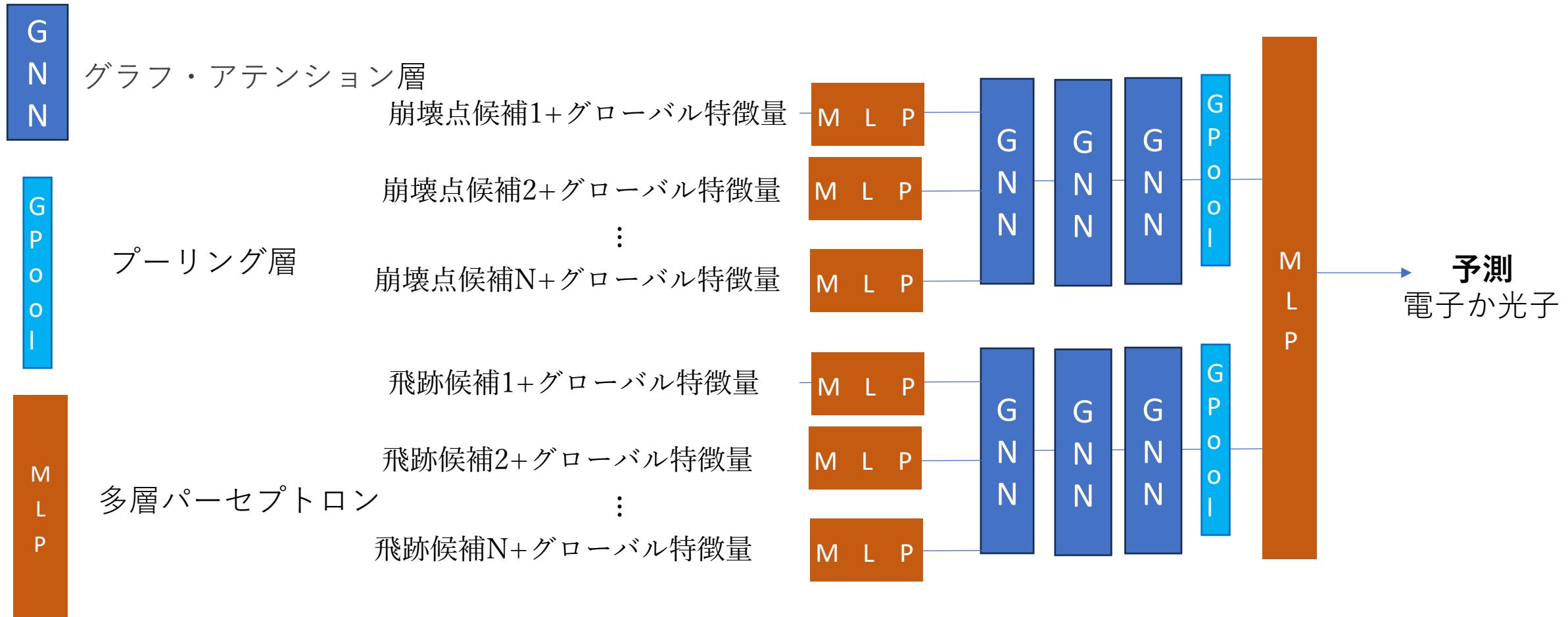
結果

GNNベースのTransformer Encoderは、MLP単独よりも優れたパフォーマンスを示しています。

ROC curve



電子/光子識別用に作成した複数GNN モデルの表現



電子/光子識別用に作成した複数GNN モデルの表現



まとめ

- ・幾つかの実験では、新たな低質量ヒッグスが存在する可能性のヒントが示されている。
- ・ $H \rightarrow \gamma\gamma$ のバックグラウンドの重要な部分は、Z粒子の崩壊に由来している。
- ・グラフは、異なるノード間の関係を表す。 $H \rightarrow \gamma\gamma$ 解析の場合、ノードがさまざまな種類の候補になる。
- ・グラフニューラルネットワークを多層パーセプトロンネットワークと組み合わせて e/γ 分類が行われる。
- ・GNN の使用の実現可能性調査に関する研究が行われている。