東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo

Beyond AI

LHCにおける
生成モデル・異常検知

2022 / 7 / 8

ML@HEP

齊藤 真彦 (ICEPP, Beyond AI)

東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo

Beyond AI

ATLAS実験における
生成モデル for カロリメータシミュレーション
異常検知 for 新粒子探索

2022 / 7 / 8

ML@HEP

齊藤 真彦 (ICEPP, Beyond AI)

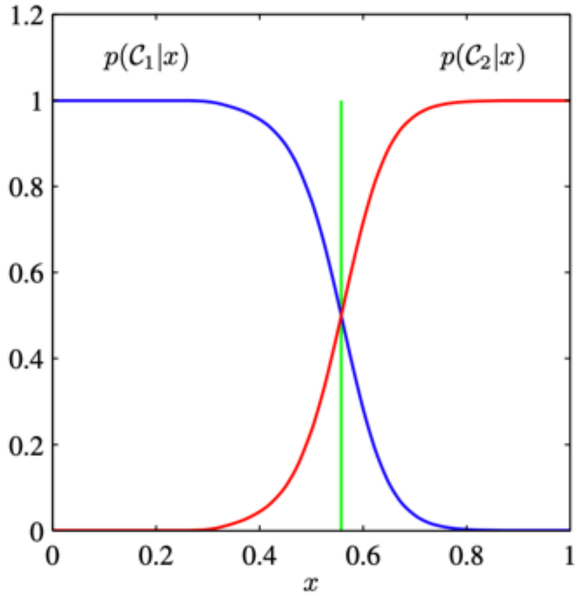
生成モデル (Generative model) と 異常検知 (anomaly detection)

- 深層学習応用の代表例は **識別モデル** (教師**あり**)
  - 入力xが与えられたときの出力yの事後確率 $p(y|x)$ をモデリング

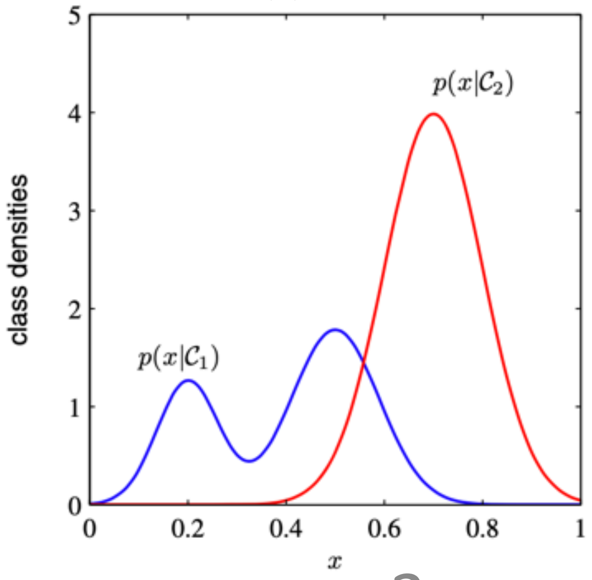
- 入力x の分布 $p(x)$ をモデリングするアプローチ (教師**なし**)
  - 生成モデル
    - $p(x)$ からサンプリングすることでデータを生成
  - 異常検知
    - $p(x)$ が小さい異常データxを見つける
    - ただし、いろいろなバリエーションあり

- ATLAS実験で論文化された研究を中心に手法を紹介
  - **GANを用いたカロリメータシャワーシミュレーション**
  - **弱教師学習を用いたDijet resonance search**

クラス確率($p(C|x)$)をモデリング



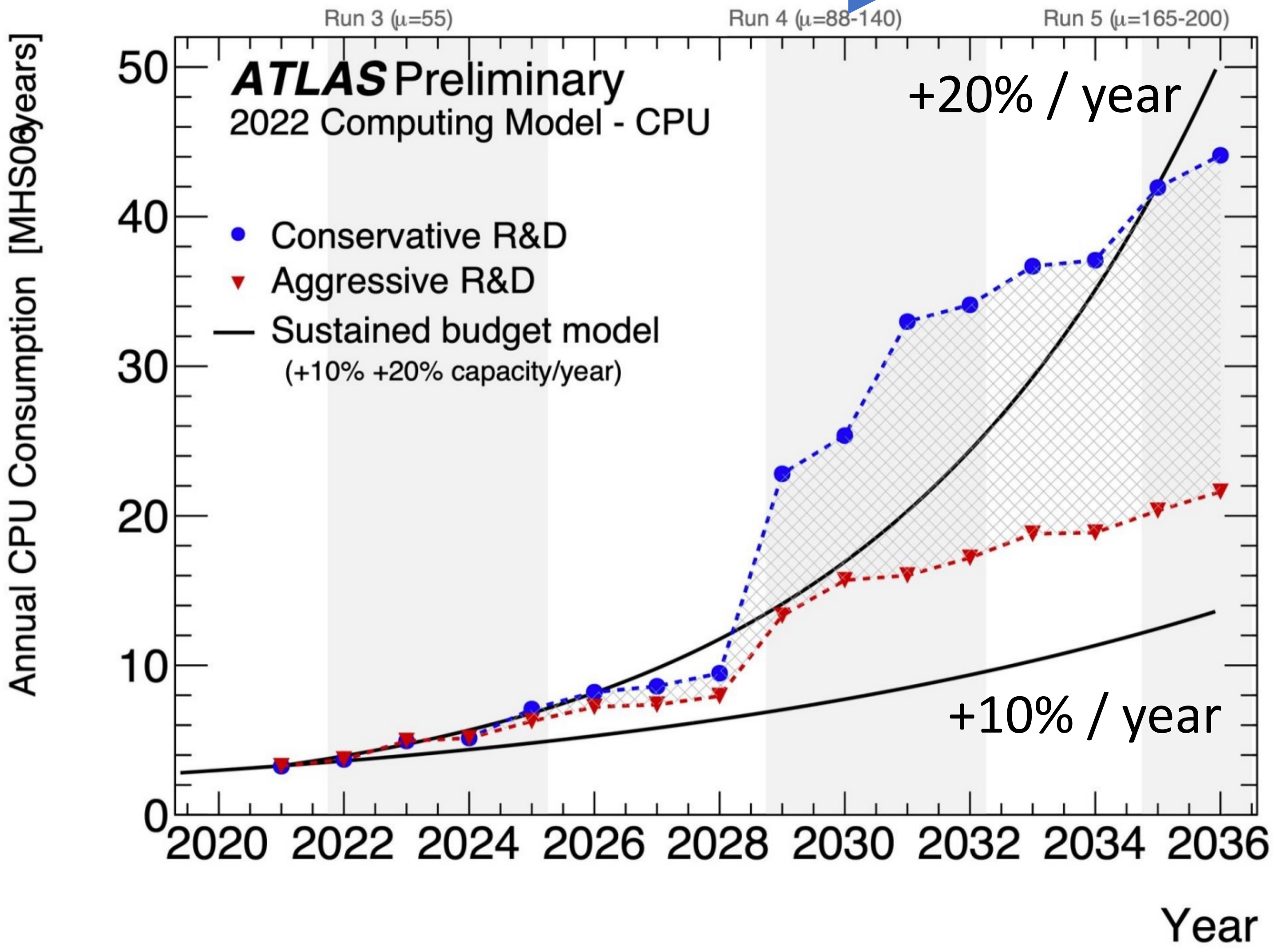
確率($p(x)$)をモデリング



生成モデル
Generative model
for calorimeter shower simulation

HL-LHC (2029-) における計算機資源 予測

HL-LHC開始

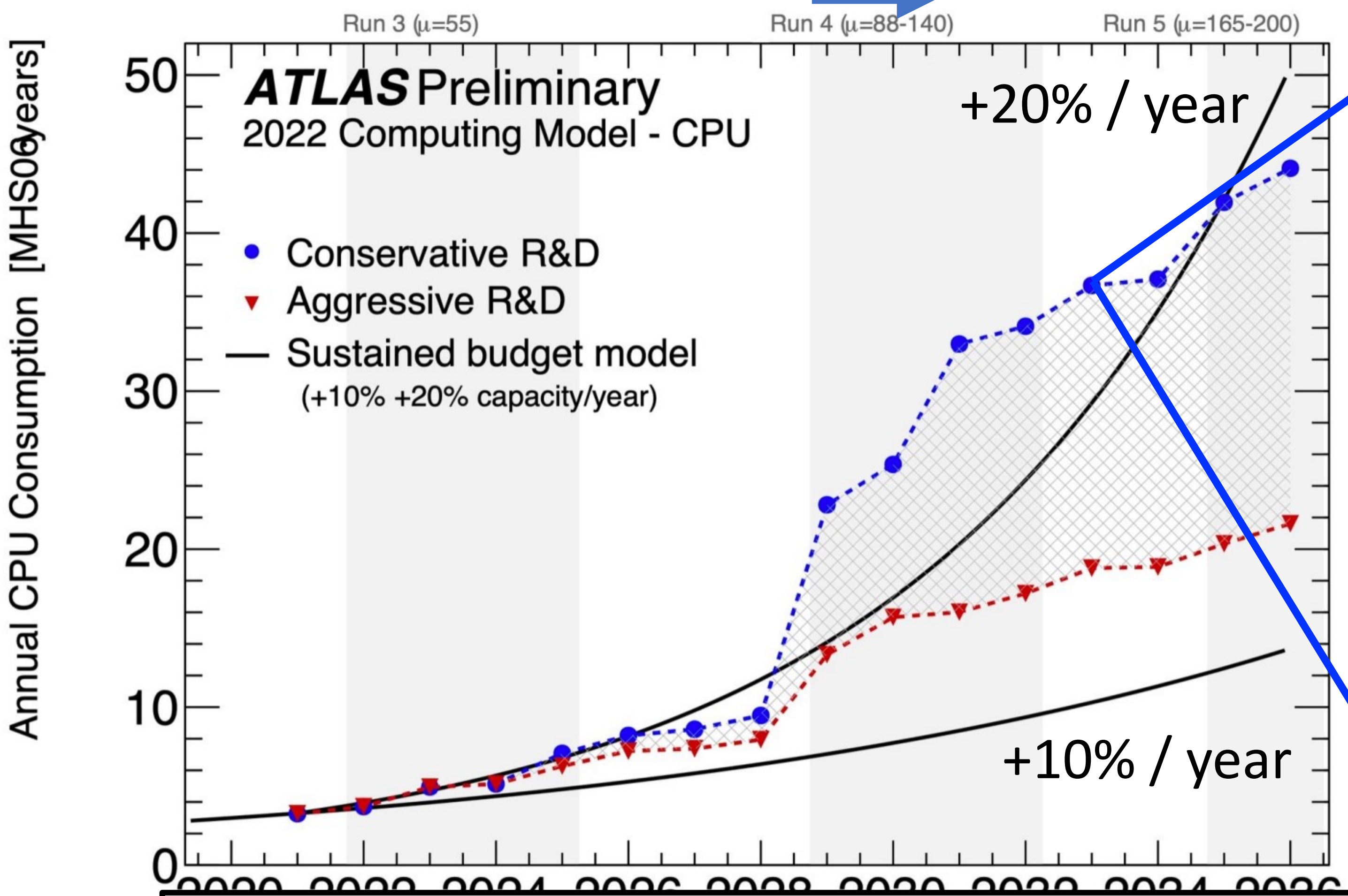


現状のFTE維持を仮定

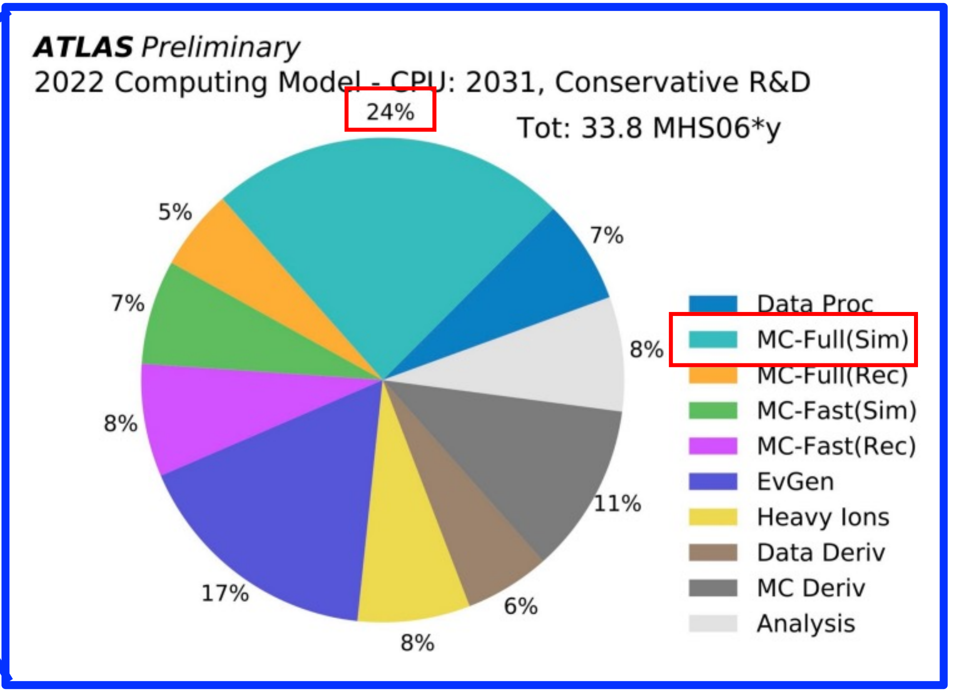
FTEの増加を仮定

HL-LHC (2029-) における計算機資源 予測

HL-LHC開始



内訳



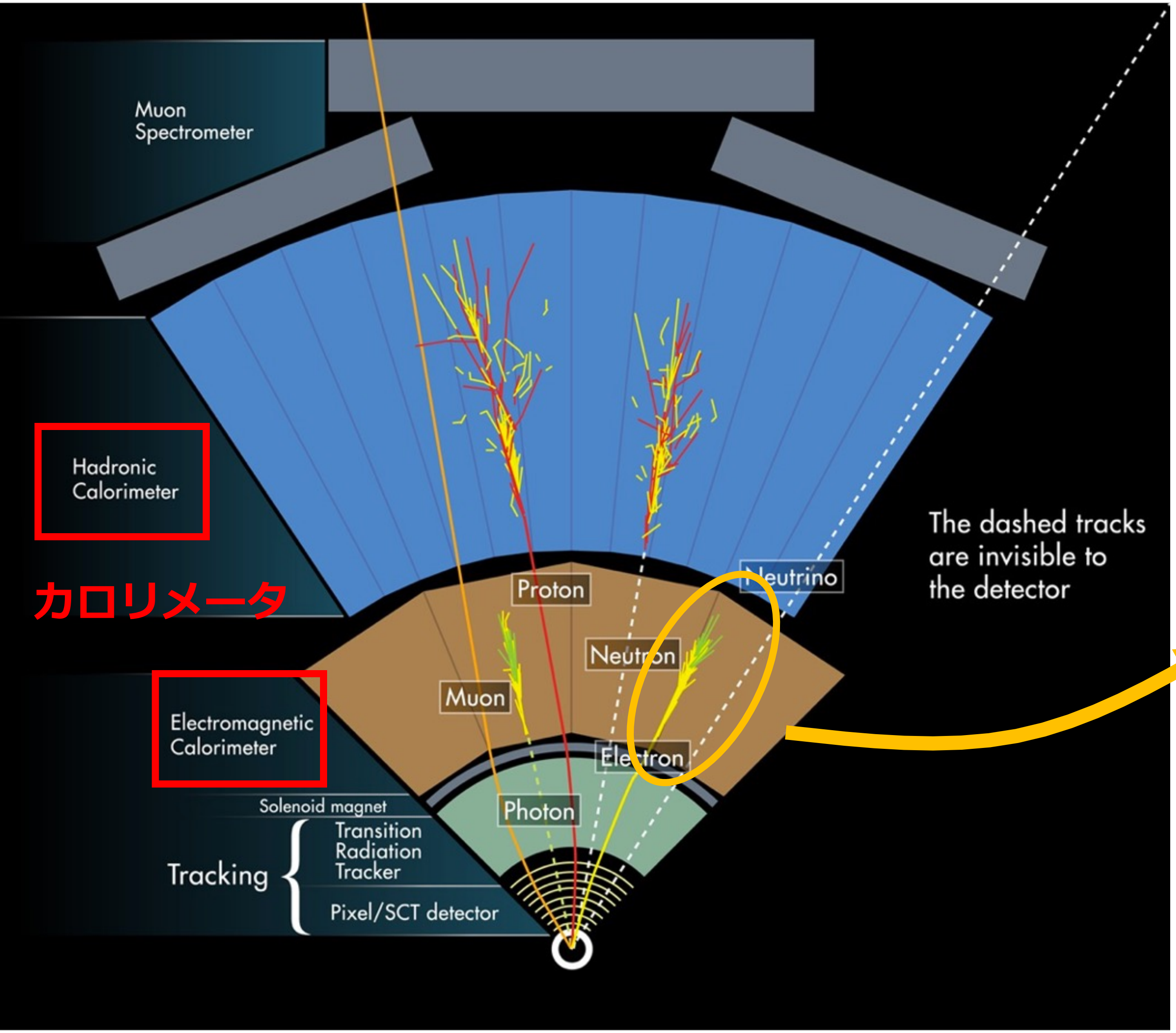
Geant4シミュレーションの割合が大きい

- Geant4を使わない高速なシミュレーション(Fast Simulation) が必要
- 深層学習を使った**生成モデル**で**高精度+高速化**を実現

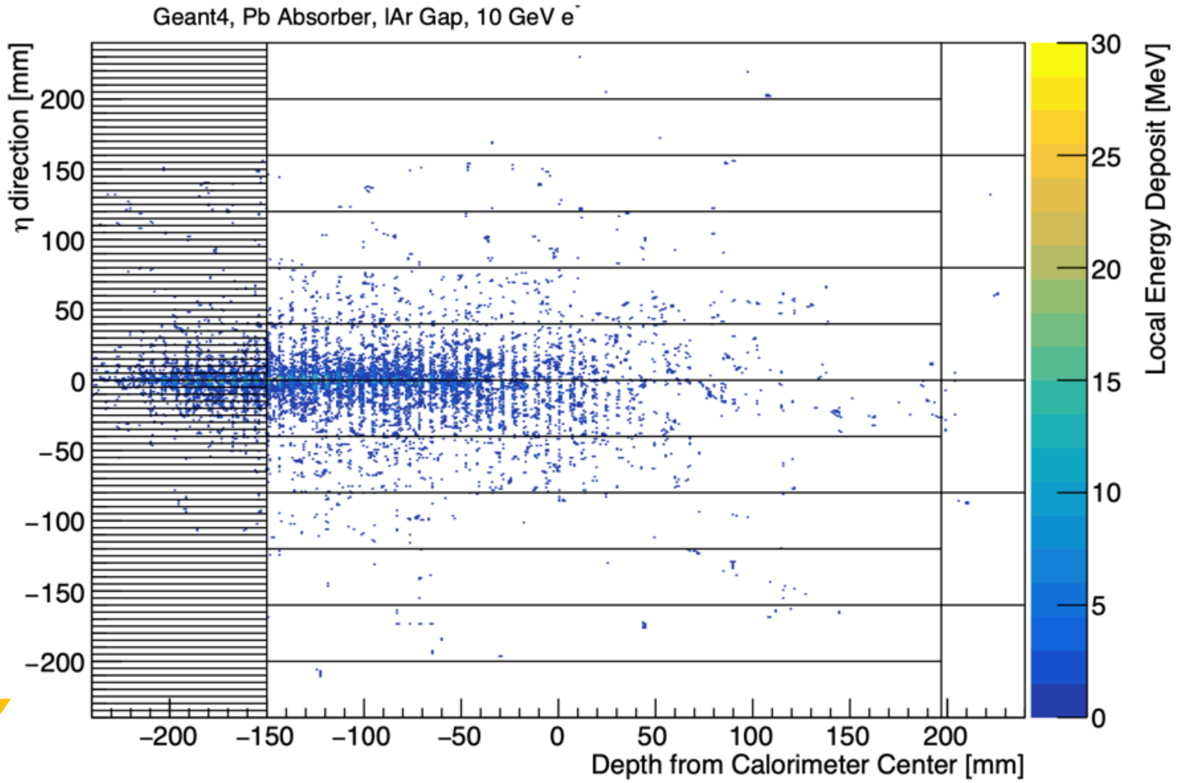
カロリメータとシャワー

Phys. Rev. D 97, 014021 (2018)

検出器



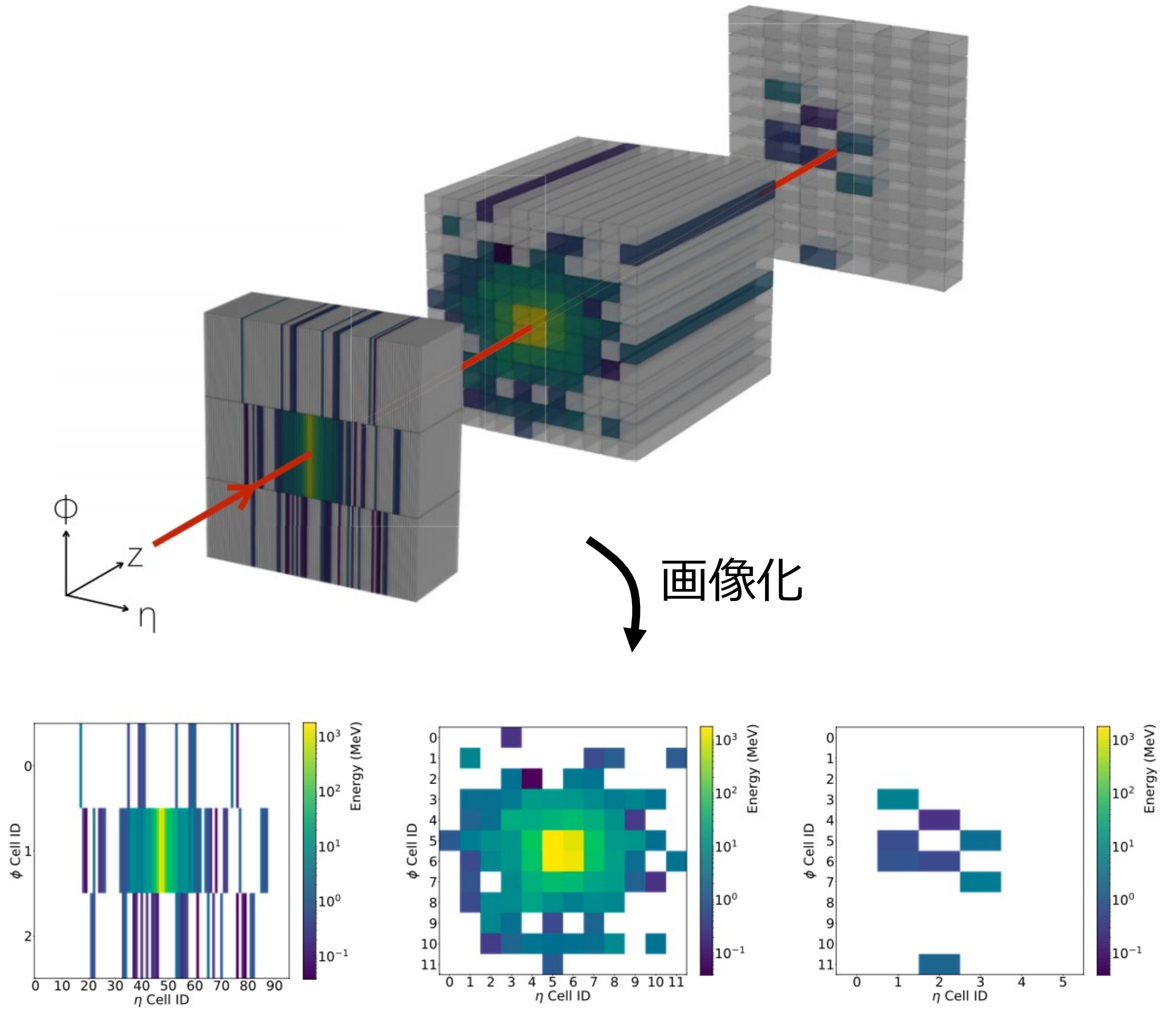
粒子のシャワー



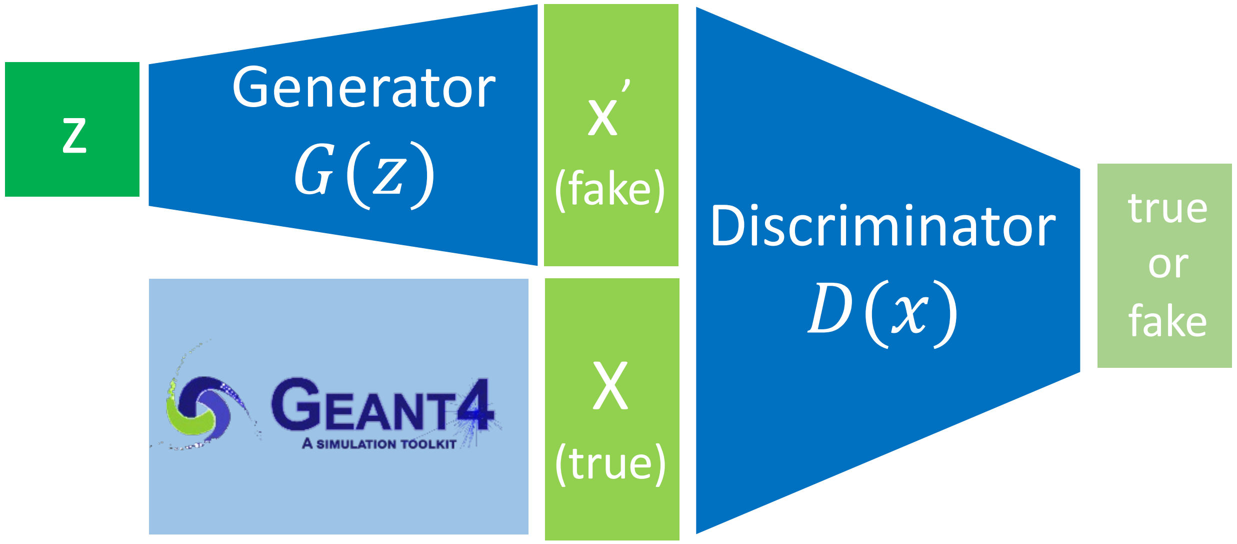
- カロリメータで粒子を止めてエネルギーを測定
  - 入射粒子が多数の粒子に崩壊 : **シャワー**
- カロリメータ中のエネルギー分布をシミュレート
  - Geant4 では一つ一つの粒子を追うため高コスト

初期の研究例: CaloGAN

Comput Softw Big Sci (2017) 1: 4
Phys. Rev. D 97, 014021 (2018)



Generative Adversarial Network (GAN)



$$\mathcal{L}_{\mathrm{adv}} = \mathbb{E}_{z\sim p(z)}\left[\log\left(1 - D(G(z))\right)\right] + \mathbb{E}_{x\sim p_{\mathrm{Geant4}}}\left[\log D(x)\right]$$

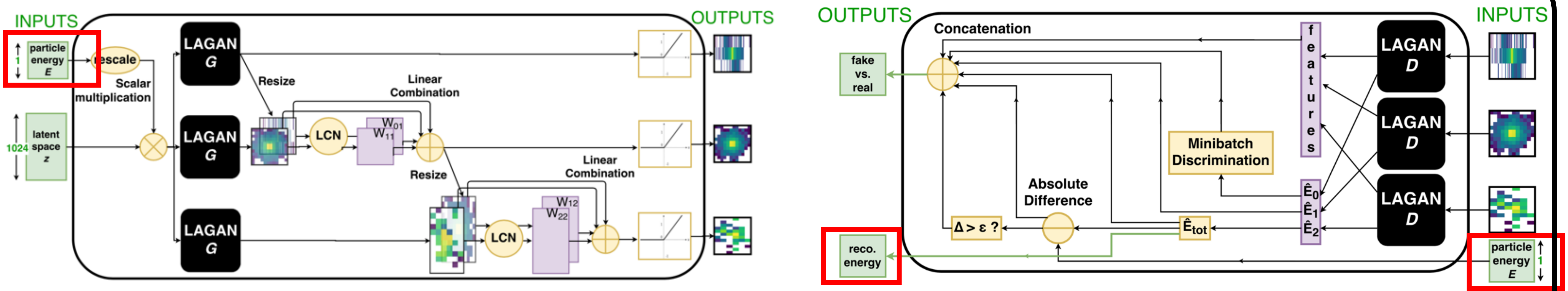
- カロリメータセルごとのエネルギー量 → 画像
- GANで3層の電磁カロリメータ(画像)を生成

CaloGAN: 概要

[Phys. Rev. D 97, 014021 (2018)](#)

モデルアーキテクチャ

Generator　　　　　　　　　　Discriminator



深層畳み込みGANをベースに改良したもの (複雑…)

後にATLASで使われるGANにも使われているアイデア

- 粒子ごと($e, \gamma, \pi^{\pm}$)に**別々のGAN**をトレーニング
- 粒子の**エネルギーで条件づけ**たGAN ($p(x|E)$)
- **総エネルギー量についての制約**をロスに追加

$$\mathcal{L}_{\text{generator}} = \lambda_E \mathcal{L}_E - \mathcal{L}_{\text{adv}}$$

$$\mathcal{L}_{\text{discriminator}} = \lambda_E \mathcal{L}_E + \mathcal{L}_{\text{adv}}$$

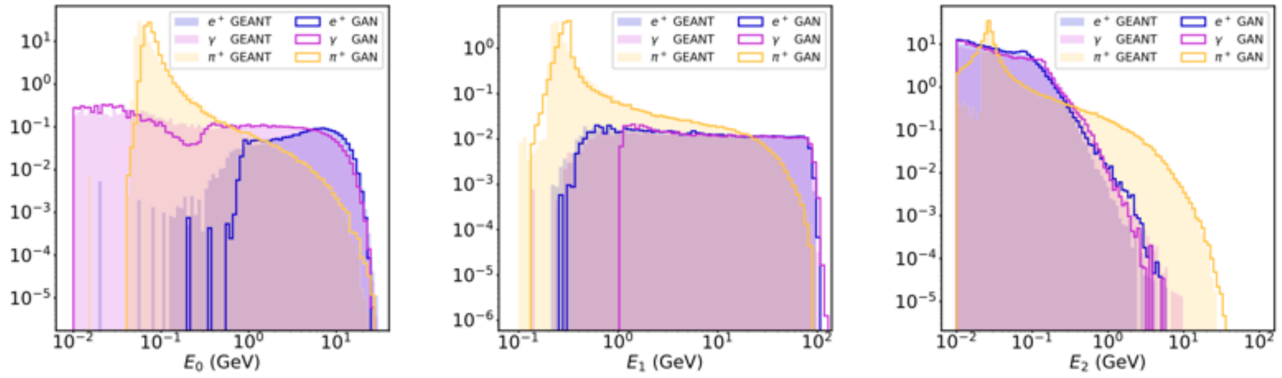
エネルギーへの制約　　通常のGANのloss

CaloGAN: 性能評価

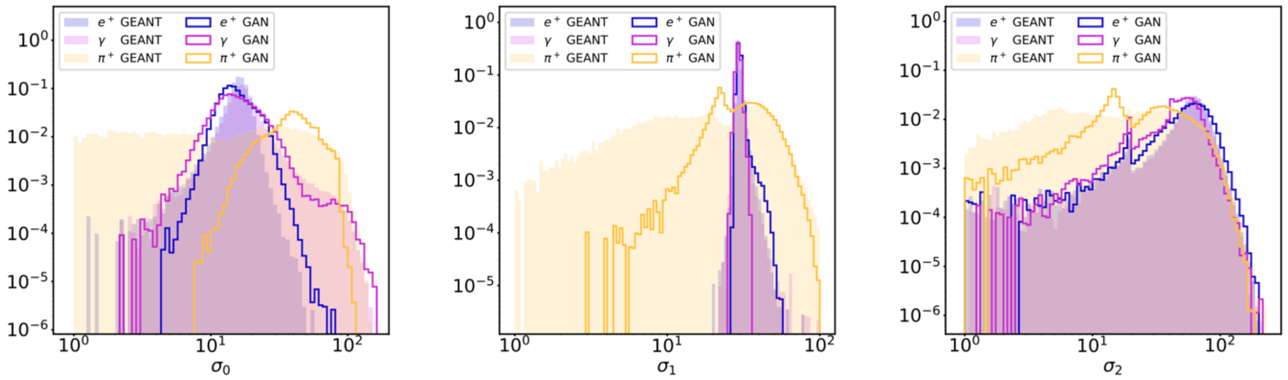
GEANT4
GAN

Phys. Rev. D 97, 014021 (2018)

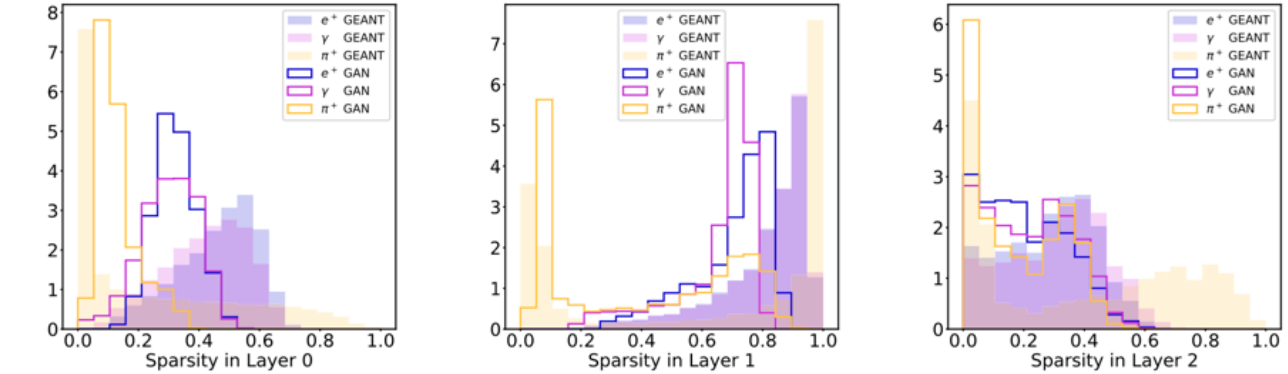
レイヤー毎のエネルギー



レイヤーごとのシャワーの横方向の幅



レイヤーごとのsparsity



Simulator	Hardware	Batch size	ms/shower
GEANT4	CPU	N/A	1772
	CPU	1	13.1
		10	5.11
		128	2.19
		1024	2.03
CALOGAN	GPU	1	14.5
		4	3.68
		128	0.021
		512	0.014
		1024	0.012

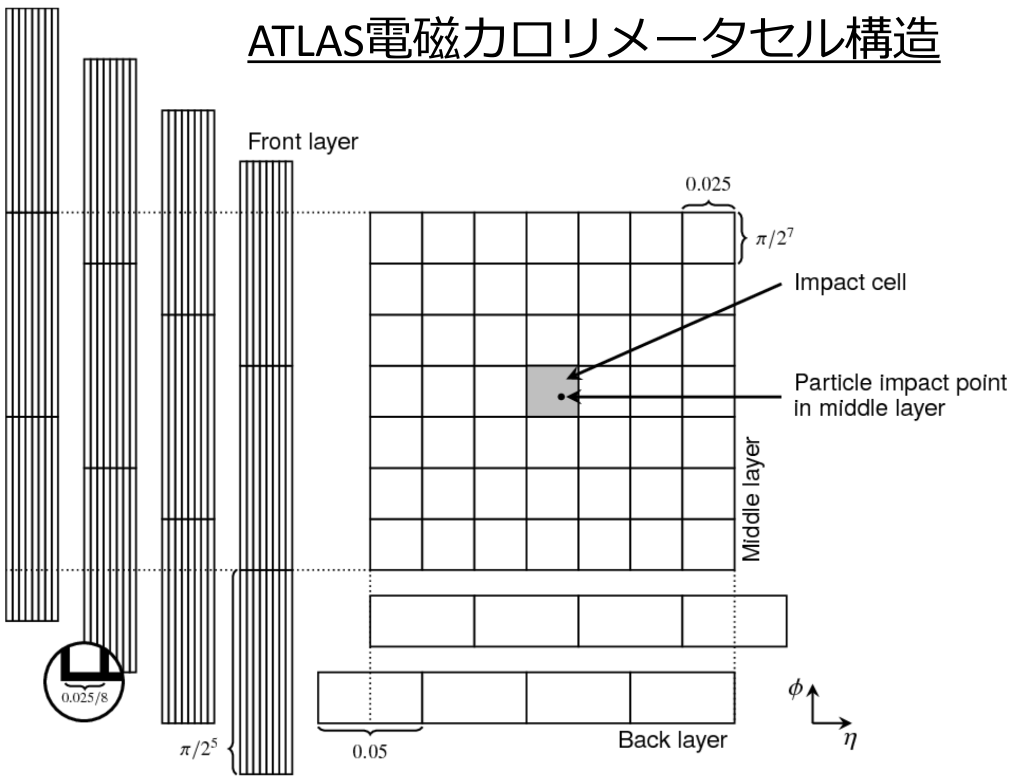
- シャワー形状を"おおよそ"再現
- **高速化**を実現
  - Geant4より **2桁** 速い

ATLASでの初期の研究: VAE と GAN

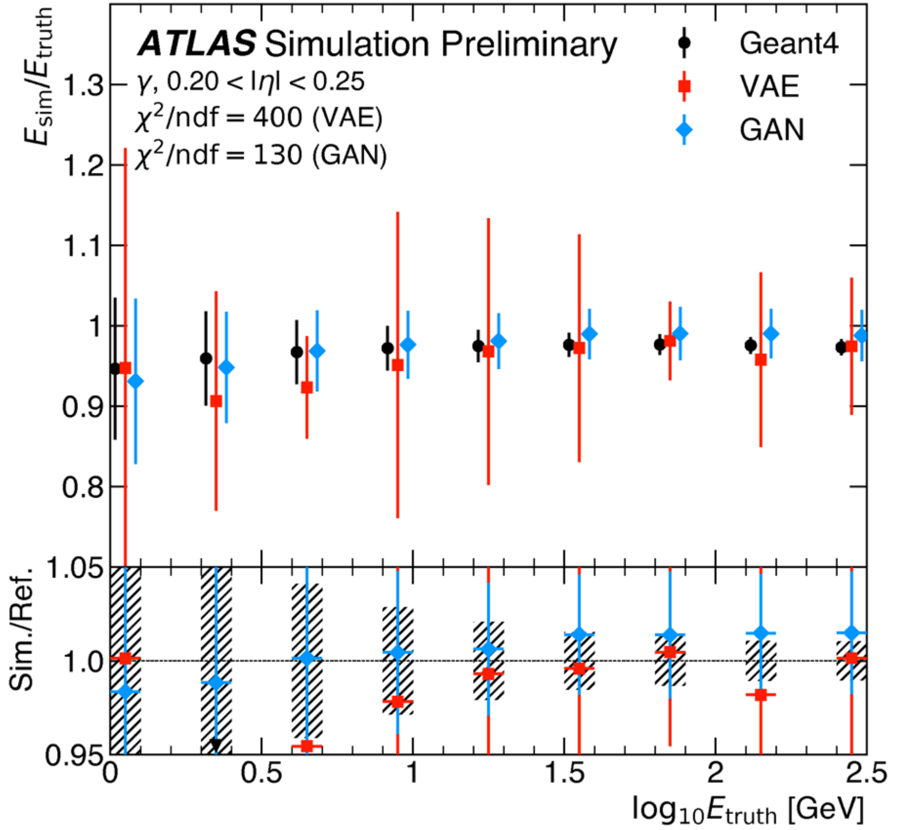
ATL-SOFT-PUB-2018-001

- **ATLAS検出器の電磁カロリメータのジオメトリ**を想定した生成モデルの研究
- 画像ではなく、266次元(セルの個数)の**1次元のベクトル**を生成
- **限られた範囲**の入射粒子だけに着目 (|η|: 0.20 ~ 0.25)
- **GANとVAE**の両方を試して比較

ATLAS電磁カロリメータセル構造



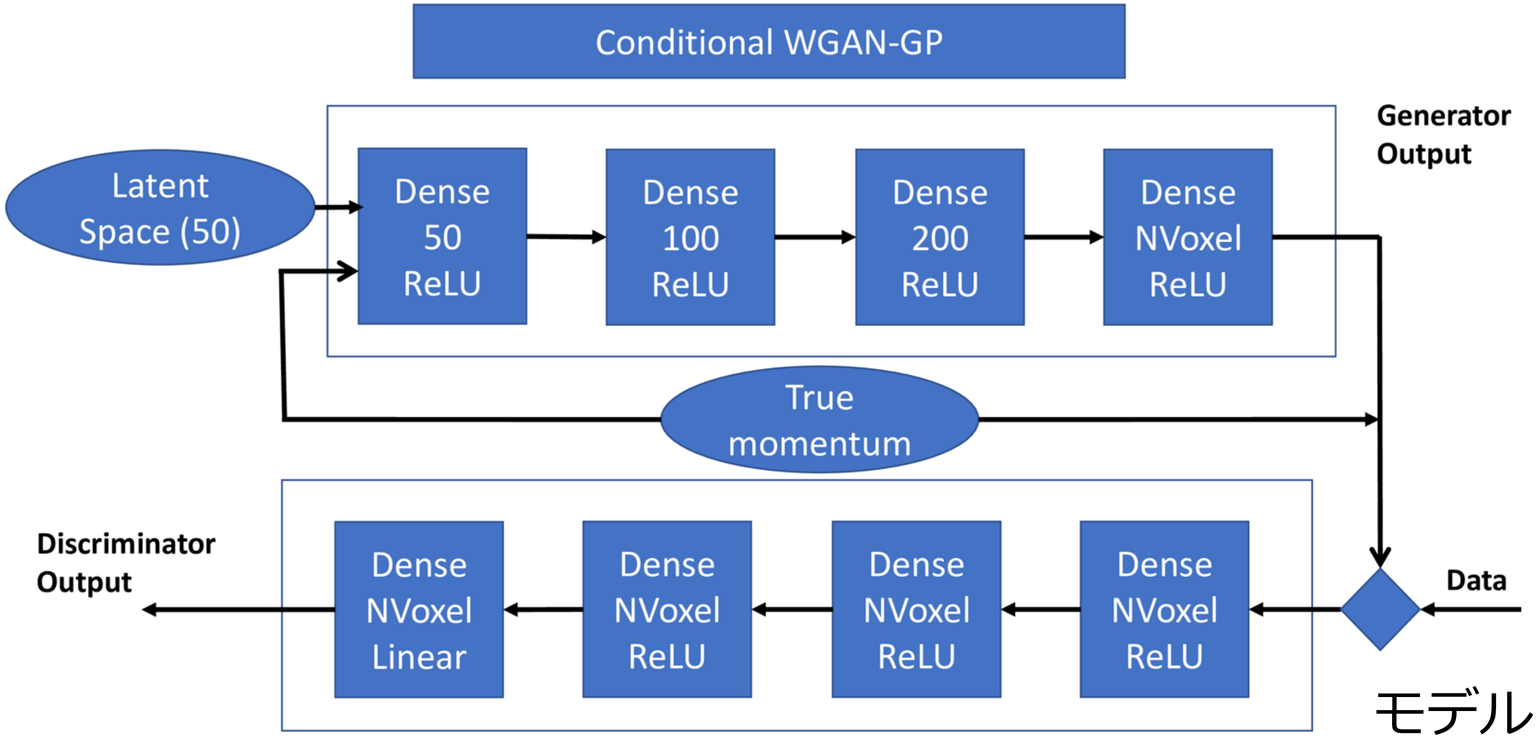
GAN と VAEの比較



- GANの方が性能が良かった
- 現在のATLASのFast simulationではGANが採用

GAN for Fast Simulation at ATLAS: 概要

Comput Softw Big Sci 6 (2022) 7

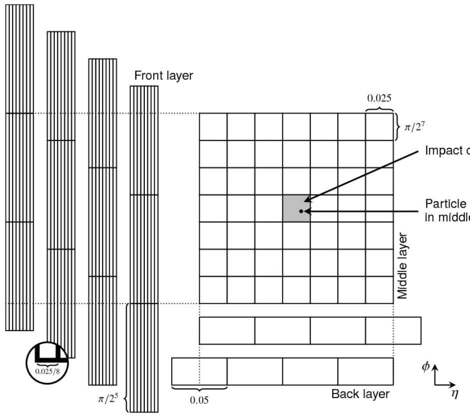


モデルはシンプル。4層のMLP

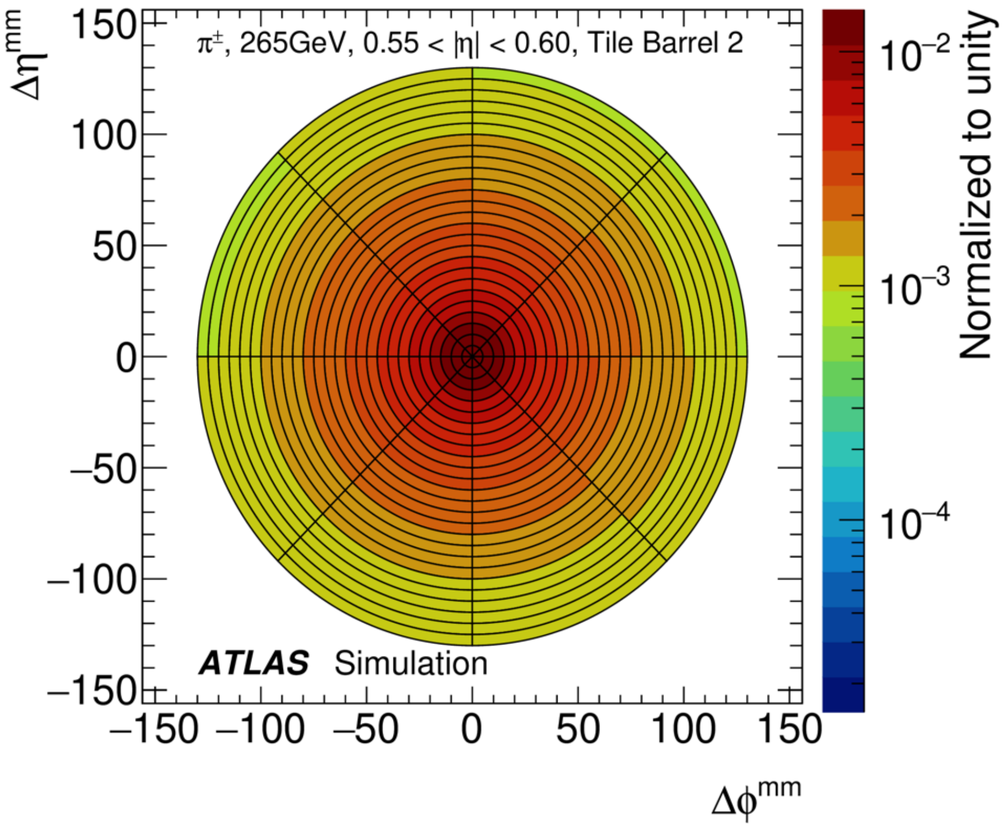
- **WGAN-GP** (Wasserstein GAN with a gradient penalty)
  - オリジナルのGANよりも学習が安定化し、モード崩壊も起こしづらい
- 粒子のエネルギーで条件付けをする (256 MeV ～ 4 TeV)
- 全 η 領域をカバー ($-5 < \eta < 5$)
- 粒子ごと(e, γ, π)、η sliceごと(100 slice) にGANを作成 (**計300 GANs**)

GAN for Fast Simulation at ATLAS: Voxelization

Comput Softw Big Sci 6 (2022) 7



Cell からVoxelへ



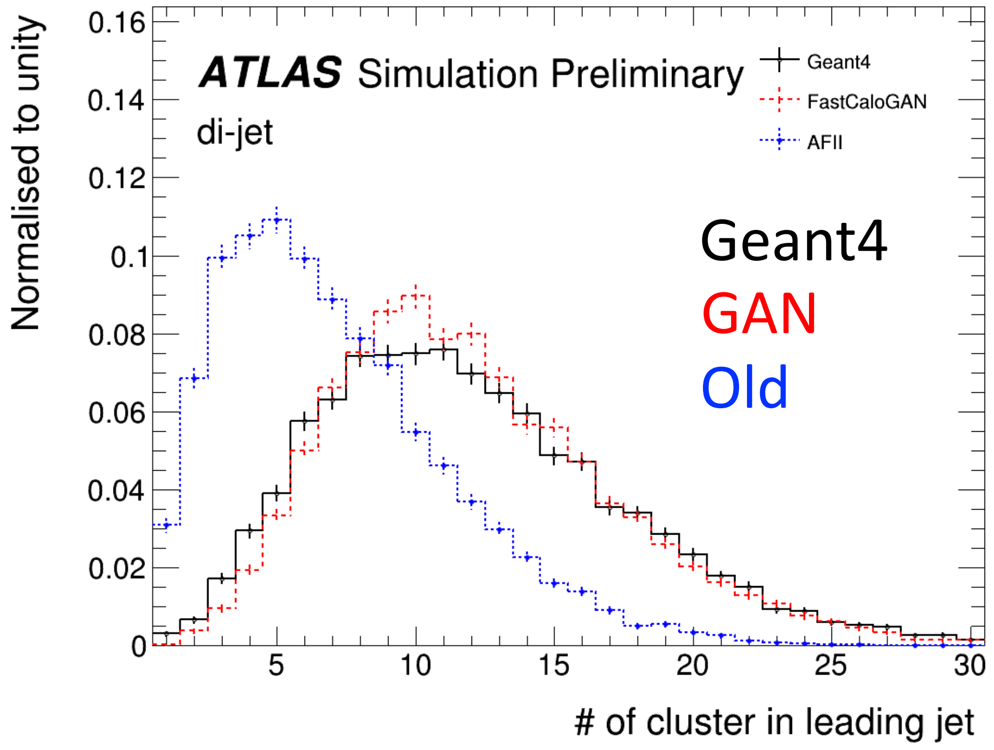
- セルベースな手法では粒子の入射位置に影響を受ける
- 入射粒子を中心とした座標系で考える
  - 生成粒子周りで円筒状のグリッド **(Voxel)**
  - Voxelはcellよりも細かい。予測後に各cellにエネルギーを割り当てる
- 粒子タイプ・η ごとにグリッドの切り方を最適化

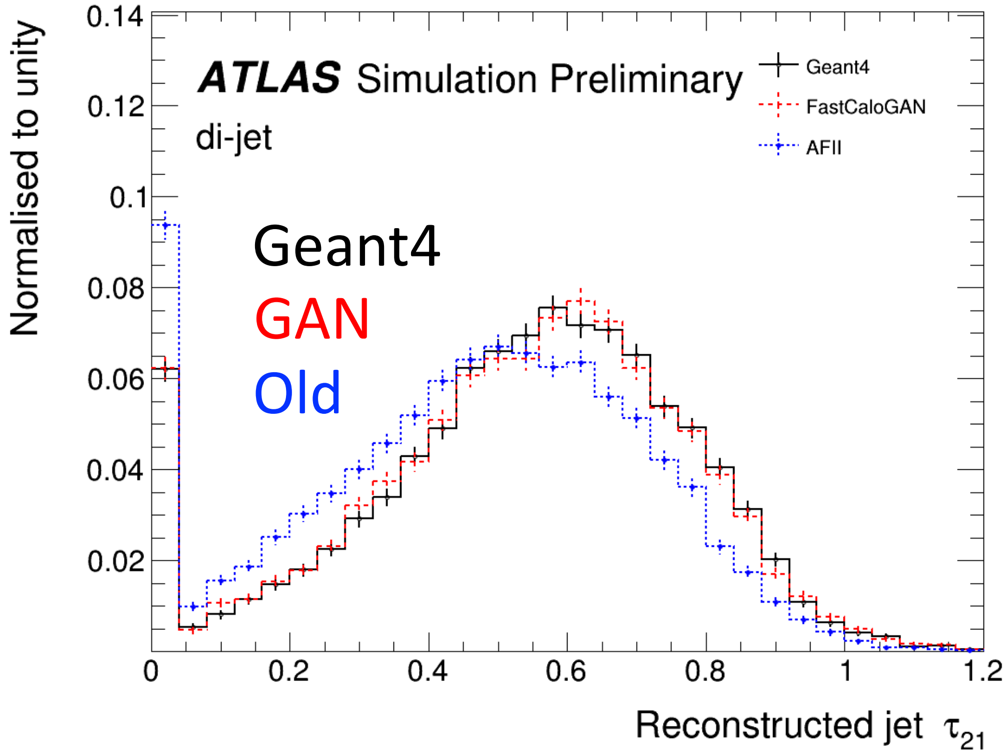
Voxelの切り方の一例

Layer	Bin boundaries in ΔR^{mm} [mm]	Number of bins in ϕ
PreSamplerB	5, 10, 30, 50, 100, 200, 400, 600	1
EMB1	1, 4, 7, 10, 15, 30, 50, 90, 150, 200	10
EMB2	5, 10, 20, 30, 50, 80, 130, 200, 300, 400	10
EMB3	50, 100, 200, 400, 600	1
TileBar0	10, 20, 30, ··· 100, 130, 160, 200, 250 ··· 400, 1000, 2000	10
TileBar1	10, 20, 30, ··· 100, 130, 160, 200, 250, ··· 400, 600, 1000, 2000	10
TileBar2	0, 50, 100, ··· 300, 400, 600, 1000, 2000	1

GAN for Fast Simulation at ATLAS: 結果

ATL-SOFT-PUB-2020-006



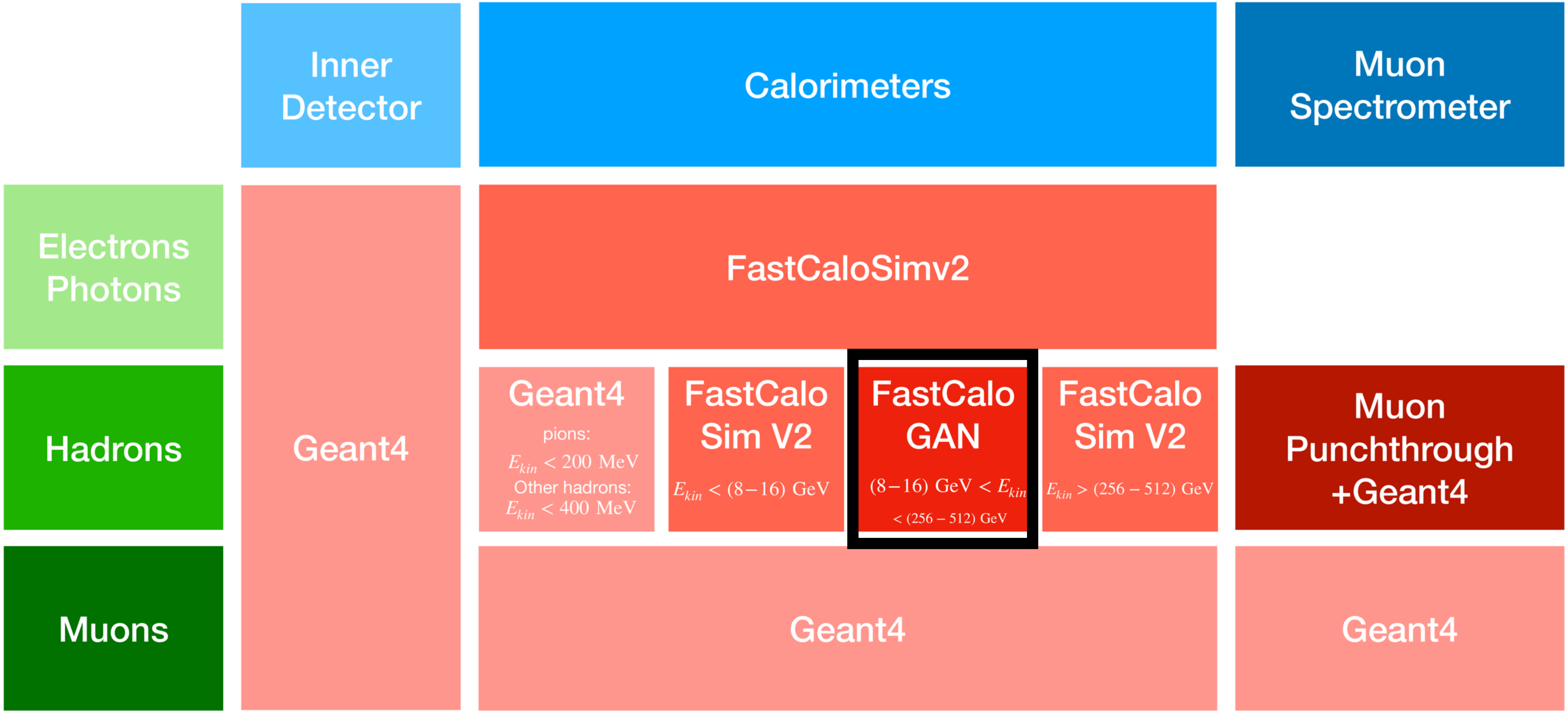


- Jet substructure variableで大きく改善
- 速度も高速
  - Geant4: 6 秒 (65 GeV pion)
    162 秒 (2 TeV pion)
  - GAN : **0.07** 秒 (pion, エネルギー依存性なし)

GAN for Fast Simulation at ATLAS: GANの位置づけ

Comput Softw Big Sci 6 (2022) 7

- 公式のソフトウェア(AtlFast3)にproductionとして導入
- ただし、使われているのは **ハドロン & 中程度のエネルギー部分のみ**
  - まだまだ改善できる余地あり

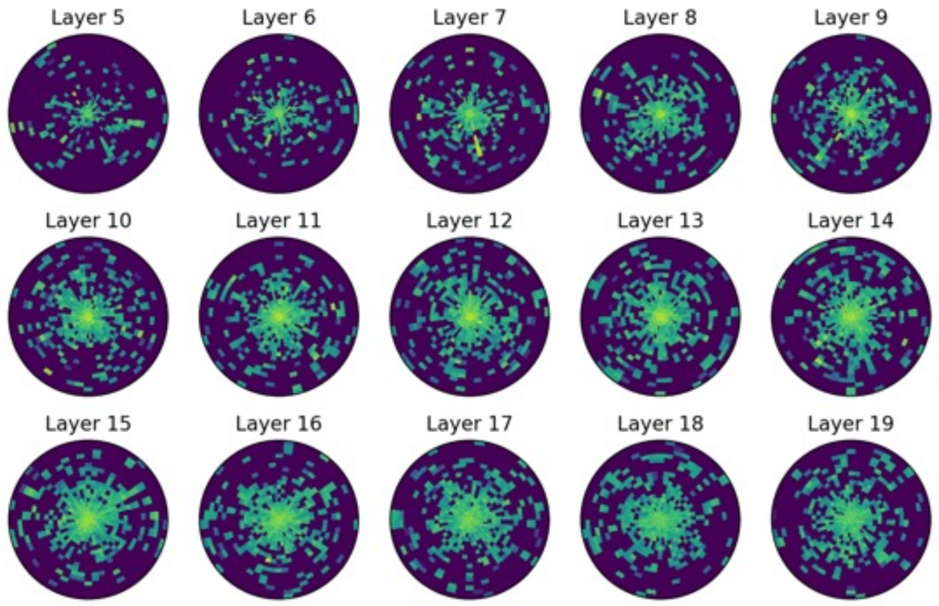


Calo Challenge 2022

Fast Calorimeter Simulation Challenge 2022

- ML4Jets 2022をターゲットにコンテストを開催
- データセット
  - 難易度の異なる3つのデータセット
  - Easy　　: Voxel数 O(100) (= ATLAS fast simulationのvoxel)
  - Medium : Voxel数 O(1000)
  - Hard　　: Voxel数 O(10000)
- 評価指標
  - 単一のものではなく多種多様な指標を使う予定
    - サンプリング時間、メモリ使用量
    - シャワー形状変数分布
    - 入射エネルギーに対する内挿がうまくできているか
    - Geant4と生成サンプルを区別するclassifier
- 今後はこのベンチマークを使って手法・モデルの比較が容易に
  - 今後の手法の発展に期待大

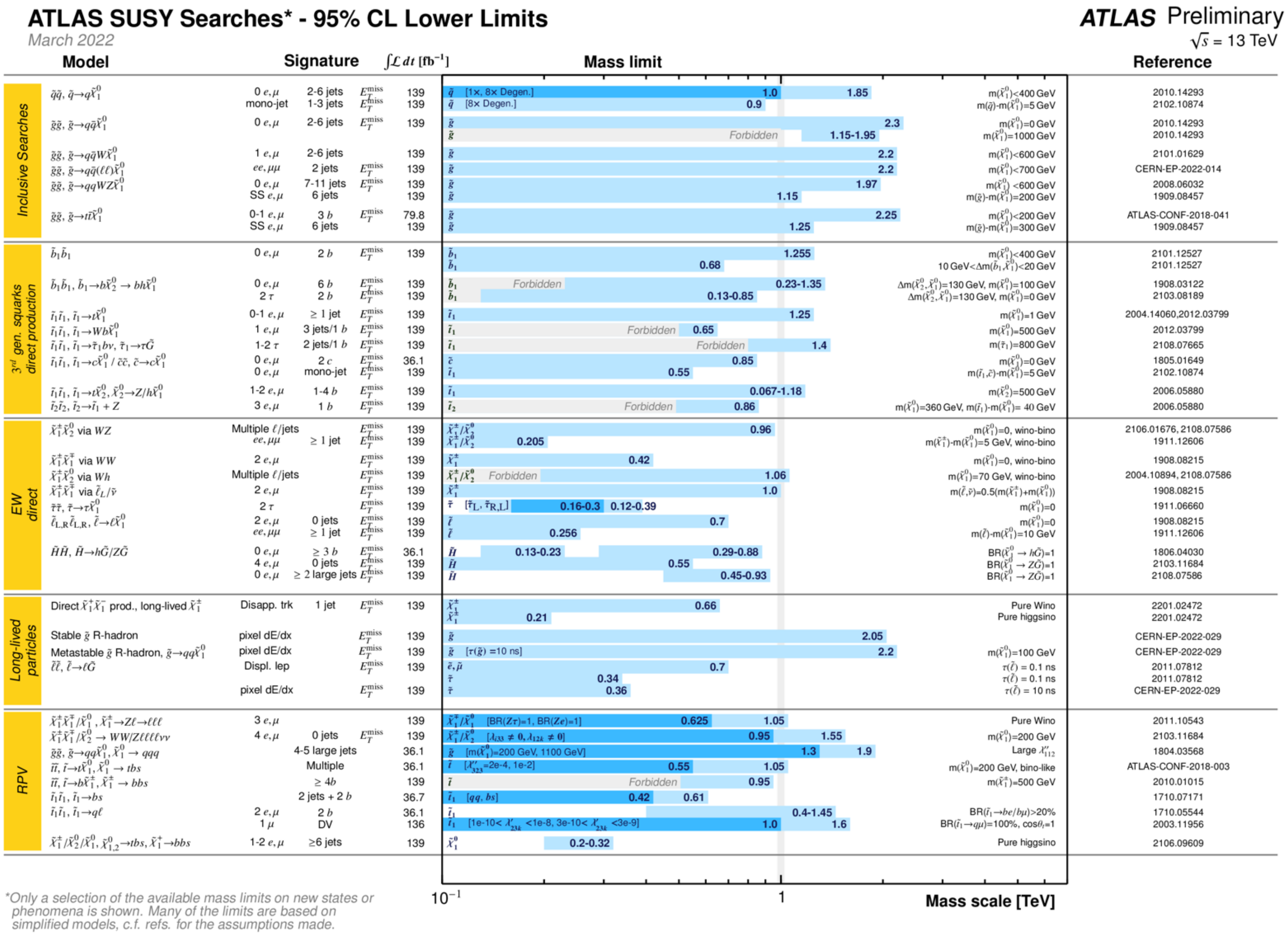
Hard datasetの入力



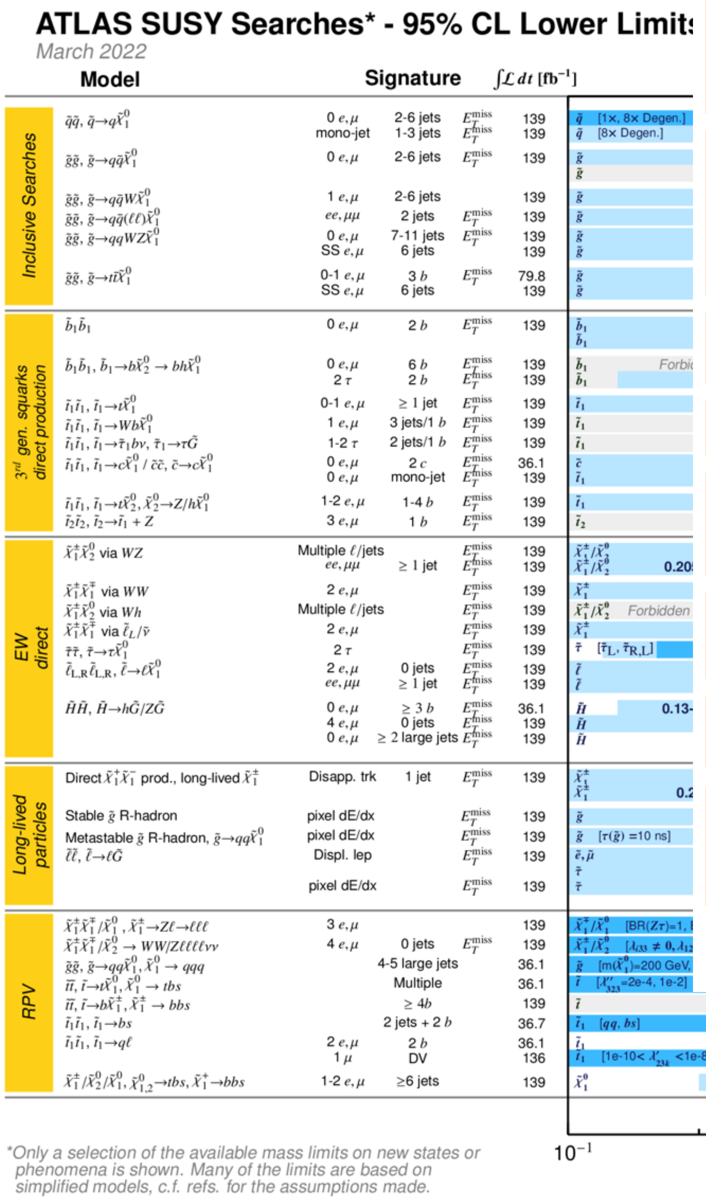
⋮

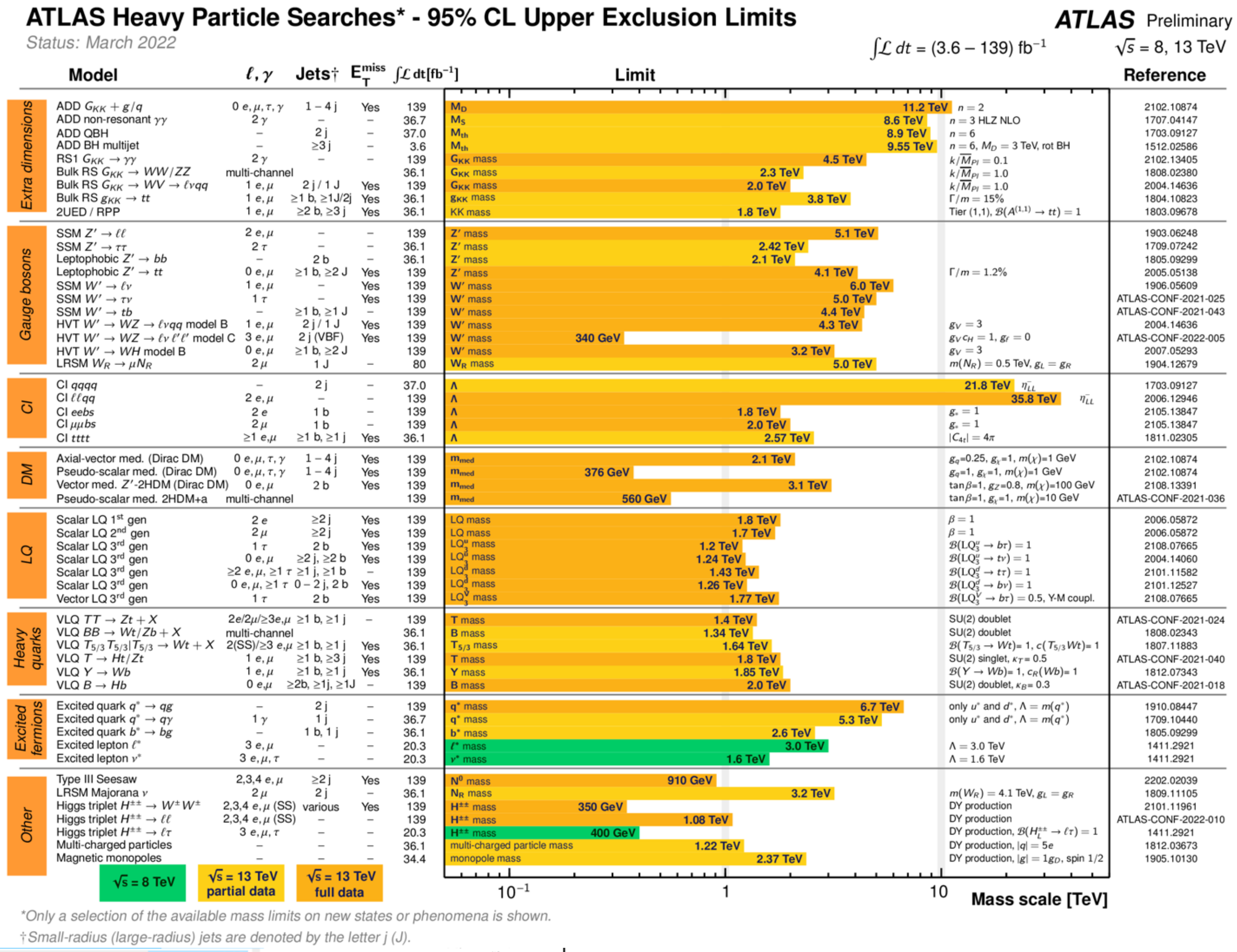
異常検知
Anomaly detection
for new particle search

ATLASでの新粒子探索

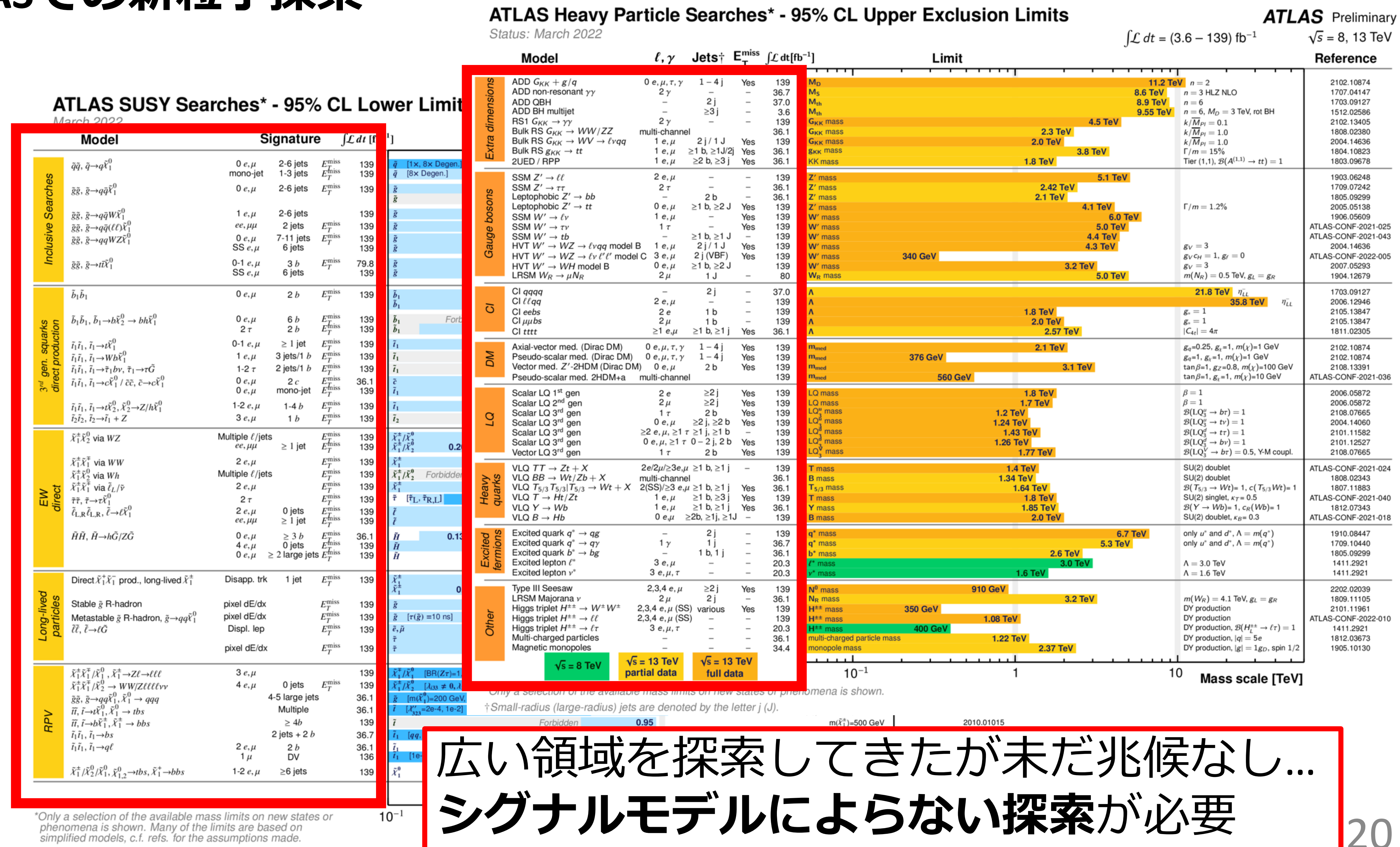


ATLASでの新粒子探索





ATLASでの新粒子探索



広い領域を探索してきたが未だ兆候なし…

シグナルモデルによらない探索が必要

シグナルモデルを仮定しない新粒子探索

モンテカルロ(MC)シミュレーションが使えるか？

YES

- **MCシミュレーションからのズレ**を網羅的に探すサーチ
- バックグラウンドは**MCシミュレーション**を使う

Model Unspecific Search in CMS (MUSiC)
ATLAS General search

NO

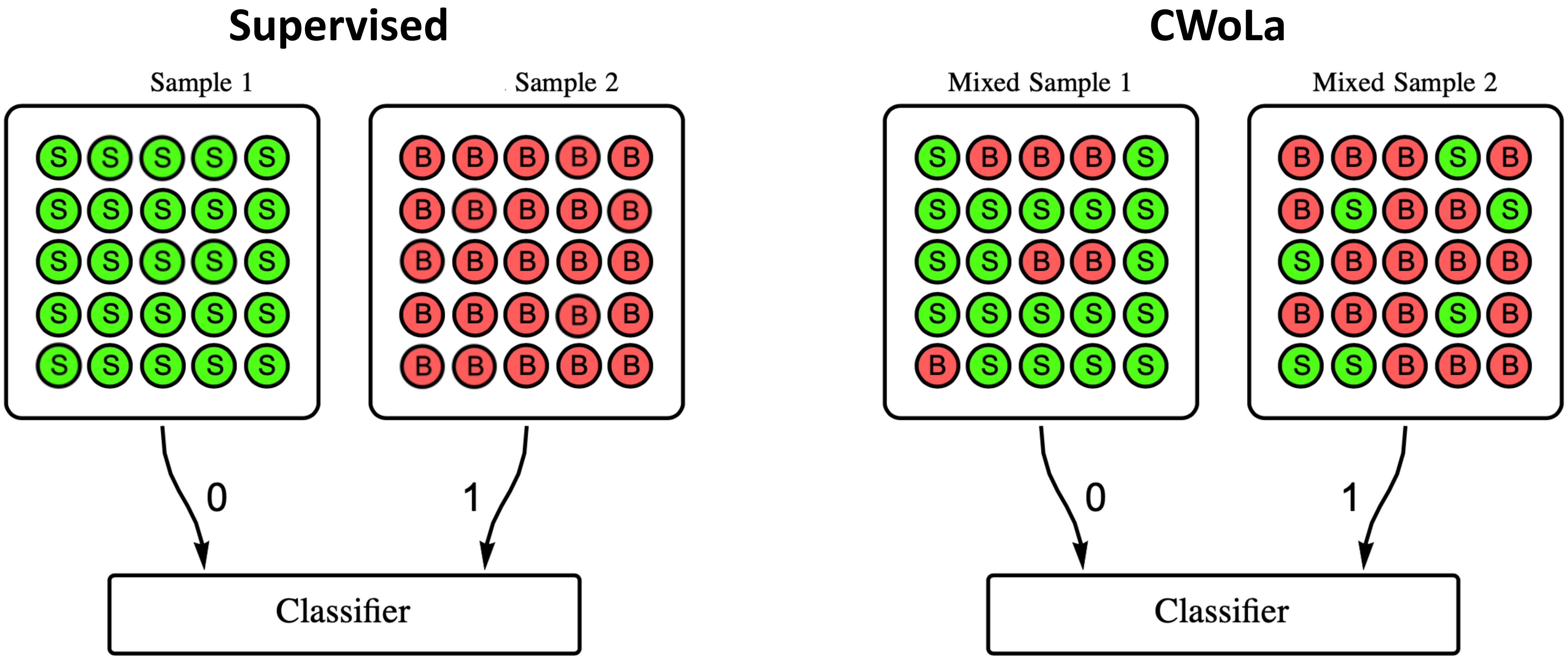
- **データドリブン**なサーチ
- バックグラウンドを**データから推定**

弱教師学習を使った新粒子探索 at ATLAS

今日のトピック

Classification Without Labels (CWoLa)

JHEP 10 (2017) 174

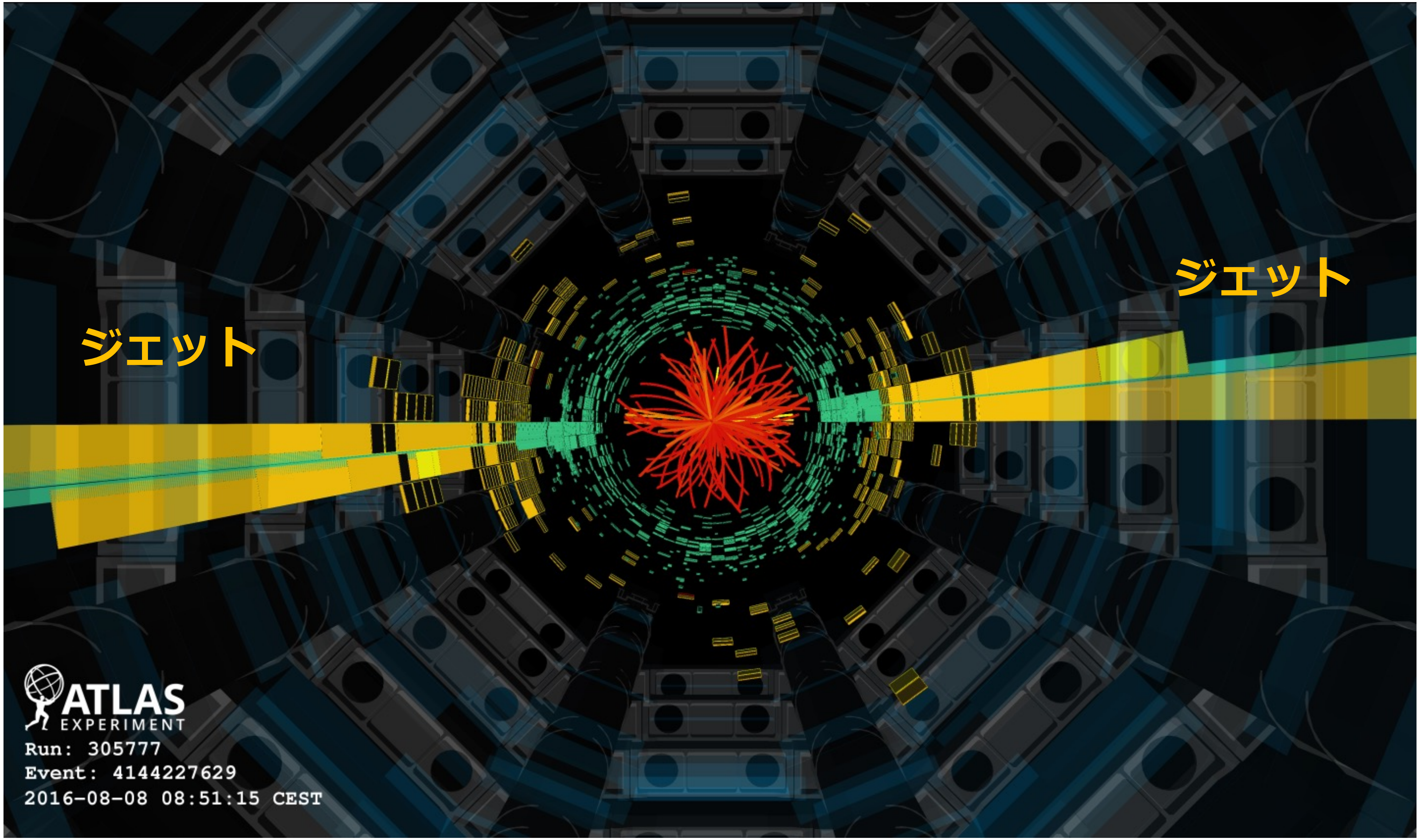


- ラベルがノイジーでも分類器はsignal/backgroundを分けるように学習される
  - ただし、統計的なふらつきの影響は受ける
- ATLAS実験で、Di-jet バンプサーチにCWoLaを適用

  => Phys. Rev. Lett. 125, 131801 (2020)

Dijet resonance search using CWoLa: Di-jet event

ATLAS-PHOTO-2019-017



Dijet resonance search using CWoLa: ターゲットモデル

Phys. Rev. Lett. 125, 131801 (2020)

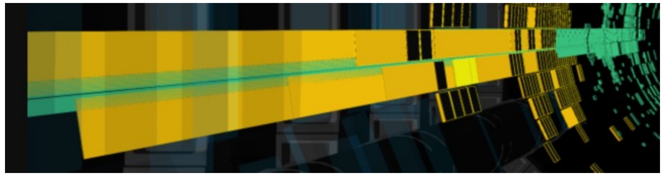
シグナルモデル

重い未知粒子
e.g. 200 GeV

重い未知粒子
e.g. 3 TeV

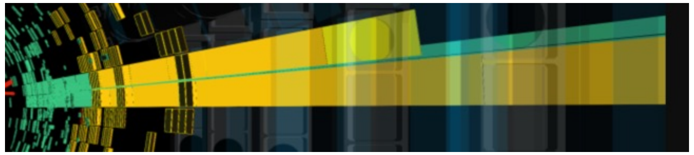
重い未知粒子
e.g. 80 GeV

大半径ジェット



B ← A → C

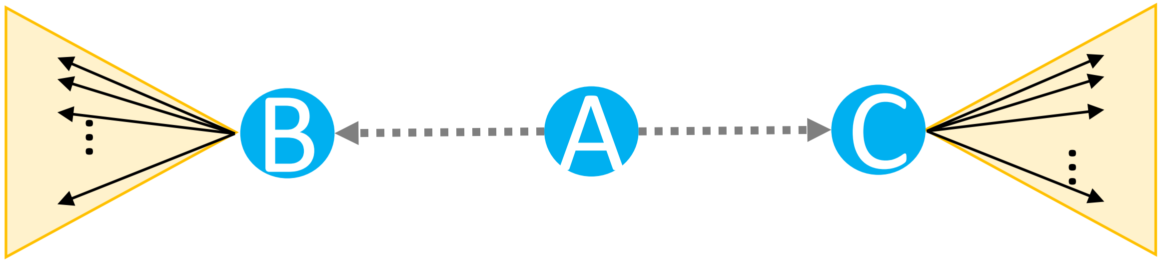
大半径ジェット

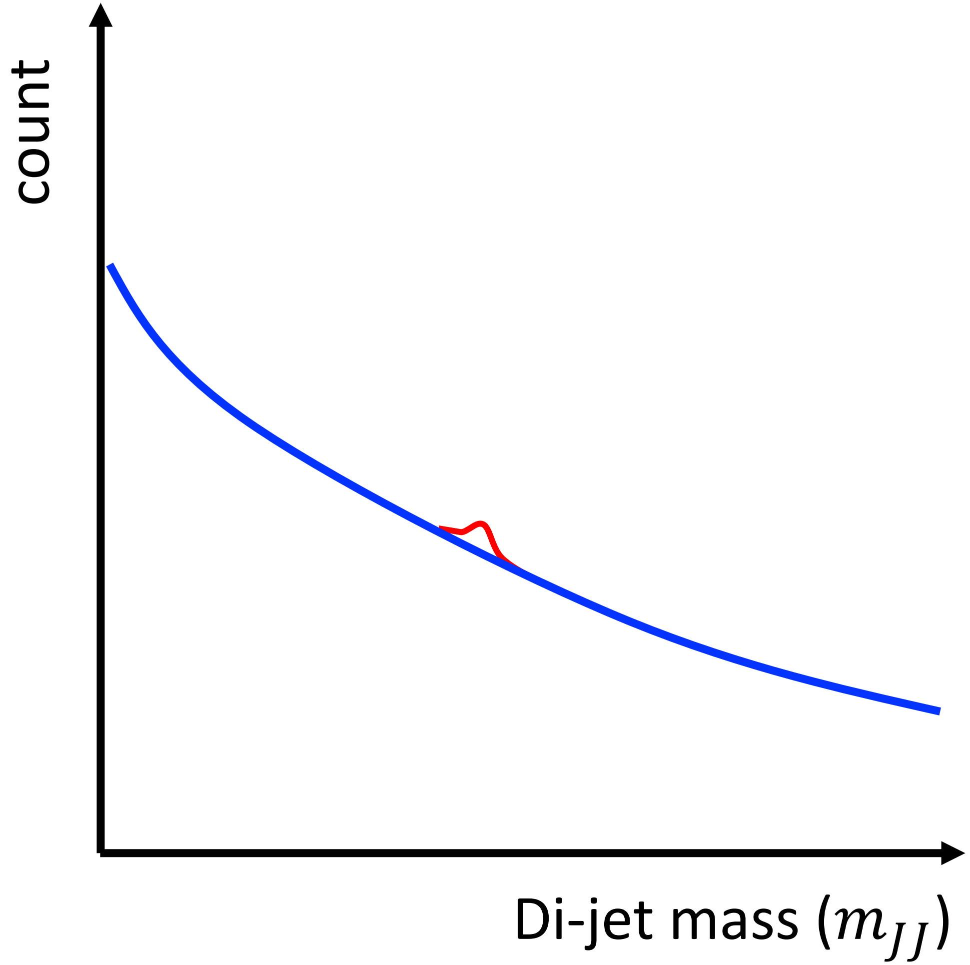


$$m_A \sim \mathcal{O}(\mathrm{TeV})$$
$$m_B, m_C \sim \mathcal{O}(100\ \mathrm{GeV})$$

Dijet resonance search using CWoLa: アルゴリズム

Phys. Rev. Lett. 125, 131801 (2020)

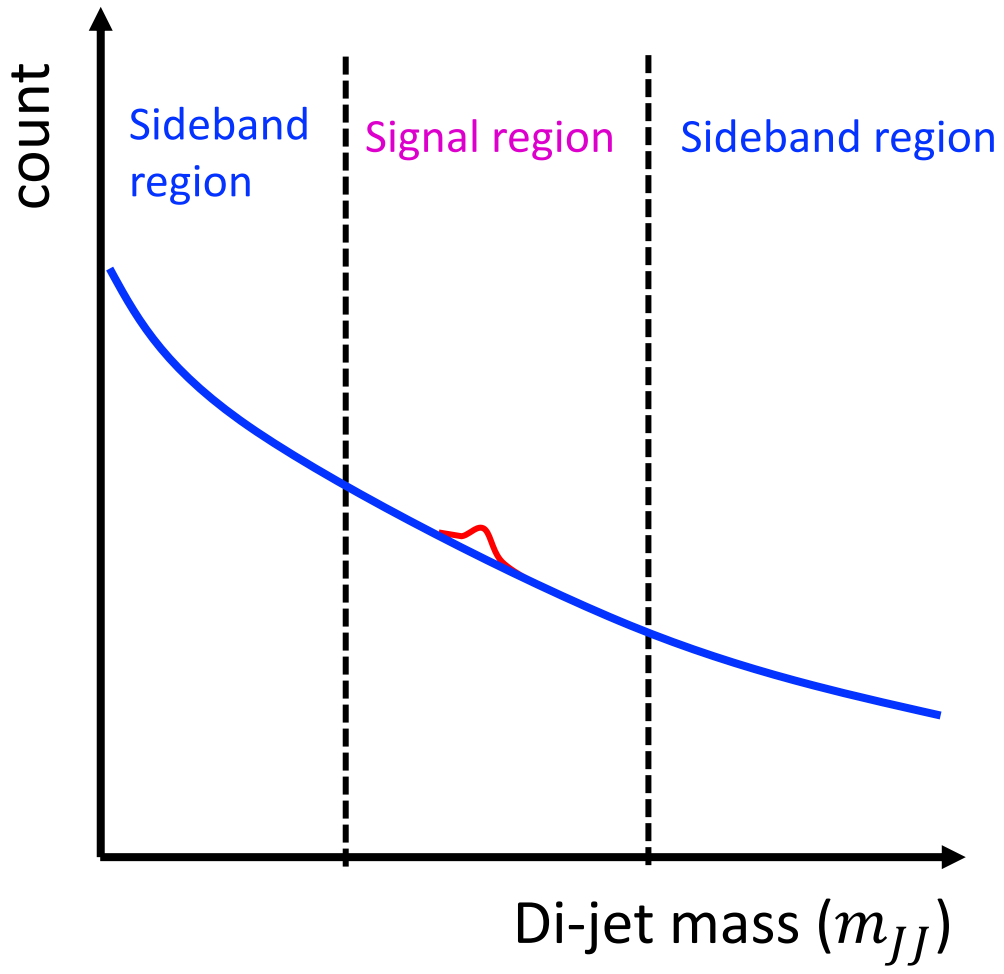




- Di-jet mass分布のバンプを探す($\mathrm{m_{JJ}} \sim m_A$)
- S/B 改善のためjet massでカットをかける
- 最適なカットはシグナルモデルに依存する
  - $m_J \sim m_B, m_C$
- CWoLaで**Data-driven**にカットを決める

Dijet resonance search using CWoLa: アルゴリズム

Phys. Rev. Lett. 125, 131801 (2020)

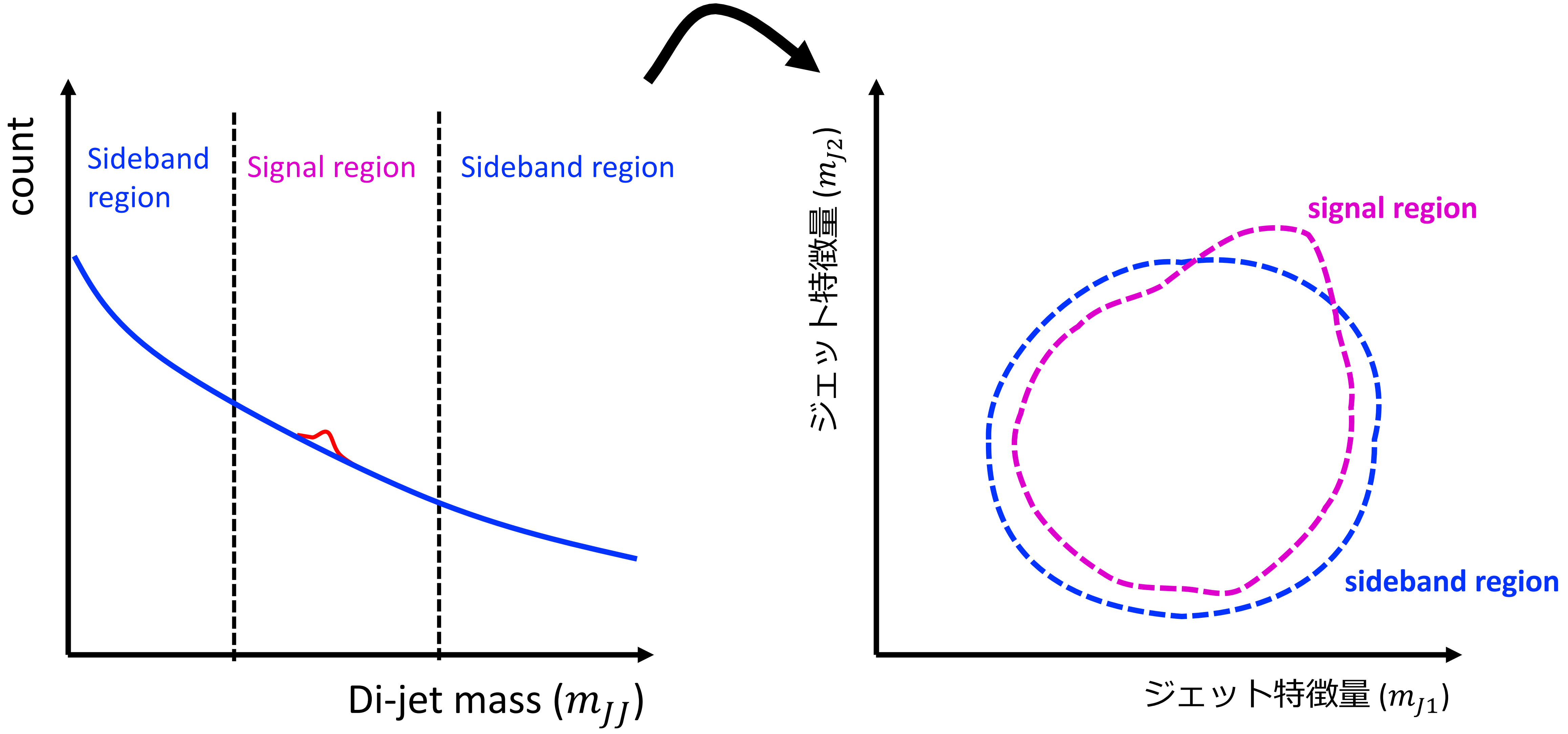


- m_{JJ}を使って
  - Signal region (y=0)
  - Sideband region (y=1)

  を定義

Dijet resonance search using CWoLa: アルゴリズム

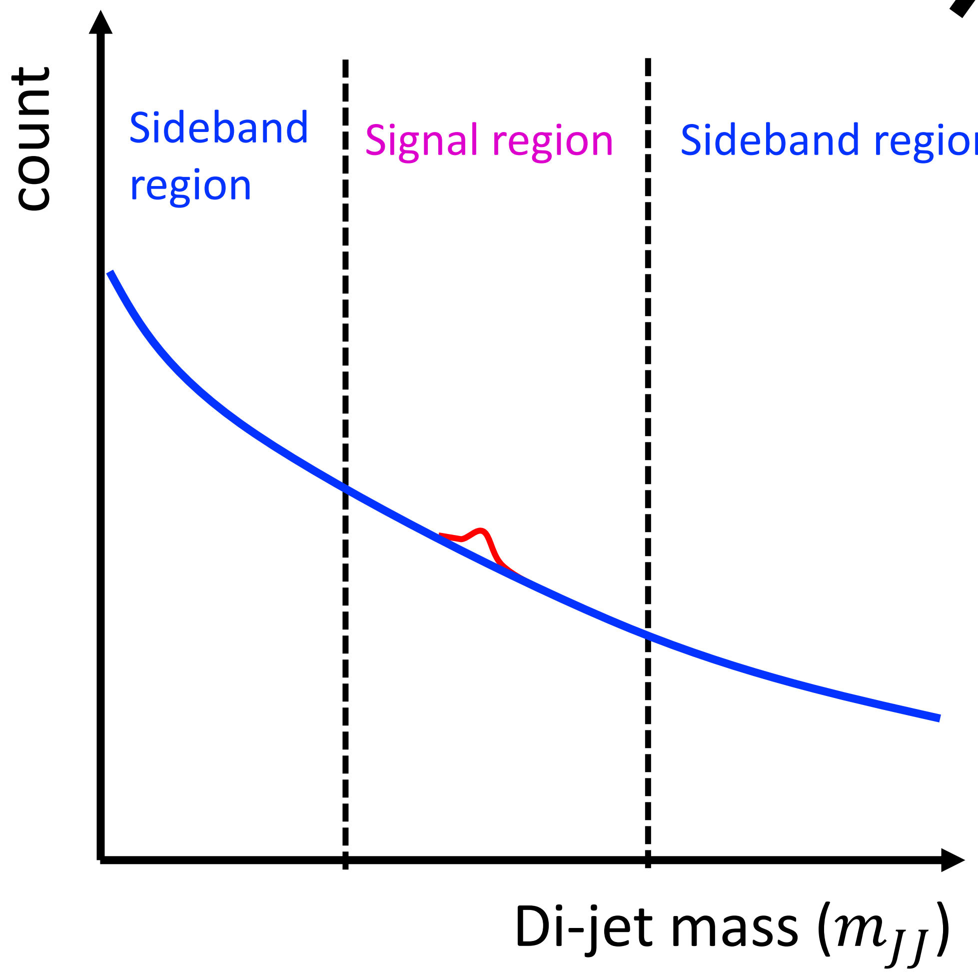
Phys. Rev. Lett. 125, 131801 (2020)

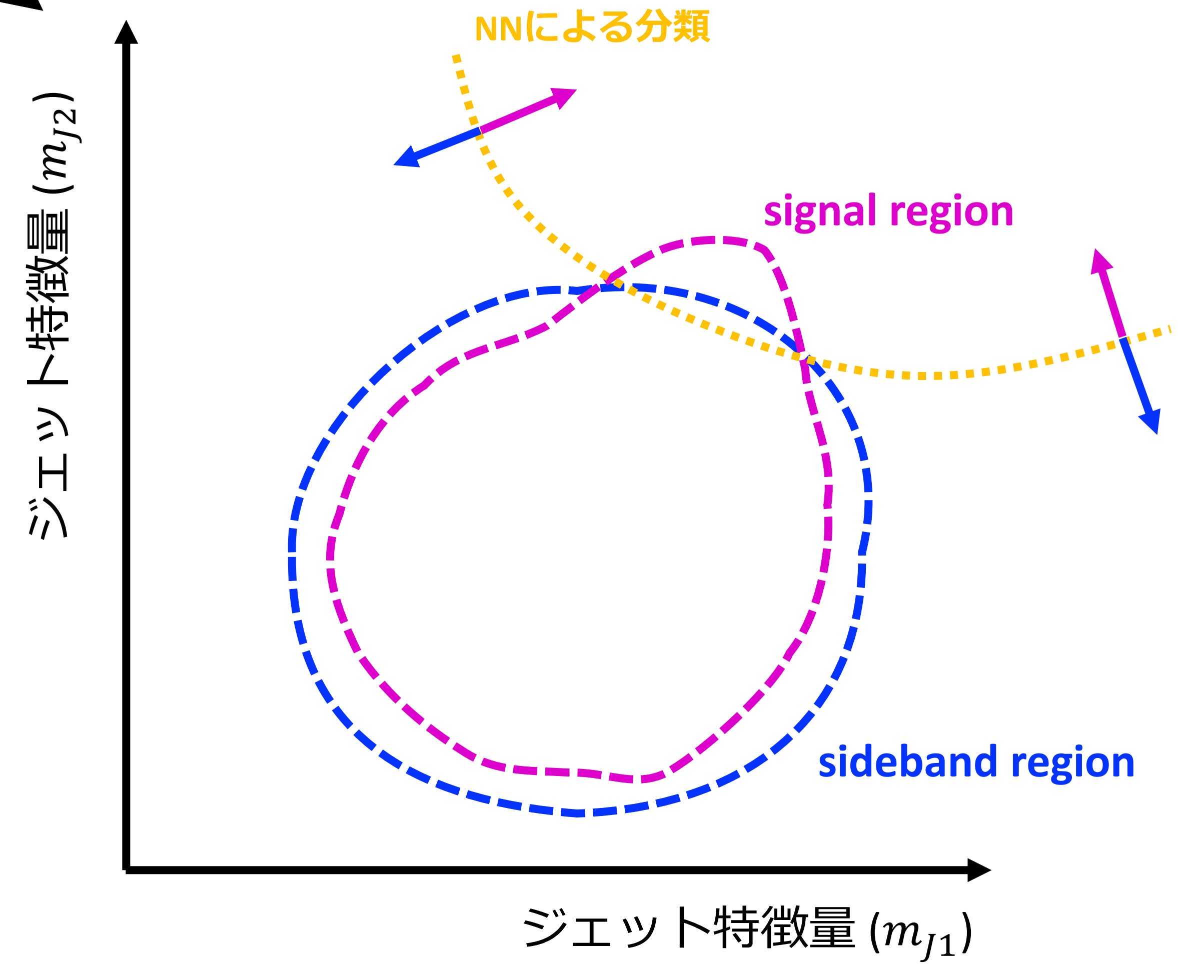


Dijet resonance search using CWoLa: アルゴリズム

Phys. Rev. Lett. 125, 131801 (2020)

- シグナルがいると、S/Bを改善するように学習
- シグナルがいないと、ランダムな分類器になる

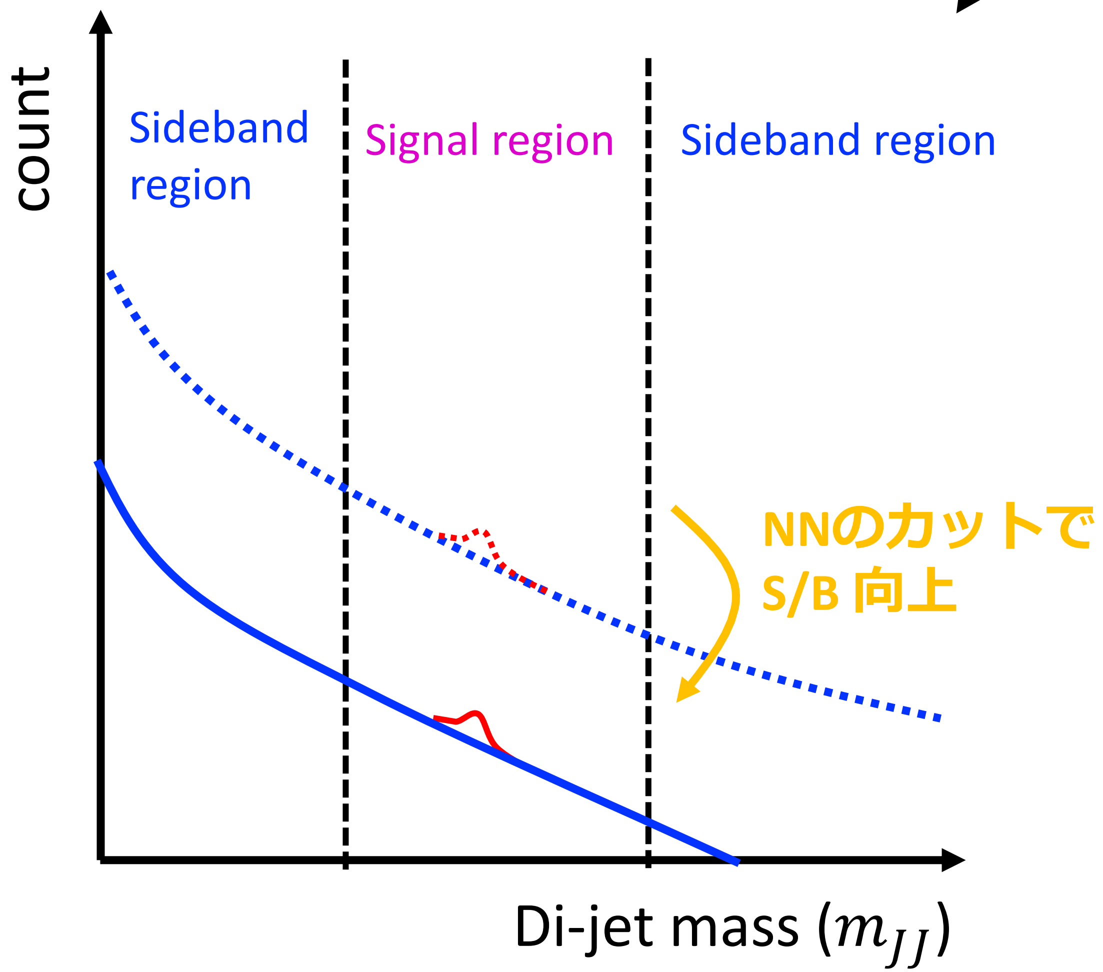


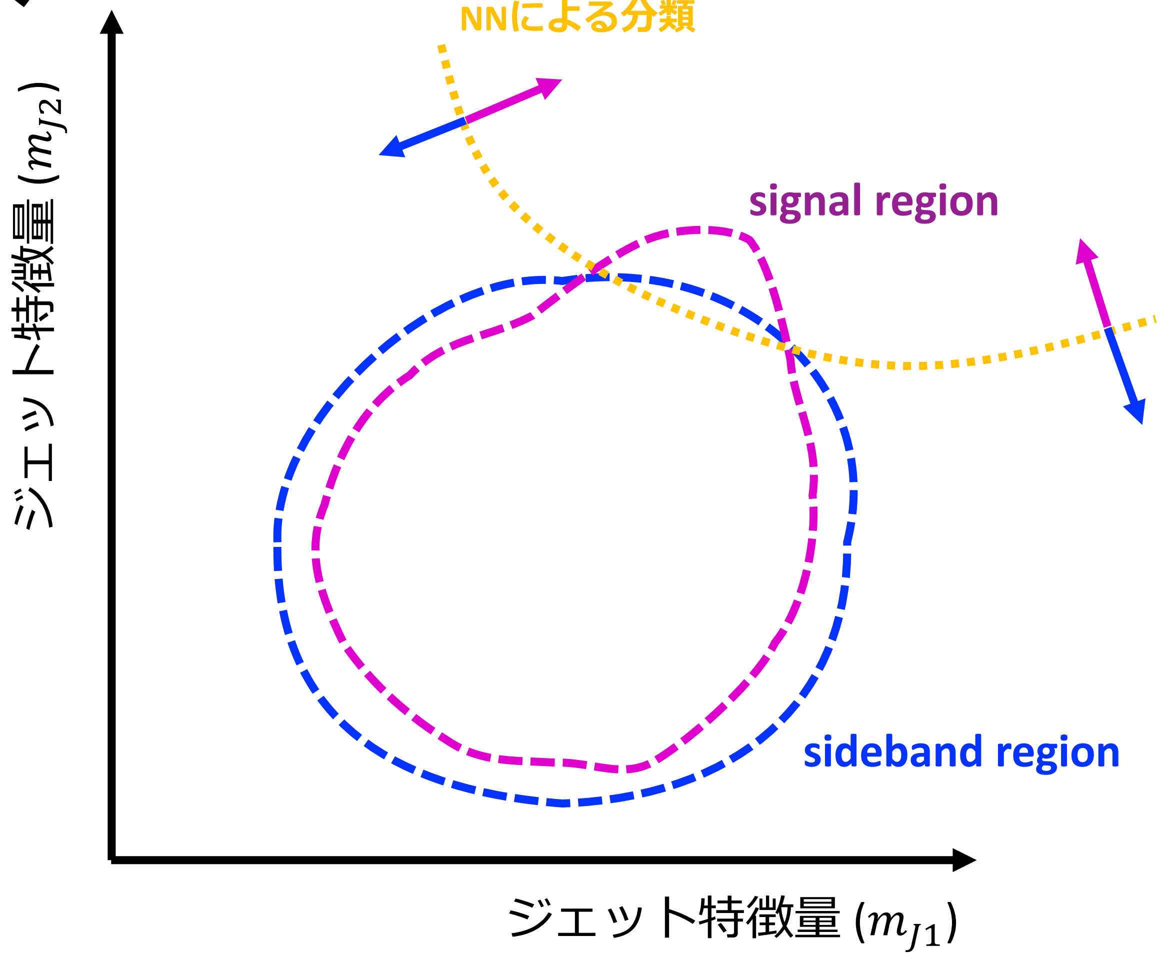


Dijet resonance search using CWoLa: アルゴリズム

Phys. Rev. Lett. 125, 131801 (2020)

学習したNNでカットをかける





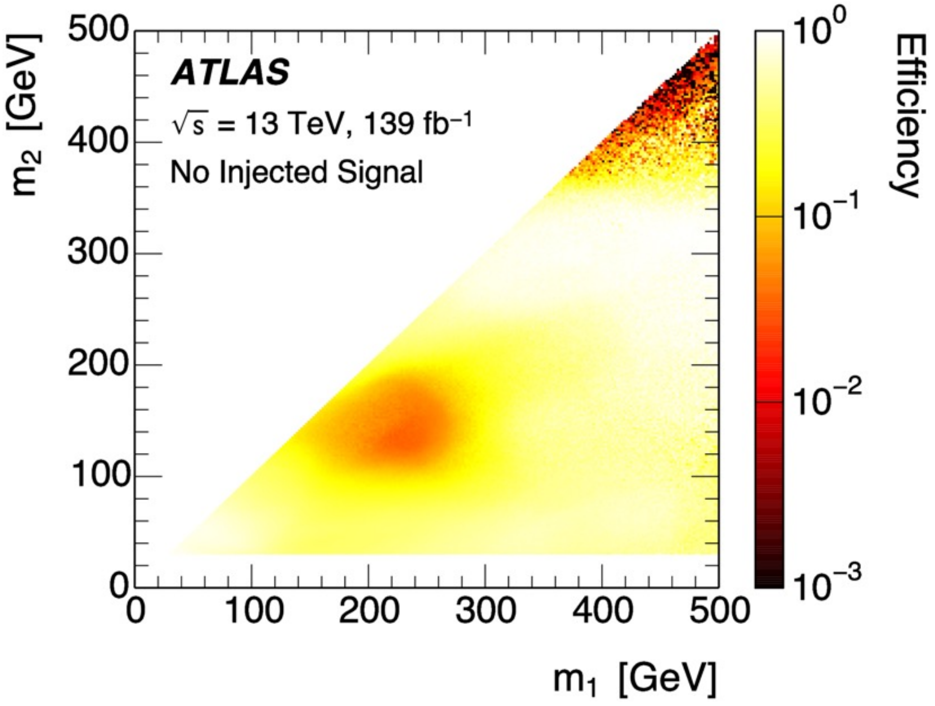
注: カット(NN)に使う変数は**di-jet massに依存しない**変数
相関があると性能が悪化する

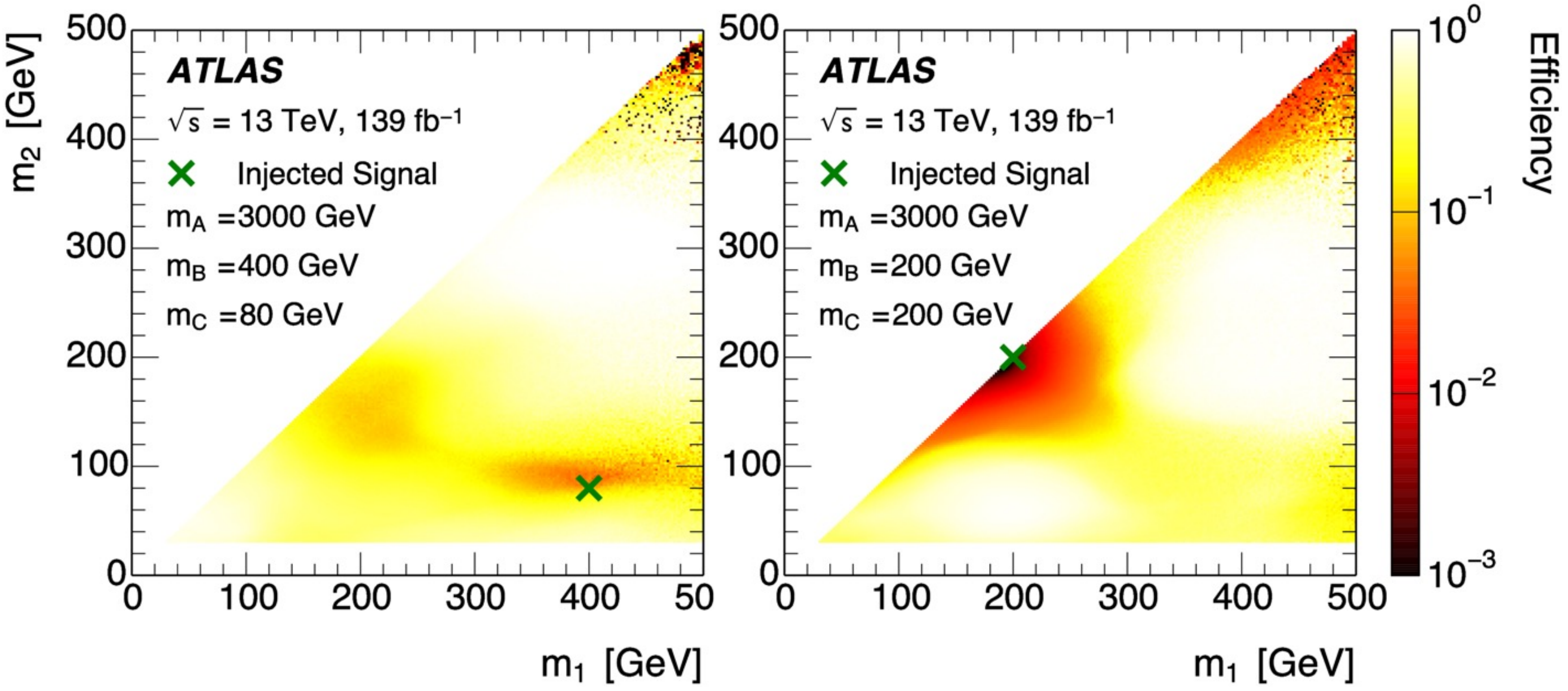
Dijet resonance search using CWoLa: NN出力

Phys. Rev. Lett. 125, 131801 (2020)

Signal injectionなし

Signal injectionあり



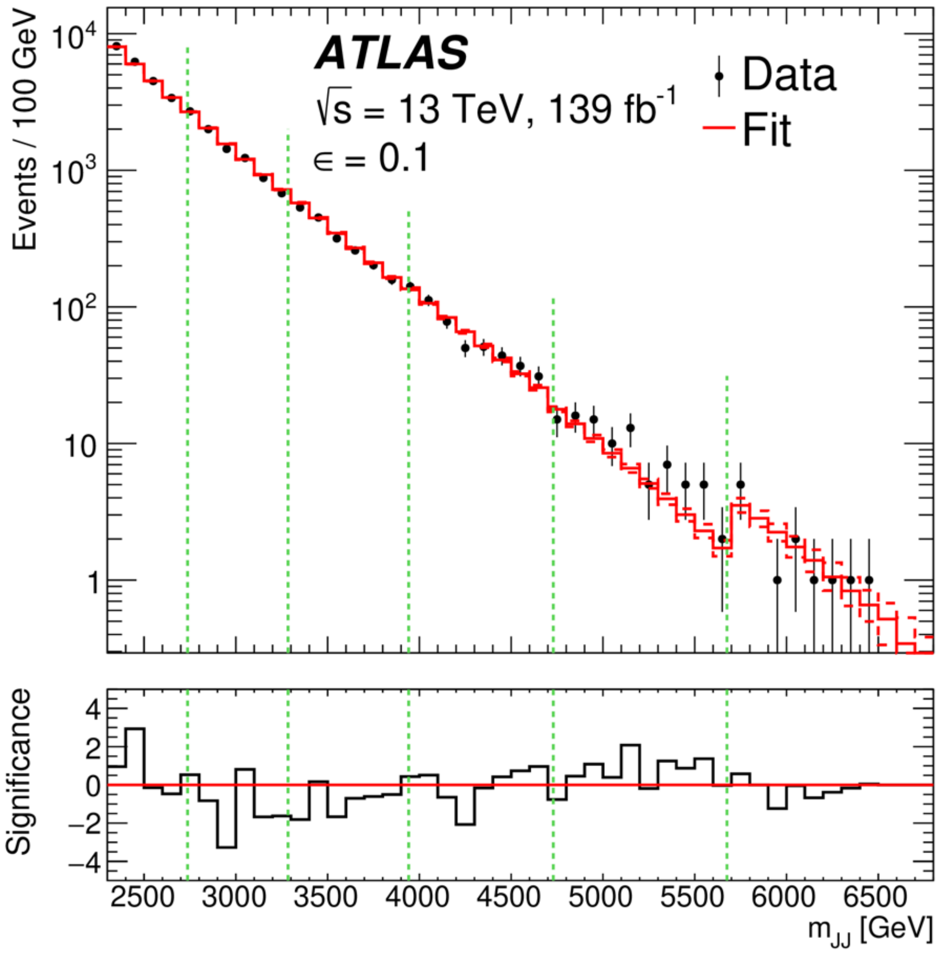


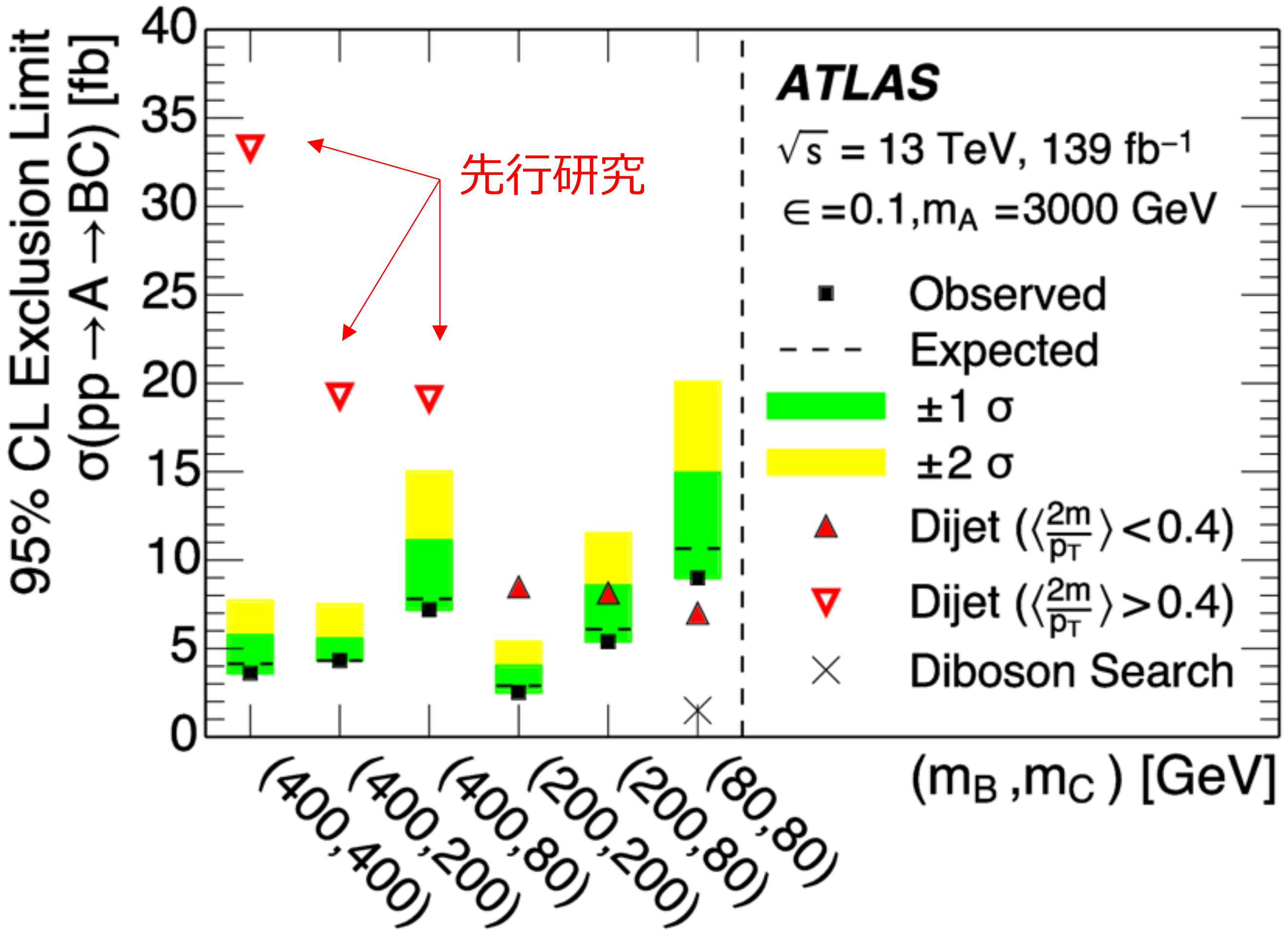
- 入力変数(m_{J1}, m_{J2})に対するNNの出力
- シグナルをインジェクトした領域をエンハンスするNN出力が得られている

Dijet resonance search using CWoLa: 結果

Phys. Rev. Lett. 125, 131801 (2020)

カット後のdi-jet mass分布





- 6つの Signal region (SR)
- シグナルがSRに少ない(or いない)ために、カットもなめらか

- シグナルモデルパラメータによっては、先行研究よりも良い結果
- モデルへの仮定が少ないことの恩恵

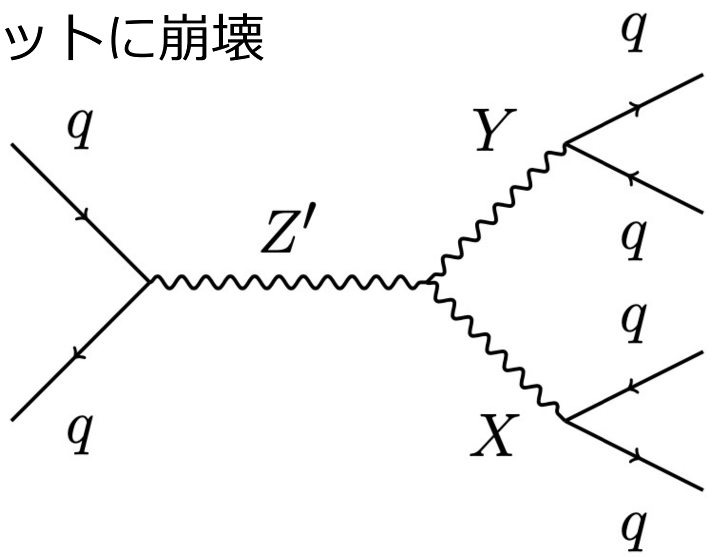
LHC Olympics 2020: 概要

2101.08320

- **シグナルモデルによらない物理探索**の研究を推進するために開催
  - 最終的に計18の提出
  - ベンチマークデータセットはZenodoで公開されており、
    先行研究との比較が可能

- 問題設定はATLAS の弱教師学習を用いたDijet resonance searchとほぼ同じ
  - 重い未知粒子(Z')が軽い未知粒子(X, Y)を通して2つの大半径ジェットに崩壊

- シグナルの質量が不明なブラックボックスサンプルから
  (1) シグナルイベント数 + uncertainty
  (2) シグナルの質量(Z', X, Y)
  等を推定



LHC Olympics 2020: データセット

2101.08320

Dataset	Signal	Background
R&D **(ラベルあり)**	$Z' \to XY \to qqqq$ (Z', X, Y) = (3500 GeV, 500 GeV, 100 GeV)	QCD di-jets
Black Box 1	$Z' \to XY \to qqqq$ (Z', X, Y) = (3823 GeV, 732 GeV, 378 GeV)	QCD di-jets (simulatorのconfigを変更)
Black Box 2	No signal (誤検知のチェック)	QCD di-jets (simulatorそのものを変更)
Black Box 3	$G \to gg$ and $G \to Rg \to ggg$ (G, R) = (4.2 TeV, 2.2 TeV)	QCD di-jets (simulatorのconfigを変更)

- BB2でシグナルを誤検知したチームもあった。今後の課題
- BB3は正解チームなし。トポロジーの変化に対応できる汎用的な手法はまだ難しい

サマリー

- 生成モデル・異常検知の応用例を紹介した
  - GANを用いたカロリメータシャワーシミュレーション
  - 弱教師あり学習(CWoLa)を使ったDijet resonance search

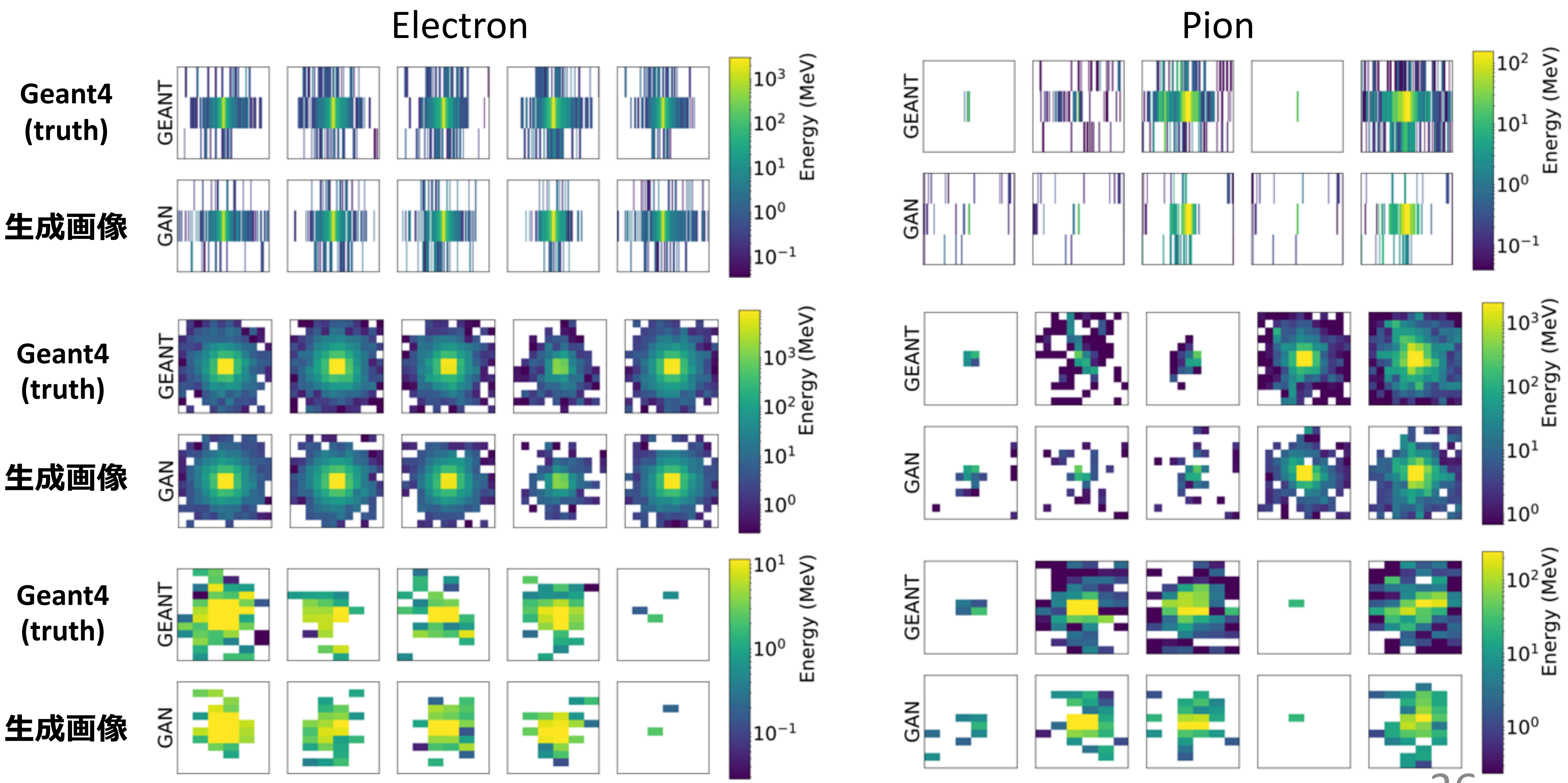
- この分野は発展途上
  - 実験で実際に使われている例はまだ少ない
  - コンテストが開催されたり、ベンチマークデータセットが整備されるなど、R&Dがしやすい環境が整ってきている

- 今後に期待

Backup

CaloGAN: 生成画像

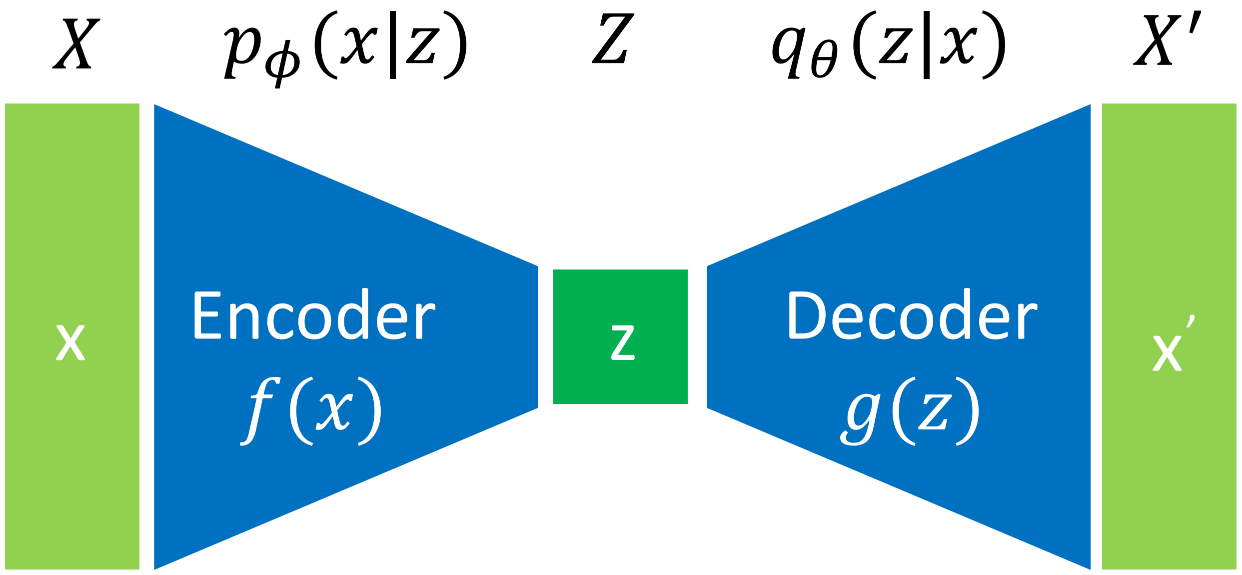
Phys. Rev. D 97, 014021 (2018)



ATLASでの初期の研究: VAE

ATL-SOFT-PUB-2018-001

Variational autoencoder (VAE)



VAE for 生成モデル

- $\|X - X'\|^2$ が小さくなるように
- Zが正規分布に従うように

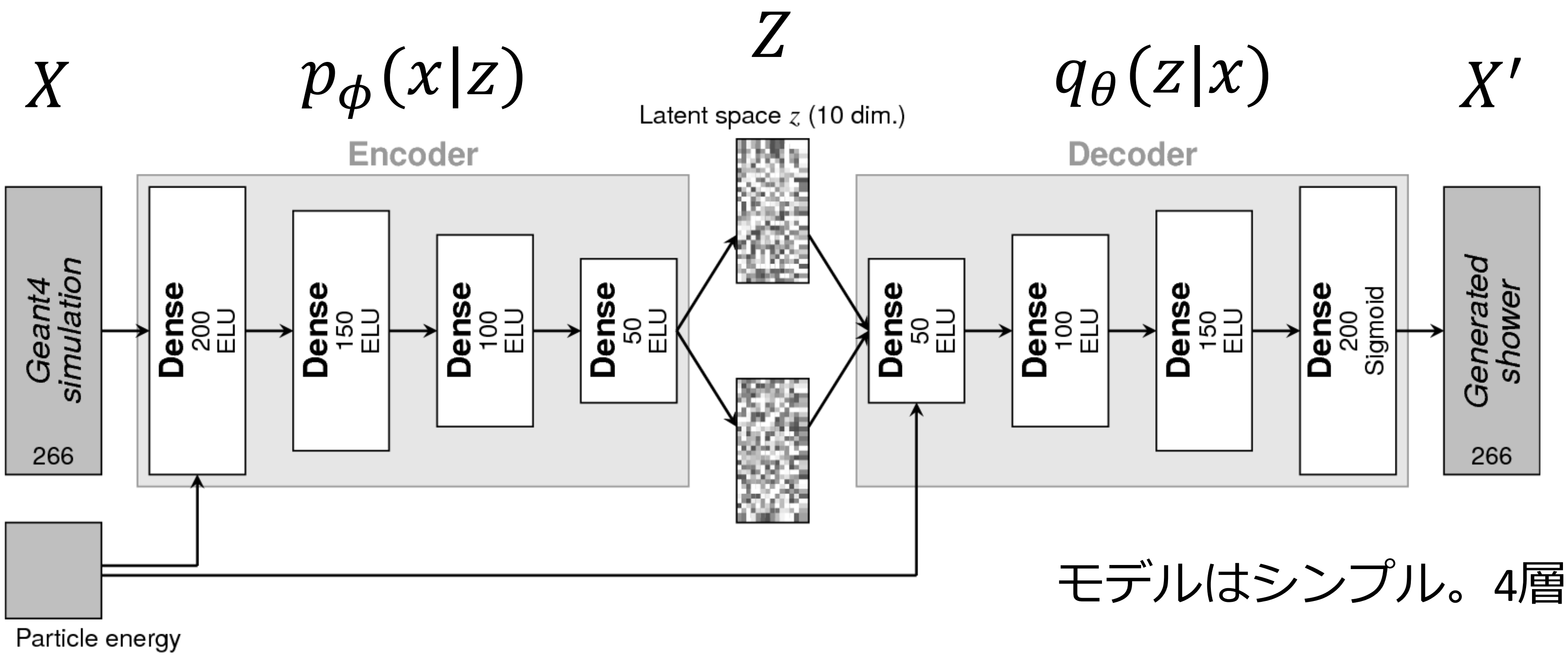
エンコーダ・デコーダを学習する

学習したデコーダを生成モデルとして使用する

$$\mathcal{L} = \|x - g(f(x))\|^2 + D_{KL}(q(z|x), p(z))$$

ATLASでの初期の研究: VAE

ATL-SOFT-PUB-2018-001



モデルはシンプル。4層のMLP

VAE for 生成モデル

- $\|X - X'\|^2$ が小さくなるように
- Zが正規分布に従うように

エンコーダ・デコーダを学習する

学習したデコーダを生成モデルとして使用する

ATLASでの工夫

- 全エネルギー
- 各レイヤーでのエネルギーの割合

もロス関数に加えることで、より詳細にコントロール

ATLASでのVAEの発展

改善が続けられている

- Normalizing Flow (IAF)の導入
- ロス関数の変更

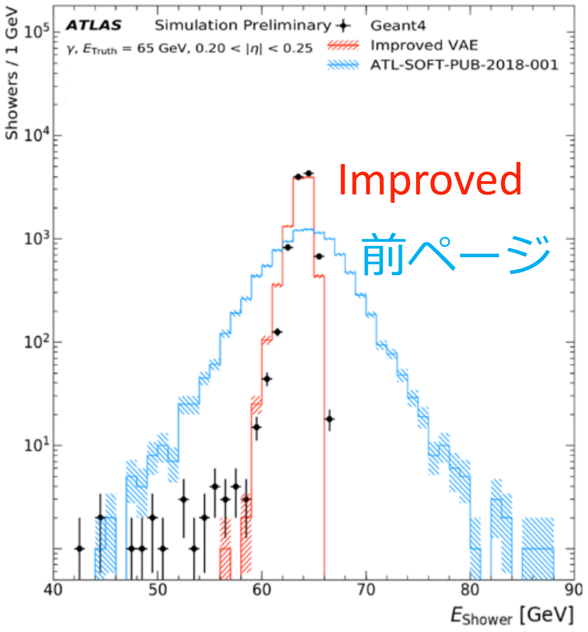
ATL-PLOTS-SIM-2019-007

ロス関数の変更

各セルのエネルギー(E_{cell}) の代わりに

エネルギー比($E_{\mathrm{total}}, E_{\mathrm{layer}}/E_{\mathrm{total}}, E_{\mathrm{cell}}/E_{\mathrm{layer}}$)

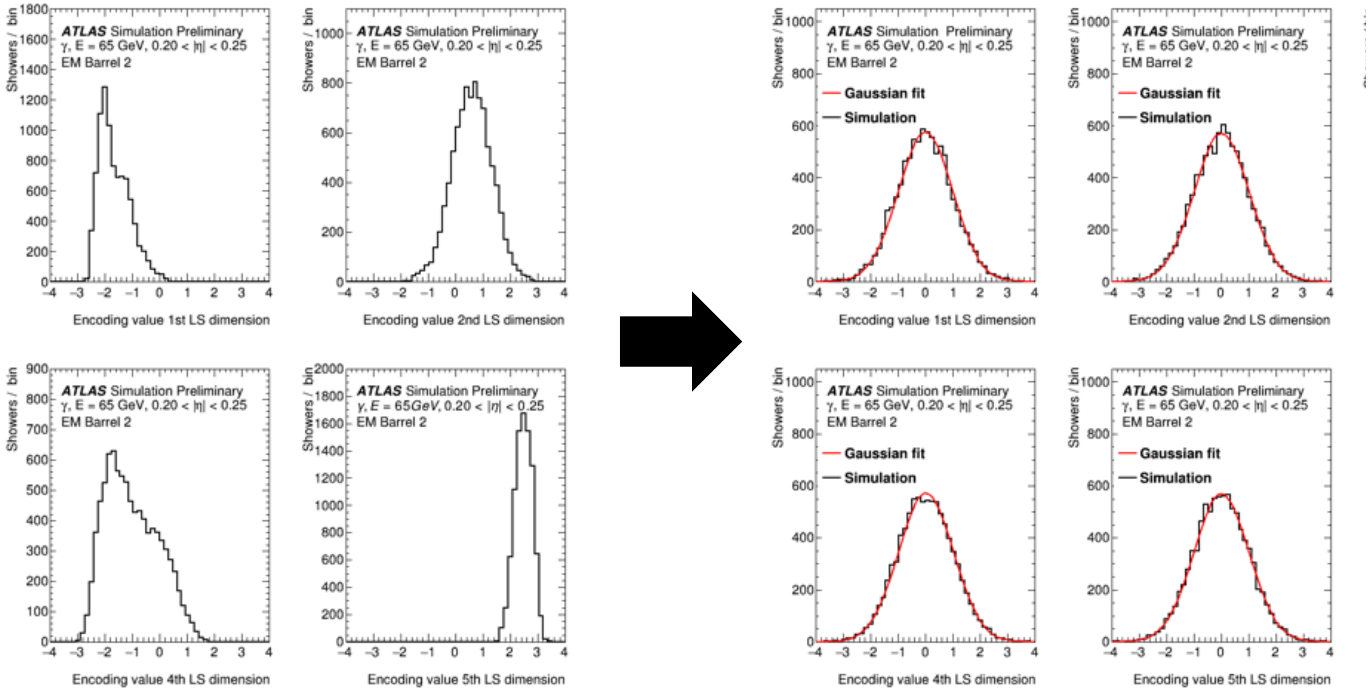
を予測



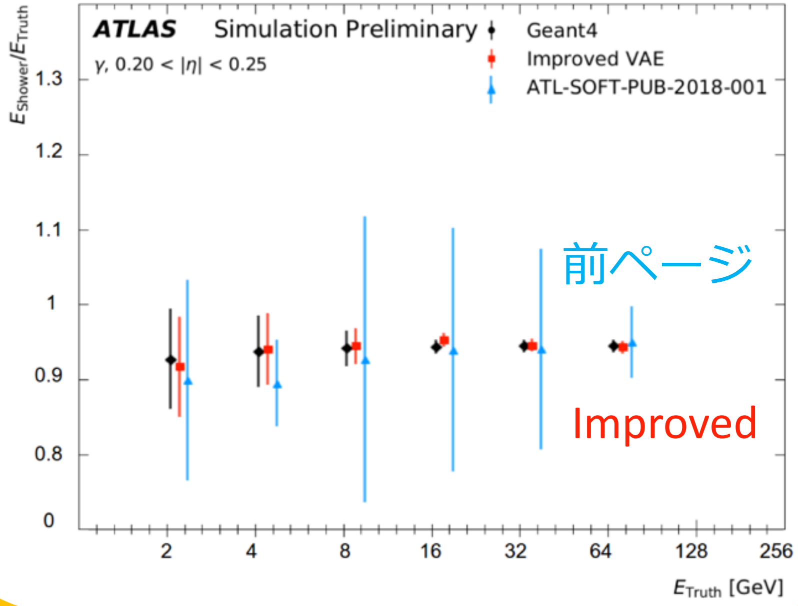
latent space(z)の分布

Original

Normalizing Flowの導入



ATL-PLOTS-SIM-2019-003



GAN for Fast Simulation at ATLAS: 学習時の工夫

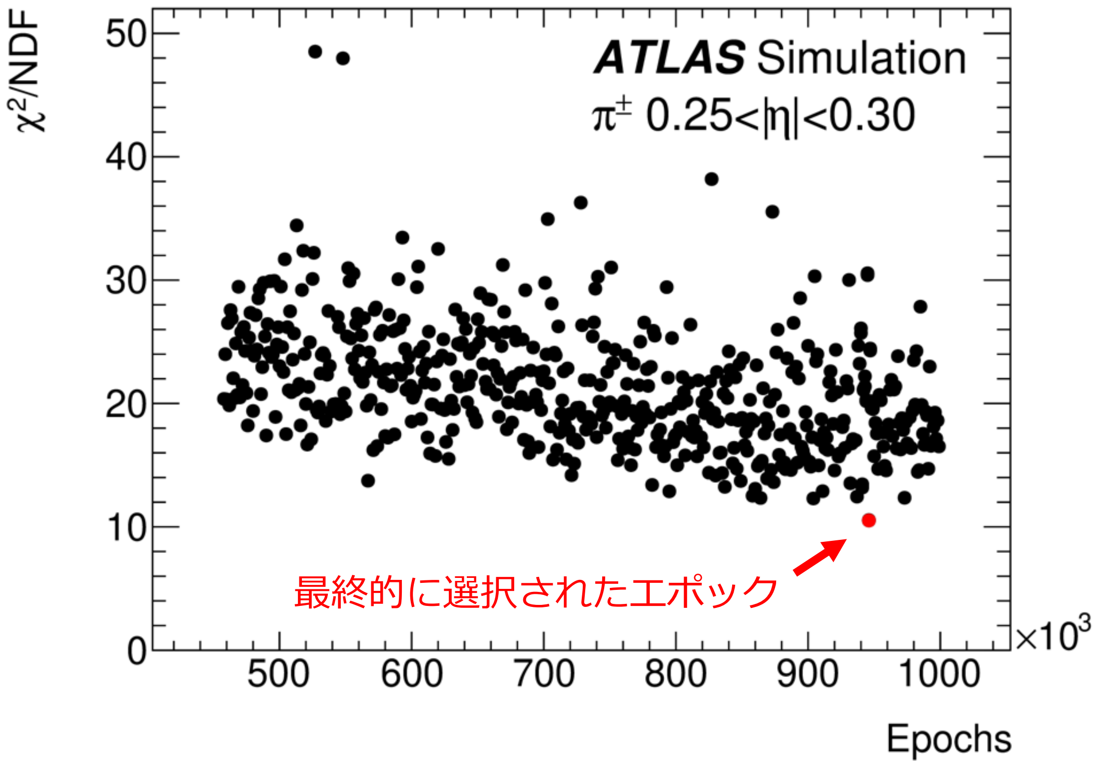
Comput Softw Big Sci 6 (2022) 7

- 学習サンプルの逐次投入
  - 全エネルギー領域(256 MeV ~ 4 TeV)を一気に学習するのではなく、1つのエネルギーポイント(32 GeV)を中心に徐々に学習サンプルを拡張
  - 全サンプル追加後は全てのサンプルを使って学習

- Best エポックの選び方
  - 最後のエポックが最適とは限らない
  - 総エネルギー量の分布がGANとGeant4でどの程度あっているかをχ^2で定義し、それを基にベストなエポックを選択

- 1つのGANモデルの学習時間は **8時間** ほど (V100使用)

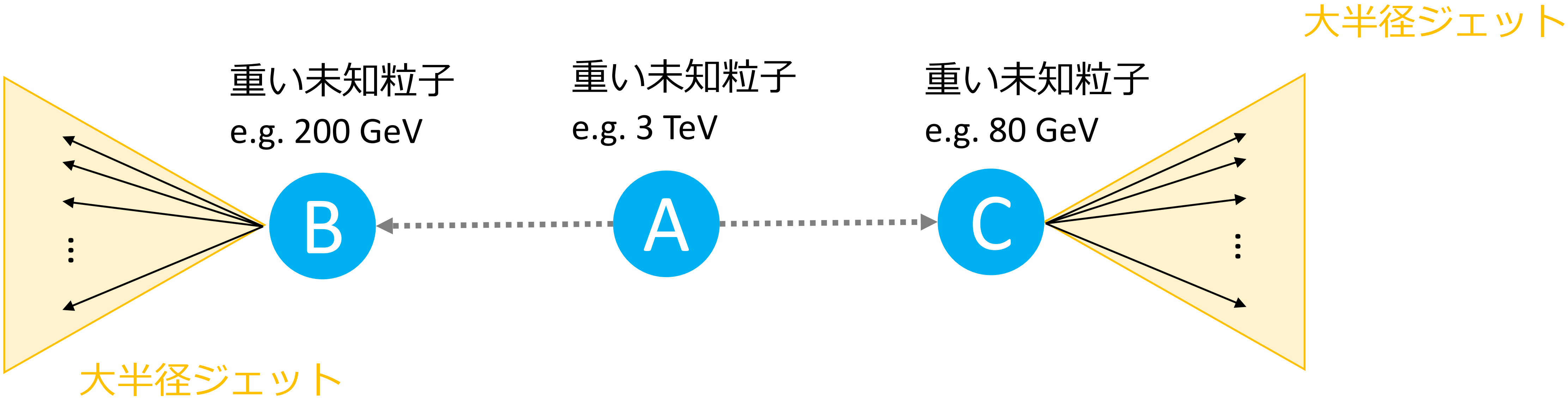
学習中の χ^2 の推移



Dijet resonance search using CWoLa: ターゲットモデル

Phys. Rev. Lett. 125, 131801 (2020)

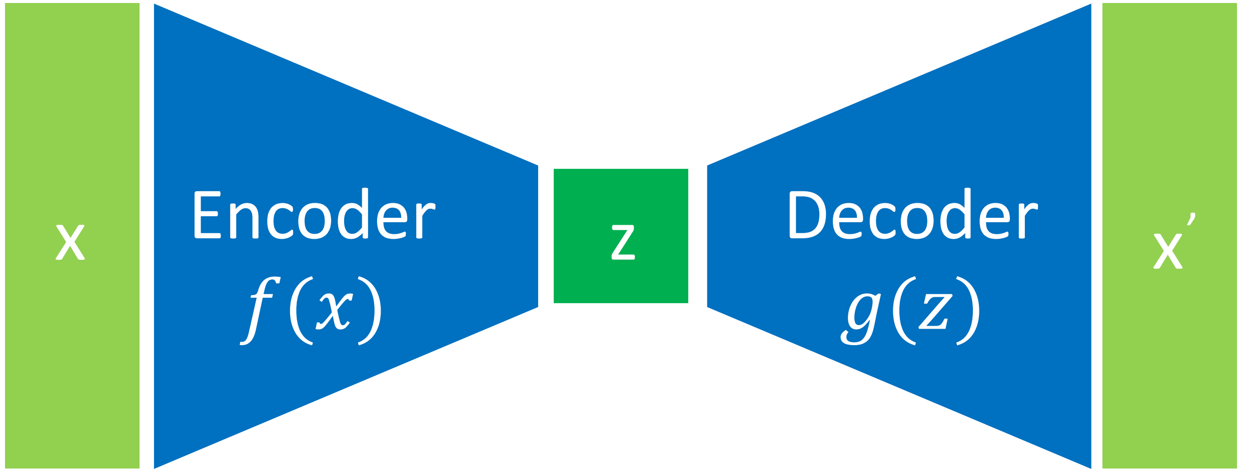
シグナルモデル



$$m_A \sim \mathcal{O}(\mathrm{TeV})$$
$$m_B, m_C \sim \mathcal{O}(100\ \mathrm{GeV})$$

LHC Olympics 2020: VAEを使った異常検知

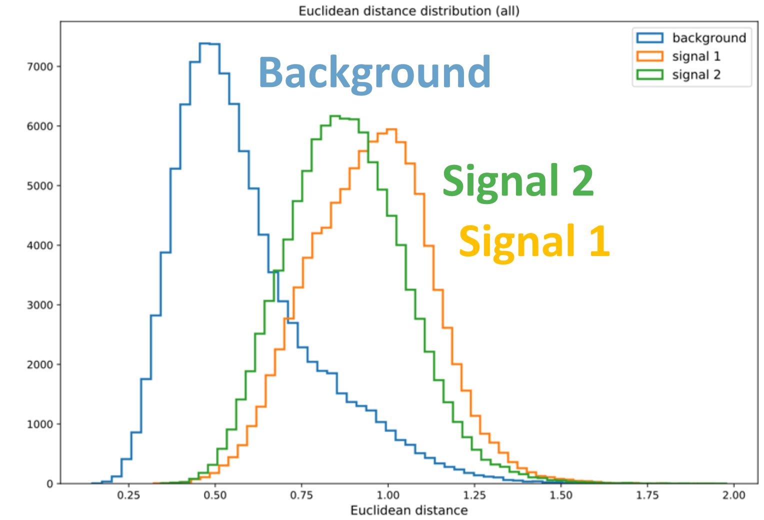
2101.08320



$$\mathcal{L} = \|x - g(f(x))\|^2 + D_{KL}(q(z|x), p(z))$$

- バックグラウンドデータ分布をVAEで学習する
- シグナルが入力されると、
  - 再構成エラー($\|X - X'\|^2$)が大きい
  - zの値($\|Z\|^2$)が大きい

  等のアノマリーとして判定できる。

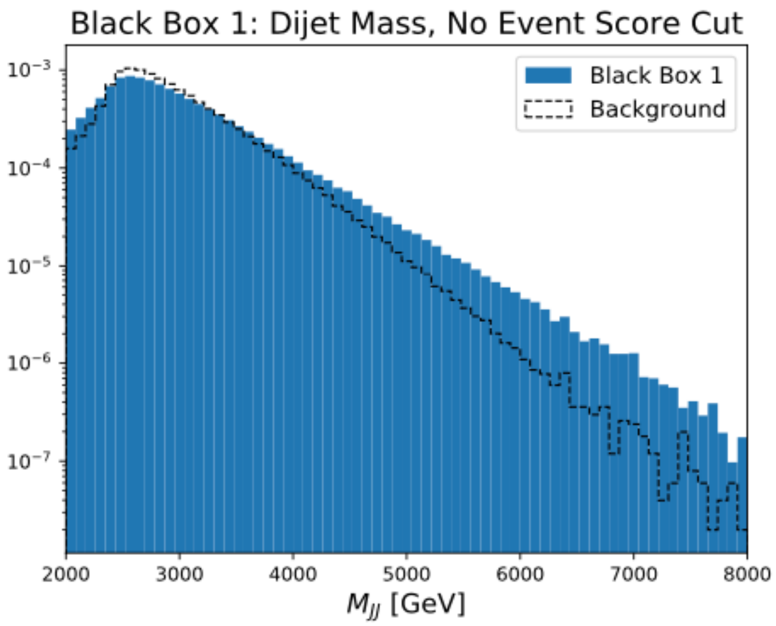


再構成エラー($\|X - X'\|^2$)

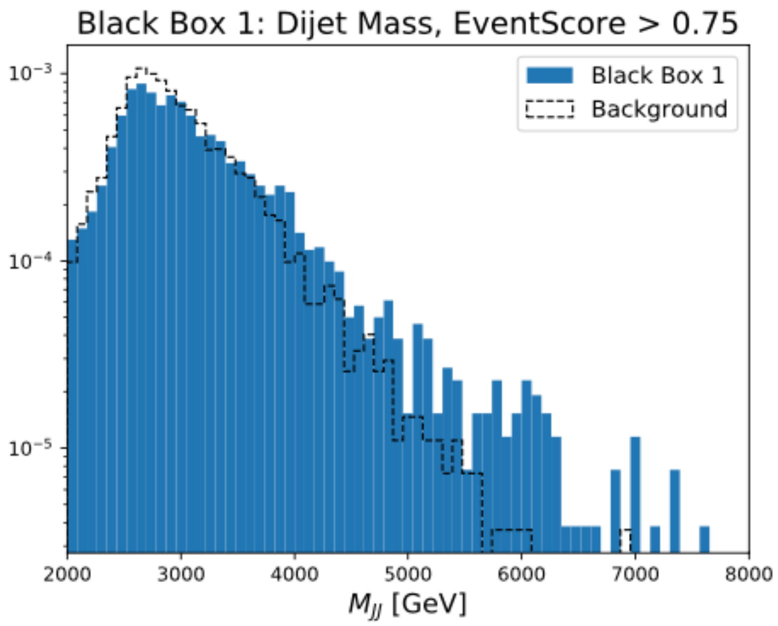
LHC Olympics 2020: VAEを使った異常検知

[2101.08320](2101.08320)

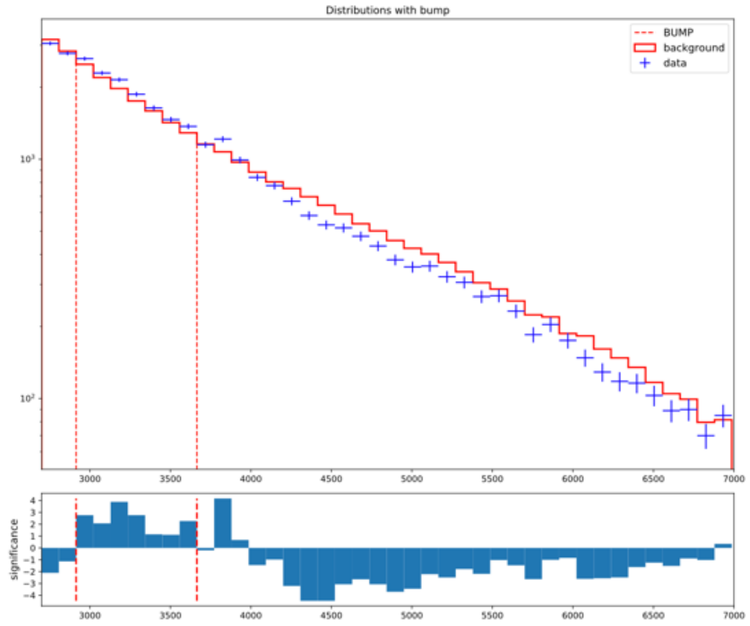
No selection



anomaly scoreでcut後



anomaly scoreでcut後

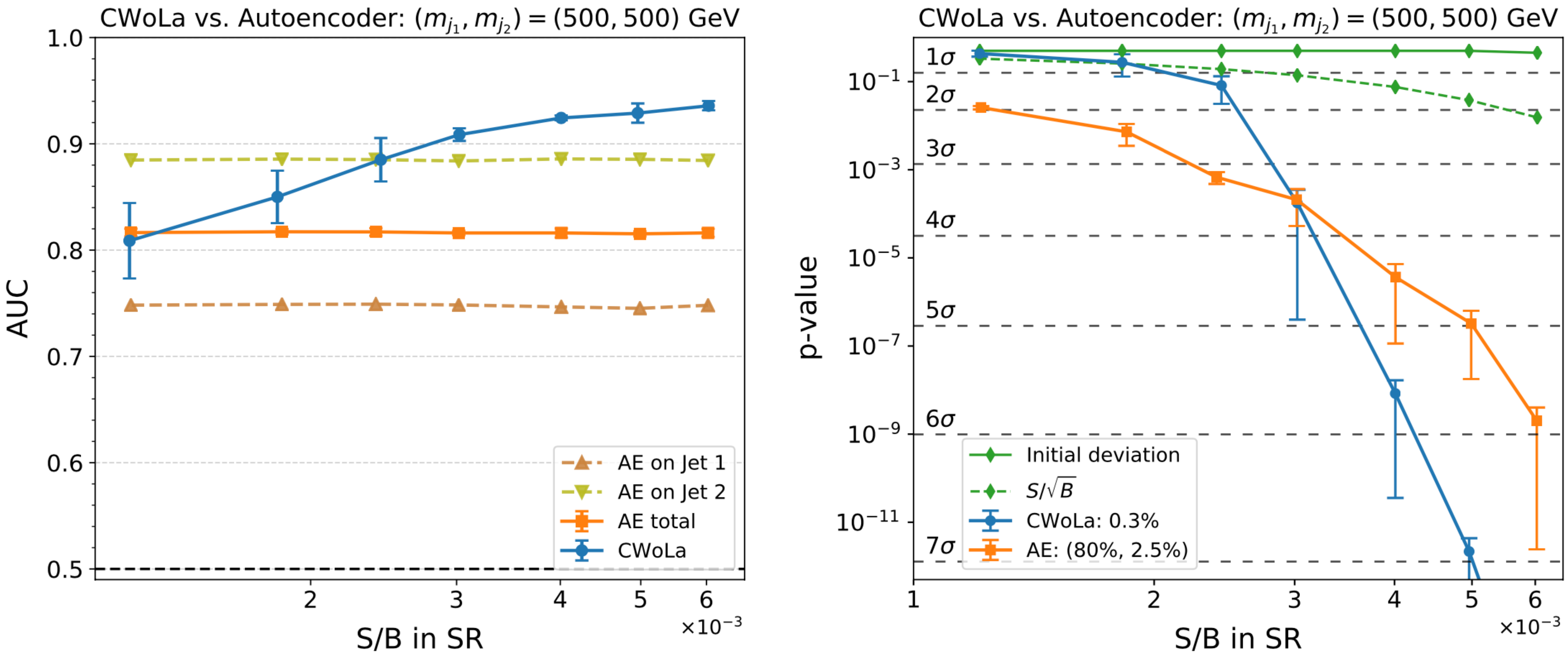


- **R&Dデータ**でVAEの学習をした例
- 学習データにバイアスがあるためうまくいかなかった

LHC Olympics 2020: VAEを使った異常検知

[2101.08320]

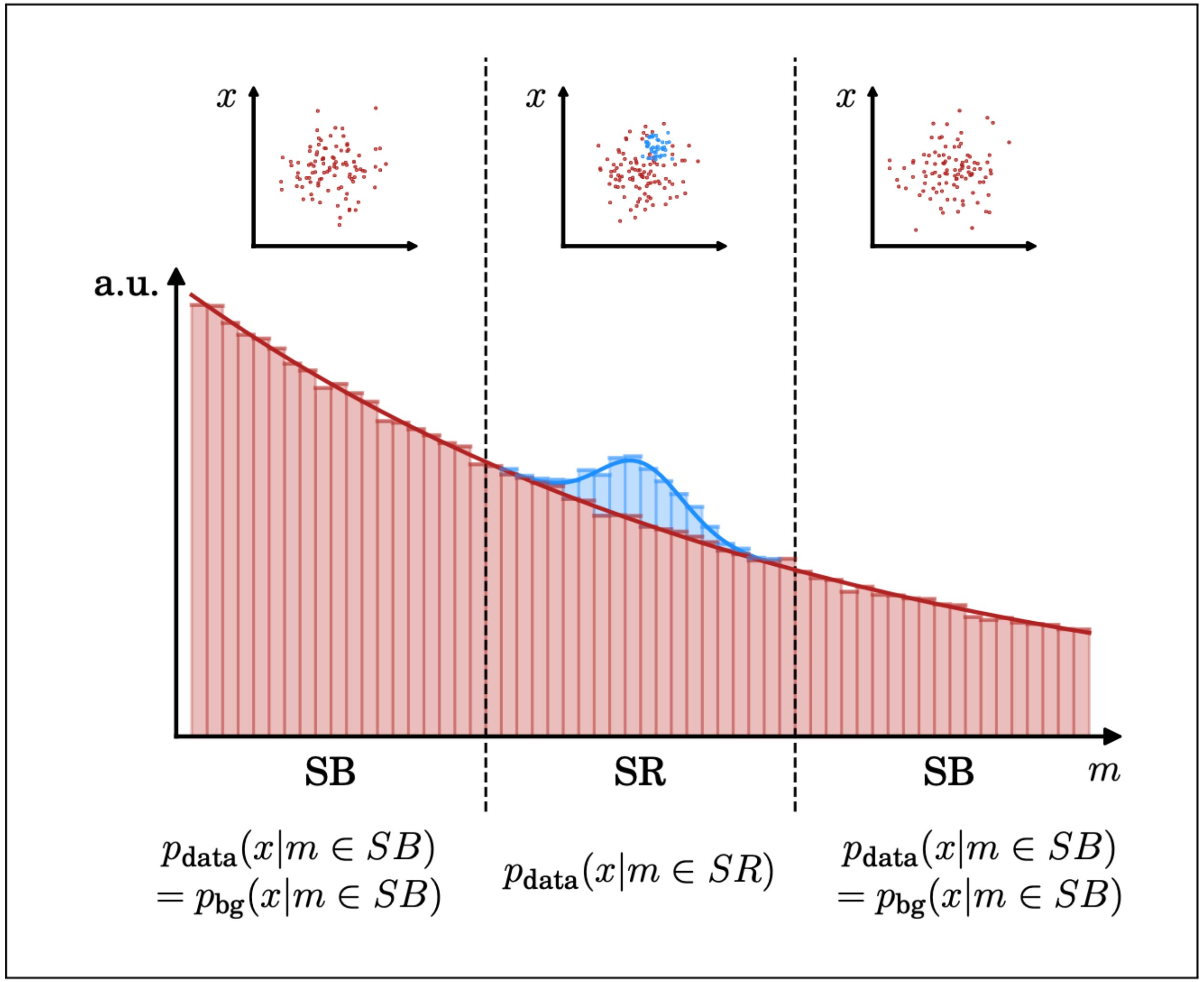
BBデータを直接使った例。S/Bが小さいためシグナルの影響が無視できた



- S/Bが小さいとCWoLaは性能が悪くなる。VAEの性能はS/Bに依存しない
  - VAE は **BG分布** をモデリングする。
  - CWoLaは **Signal分布とBG分布の差** をモデリングする。

LHC Olympics 2020: 密度推定を使った異常検知

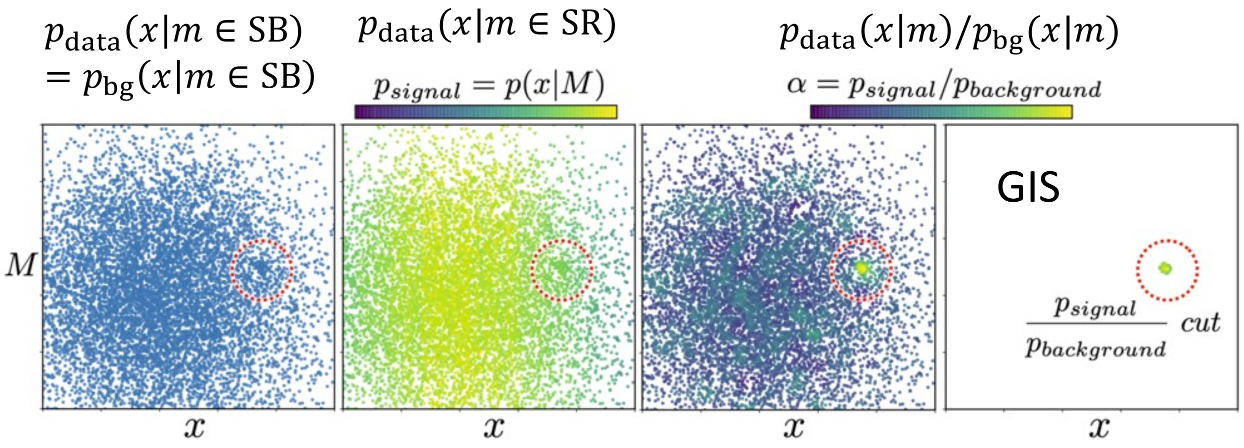
2101.08320



2109.00546

手法

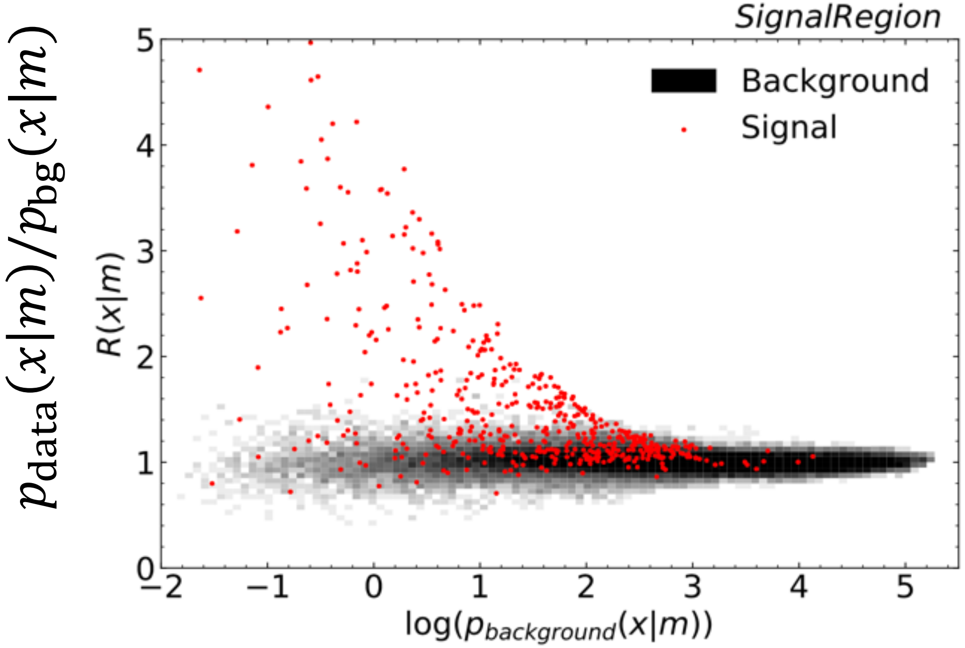
1. NNでSRの分布 $p_{\text{data}}(x|m \in \text{SR})$ フィット
2. NNでSBの分布 $p_{\text{data}}(x|m \in \text{SB})$ フィット
   - $p_{\text{data}}(x|m \in \text{SB}) = p_{\text{bg}}(x|m \in \text{SB})$
3. SB で推定した分布をSRに内挿
4. $p_{\text{data}}(x|m)/p_{\text{bg}}(x|m)$ を分類器として使う



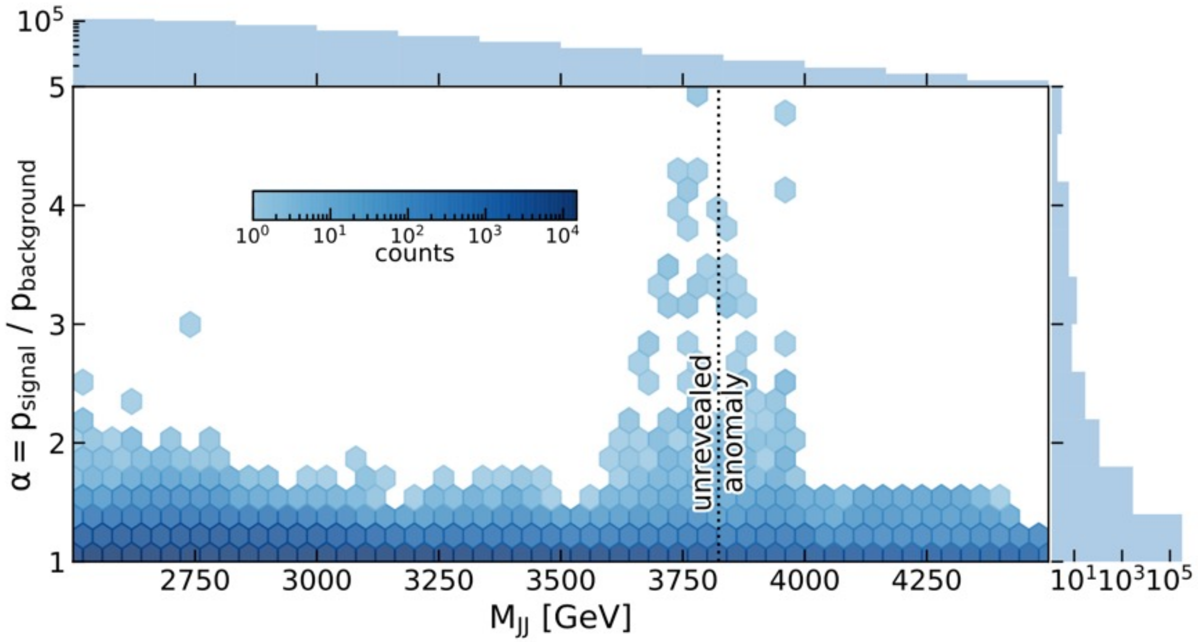
LHC Olympics 2020: 密度推定を使った異常検知

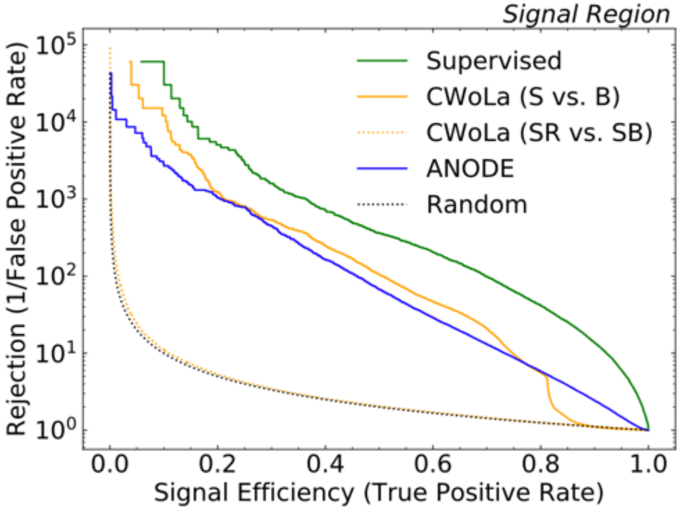
2101.08320

Masked Autoregressive Flow (MAF)を使った例



Gaussianizing Iterative Slicing (GIS) を使った例



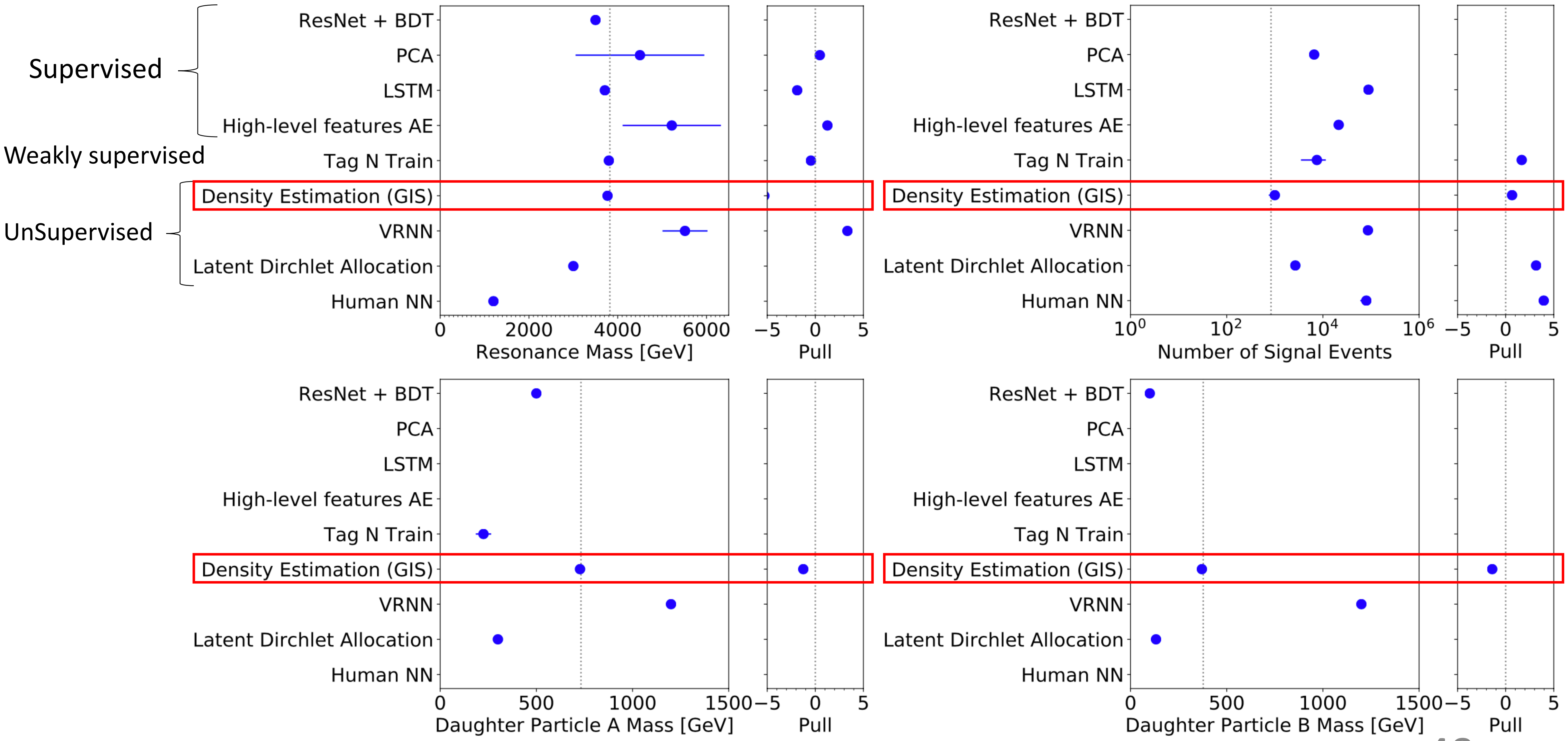


- Normalizing Flowで密度推定
- CWoLaより少し結果が悪い
  - CWoLaは比(p_S/p_B)を直接モデリングしているが、密度推定では2つの分布を独立にモデリング
- CWoLaと異なり、massへの**相関にロバスト**

LHC Olympics 2020

(i): blind datasetでの結果
(ii): unblind後の結果
(iii): R&D のみ

Section	Short Name	Method Type	Results Type
3.1	VRNN	Unsupervised	(i) (BB2,3) and (ii) (BB1)
3.2	ANODE	Unsupervised	(iii)
3.3	BuHuLaSpa	Unsupervised	(i) (BB2,3) and (ii) (BB1)
3.4	GAN-AE	Unsupervised	(i) (BB2-3) and (ii) (BB1)
3.5	GIS	Unsupervised	(i) (BB1)
3.6	LDA	Unsupervised	(i) (BB1-3)
3.7	PGA	Unsupervised	(ii) (BB1-2)
3.8	Reg. Likelihoods	Unsupervised	(iii)
3.9	UCluster	Unsupervised	(i) (BB2-3)
4.1	CWoLa	Weakly Supervised	(ii) (BB1-2)
4.2	CWoLa AE Compare	Weakly/Unsupervised	(iii)
4.3	Tag N' Train	Weakly Supervised	(i) (BB1-3)
4.4	SALAD	Weakly Supervised	(iii)
4.5	SA-CWoLa	Weakly Supervised	(iii)
5.1	Deep Ensemble	Semisupervised	(i) (BB1)
5.2	Factorized Topics	Semisupervised	(iii)
5.3	QUAK	Semisupervised	(i) (BB2,3) and (ii) (BB1)
5.4	LSTM	Semisupervised	(i) (BB1-3)



Supervised
Weakly supervised
UnSupervised
ResNet + BDT
PCA
LSTM
High-level features AE
Tag N Train
Density Estimation (GIS)
VRNN
Latent Dirchlet Allocation
Human NN
Resonance Mass [GeV]
Pull
Number of Signal Events
Daughter Particle A Mass [GeV]
Daughter Particle B Mass [GeV]

WGAN-GP

Original GAN : $L(G,D) = \mathbb{E}_{\tilde{x}\sim p_{\mathrm{gen}}}\left[\log\left(1 - D(\tilde{x})\right)\right] + \mathbb{E}_{x\sim p_{\mathrm{Geant4}}}[\log D(x)]$

WGAN-GP : $L(G,D) = \mathbb{E}_{\tilde{x}\sim p_{\mathrm{gen}}}[D(\tilde{x})] - \mathbb{E}_{x\sim p_{\mathrm{Geant4}}}[\,D(x)] + \lambda\mathbb{E}_{x\sim p_{\hat{x}}}[(\,\|\Delta_{\hat{x}}D(\hat{x})\|_2 - 1)^2]$

Original GAN ではJSD(Jensen-Shannon divergence)を最小化する
WGANではWasserstein distanceを最小化する