Search for new physics in final states with **semi-visible jets or anomalous signatures** using the ATLAS detector

August 27th, 2025

Ki Ryeong Park

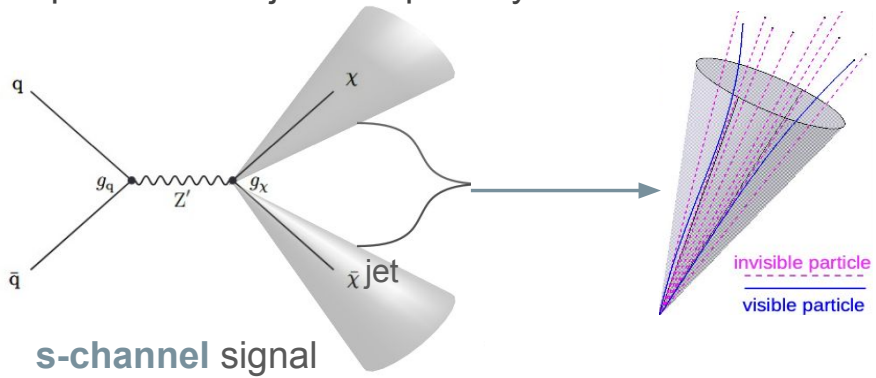
on behalf of the ATLAS Collaboration

at Lepton-Photon 2025 Symposium

What is SVJ?

- Semi-visible jet (SVJ): potential dark matter candidate signature
  - predicted w/ **minimal assumptions** to **avoid blind spots**
    - **heavy mediators** (Z') between **SM** + **dark sectors**
    - dark quarks shower + hadronize to dark hadrons (dark **QCD** in Hidden Valley)
    - **some** dark hadrons **promptly** decay to SM inside detector while others stable

      → invisible + visible particles in a jet from primary vertex → **semi-**visible

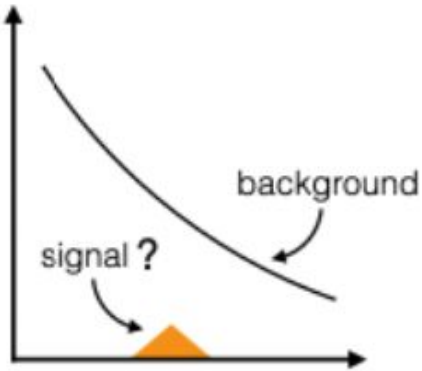


s-channel signal

event w/ 2 back-to-back SVJs

Why Is This Search Interesting?

- Naive strategy for s-channel
  - look for a resonance peak of 2 SVJs at $m_{Z'}$
    - from m_T distribution
  - enhance the peak → remove background by finding a discriminating variable



$$m_{\mathrm{T}}^2 = \left[E_{\mathrm{T,JJ}} + E_{\mathrm{T}}^{\mathrm{miss}}\right]^2 - \left[\vec{p}_{\mathrm{T,JJ}} + \vec{p}_{\mathrm{T}}^{\mathrm{miss}}\right]^2$$
$$= m_{\mathrm{JJ}}^2 + 2p_{\mathrm{T}}^{\mathrm{miss}}\left[\sqrt{m_{\mathrm{JJ}}^2 + p_{\mathrm{T,JJ}}^2} - p_{\mathrm{T,JJ}}\cos(\phi_{\mathrm{JJ,miss}})\right]$$

m_T → invariant mass of 2 jets + E_T^{miss}

Why Is This Search Interesting?

- Challenge:
  - **cannot** use standard, **high-level** variables alone as discriminating variables e.g.
    - E_T^{miss} → can be large due to mismeasurement even for non-SVJ
    - jet-related → not all tracks are visible in SVJ
  - require a tool that is sensitive to a **variety** of **anomalous signatures** to **avoid blind spots**
    - many **free** parameters b/c model is minimally biased e.g. n_{fqD}, n_{cD}, Λ_D, etc
    - previous search with specific parameter settings (paper)

→ Opportunity for a novel **anomaly detection** tool to search for SVJ + anomalous signatures

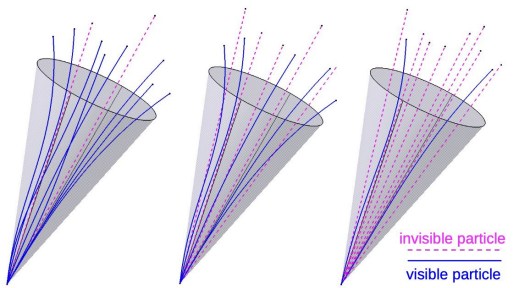
using **low-level** (track) variables

→ new ATLAS paper **result in 2025**

Analysis Samples

- Data from the ATLAS detector (paper) from pp collision at LHC
  - large √s + data size (13 TeV + 140 fb^{-1} in run 2) → ideal for **heavy** mediator signal
  - estimate background + find signal strength
- Simulated samples:
  - signal: events w/ SVJ(s) w/ varying **$m_{Z'}$ + R_{inv}**
    - benchmark: $m_{Z'}$ = 2.0 - 5.0 TeV + R_{inv} = [0.2, 0.4, 0.6]
  - background:
    - **QCD multijets dominates** (~80%)
      - b/c mismeasurement can mimic SVJ
  - develop analysis strategy + tool → minimize bias
    - background estimation → data driven instead of QCD MC

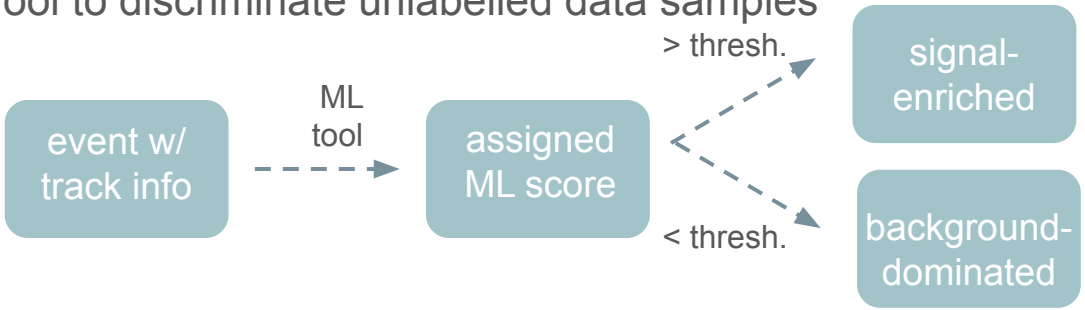
$$R_{inv} = \frac{\text{number of invisible hadrons}}{\text{number of all hadrons}}$$

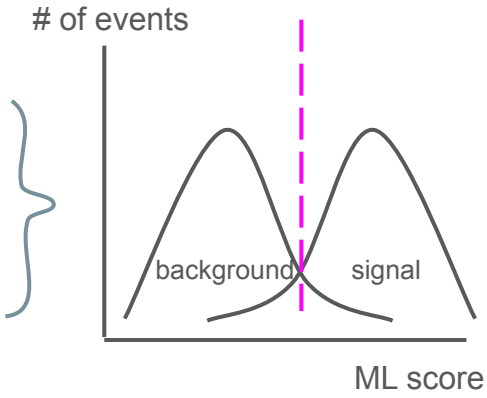


more **invisible** SVJ -> **higher** R_{inv}

Machine Learning (ML) Analysis Strategies

- Tool to discriminate unlabelled data samples





- Requirements for tool (for reasons described earlier)
  - permutation-invariant — no physical meaning in track order
  - variable length (of inputs) — different number of tracks for each jet
  - unsupervised training — many free parameters in signal model
  - handle out-of-distribution anomalies — avoid misclassifying unseen background

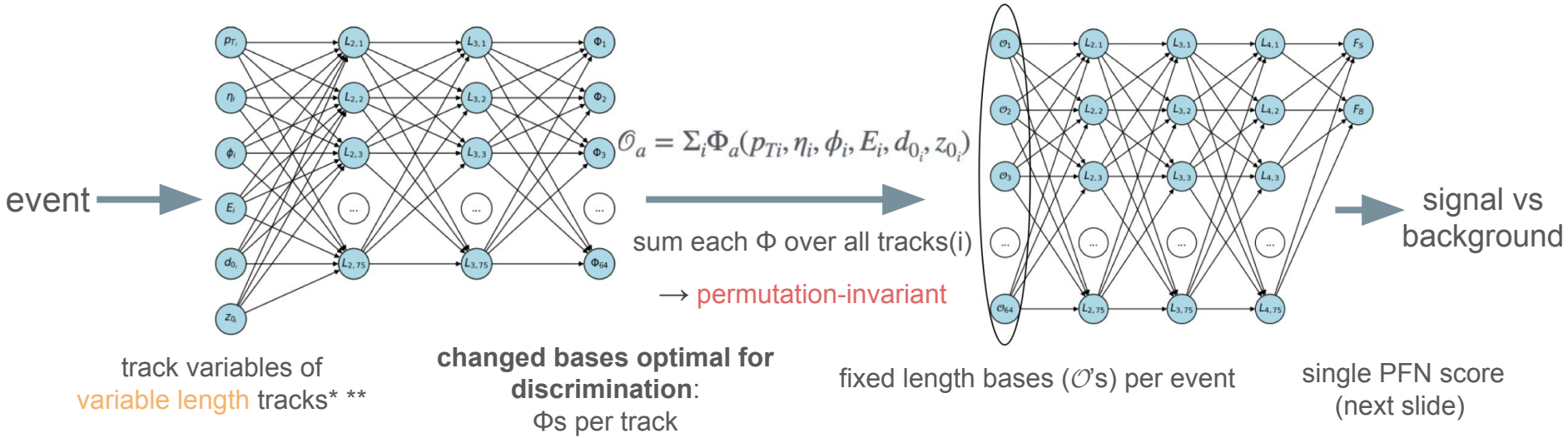
no conventional ML architectures meet all these requirements

→ develop custom ML tool using 2 architectures

* maximum 80 highest pT generally includes most tracks of each jets
** 2 leading jets are used for tracks so x 2 jets

ML Analysis Strategies - PFN

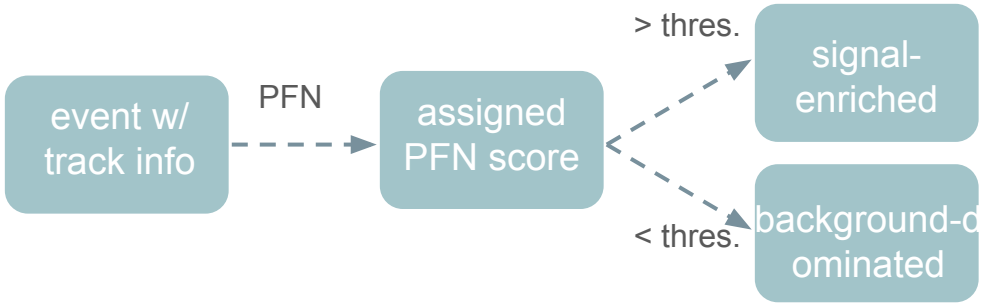
- Chose + adapted a neural network architecture called Particle Flow Network (**PFN)**
  - already developed for using sets to be permutation invariant + handle variable length for collider events (paper)



*about 500k events each

ML Analysis Strategies - PFN

- PFN is trained w/ labelled **sim. events** (signal vs QCD*) → **supervised**



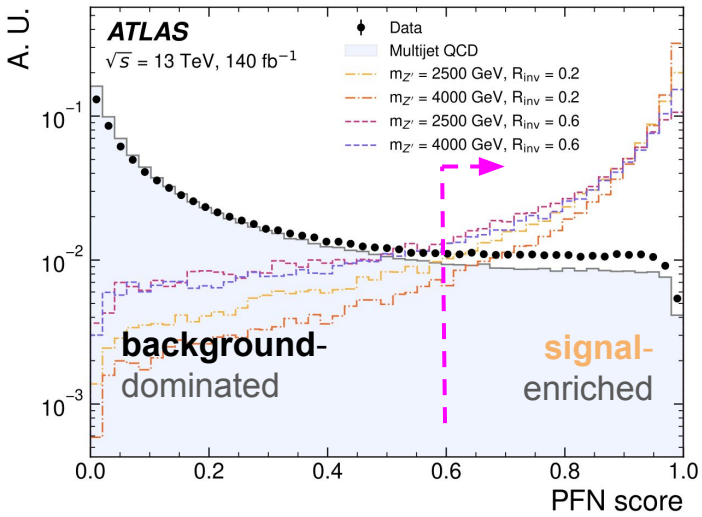
PFN (model-dependent)

well-defined separation for **specific signals**

→ use supervised PFN

for **model-dependent** strategy

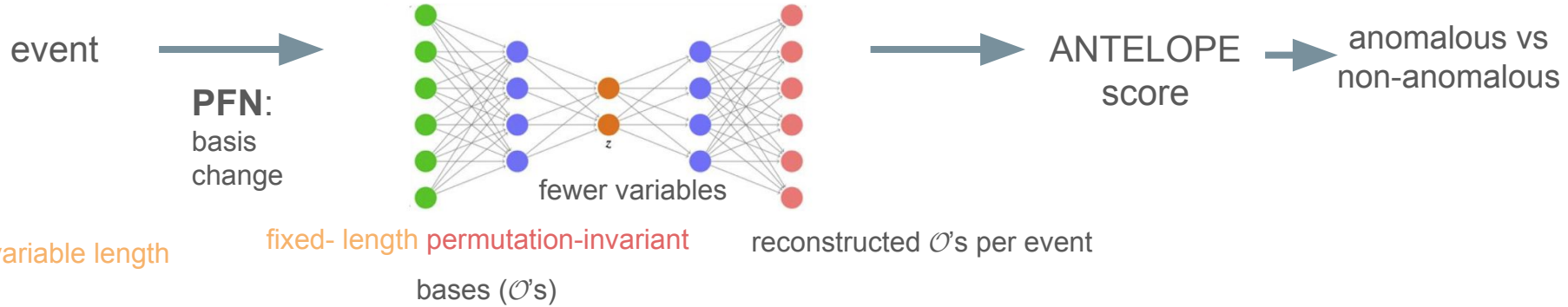
for higher sensitivities for specific signals



* about 200k events each
** 64 to 12
*** mean squared error + KL divergence term

ML Analysis Strategies - ANTELOPE

- ANTELOPE (**An**omaly de**te**ction on partic**l**e fl**o**w latent s**p**ac**e**) (paper)
  - **train on data** events*, but PFN (already trained w/ sim.) used for basis change → **semi-supervised**
  - incorporates variational autoencoder **(VAE)**:
    - autoencoder: identify anomalous events by learning to reduce dimensions** and reconstruct
      - anomalous events → reconstruct worse → **higher** anomaly/ANTELOPE score***
    - variational: reduced dimensions are continuous → handle out-of-distribution anomalies



* a scalar decays into SM quark and dark quark; dark quarks decays into emerging jet

ML Analysis Strategies - ANTELOPE

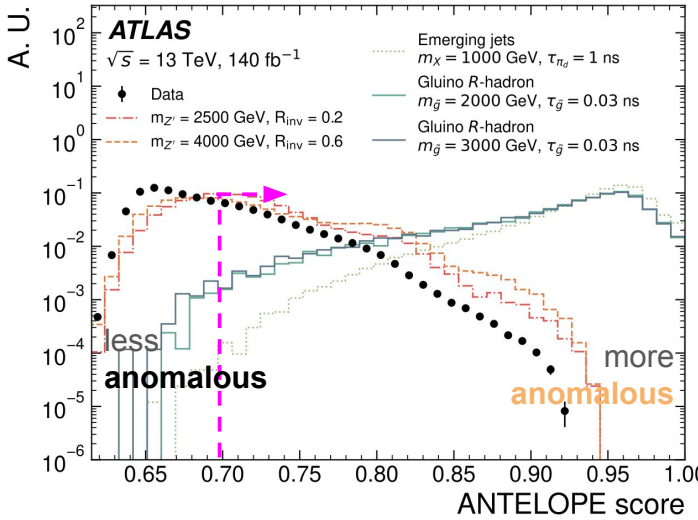
- **SVJ** benchmark signals **more anomalous** than **data** as expected
  - can we see if ANTELOPE can generalize beyond SVJ?

    → compare w/ other BSM signals w/ resonance in m_T
    - Emerging Jets — dark QCD similar to SVJ but displaced*
    - Gluino R-hadron — split-SUSY model w/ long-lived gluinos decaying to R-hadrons

→ anomalous signatures w/ displaced tracks + large E_T^{miss}

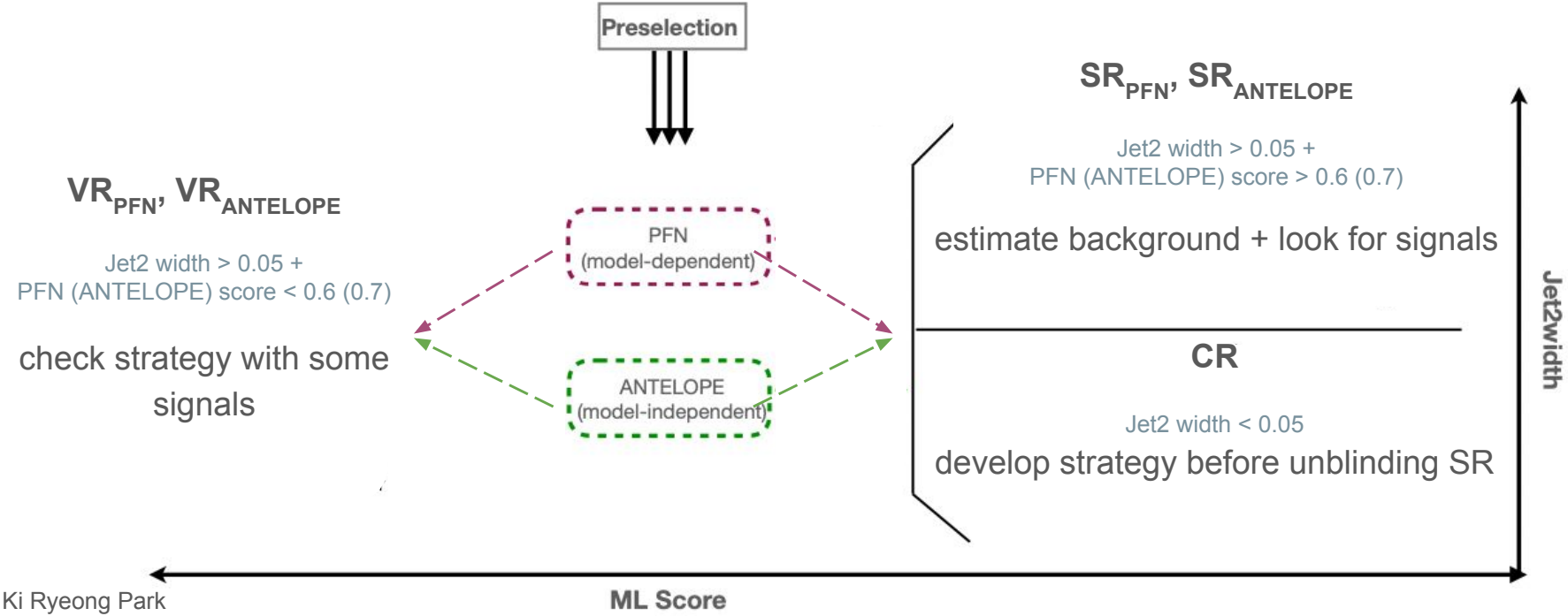
ANTELOPE (model-independent)

well-defined separation for other anomalous signatures
→ semi-supervised ANTELOPE
generalize beyond SVJs
→ can be
used for model-independent strategy



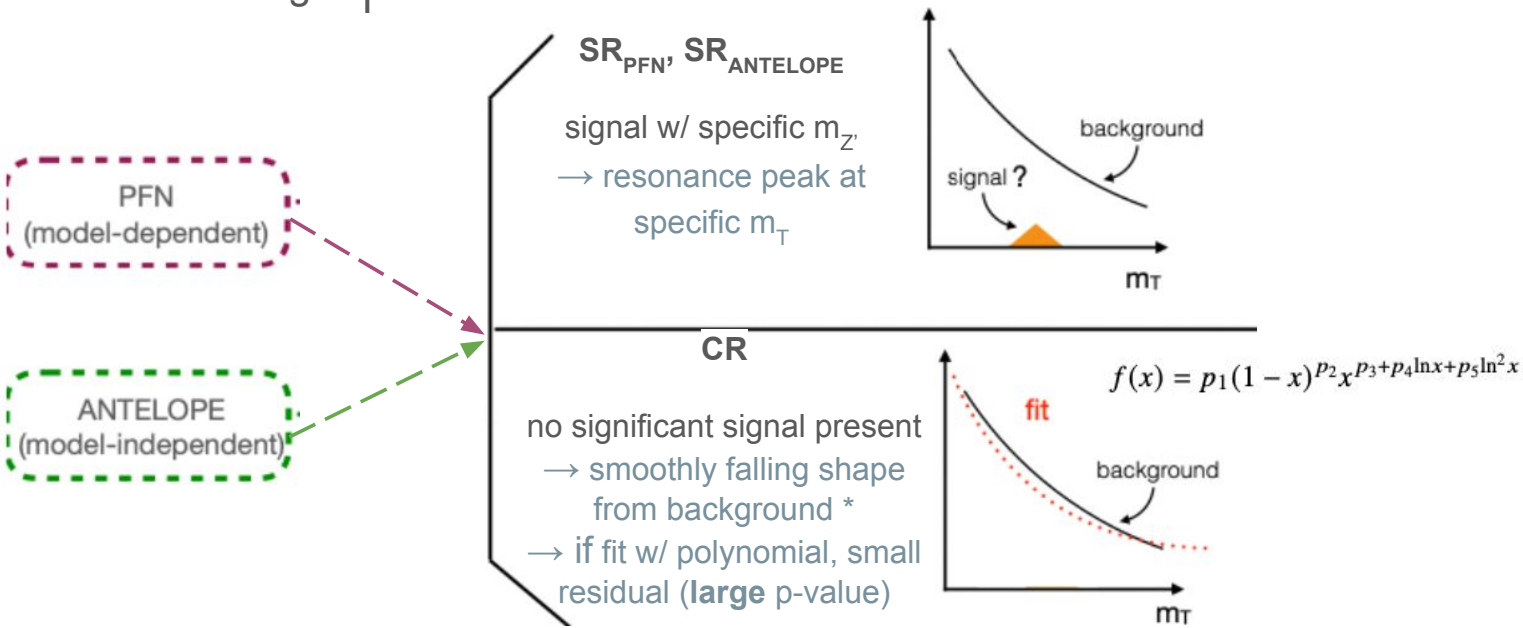
Analysis Strategies

- Both strategies have signal, control, validation region determined by
  - PFN/ ANTELOPE score
  - Jet2 Width (calorimeter width)*



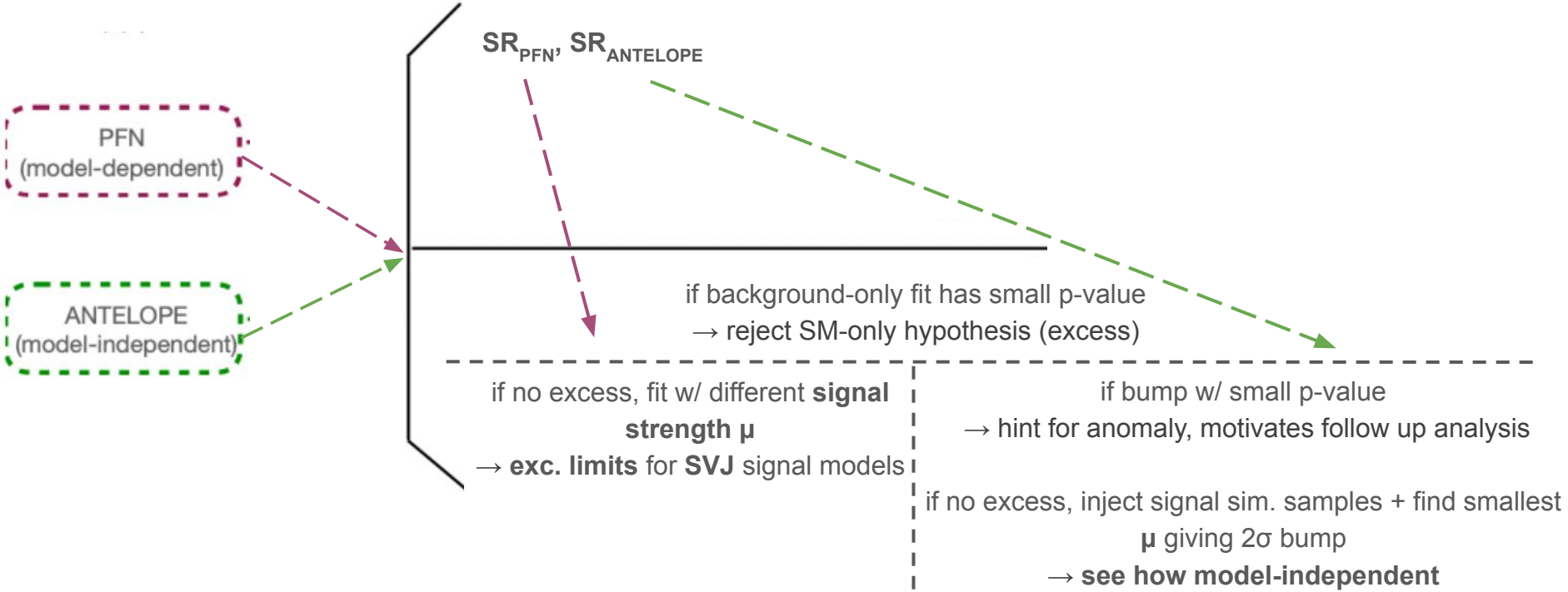
Analysis Strategies

- Developed background estimation strategy + estimate background + look for signal by fitting **m_T distribution**



Analysis Strategies

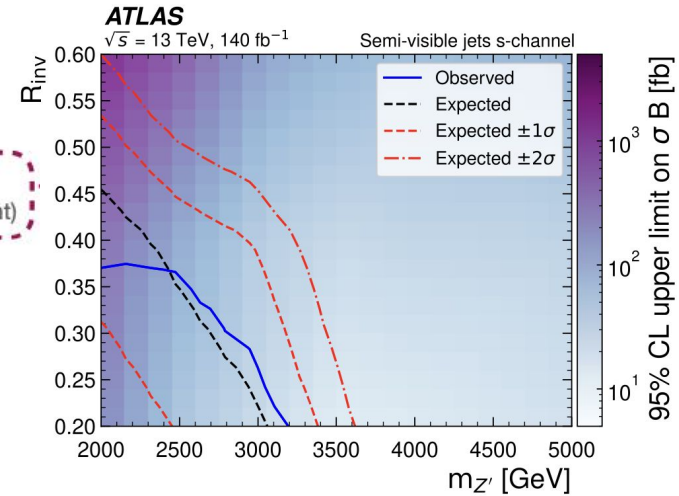
- To identify if there exist a resonance peak in **SR,** we can check



Results

- w/ syst. uncertainties (considered for PFN only) where signal theory syst. dominant (~20%)
  - **no excess** beyond the SM found (0.26 and 0.74 p-values)

PFN (model-dependent)

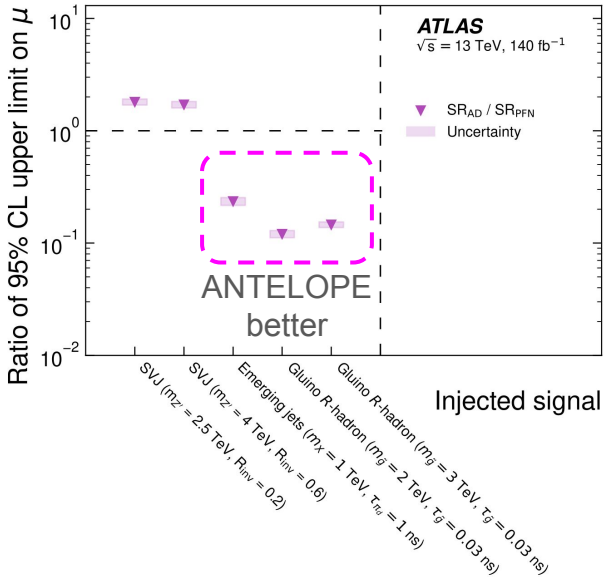


excluded $m_{Z'}$ = 2 - 3.2 TeV w/ varying R_{inv} = 0.2 - 0.37

→ **higher** R_{inv} **less excluded**

→ more invisible: innately harder signature

ANTELOPE (model-independent)



higher (lower) **ratio** of upper limits of $\mu_{ANTELOPE}$ /μ_{PFN} for SVJ (other BSM)

→ PFN better for **SVJ signals**

ANTELOPE **expands the reach** of the PFN region

Conclusion

- No excess beyond the SM prediction after unblinding
  - **excluded** $m_{Z'}$ = 2000-3200 GeV w/ varying Rinv = 0.2 - 0.37 (PFN)
- ANTELOPE **broadens the reach** of the analysis via anomaly detection
  - **novel** approach to use an **ensemble** of ML tools
    - **future ideas**: compare to GNN/transformer for feature extraction + variational autoencoder for anomaly detection, multiclass classification for feature extraction incl. other SM background processes, etc

→ **interest to** analyses in need of **custom** tools + **BSM** search community

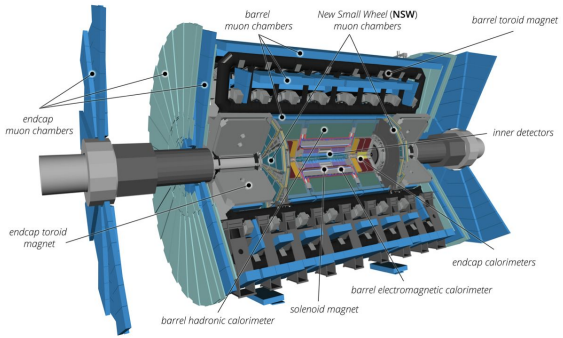
More details on:

- backup slides
- ANTELOPE paper: Semi-supervised permutation invariant particle-level anomaly detection
- ATLAS s-channel SVJ paper: Search for new physics in final states with semi-visible jets or anomalous signatures using the ATLAS detector

Backup

Content

- What is SVJ?
- Why is this search interesting?
- Analysis Samples
- ML Analysis Strategies
- Analysis Strategies
- Results
- Conclusion



- Backup
  - Content
  - Signal Benchmark
  - Selections
    - Object Selections
    - Trigger-level Selections
    - Event-level Selections
  - Analysis Variables
  - Machine Learning Tools - Comparison
  - Machine Learning Tool - Detail
  - Analysis Strategies
  - Systematics
  - Statistical Treatment
  - Limits

* e.g. snow mass (p54) says N_{fD}=4 but chose 2 for simplicity -> minimal impact b/c R_{inv} is directly controlled as an input for simulation

Signal Benchmarks

- Dark QCD parameters
  - motivated by Snowmass 2021 proposed benchmarks (paper) + recommendations from some theorists*
  - m_{qD} =10 GeV, $\rho_D \rightarrow \pi_D\pi_D$, running α_D
  - mass hierarchy set to reduce # of parameters to scan and inspired by SM QCD
    - $m_{\rho D} > \Lambda_D$ similar to SM
    - $m_{\rho D} > m_{\pi D}$ dark ρ decays rapidly into π before decay to SM hadrons (similar to SM)
    - $m_{\pi D} < 2m_{qD}$ dark π cannot decay into two dark quarks -> can be shown in detector
    - $\Lambda_D \sim m_{qD}$ ensures bound states to be formed in dark QCD
  - signal grid:
    - $m_{Z'}$ = [1.0, 1.25, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0]
    - R_{inv} = [0.2, 0.4, 0.6]

Parameter	N_{cD}	N_{fD}	Λ_D [GeV]	m_{π_D} [GeV]	m_{ρ_D} [GeV]	m_χ [GeV]	g_q	g_χ
Value	3	2	10	17	31.77	10	1	0.1

Object-level Selections

To be sensitive to **SVJ signature and topology**,

- use R=0.4 jet — small RJ
  - use track+ calorimeter clusters info (ParticleFlow) — suppress pileup + avoid double-counting
- track
  - ghost-associated tracks — ML tool input
    - reconstruct jets w/ track pT set as 1 GeV — jet axis or energy change miniscule
    - instead of geometric matching (ΔR) — cluster info more inclusive for R_{inv} >0
  - pT > 1 GeV — suppress pileup + fake tracks
  - |η| < 2.1 — within inner detector (< 2.5) with small RJ radius
- leptons
  - electrons + muons — overlap removal for jets
- E_T^{miss}
  - calculated using jets + electrons + muons — suppress background (exp. falling QCD)

Trigger-level Selections

- **Lowest pT, unprescaled** single-jet trigger:
  - 2015: HLT_j360 (pT >=360 GeV)
  - 2016-17: HLT_j380 (pT >=380 GeV)
  - 2017: HLT_j225_gsc380_boffperf_split (pT >=380 GeV w/ GSC calibration for jets w/ pT > 225 GeV)
  - 2017-18: HLT_j225_gsc420_boffperf_split (pT >=420 GeV w/ GSC calibration for jets w/ pT > 225 GeV)

*non-collision background: calorimeter noise, cosmic lays, etc -> tend to have low pT + small Φ (close to beamline)

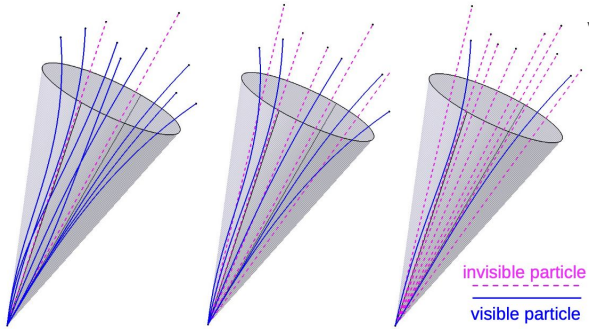
Event-level Selections

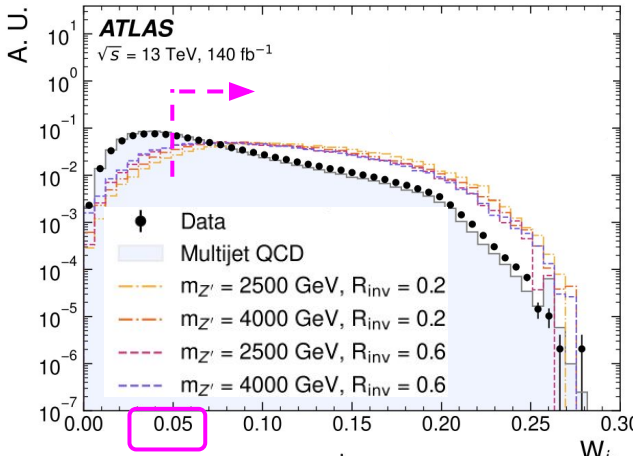
- Preselections
  - trigger
    - leading jet (j1) pT > 450 GeV — trigger efficiency plateau
  - quality cuts
    - leading, subleading jet |η| < 2.1 — within inner detector (< 2.5) with small RJ radius
    - m_T > 1.5 TeV — smoothly falling m_T for fitting
    - 3+ tracks for j1 and jet2 (each) — sufficient information for ML algorithm
  - enhance signal
    - 2+ jets — 2 SVJs
    - 0 leptons — no leptons for SVJ
    - $\Delta Y_{j1,j2} < 2.8$ — hard scatter process
  - reduce background
    - subleading jet (j2) pT > 150 GeV — remove non-collision background*
    - ΔΦ(j1, j2) > 0.8 — remove non-collision background
    - E_T^{miss} > 200 GeV — remove QCD background

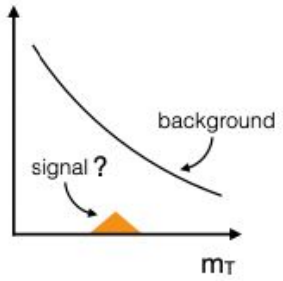
*$\Delta R_{i,\,jet}$ = distance between cluster + jet axis
**jet with second highest pT in the event

Analysis Variables

- Notable variables: R_{inv}, jet2* width, m_T







$$R_{inv} = \frac{\text{number of dark hadrons}}{\text{number of all hadrons}}$$

R_{inv} **higher** for more **invisible** SVJ

search grid for different signals

$$W_j = \frac{\Sigma_i p_T^i \times \Delta R_{i,\text{jet}}}{\Sigma_i p_T^i}$$

jet2** width **lower** for more **background**
b/c QCD jet is more homogenous

reduce background

$$m_T^2 = \left[E_{T,JJ} + E_T^{miss}\right]^2 - \left[\vec{p}_{T,JJ} + \vec{p}_T^{miss}\right]^2$$
$$= m_{JJ}^2 + 2p_T^{miss}\left[\sqrt{m_{JJ}^2 + p_{T,JJ}^2} - p_{T,JJ}\cos(\phi_{JJ,miss})\right]$$

m_T (inv. mass of 2 jets and E_T^{miss}) **larger** for signals with **higher** $m_{Z'}$

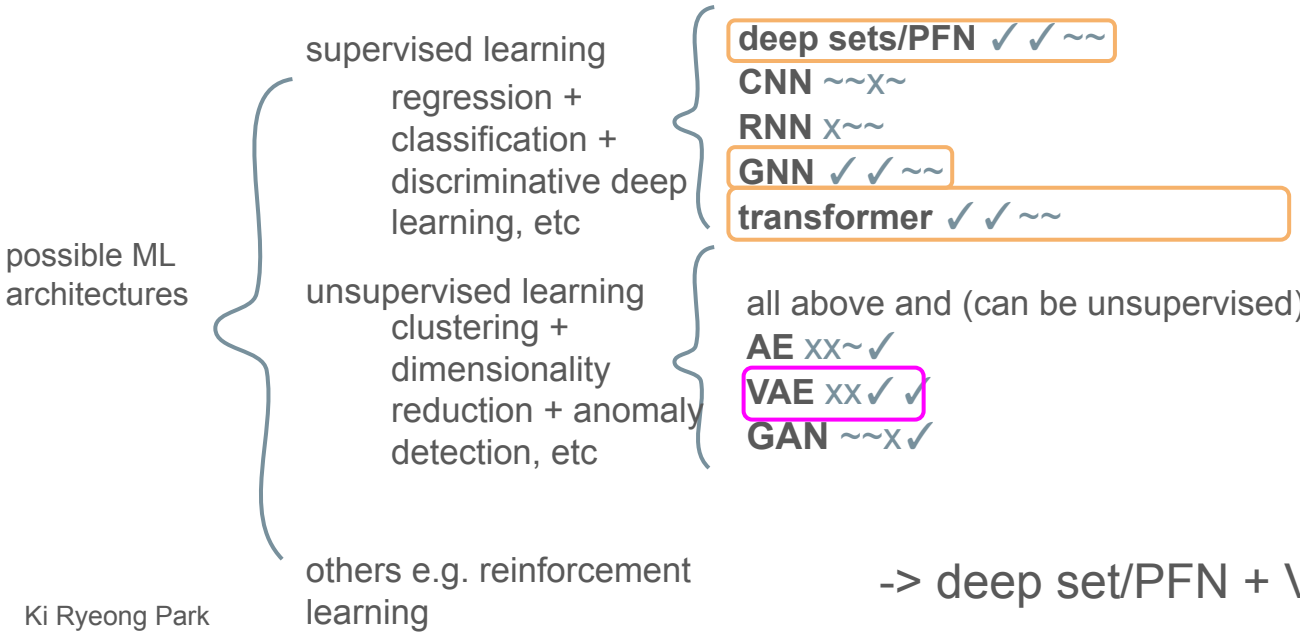
search for resonance of a signal w/ specific $m_{Z'}$

Machine Learning Architectures - Brief Comparison

- permutation-invariant + variable length

  + minimal reliance on specific training samples + handle out-of-distribution anomalies

  -> 4 criteria: ✓: generally true, x: generally false, ~: with variations yes



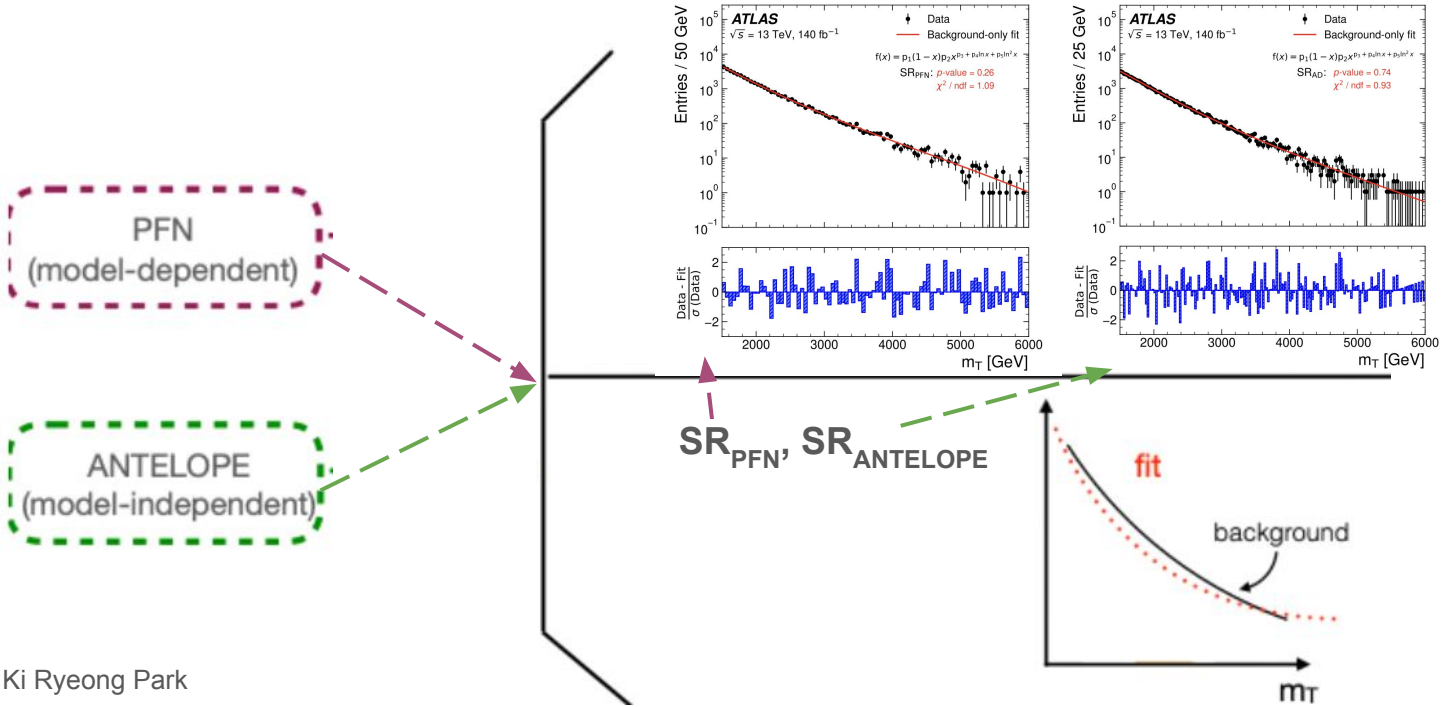
-> deep set/PFN + VAE chosen for SVJ

Machine Learning Architecture - Details

Parameter	PFN	VAE
Batch Size	500	32
Number of Epochs	100	50
Activation Function	ReLU	Leaky ReLU
Optimizer	Adam [43]	Adam
Learning Rate	1×10^{-3}	1×10^{-5}
Loss	Categorical Cross Entropy	MSE + D_{KL}

Analysis Strategies

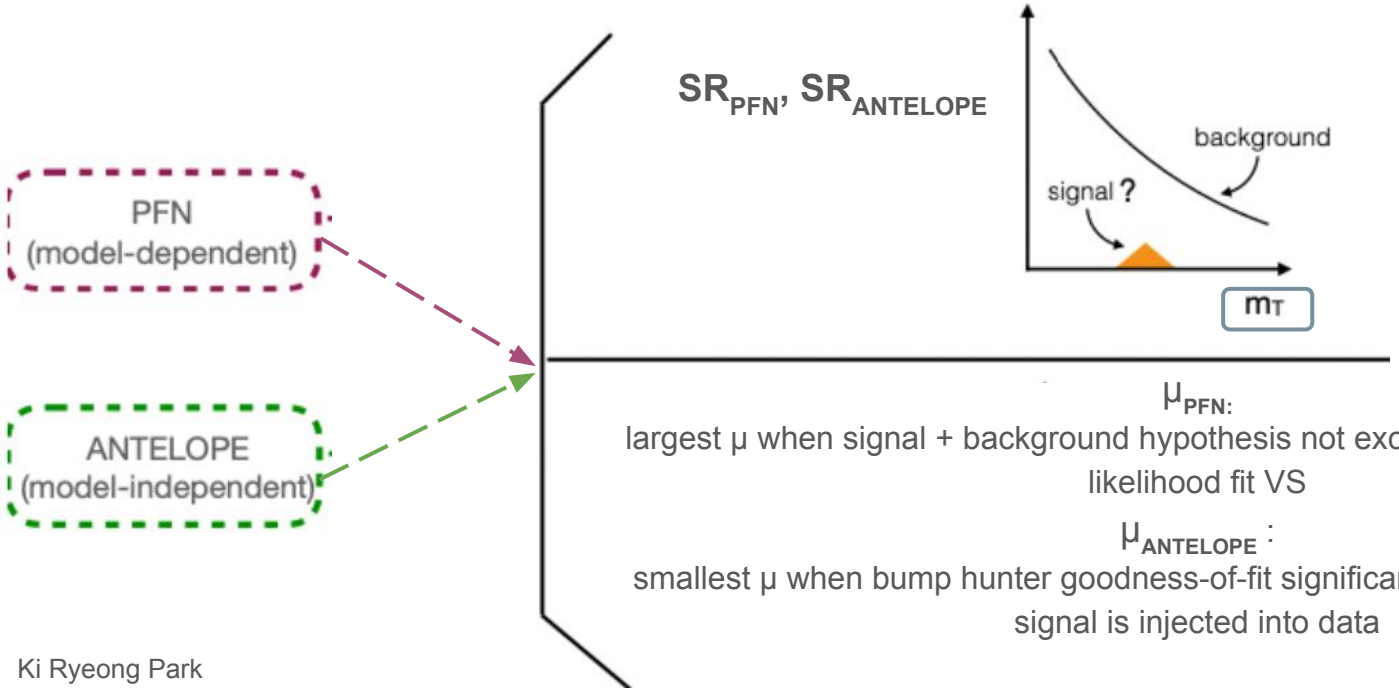
- 2 strategies w/ SRs
  - for $SR_{ANTELOPE}$ bump hunter's largest deviation found is 1.7- 1.9 TeV at p-value of 0.81



background only fit

Analysis Strategies

- 2 strategies w/ SRs



μ_{PFN}:
largest μ when signal + background hypothesis not excluded at 95% CL in profile likelihood fit VS

$\mu_{ANTELOPE}$:
smallest μ when bump hunter goodness-of-fit significance reaches 2σ when MC signal is injected into data

Systematics

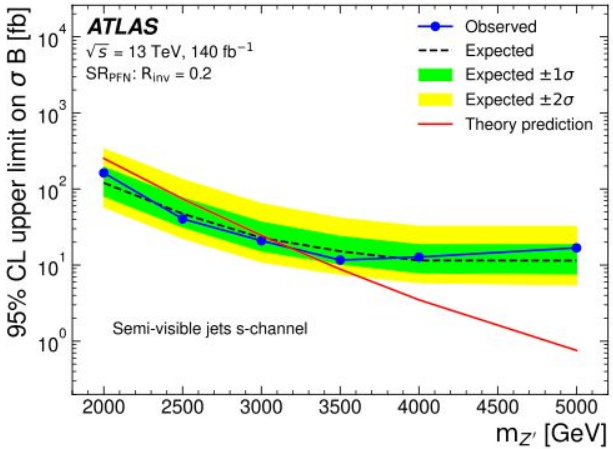
- Background estimates: relies on **mT distribution** of background and signal
  - uncertainties on mT distribution of **background:** → expected to be small and absorbed by the background fit → not considered
  - uncertainties on mT distribution of **signal** for each signal point → incorporated as nuisance parameter for signal + background fit
    - **shape**
      - jet pT scale: small, but still considered
    - **normalization (total # of signal events)**
      - **luminosity**:
        - **0.83%** for each signal points
      - **spurious signal**: if there is no true signal, can fit absorb the statistical fluctuation?
        - $\mathcal{O}$**(1-10)%** dep. on signal point
      - **theory uncertainty**: if signal events were simulated with different parameter settings, how many more/less events? (tech. shape change but conserv. norm)
        - initial-state + final-state radiation -> **20%** for each signal points
        - PDF, α_S -> **20%** for each signal points

Statistical Treatment

- Testing if fit can work even with stat. fluctuations in mT distribution? aka. can it differentiate stat. fluctuations vs actual signal presence?
  - **background-only fit** for sim. CR + VR + SR; for data CR + VR
    - yield matched w/ total yield of SR data from downsampling (not biasing b/c total yield does not show sensitive mT shape info) → statistically identical as SR
    - check if fit succeeds and w/ large p-value → **can fit succeed w/ SR data yield and corresponding stat. fluctuations?**
  - **background-only fit** in pseudo data (a.k.a. many fits)
    - smoothed data from CR + VR data with varying level of stat. fluctuations
    - check if fit succeeds and w/ varying p-value → **can fit be balanced w/ flexible to fit w/ varying amt of stat. fluctuations but not to flexible to absorb actual signals?**
  - **signal-and-background fit** in smoothed data from CR w/ x amount of sim. signals (input) **injected**
    - check if there is **linearity** w/ the injected amount and fit prediction→ **can fit accurately predict how much signals are present?**

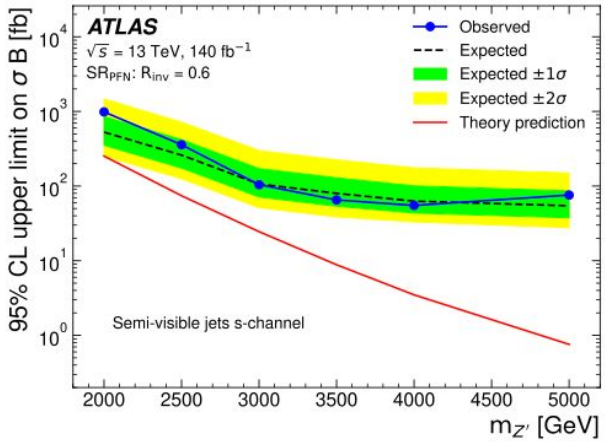
Limits

- Observed, expected limits at 95% CL
  - excluded $m_{Z'}$ = 2000-3200 GeV w/ varying R_{inv} = 0.2 - 0.37



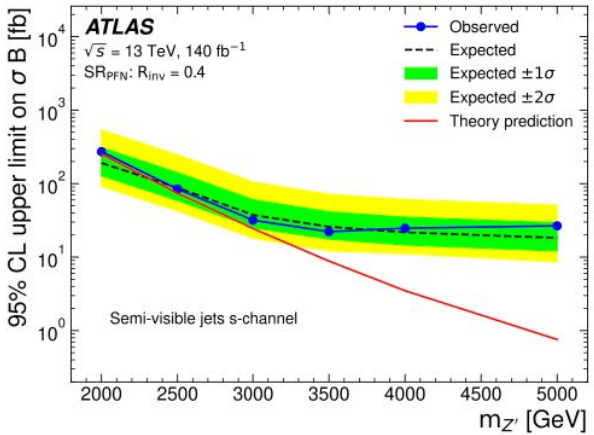
trigger limited

R_{inv} = 0.2



lower track efficiency for boosted regime

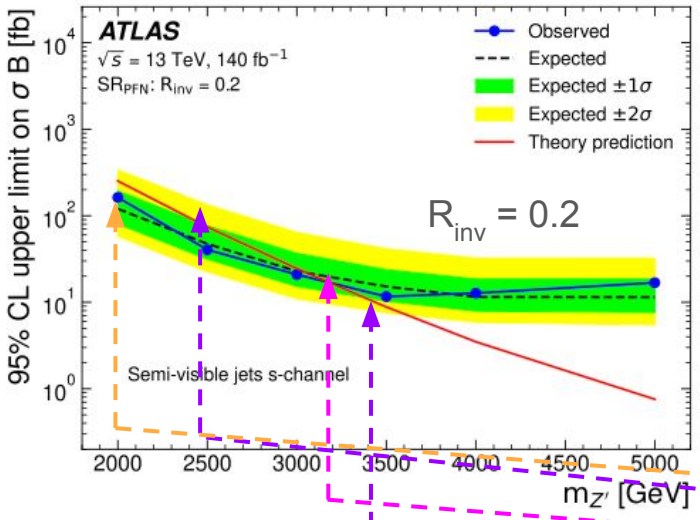
R_{inv} = 0.4



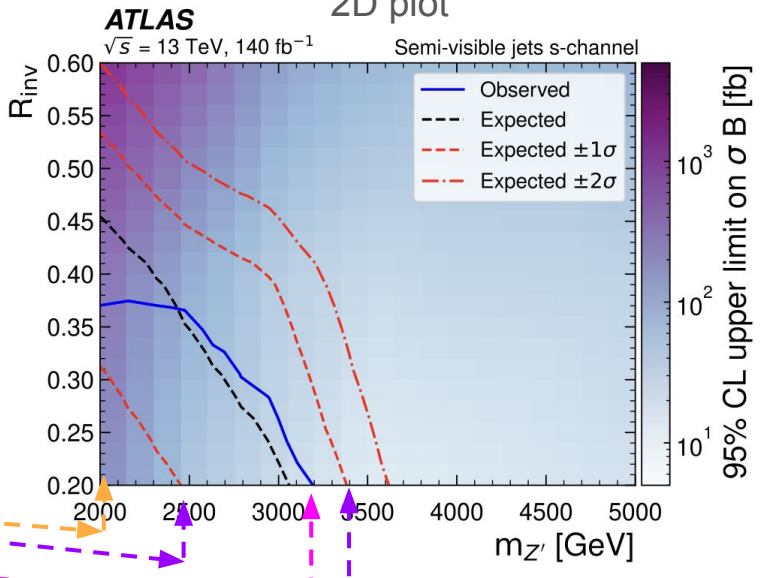
R_{inv} = 0.6

Limits

1D plot for R_{inv}=0.2



2D plot



σ x BR of observed

expected +1σ and theory intersect

observed and theory intersect

expected -1σ and theory intersect

excluded when observed < theory

- **colored** background: σ x BR of observed (interpolated between the observed in 1D)
- observed in 2D: where observed in 1D < theory in 1D
- expected in 2D: where expected in 1D < theory in 1D

e.g. observed- what we exclude e.g. 3200 GeV upto 0.2 is excluded
expected - what could have been excluded (sensitivity) e.g. 3200 GeV