

Learning the Propagation of Cosmic Rays from Dark Matter with Neural Networks

Lena Rathmann

In collaboration with Jan Heisig, Felix Kahlhoefer,
Adrian Mulas

TeVPA 2026

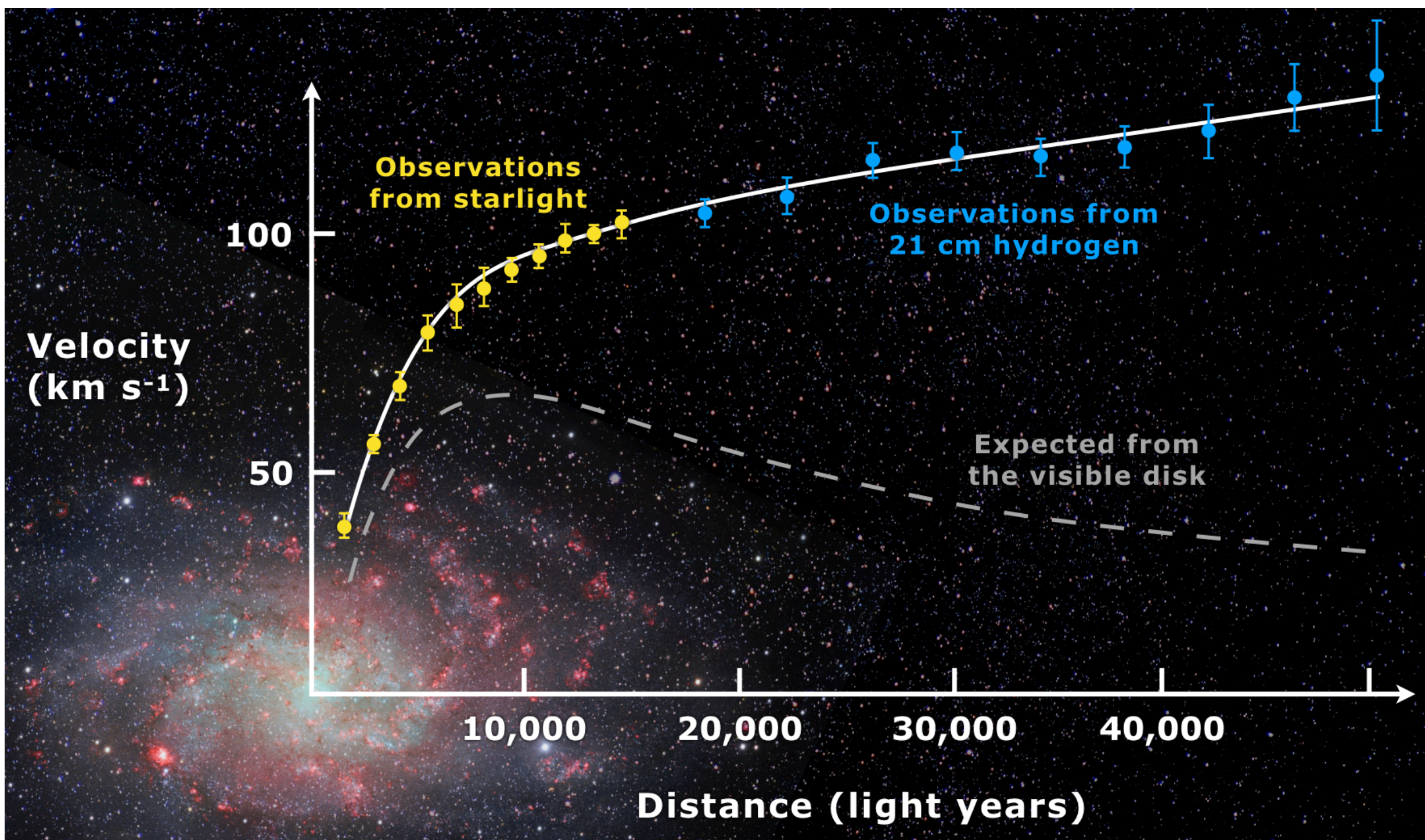
03.09.2026

Evidence for Dark Matter

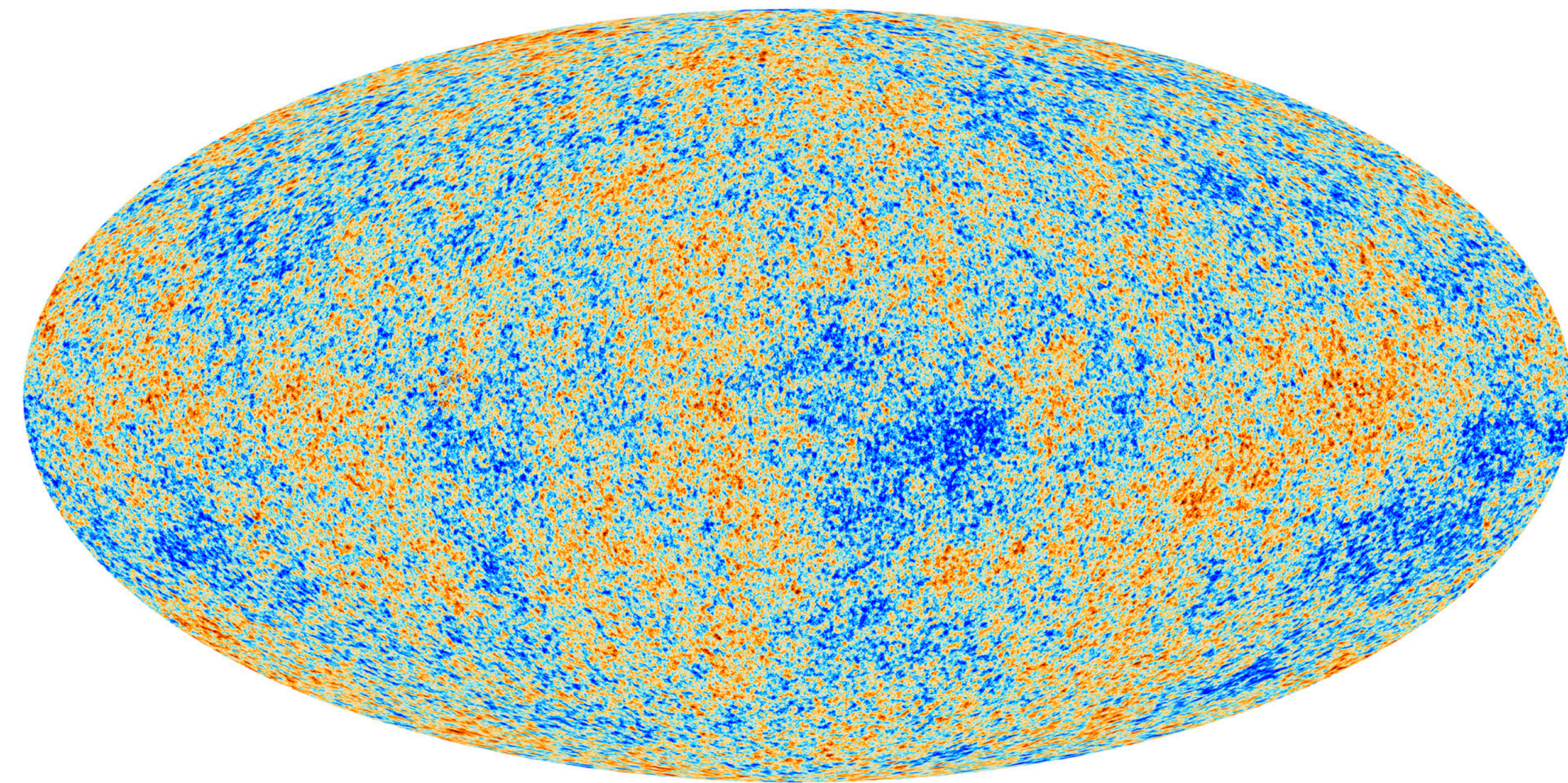


Galaxy rotation curves

CMB anisotropies

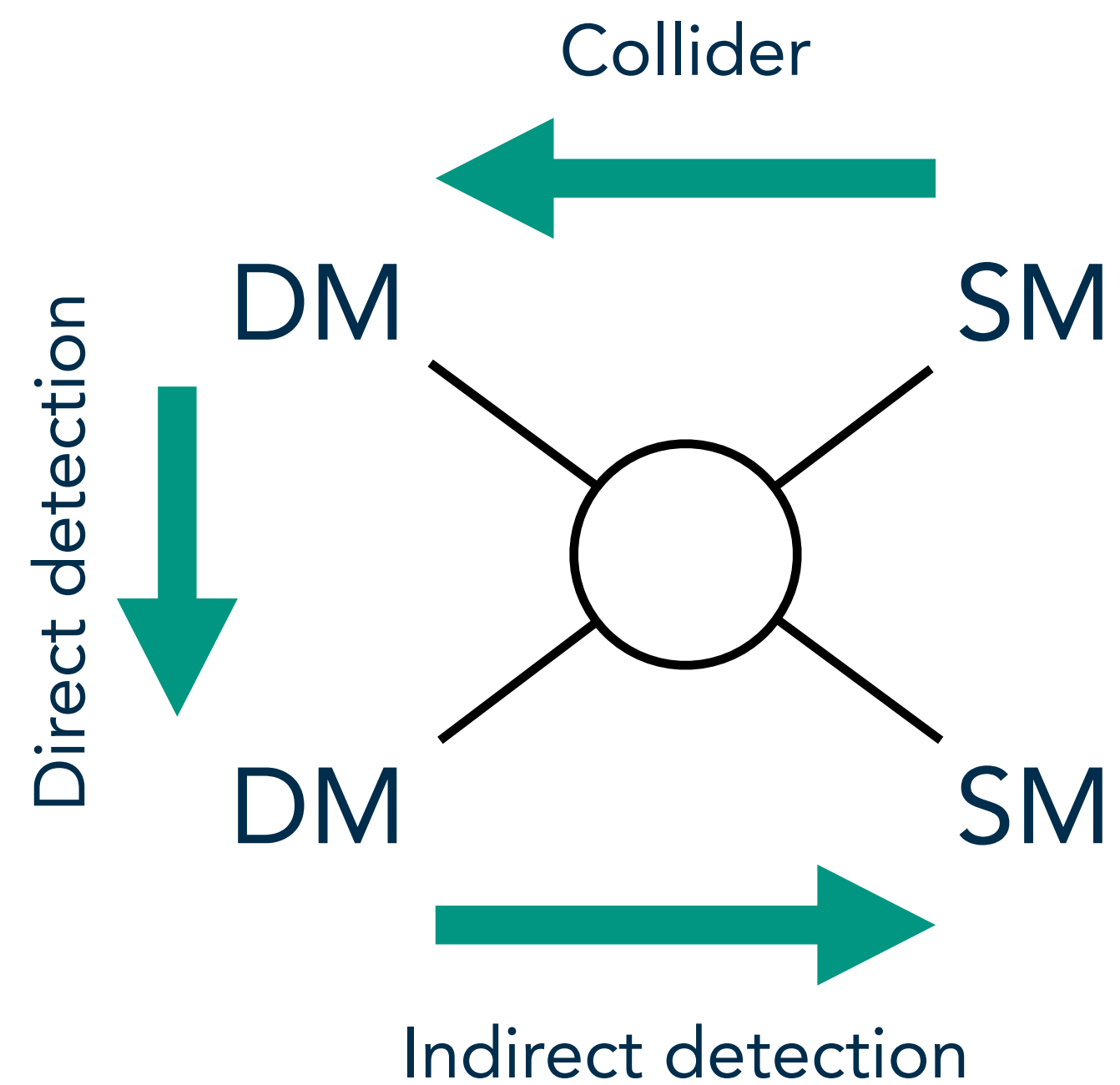


[Wikipedia]

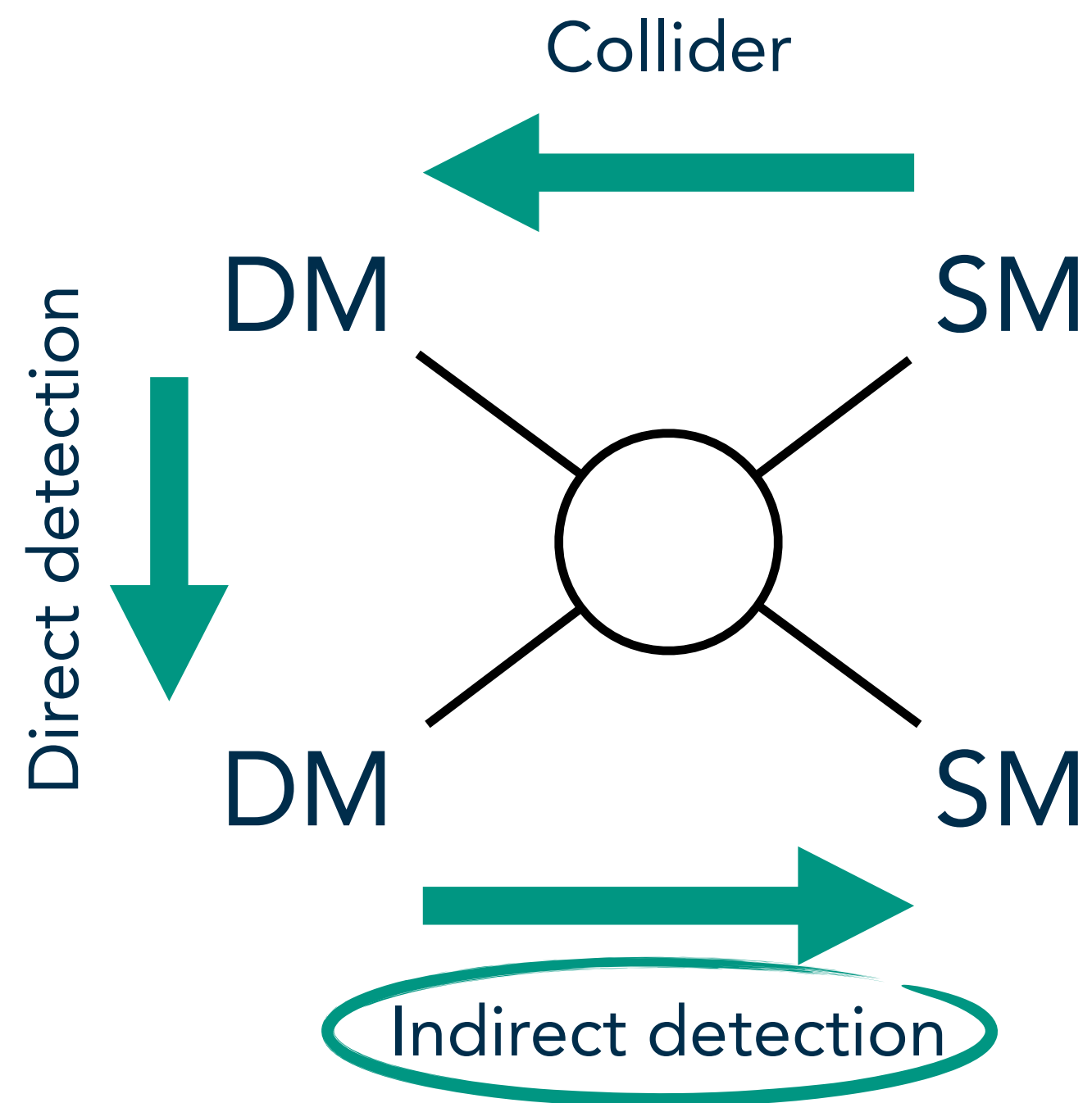


[ESA]

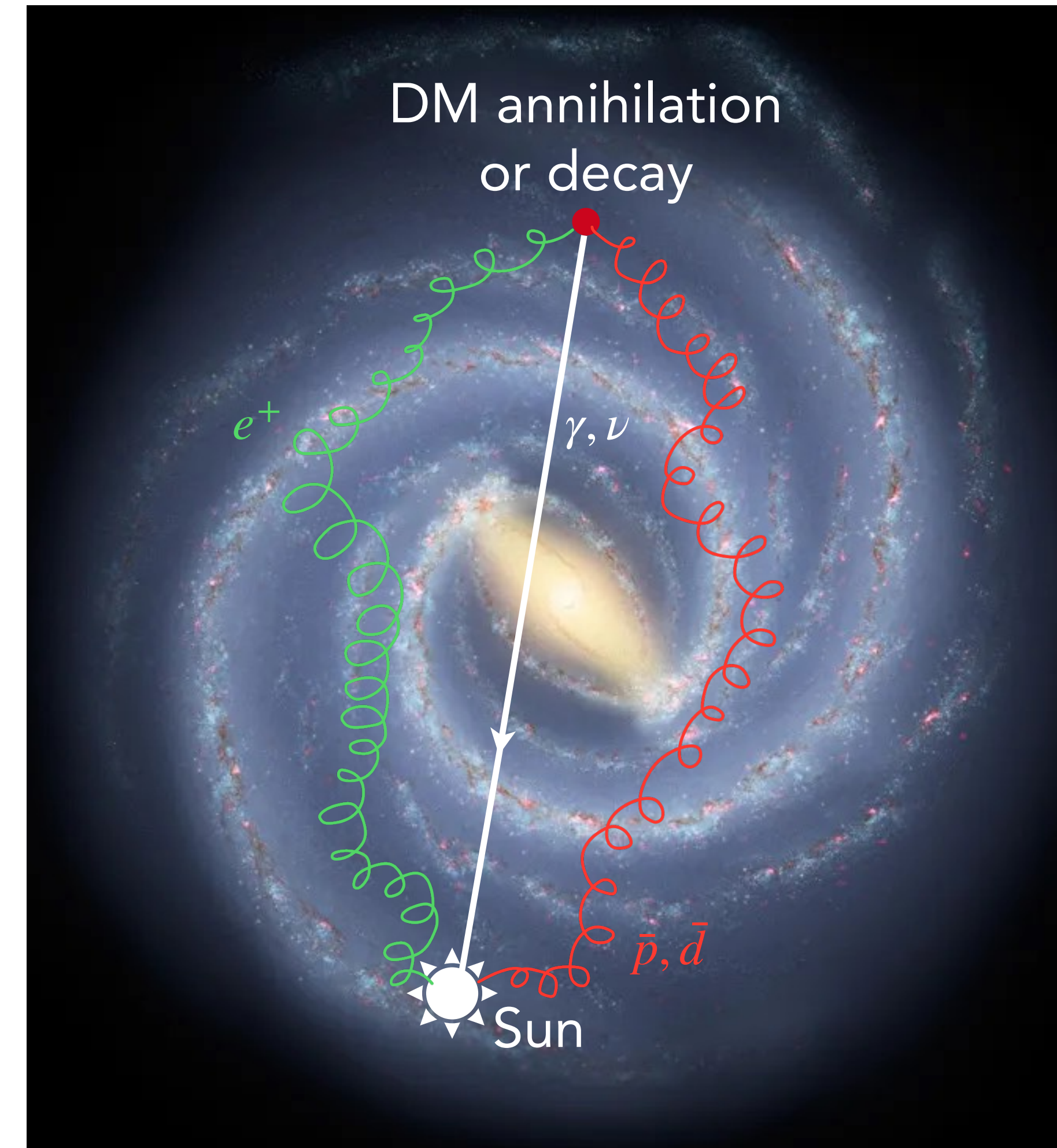
Search for Dark Matter



Search for Dark Matter



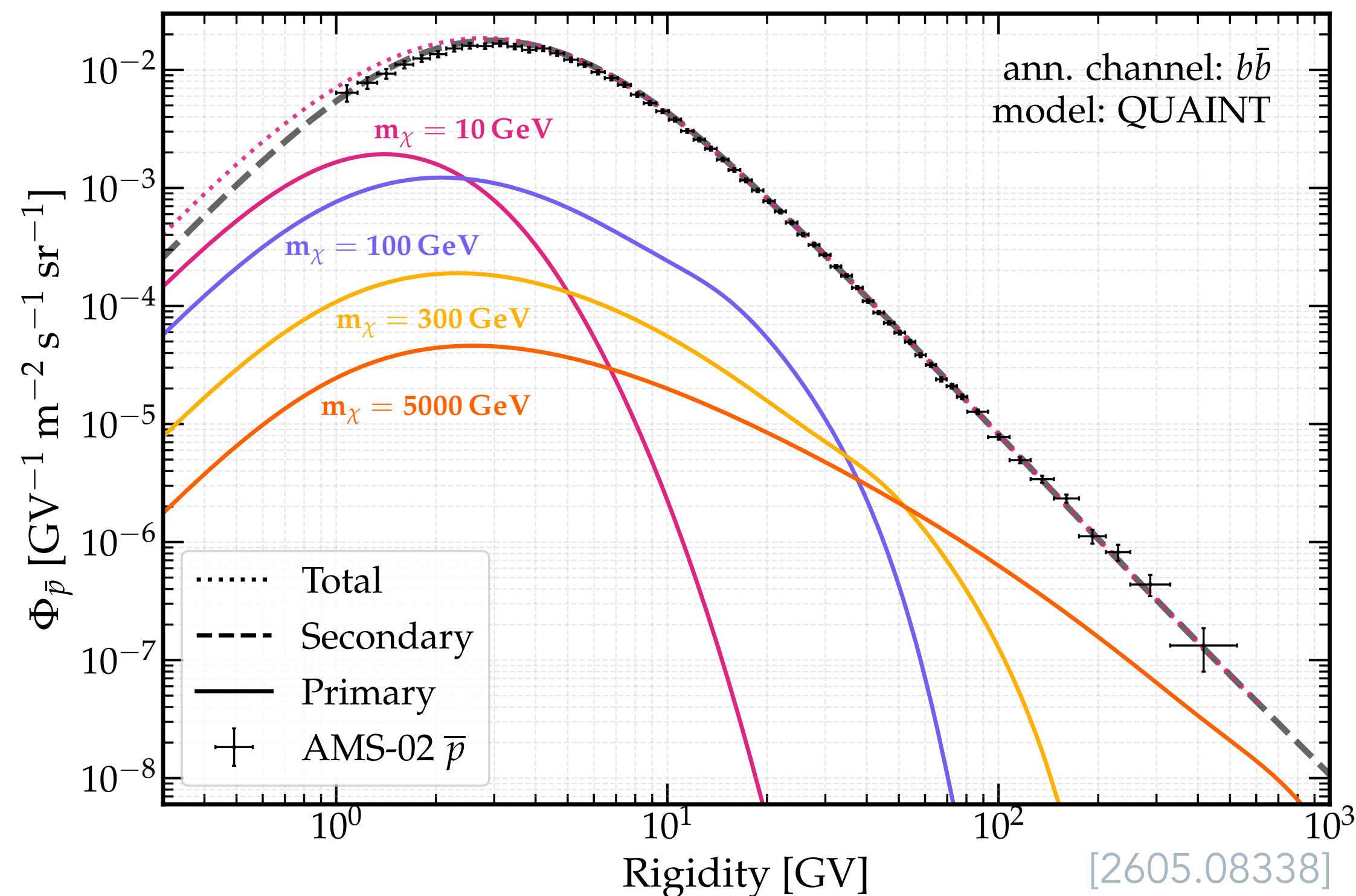
- DM annihilates or decays into SM particles in galaxy
- **Detect resulting cosmic rays** on Earth with AMS-02, Fermi LAT, H.E.S.S., ...
- Look for excess over background
- **Promising targets:** antinuclei



Antinuclei as a Promising Signature

Antiprotons

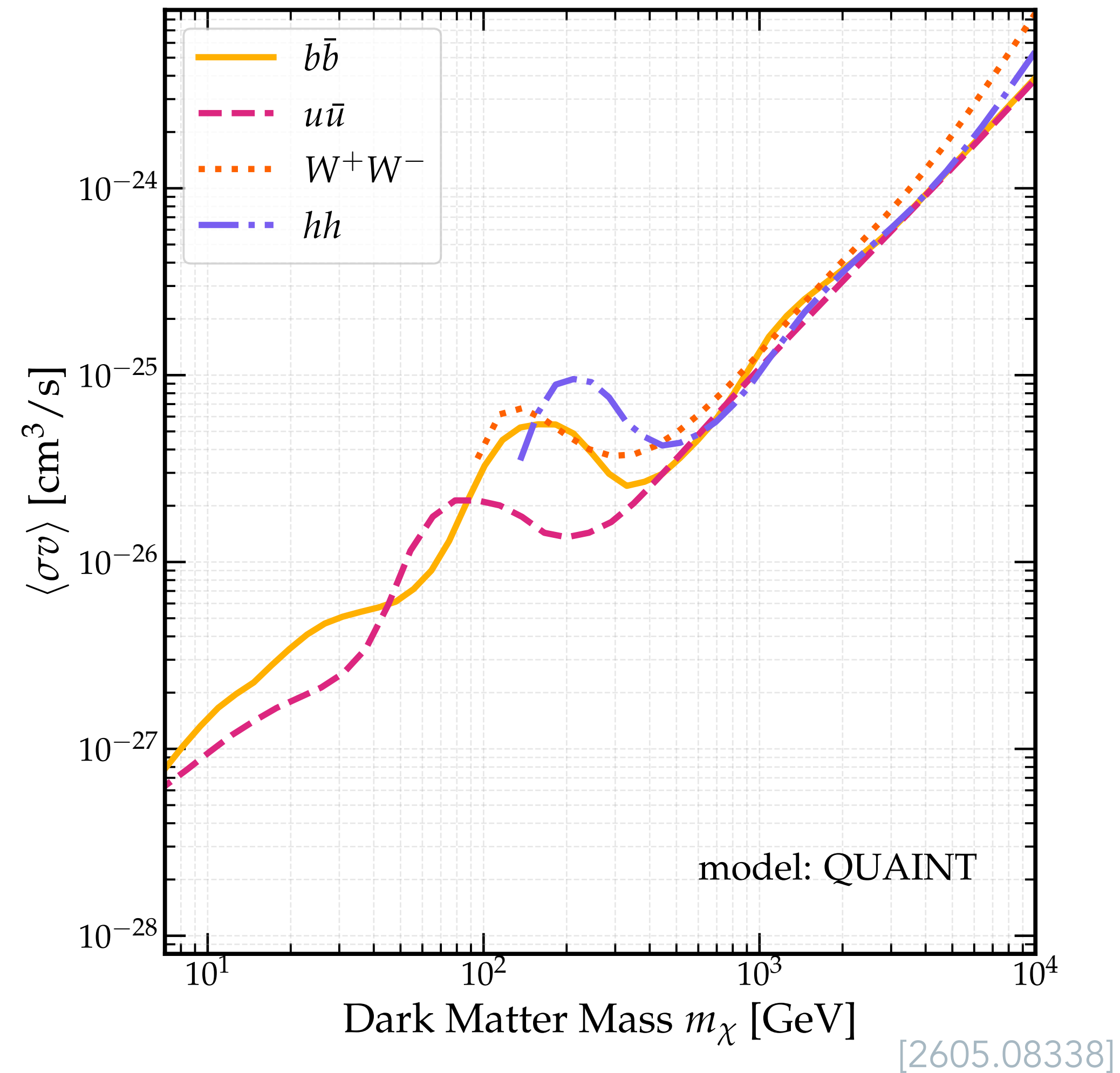
- Predicted antiproton flux can be constrained with **AMS-02 measurements**



Antideuteron

Antinuclei as a Promising Signature

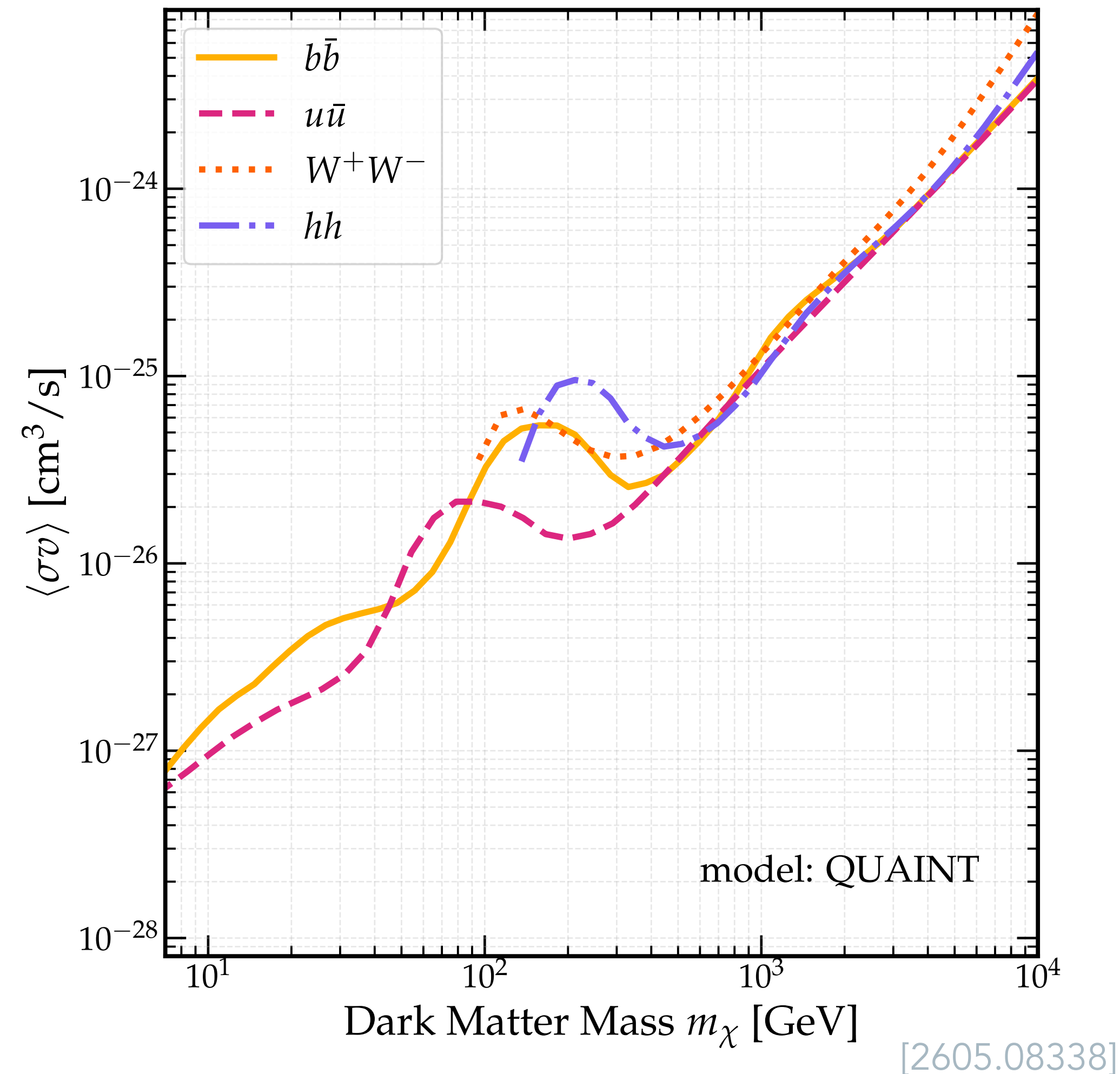
Antiprotons



Antideuteron

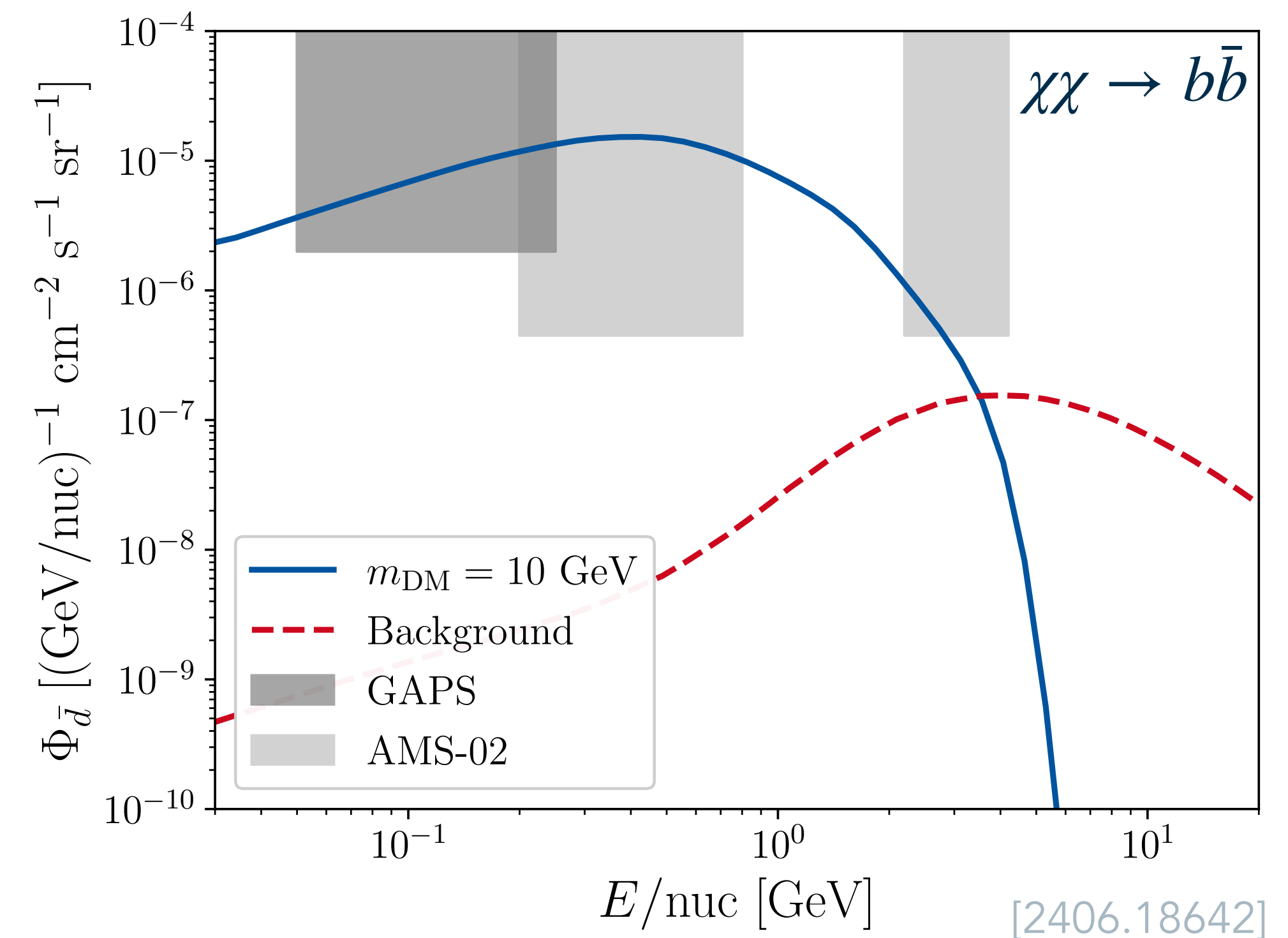
Antinuclei as a Promising Signature

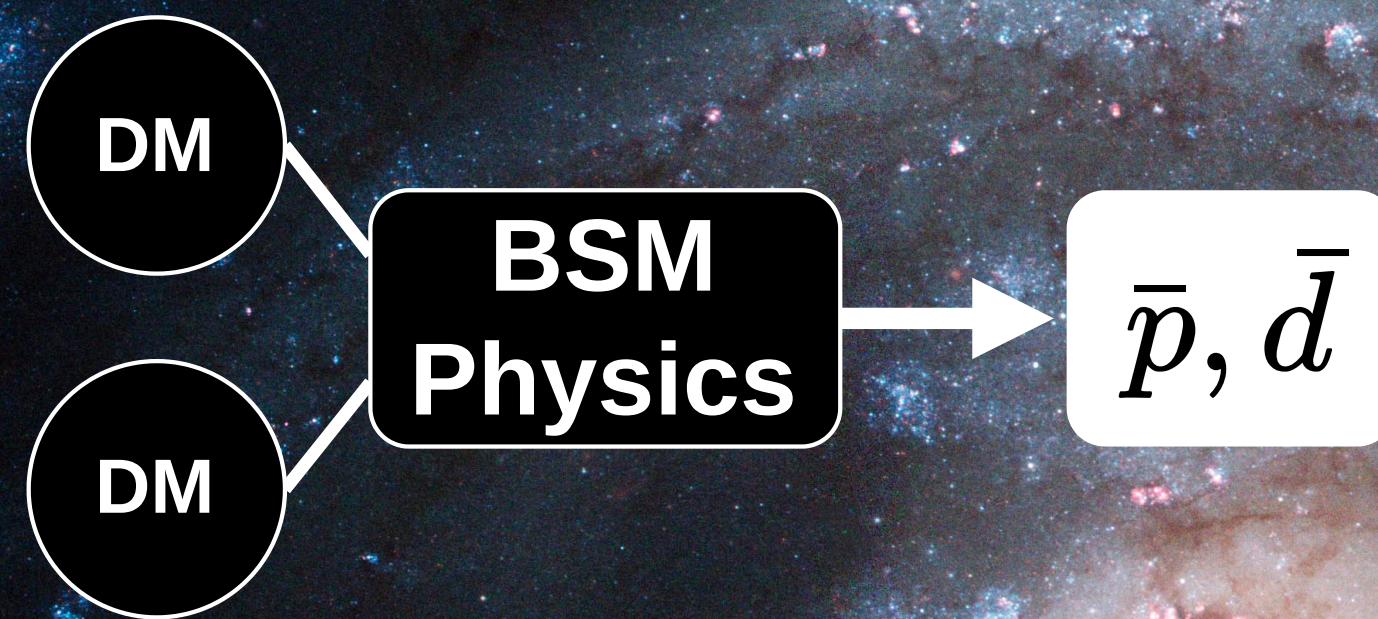
Antiprotons



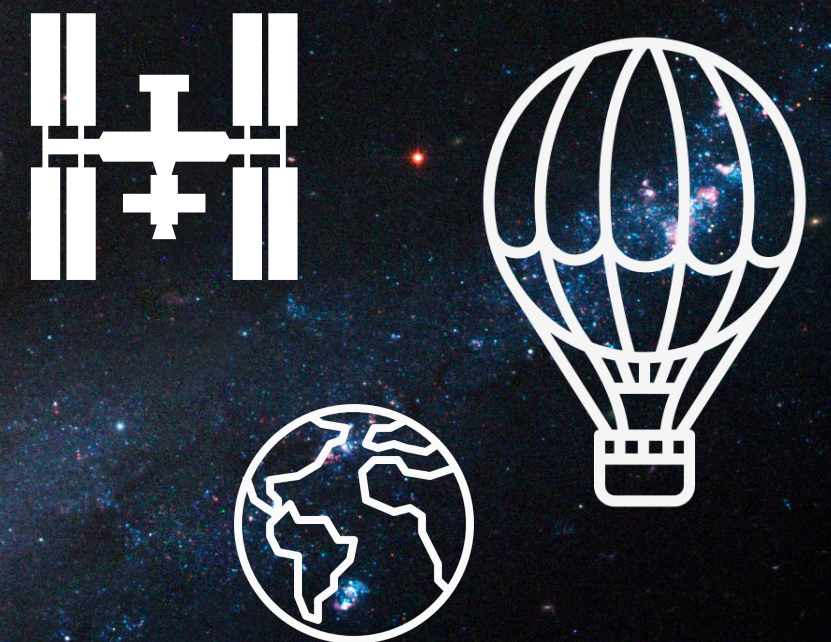
Antideuteron

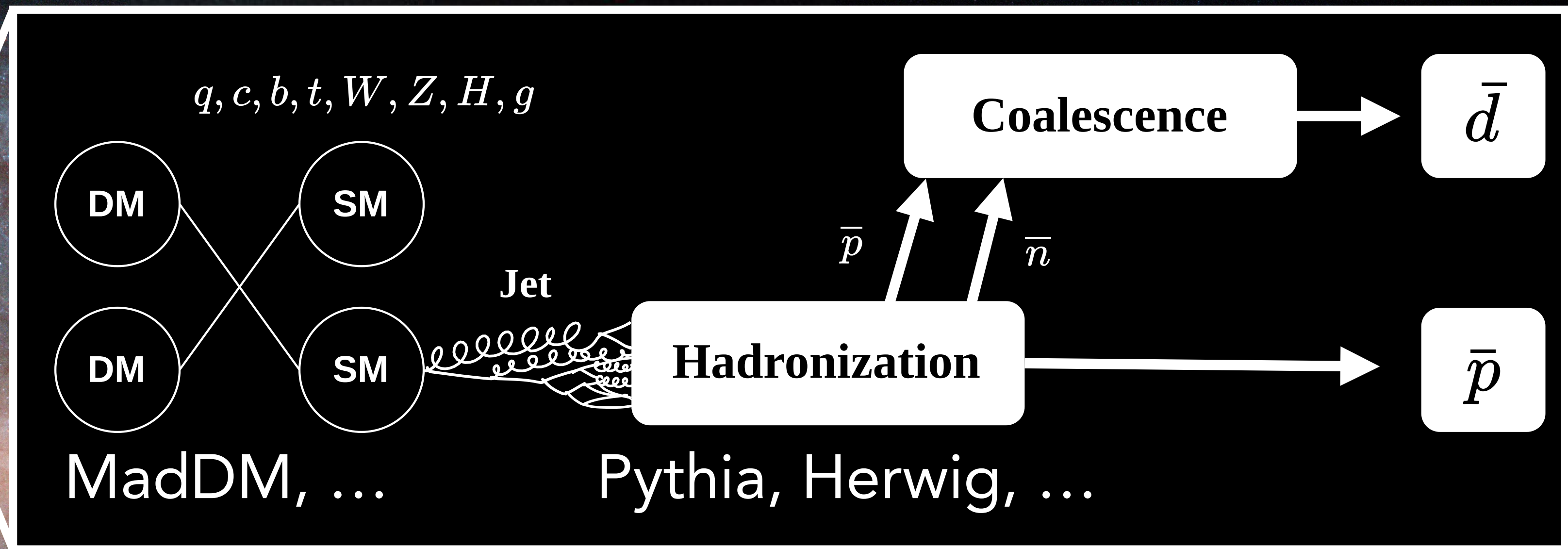
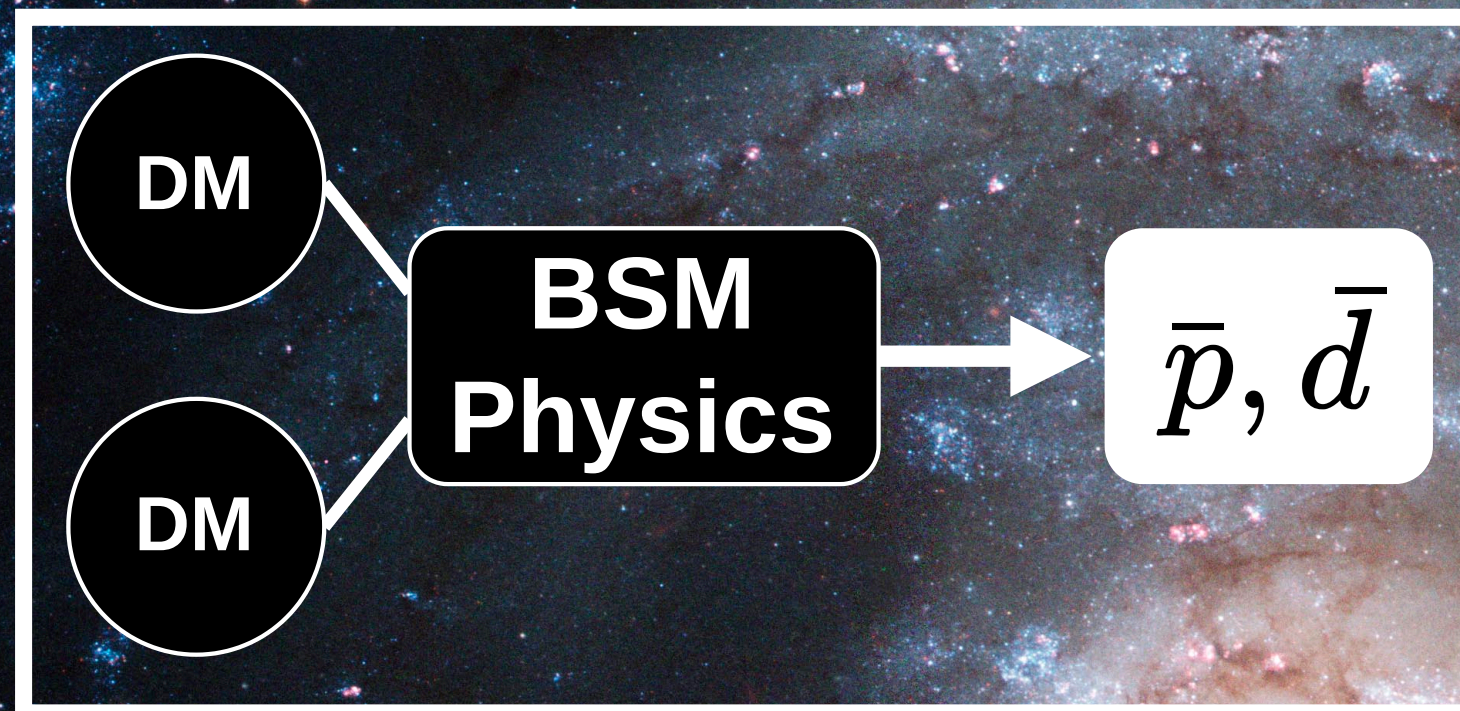
- Flux peaks at low energies where secondary flux is strongly suppressed



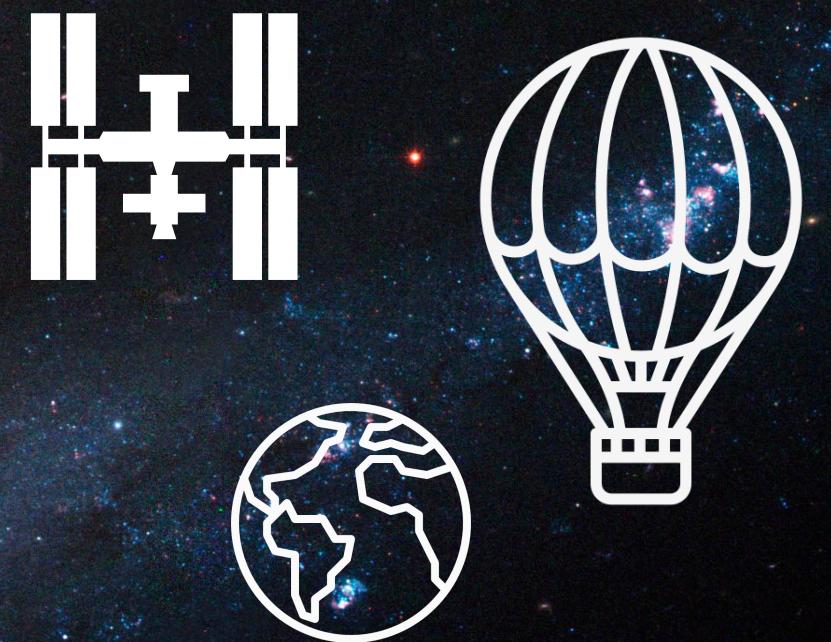


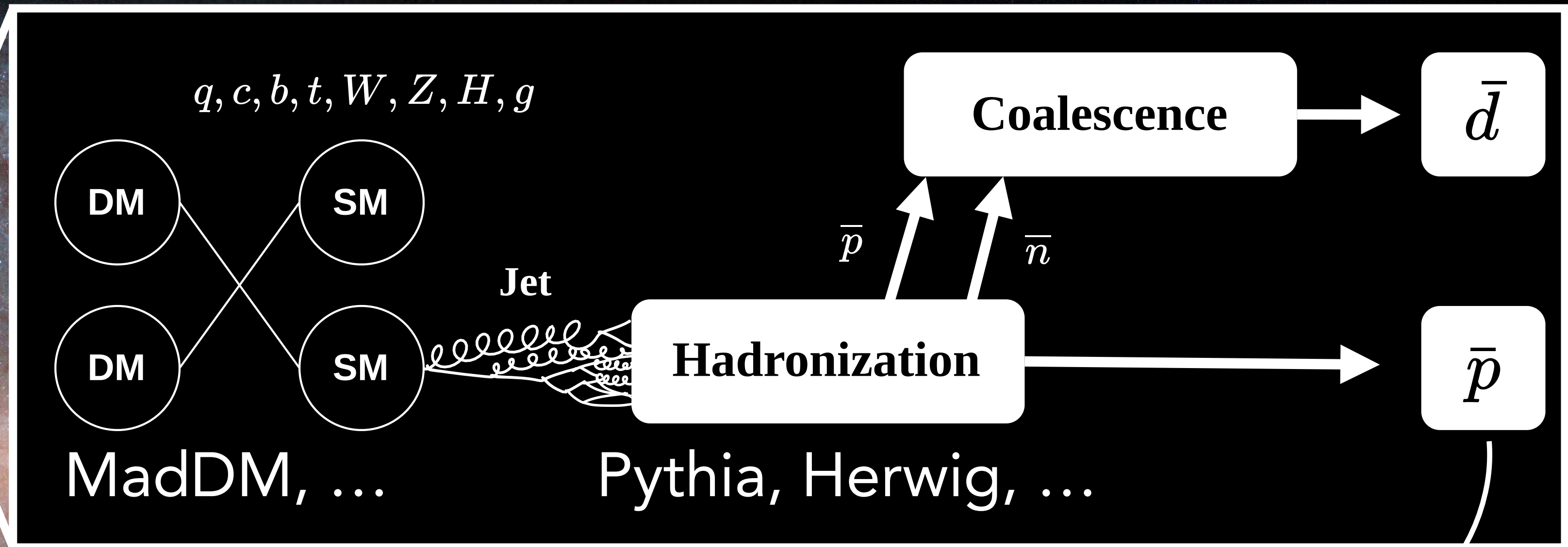
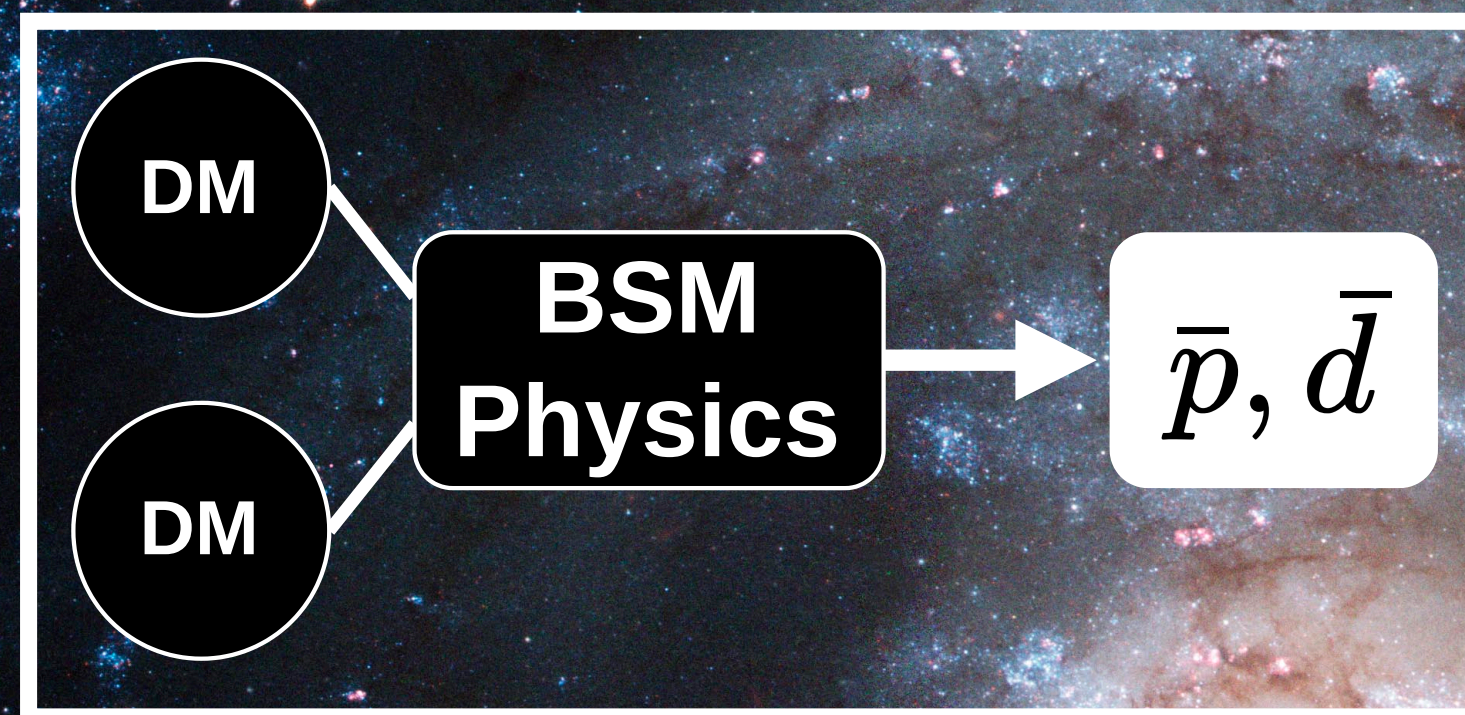
How do we simulate the production and propagation of antinuclei?



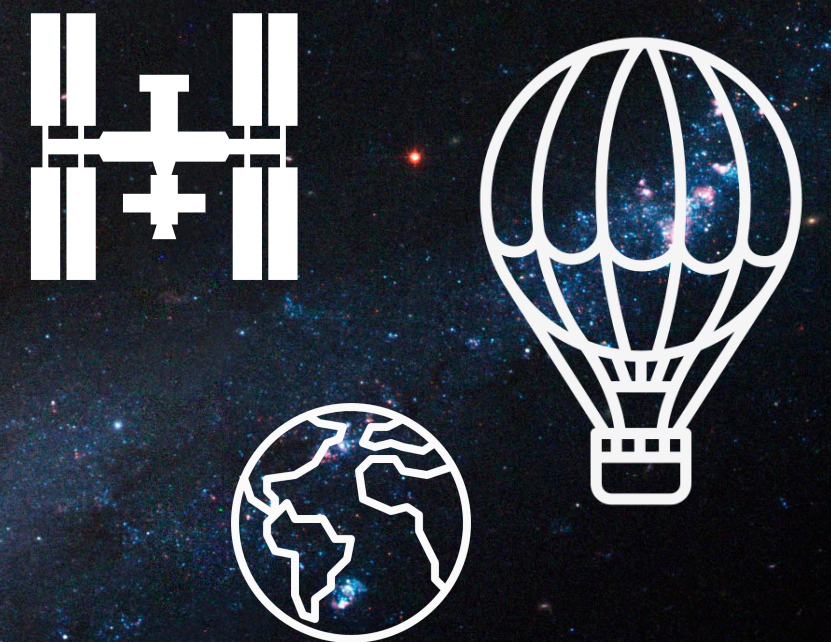
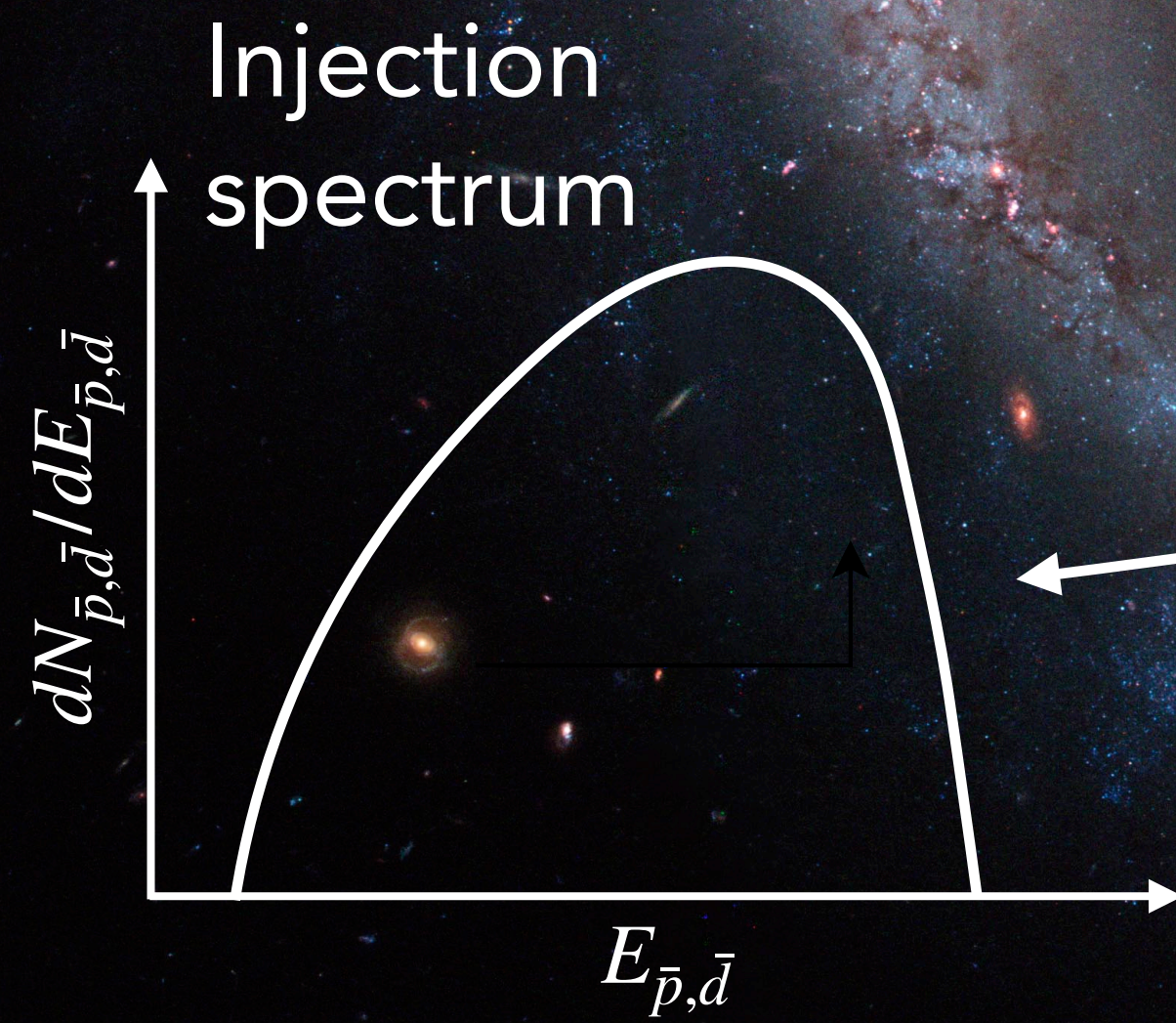


Production

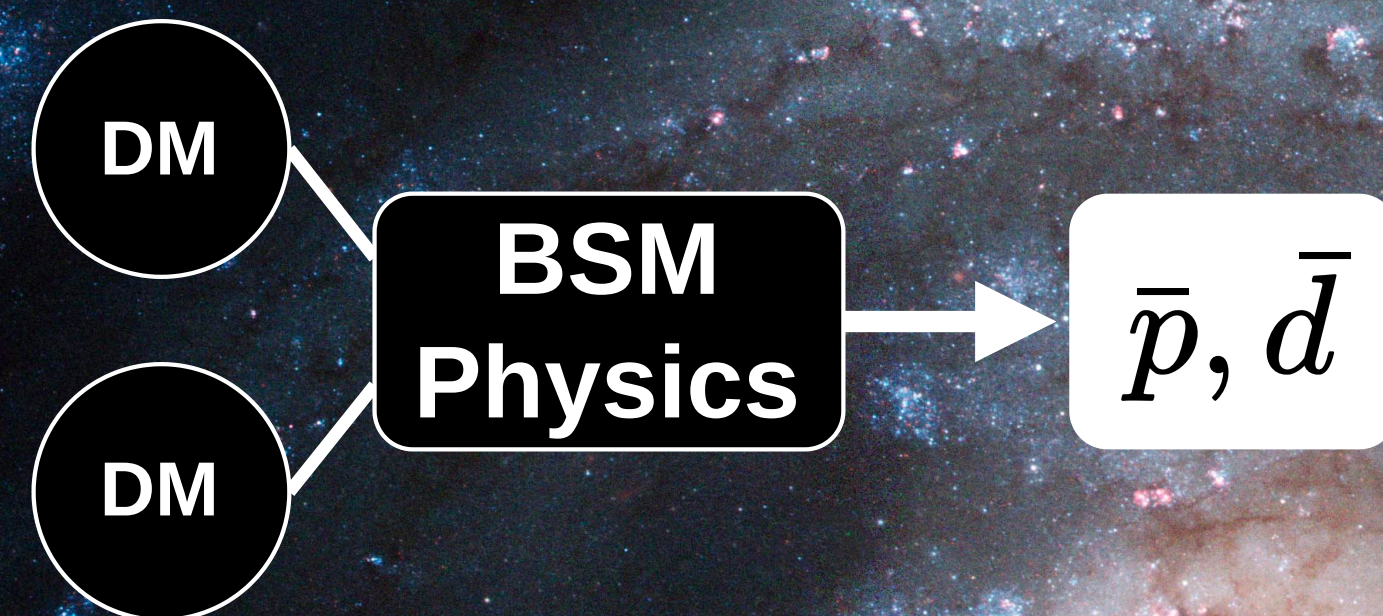




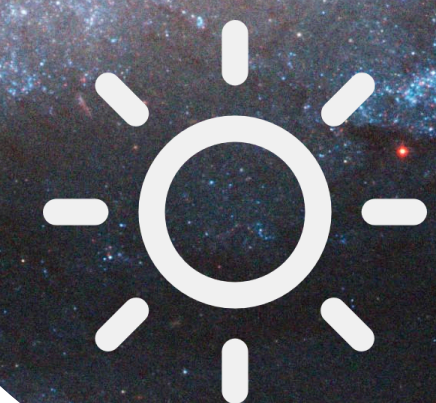
Production



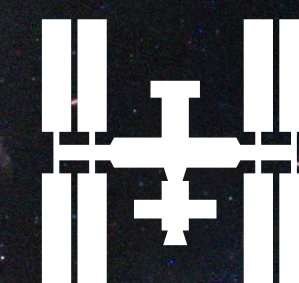
Propagation

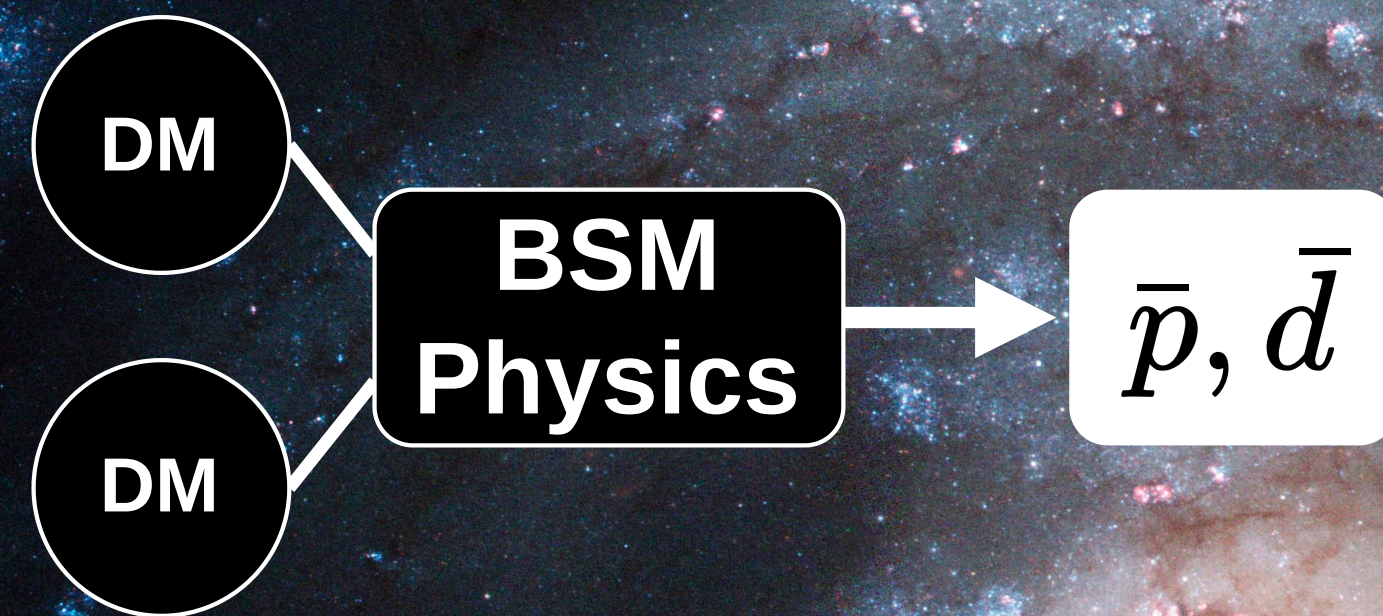


Galactic
Propagation
Galprop, Dragon, ...



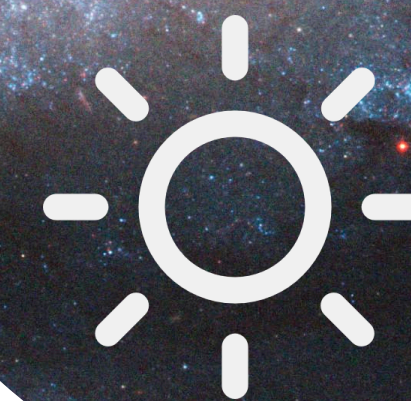
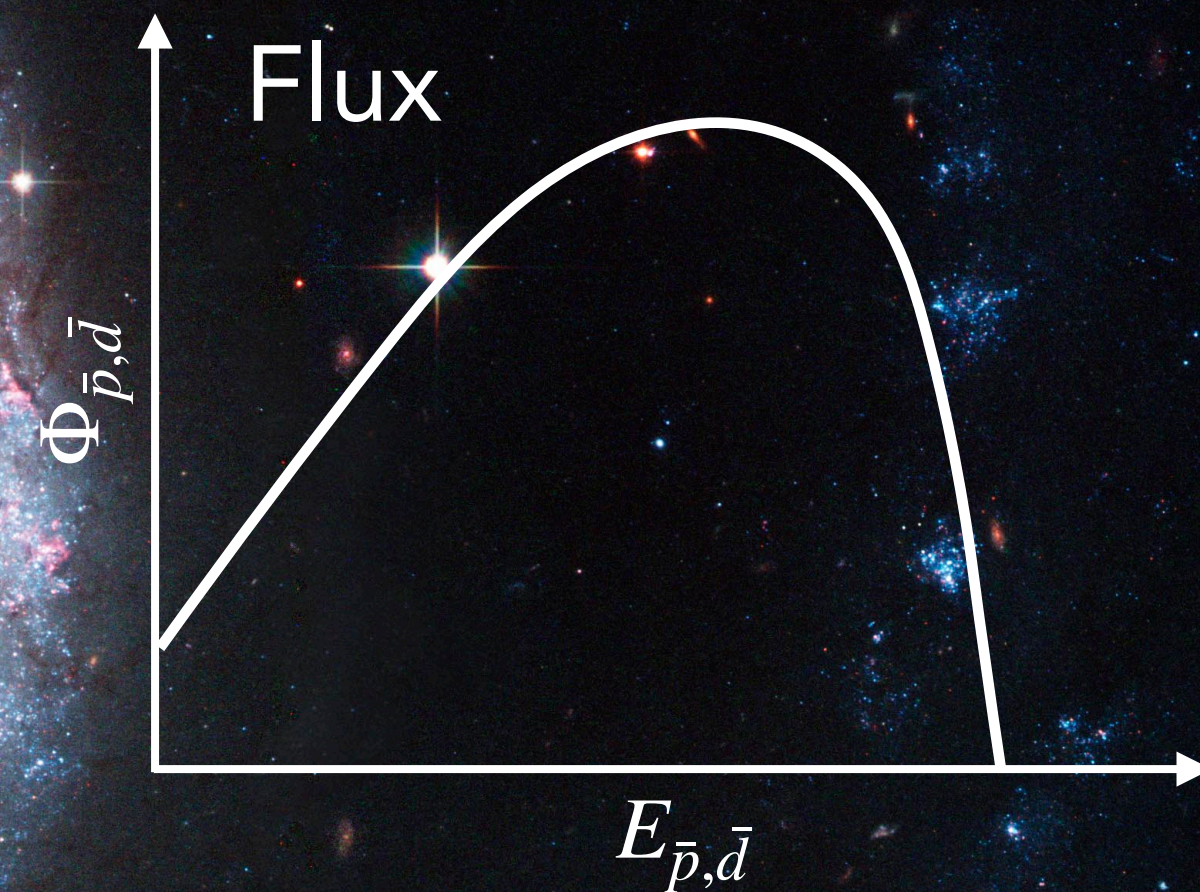
Solar
Modulation



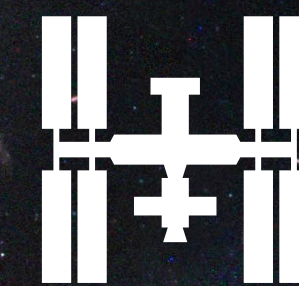


Propagation

Galactic Propagation
Galprop, Dragon, ...

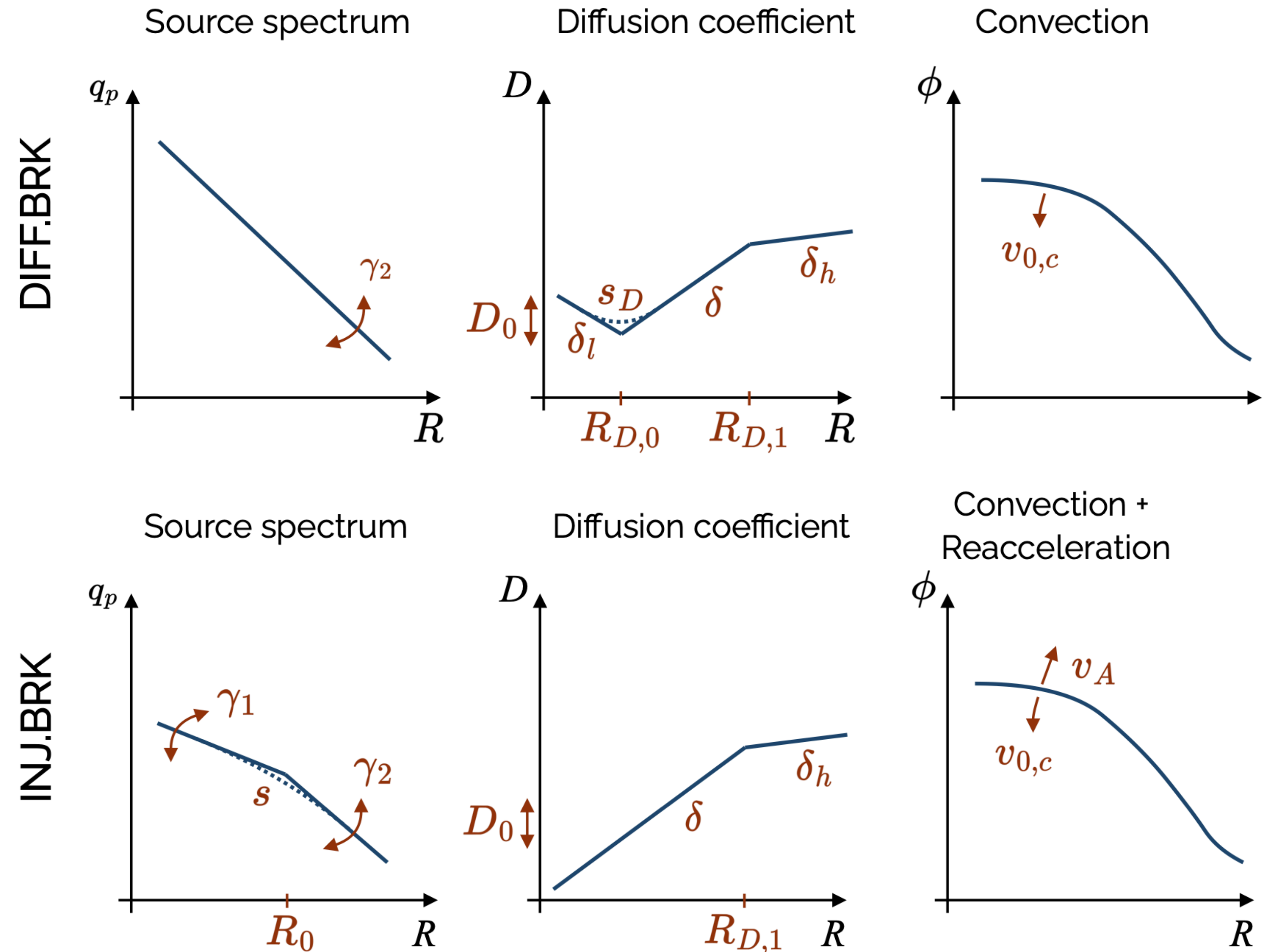


Solar Modulation



Propagation Setups

- Use **diffusion break** and **injection break** models
[2112.08381]
- Used **Galprop** for computation
- In total: **10 to 12 parameters** θ_{prop} to be fitted to data
→ marginalizing over θ_{prop} to obtain limit on $\langle\sigma v\rangle$ becomes **very expensive**

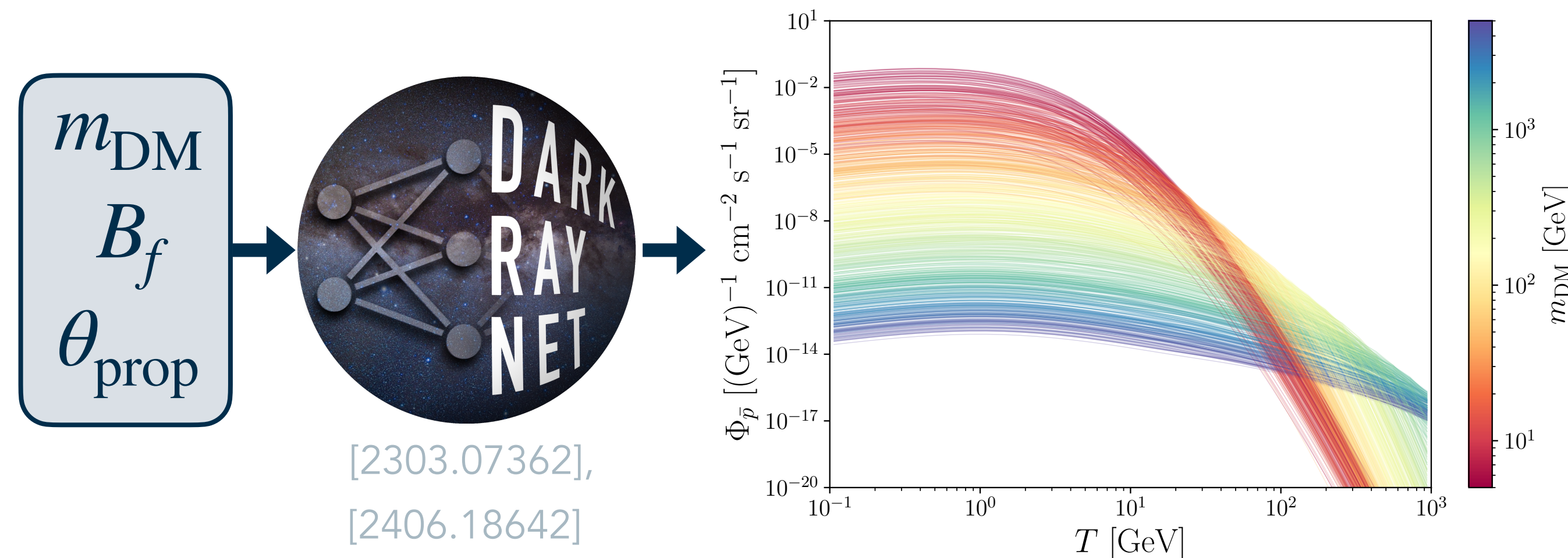


Obtaining Limits on $\langle\sigma v\rangle$

- Limits on $\langle\sigma v\rangle$ exist, computed with **fixed annihilation spectra and likelihood** only for pure annihilation channels like $b\bar{b}$, W^+W^- , ...

Obtaining Limits on $\langle\sigma v\rangle$

- Limits on $\langle\sigma v\rangle$ exist, computed with fixed annihilation spectra and likelihood only for pure annihilation channels like $b\bar{b}$, W^+W^- , ...
- Quick computation of propagated $\bar{p}$ and $\bar{d}$ flux for fixed annihilation channels can be done with **DARKRAYNET** (based on recurrent neural networks)



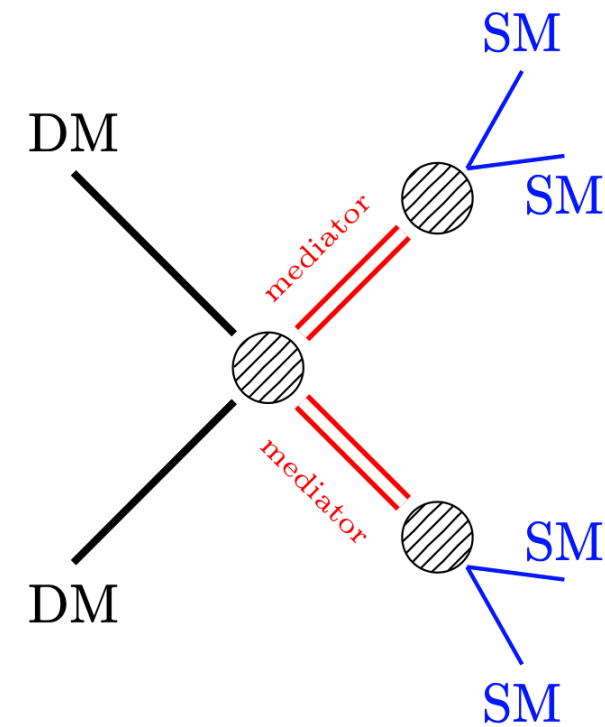
<https://github.com/kathrinnp/DarkRayNet>

- Relative error $\mathcal{O}(10^{-2})$
- Computes marginalized flux ($\mathcal{O}(10^4)$ fluxes) in $\mathcal{O}(1\text{s})$ instead of hours
- Compute $\bar{p}$ limit on $\langle\sigma v\rangle$ yourself or see if AMS-02/GAPS are sensitive to $\bar{d}$ flux

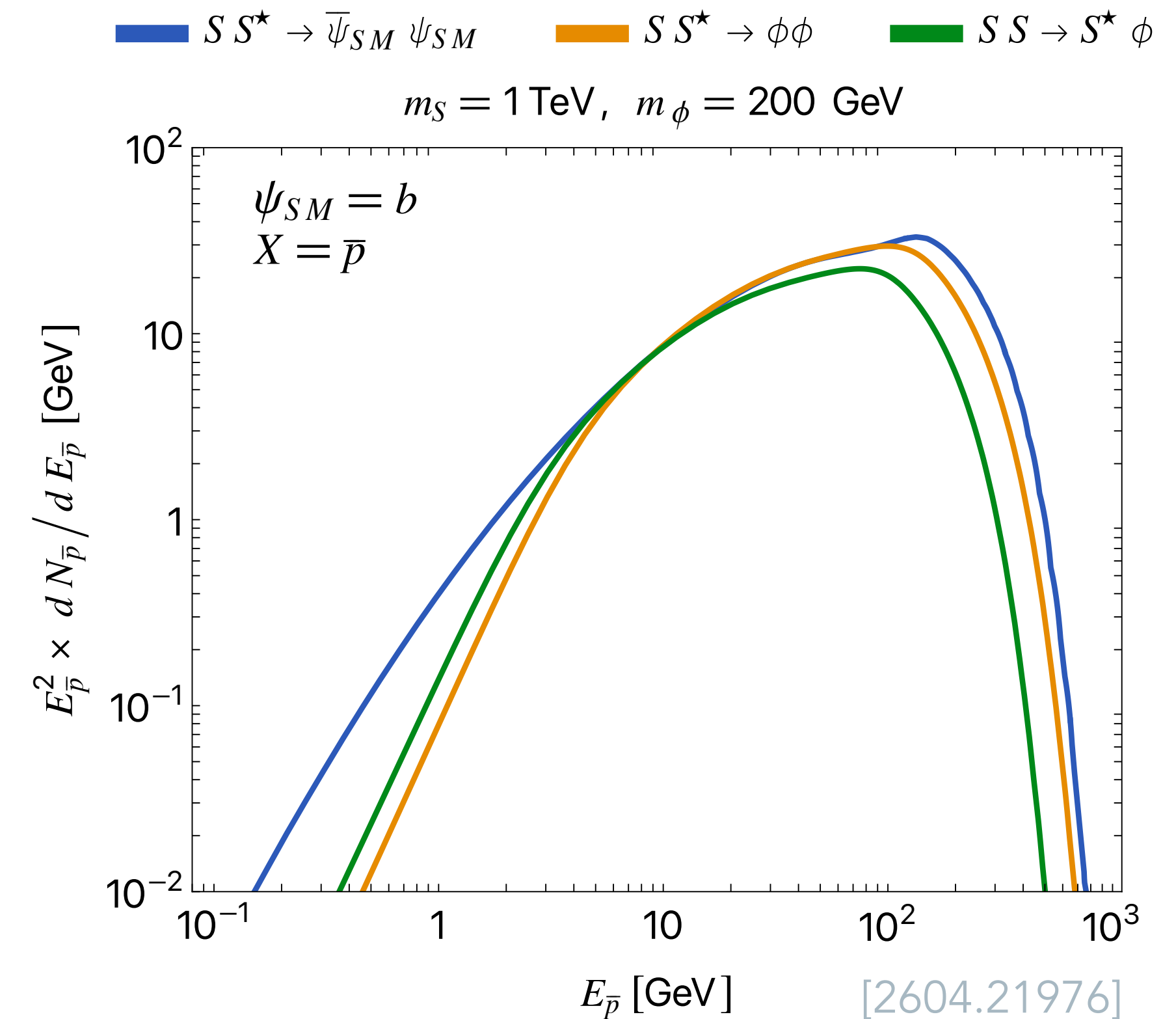
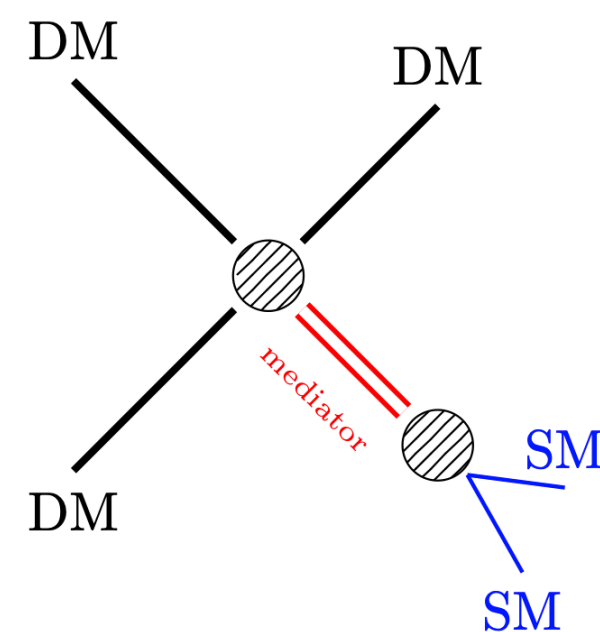
Limitations

- No limits on non-standard annihilation spectra, for example mixed final states or cascade annihilations:

One-step cascade annihilation



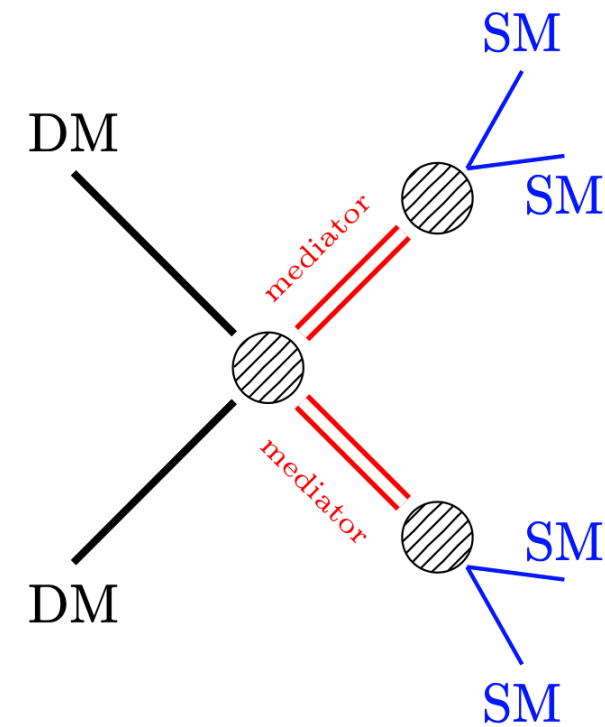
One-step cascade semi-annihilation



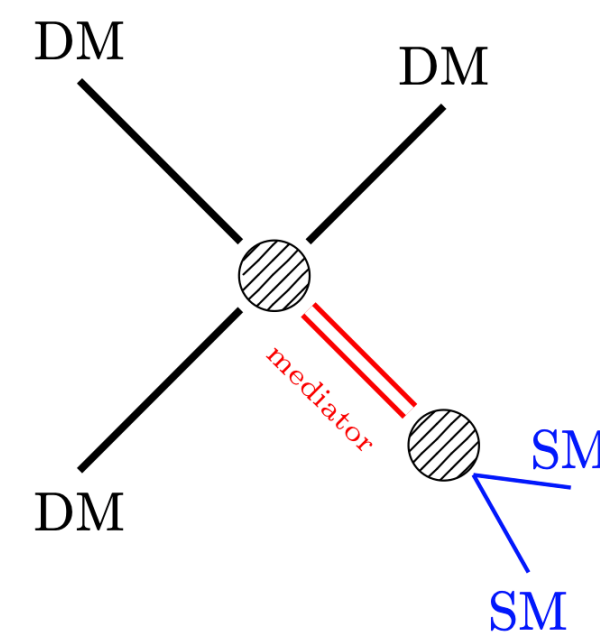
Limitations

- **No limits on non-standard annihilation spectra, for example mixed final states or cascade annihilations:**

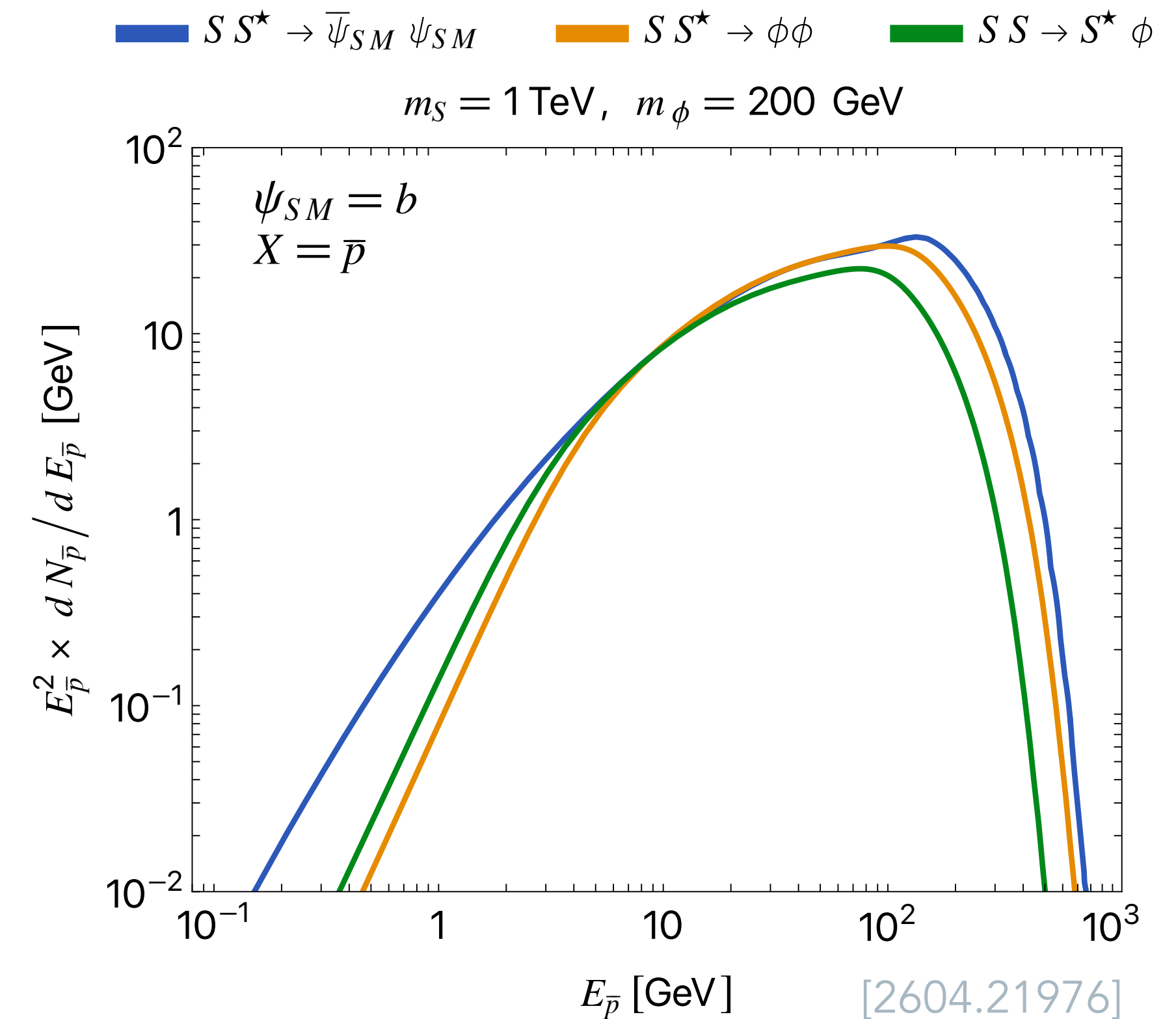
One-step cascade annihilation



One-step cascade semi-annihilation



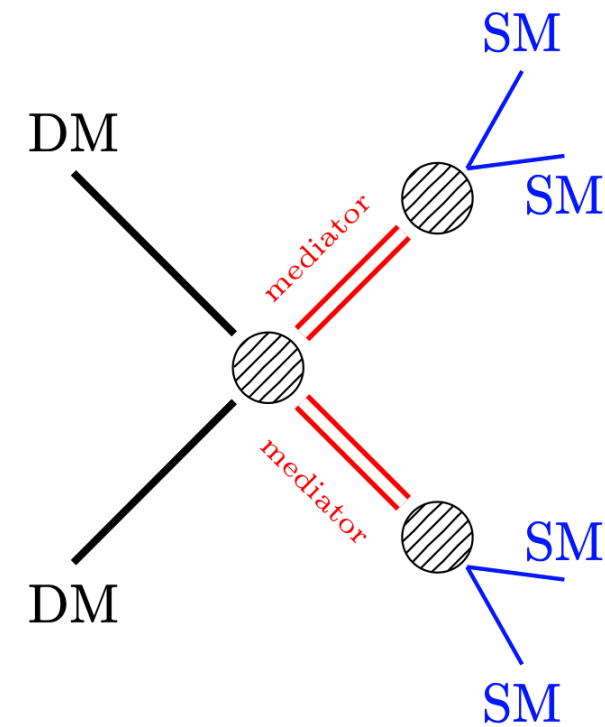
- **Theoretical uncertainties in different event generators and in coalescence mechanism of $\bar{d}$ production can't be investigated quickly**



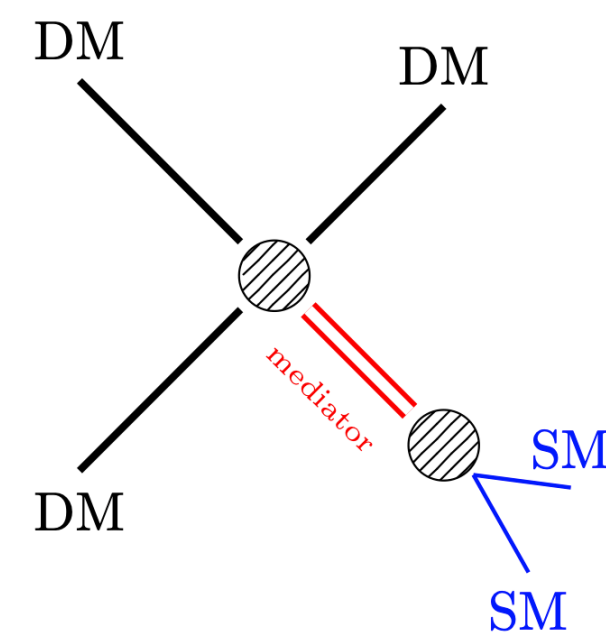
Limitations

- **No limits on non-standard annihilation spectra, for example mixed final states or cascade annihilations:**

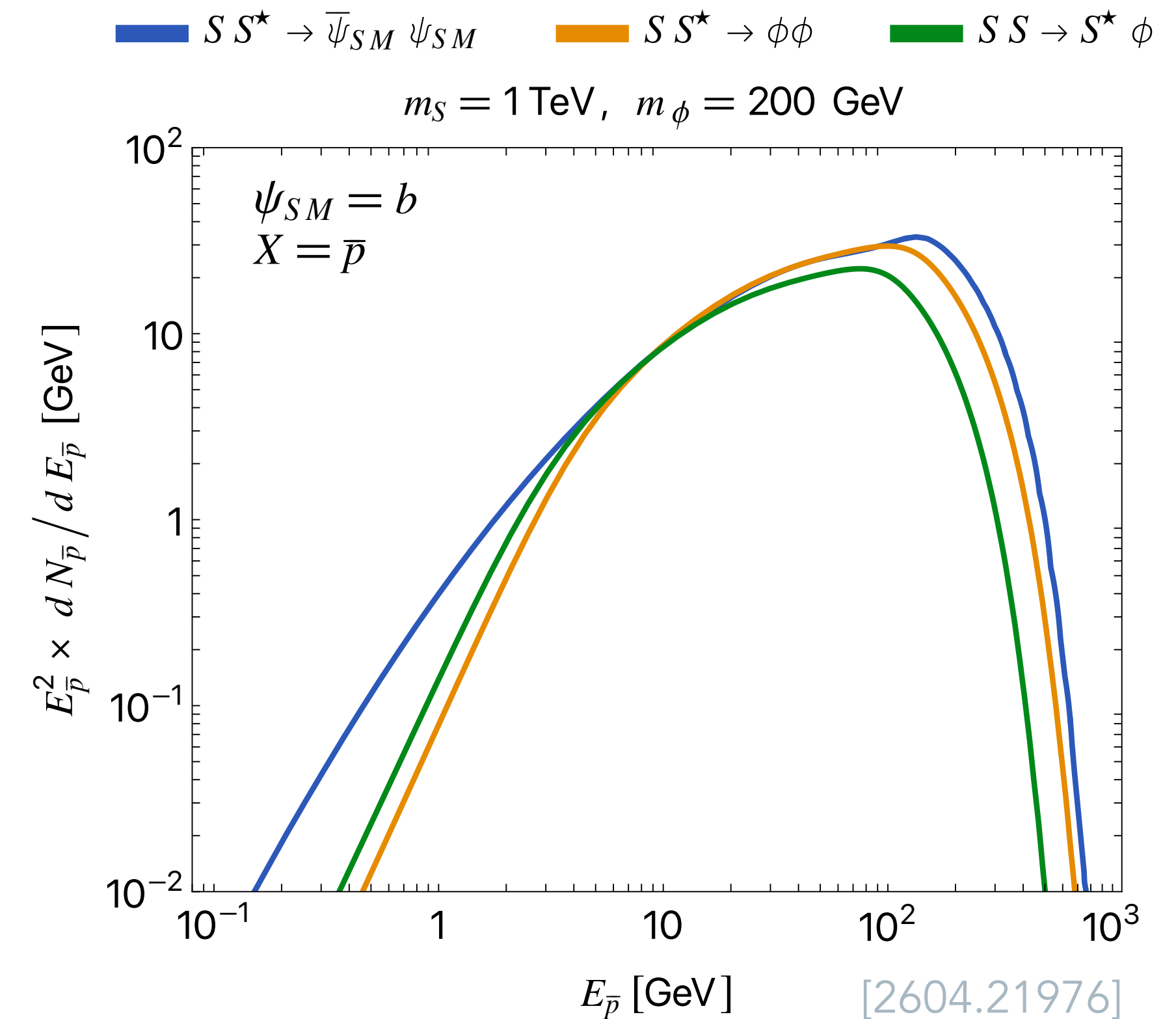
One-step cascade annihilation



One-step cascade semi-annihilation

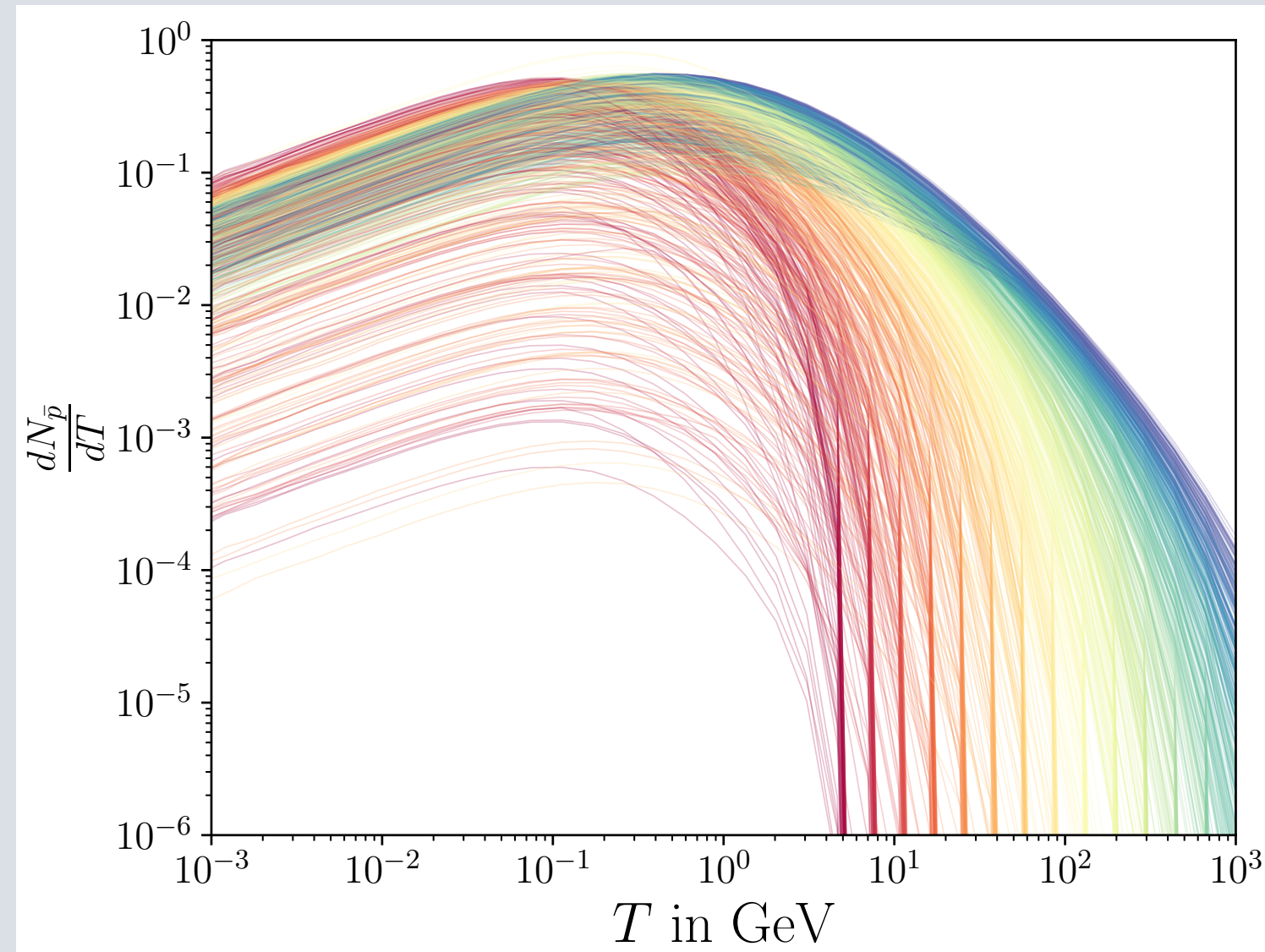


- **Theoretical uncertainties in different event generators and in coalescence mechanism of $\bar{d}$ production can't be investigated quickly**
→ Every new injection spectrum requires a new propagation calculation

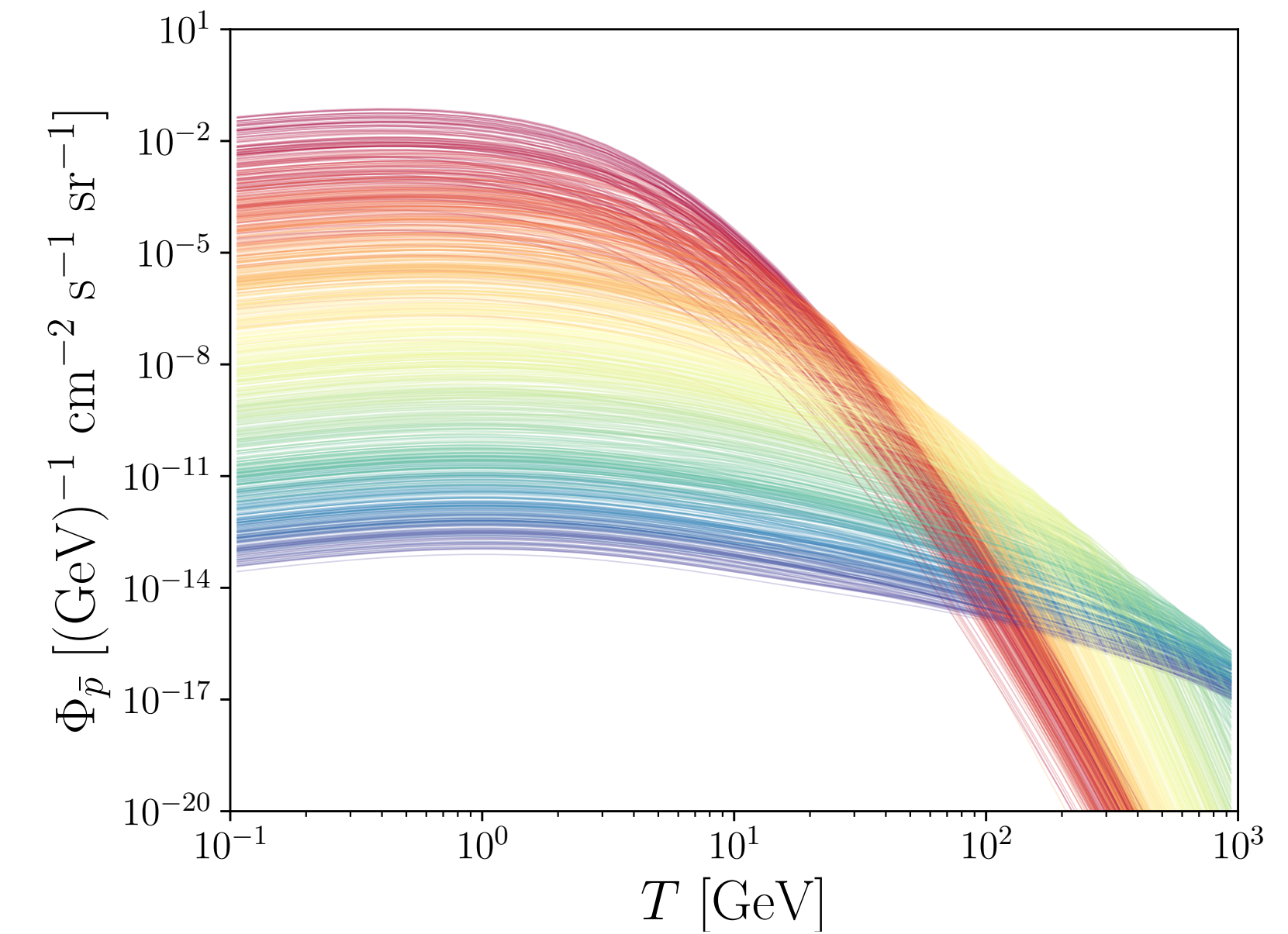


Spectrum-Agnostic Emulator

[2112.12745]

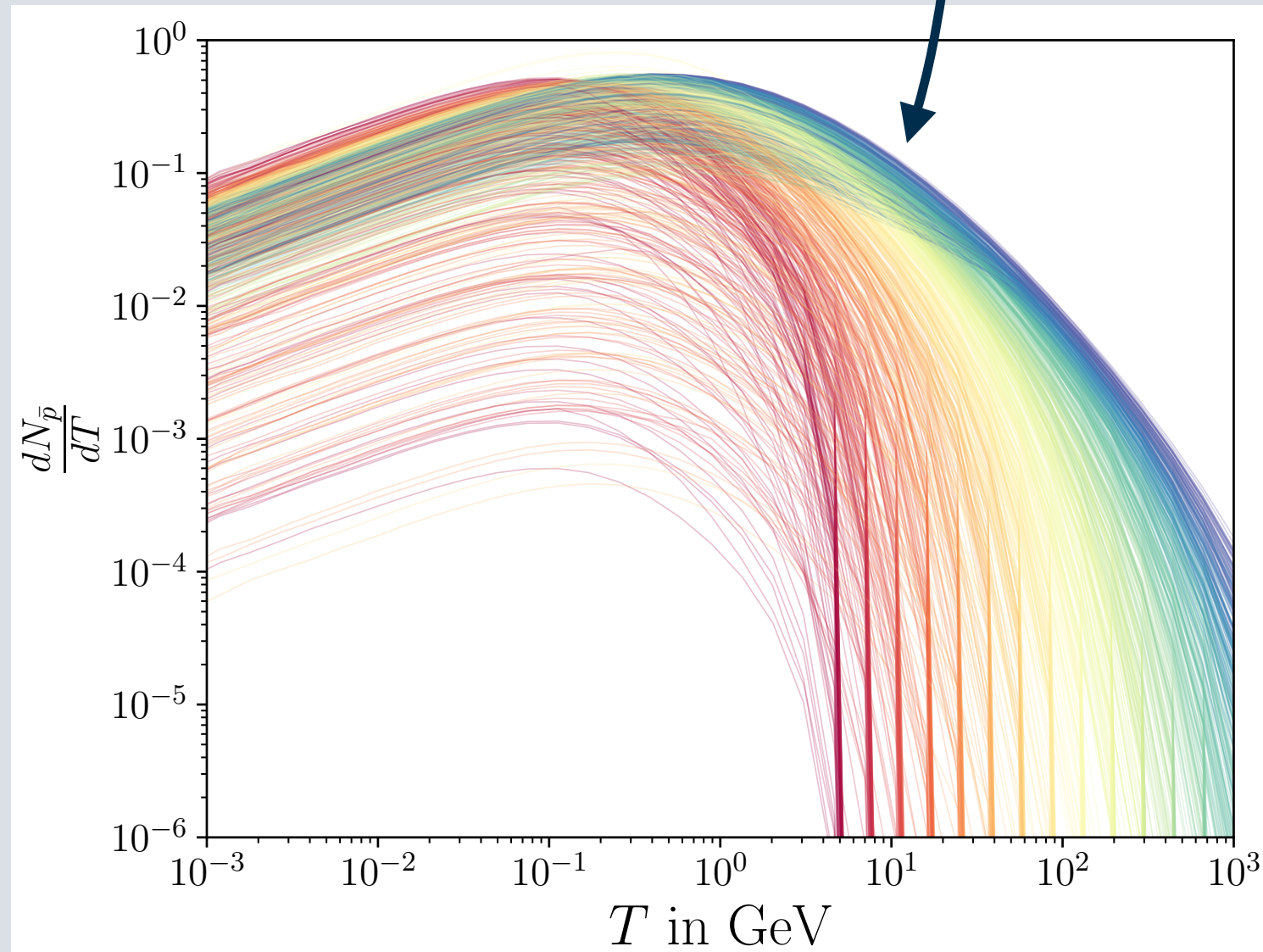


+
 θ_{prop}



Spectrum-Agnostic Emulator

$\bar{p}$ injection spectra
from PPPC tables [1012.4515]

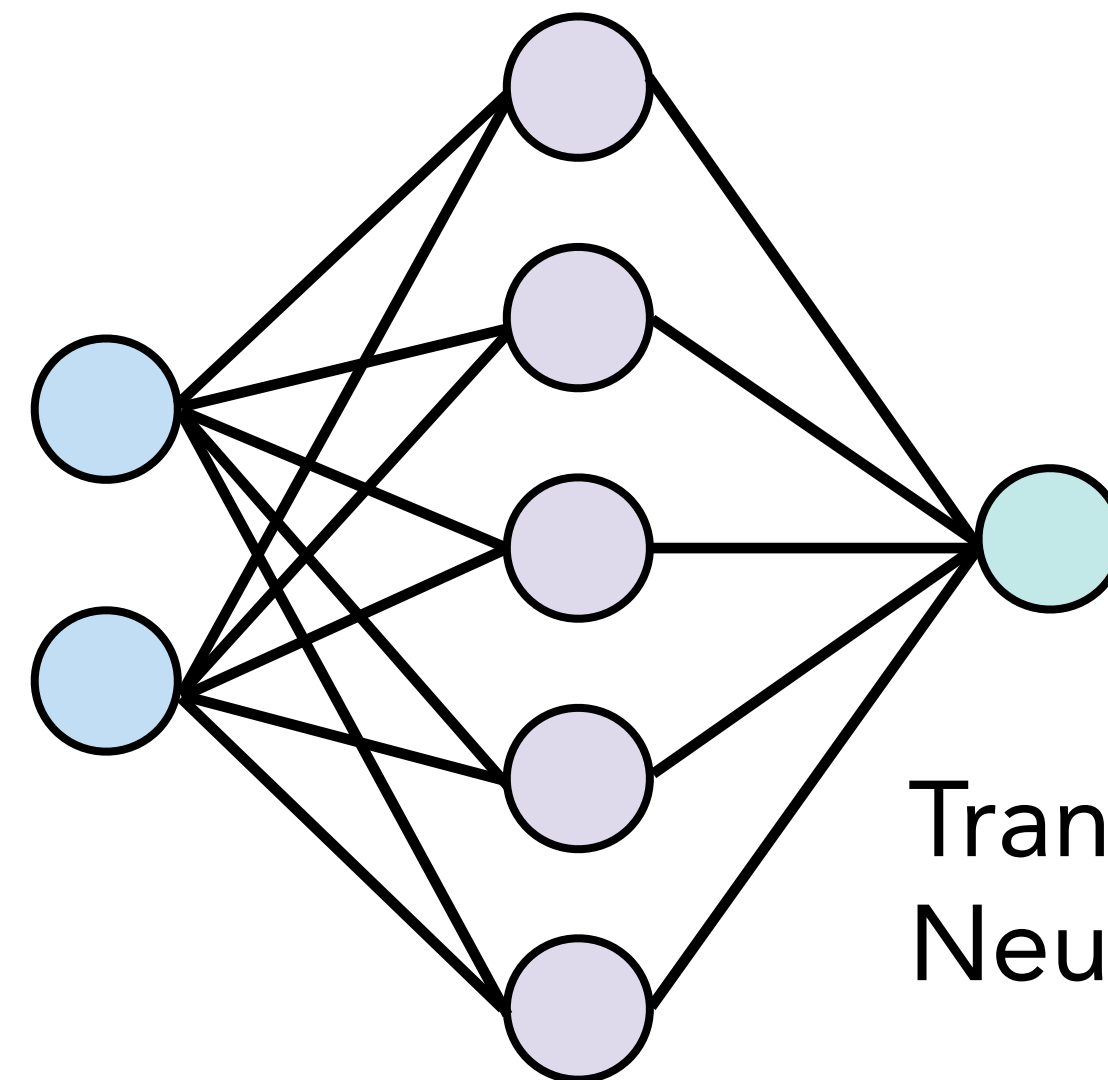


+
 θ_{prop}

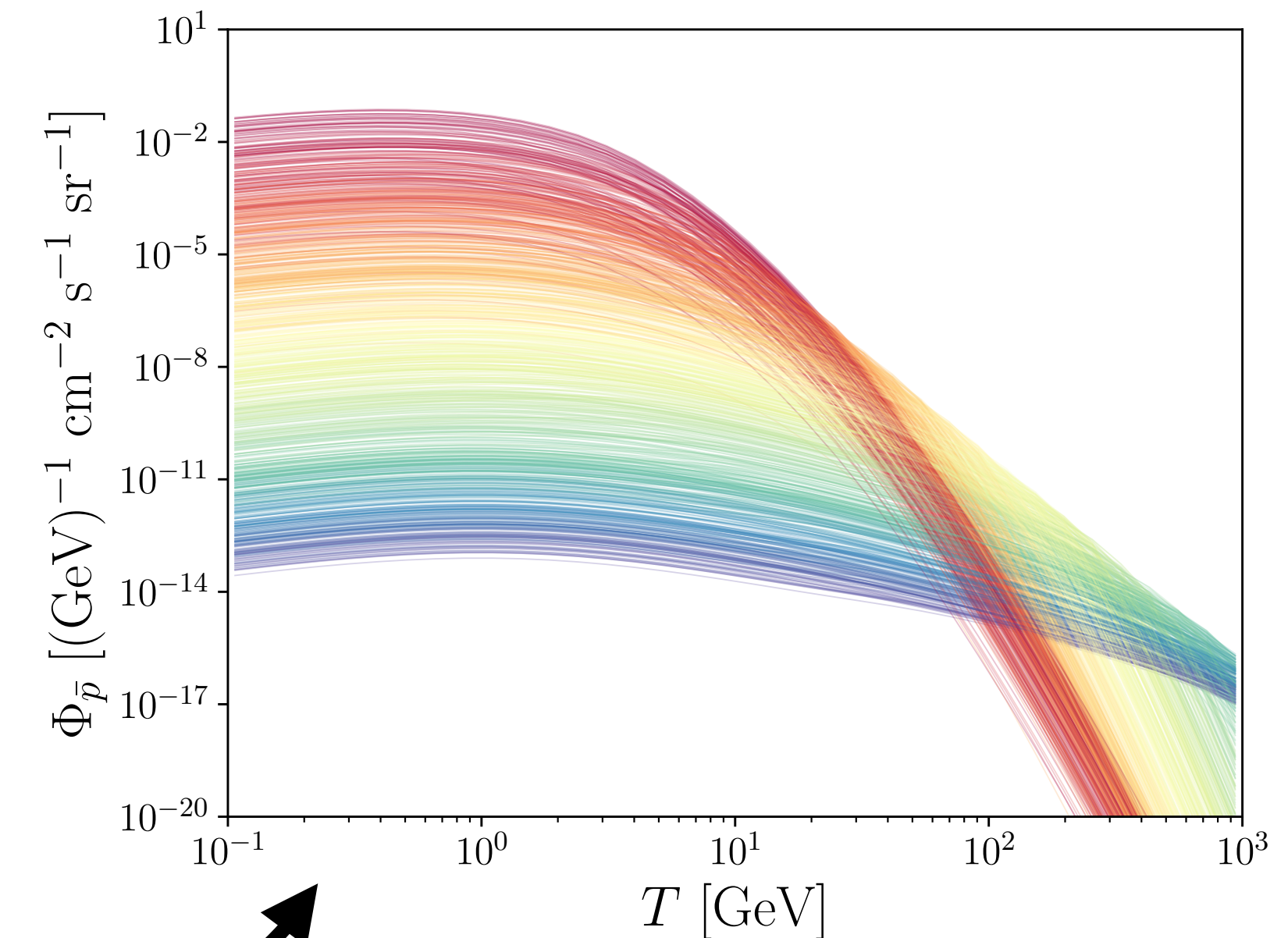
Both for INJ.BRK and
DIFF.BRK models



Train

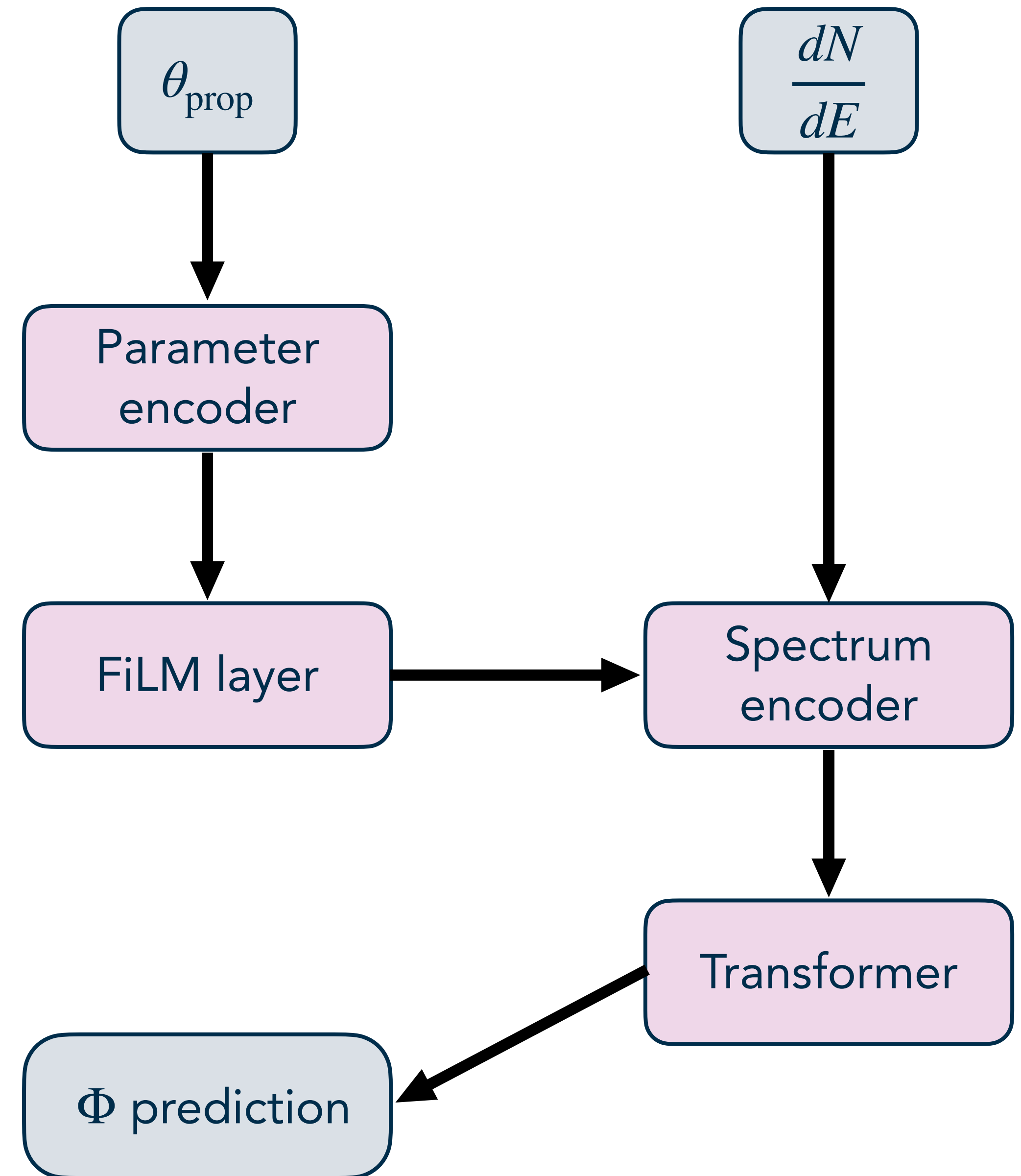


Transformer
Neural Network



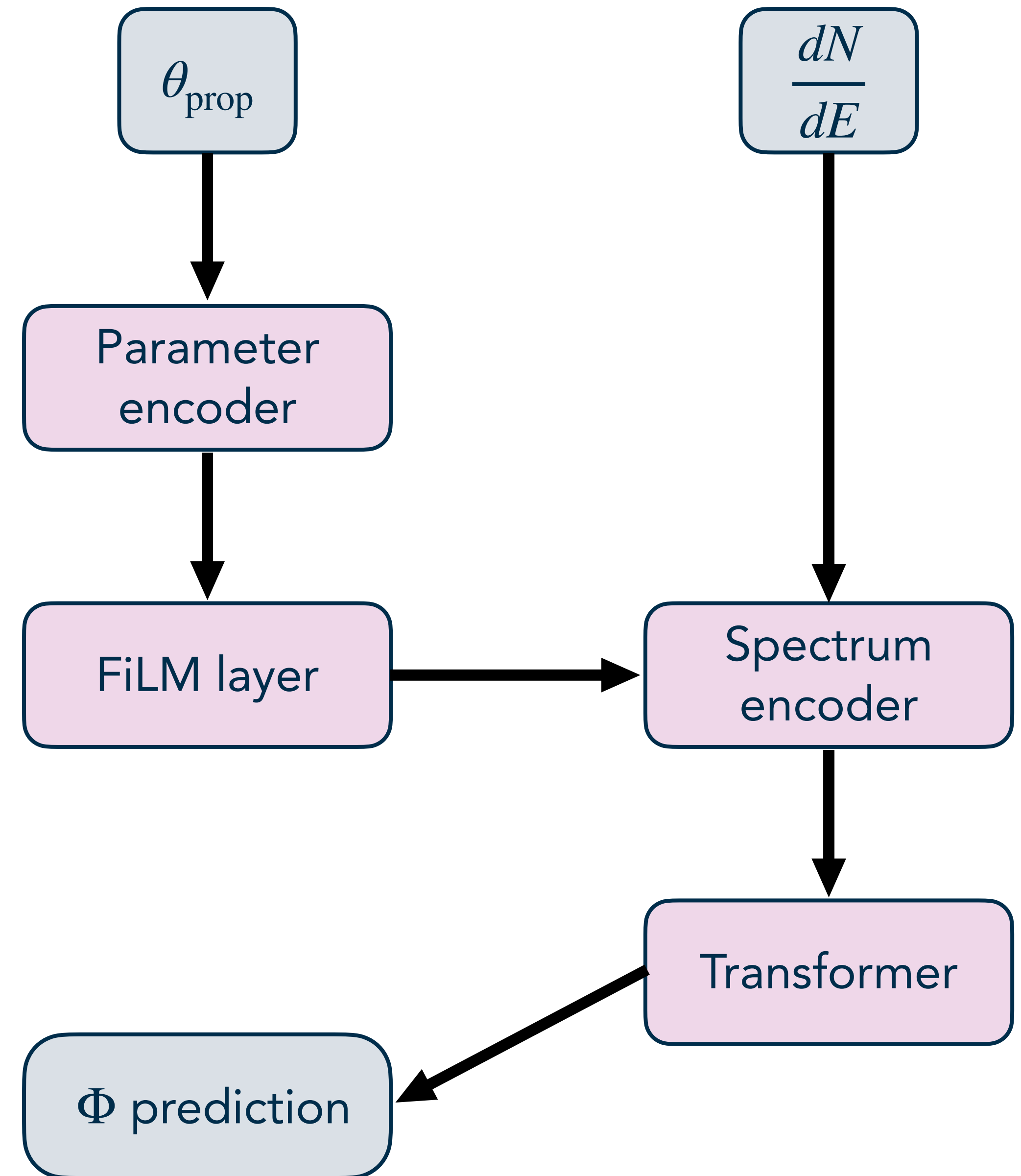
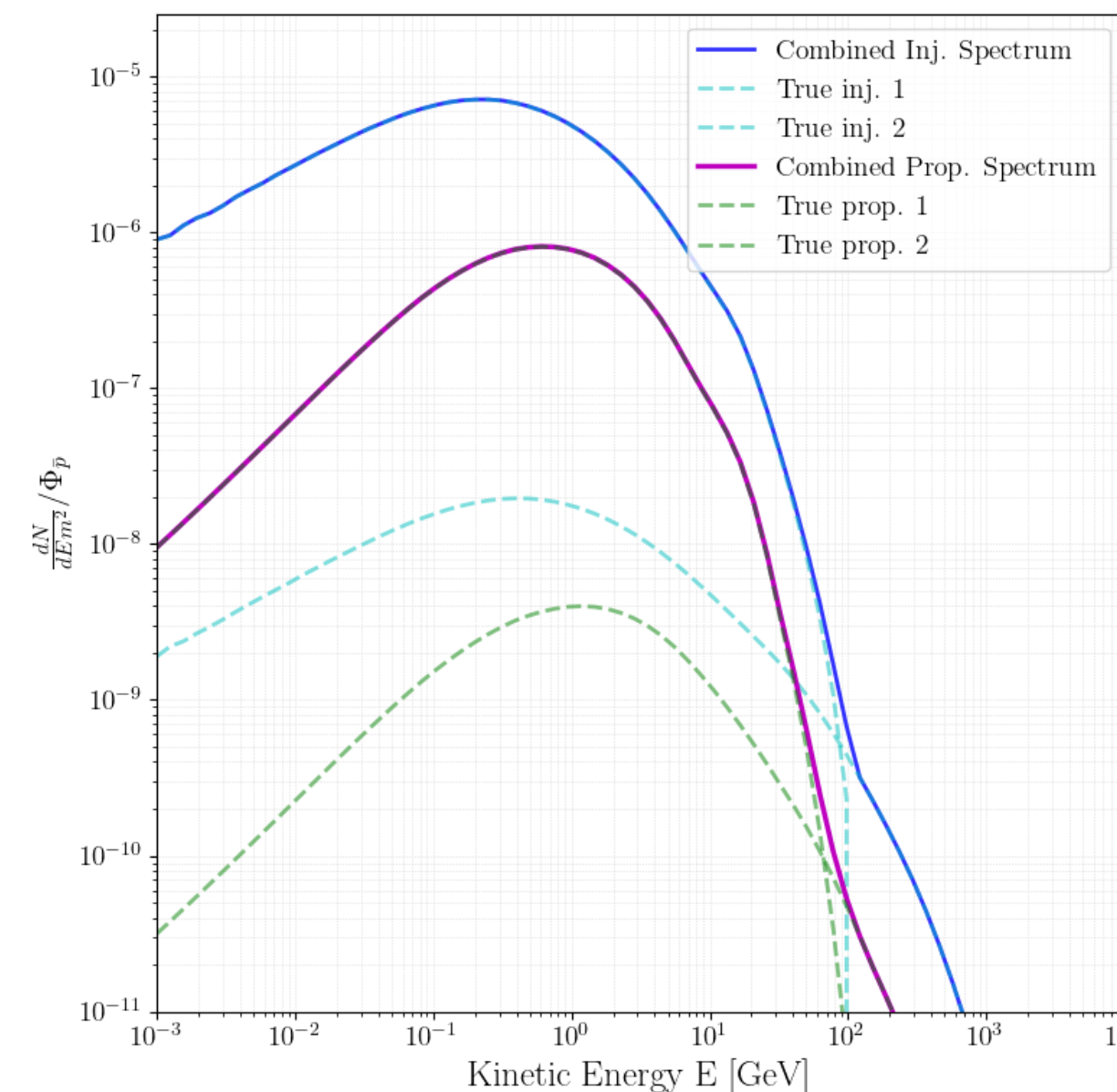
Network Architecture

- Use **FiLM** (Feature-wise Linear Modulation) **layer** to condition spectrum representation by scaling and shifting it
- **Update architecture** to state-of-the-art **transformer** with information exchange through self-attention



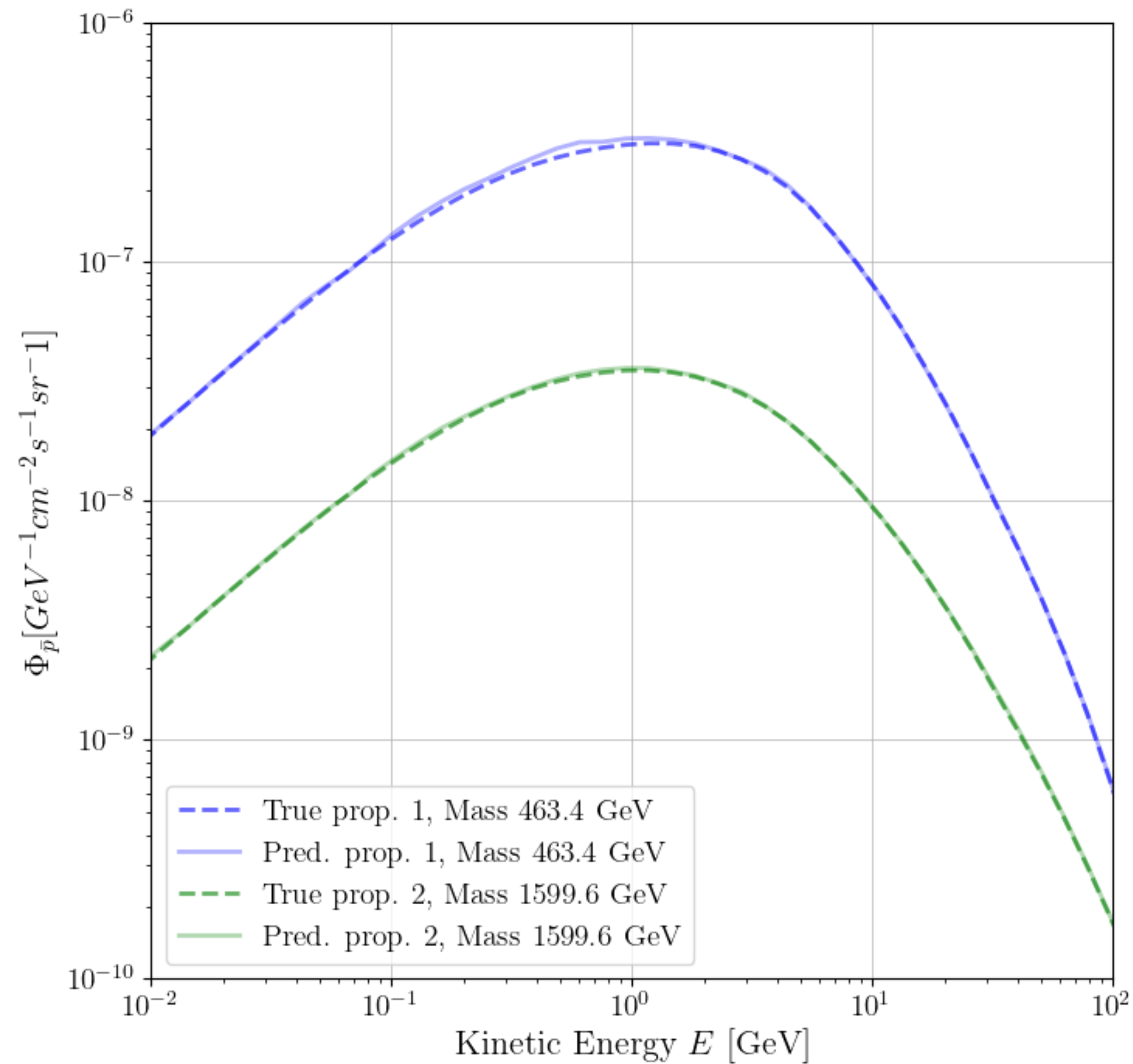
Network Architecture

- Use **FiLM** (Feature-wise Linear Modulation) **layer** to condition spectrum representation by scaling and shifting it
- Update architecture to state-of-the-art **transformer** with information exchange through self-attention
- Add superposition of different injection spectra to training set



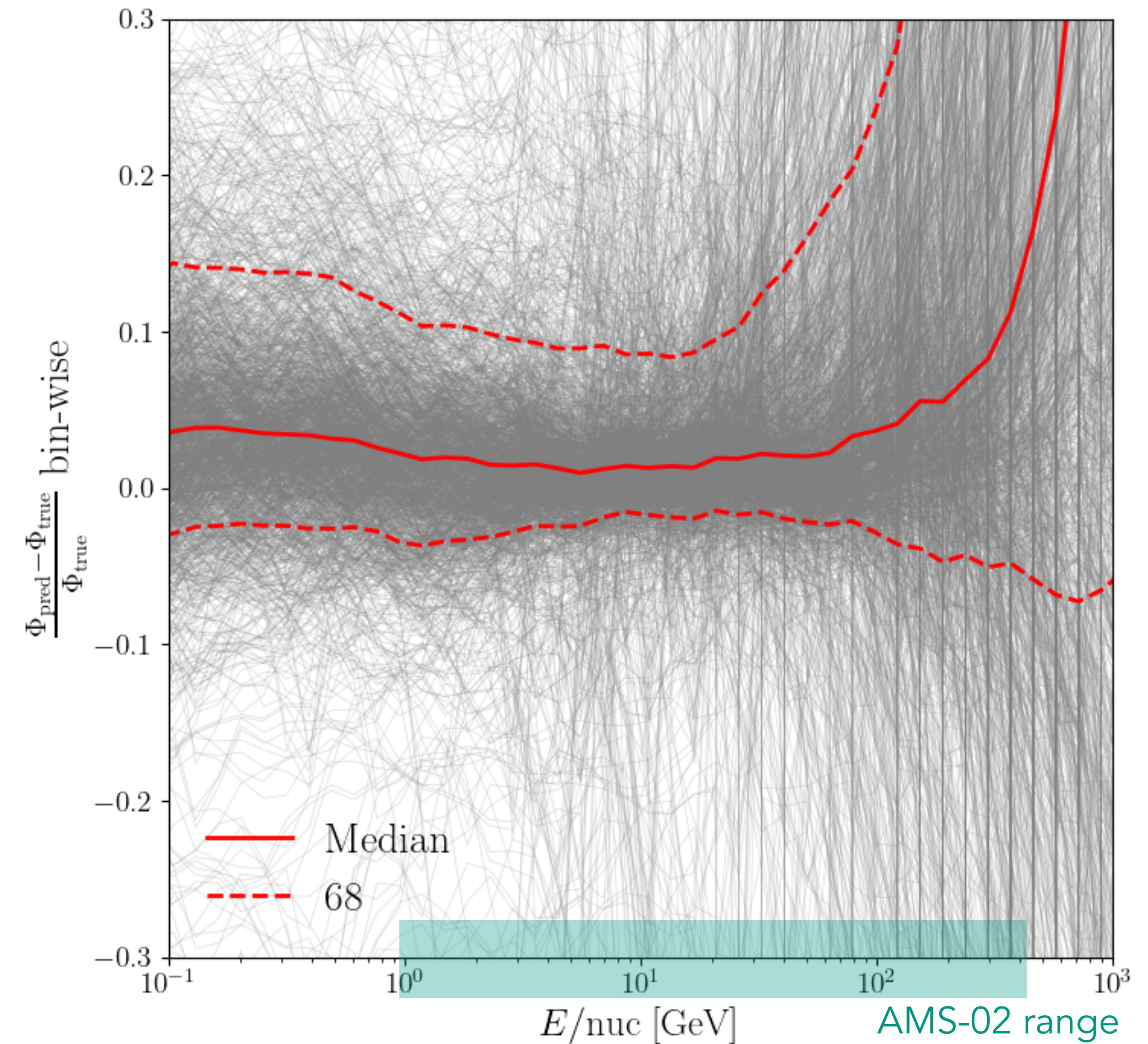
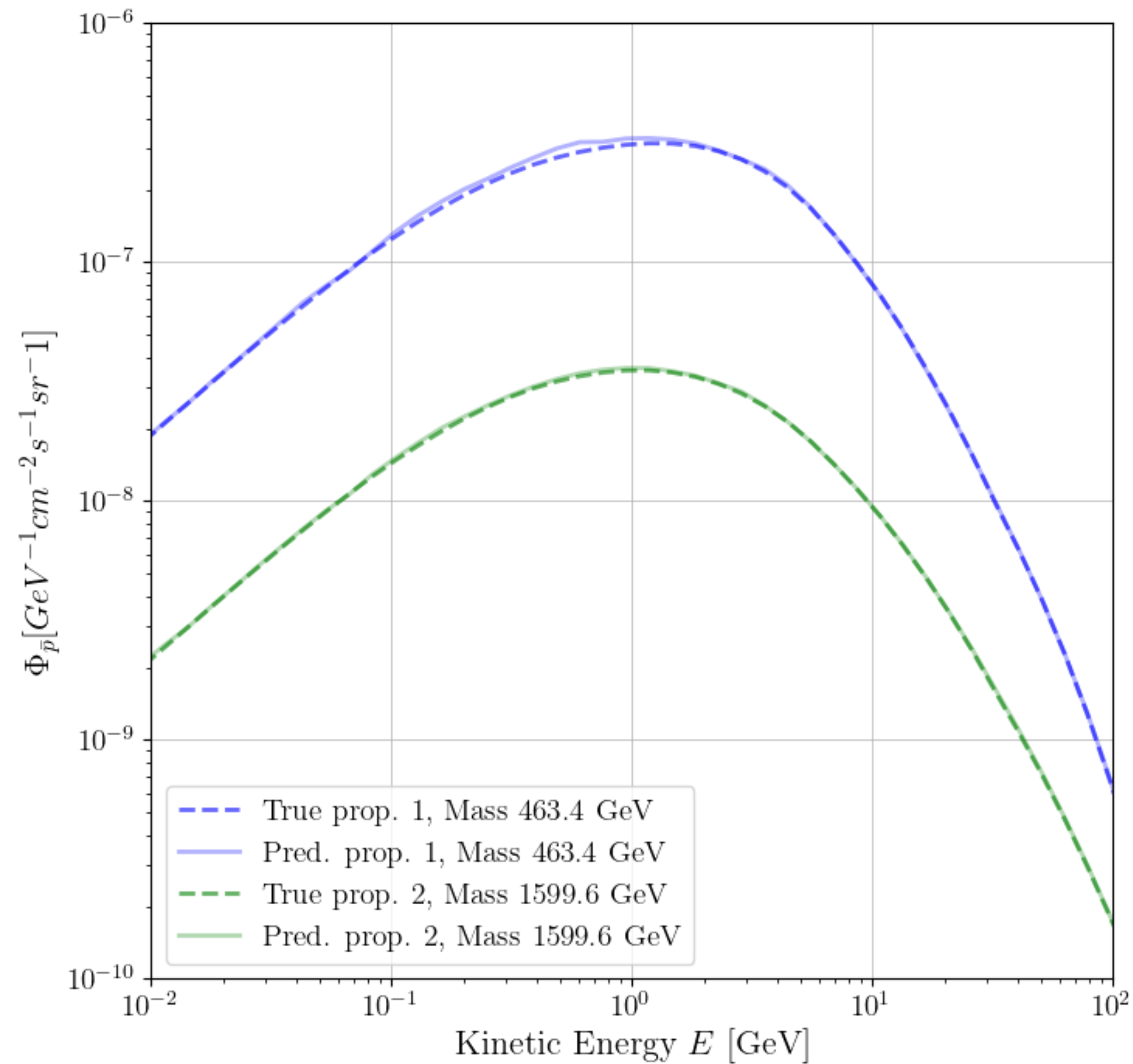
Network Performance (Antiprotons + DIFF.BRK)

Preliminary



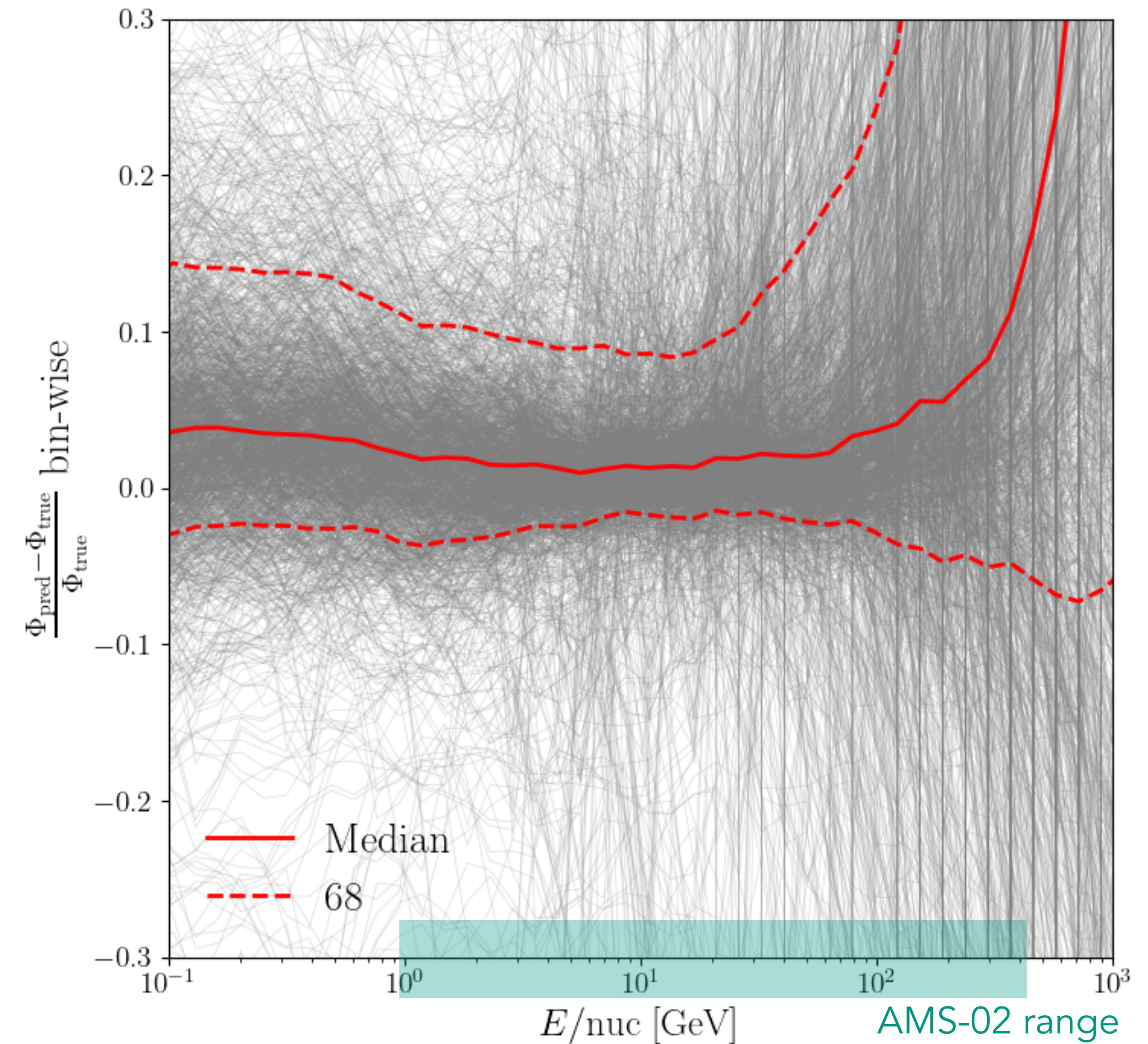
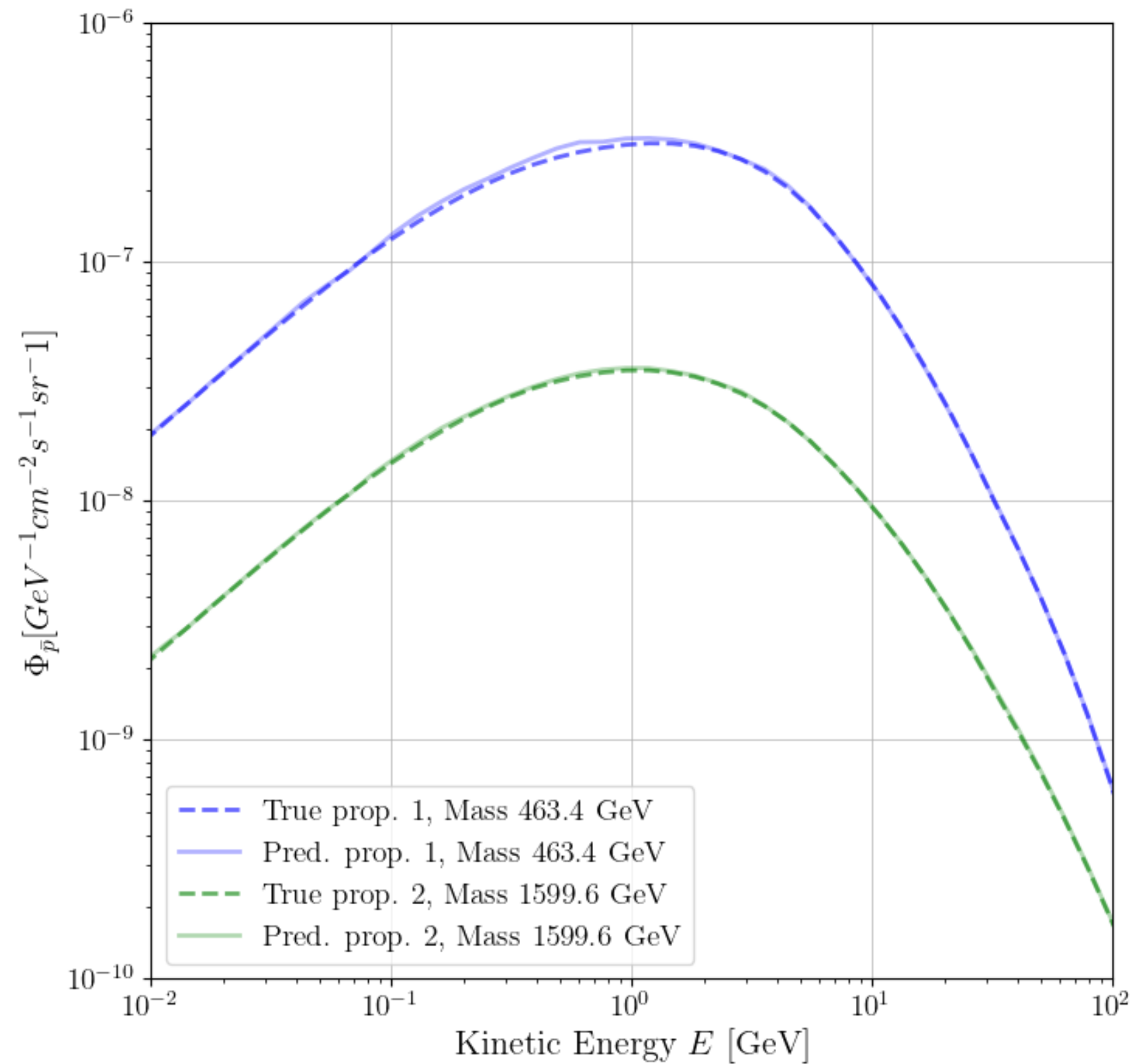
Network Performance (Antiprotons + DIFF.BRK)

Preliminary



Network Performance (Antiprotons + DIFF.BRK)

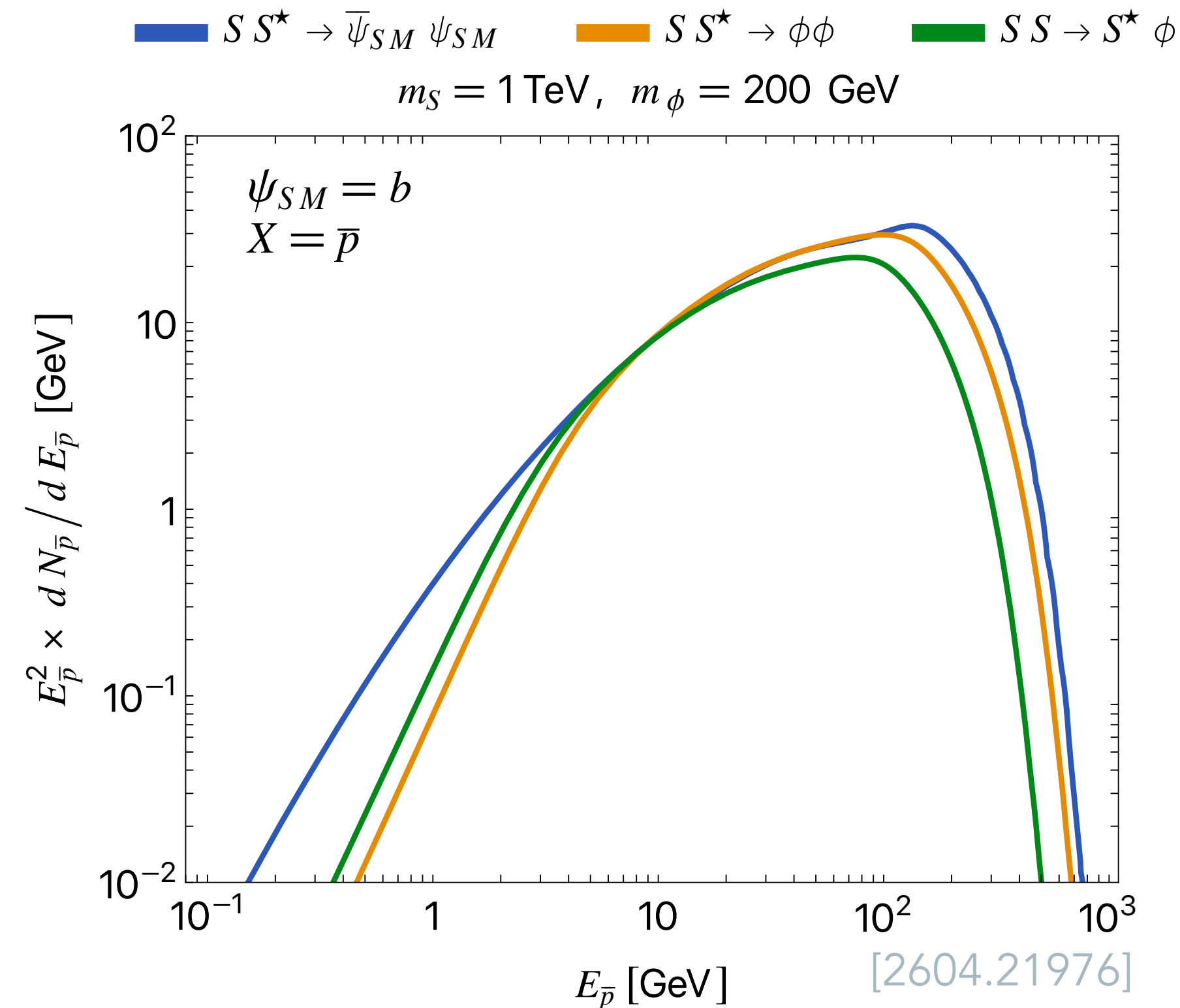
Preliminary



- Computes **1000 fluxes** in $\mathcal{O}(1\text{s})$ instead of $\mathcal{O}(10\text{s})$ for one flux with Galprop

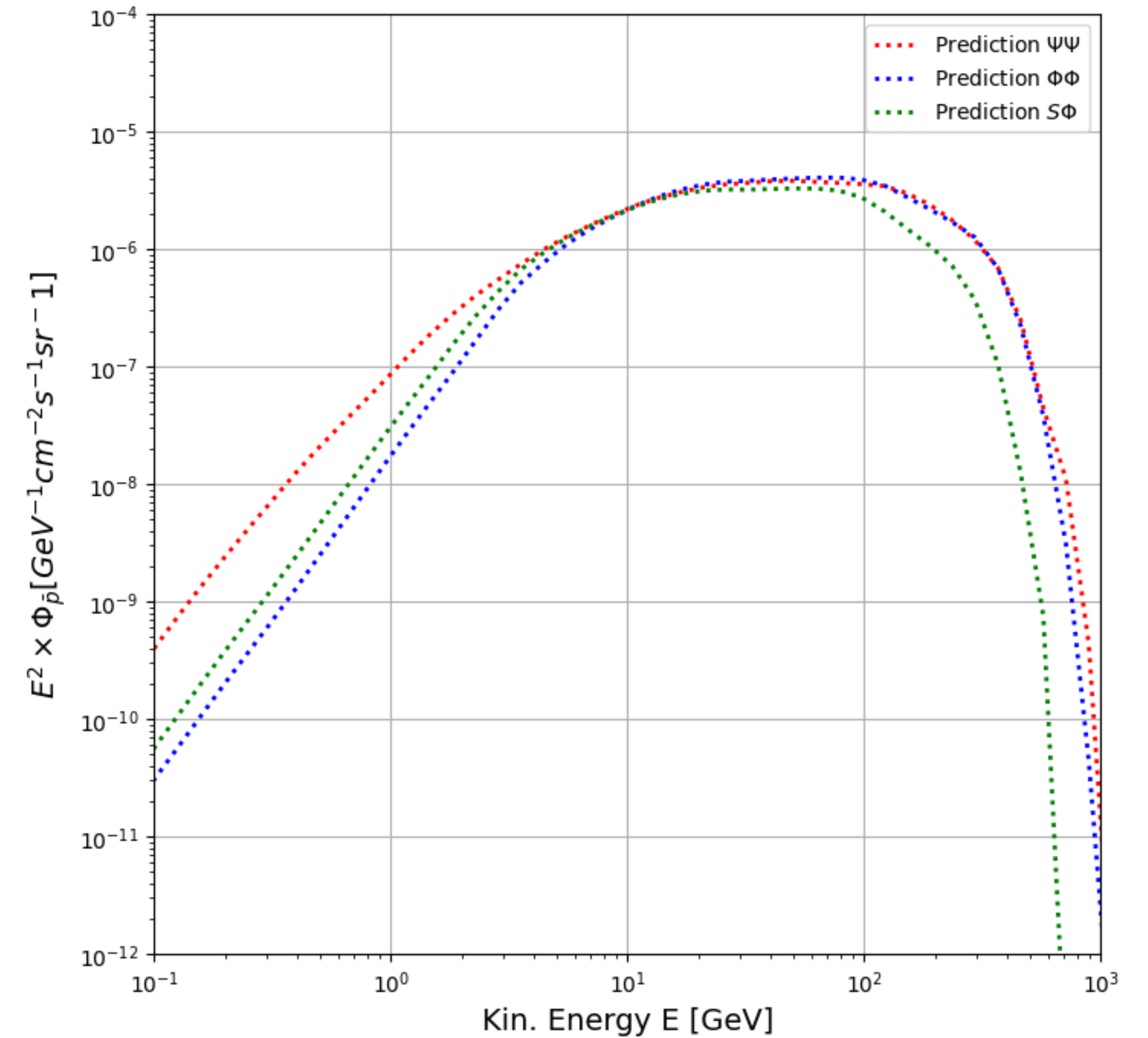
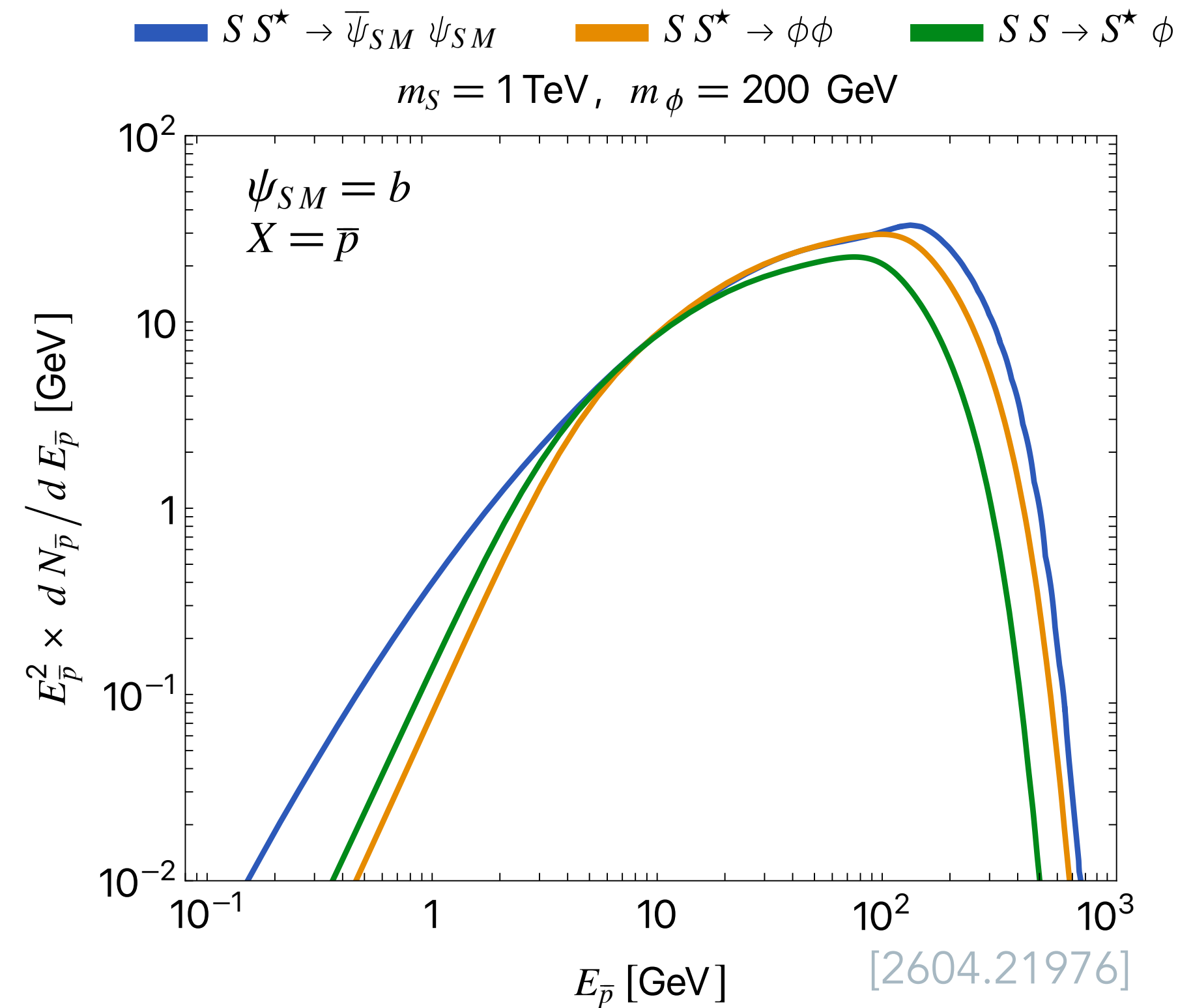
Application to Non-Standard Spectra

- Apply emulator to one-step cascade (semi-)annihilation injection spectra for best-fit propagation parameters (flux has not been computed in 2604.21976)



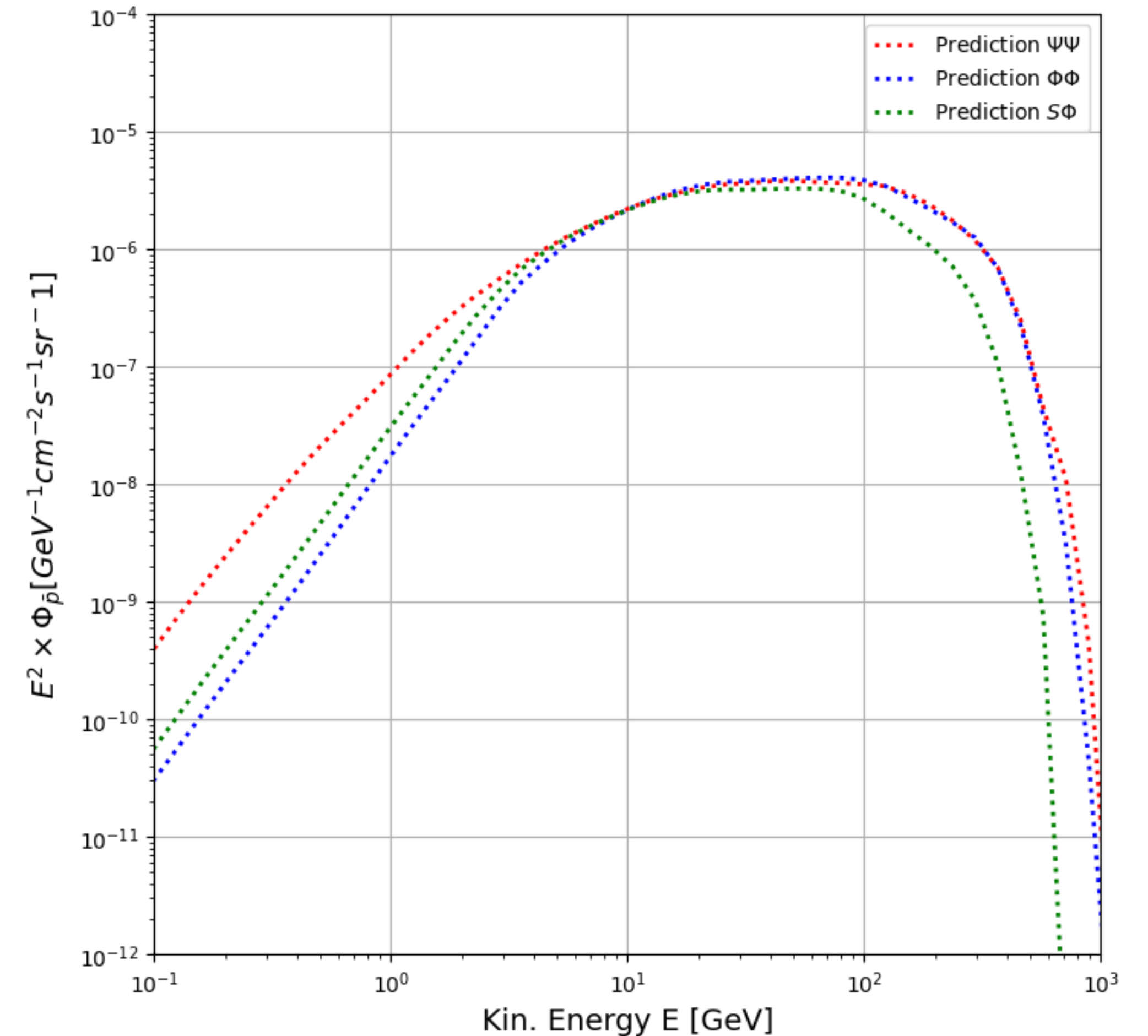
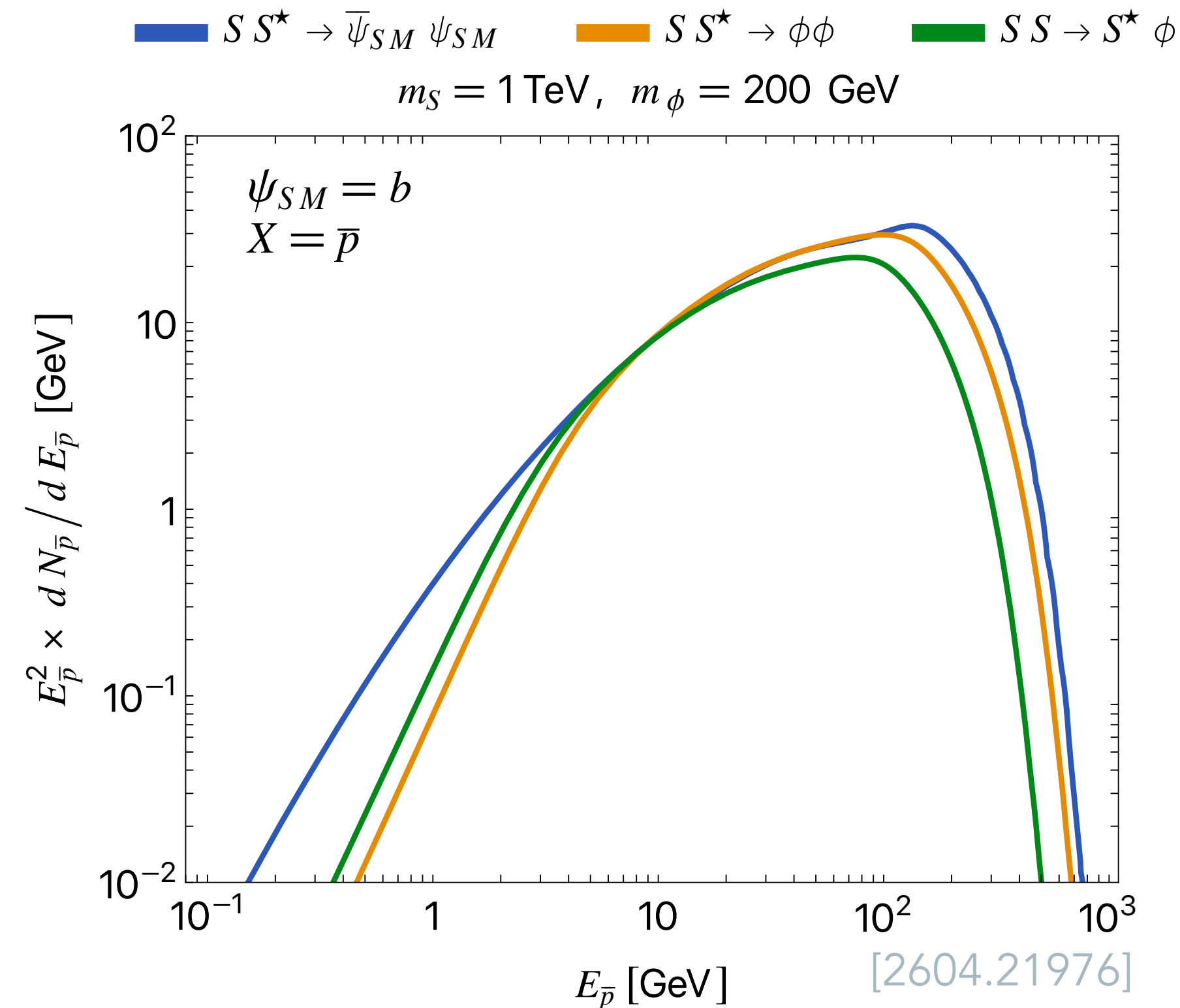
Application to Non-Standard Spectra

- Apply emulator to one-step cascade (semi-)annihilation injection spectra for best-fit propagation parameters (flux has not been computed in 2604.21976)



Application to Non-Standard Spectra

- Apply emulator to one-step cascade (semi-)annihilation injection spectra for best-fit propagation parameters (flux has not been computed in 2604.21976)



→ Possible to also compute marginalized flux in seconds

Conclusions

- Indirect detection can place strong bounds on $\langle\sigma v\rangle$, limits usually published for annihilation into pair of SM particles
- **Numerical propagation becomes expensive** when exploring non-standard injection spectra for many propagation parameters
- Developed a **spectrum-agnostic neural-network emulator** that predicts $\bar{p}$ (and soon also $\bar{d}$) flux for arbitrary DM annihilation injection spectra
- Speed-up of $\mathcal{O}(10^4)$ with precision of $\mathcal{O}(10\%)$
- Used emulator to **predict flux for one-step cascade (semi-)annihilation injection spectra** which can be used to compute limits on $\langle\sigma v\rangle$

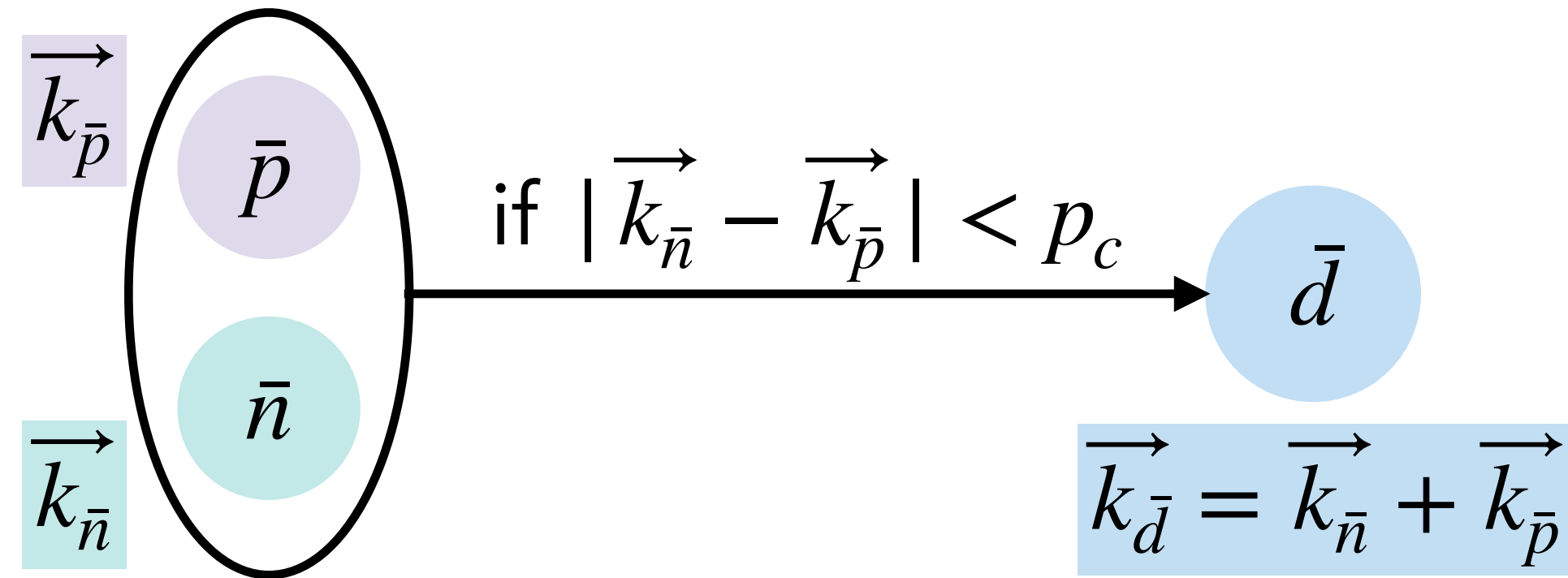
Conclusions

- Indirect detection can place strong bounds on $\langle\sigma v\rangle$, limits usually published for annihilation into pair of SM particles
- **Numerical propagation becomes expensive** when exploring non-standard injection spectra for many propagation parameters
- Developed a **spectrum-agnostic neural-network emulator** that predicts $\bar{p}$ (and soon also $\bar{d}$) flux for arbitrary DM annihilation injection spectra
- Speed-up of $\mathcal{O}(10^4)$ with precision of $\mathcal{O}(10\%)$
- Used emulator to **predict flux for one-step cascade (semi-)annihilation injection spectra** which can be used to compute limits on $\langle\sigma v\rangle$

Thank you!

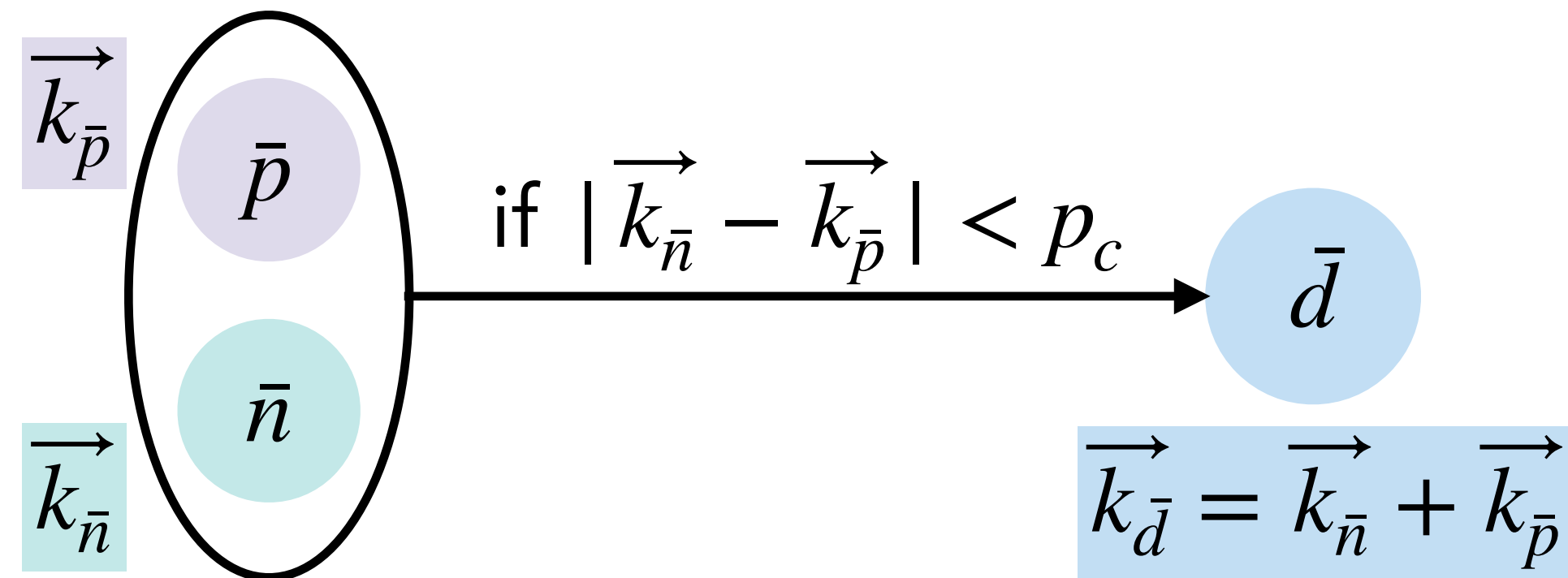
Backup Slides

Antideuteron Production

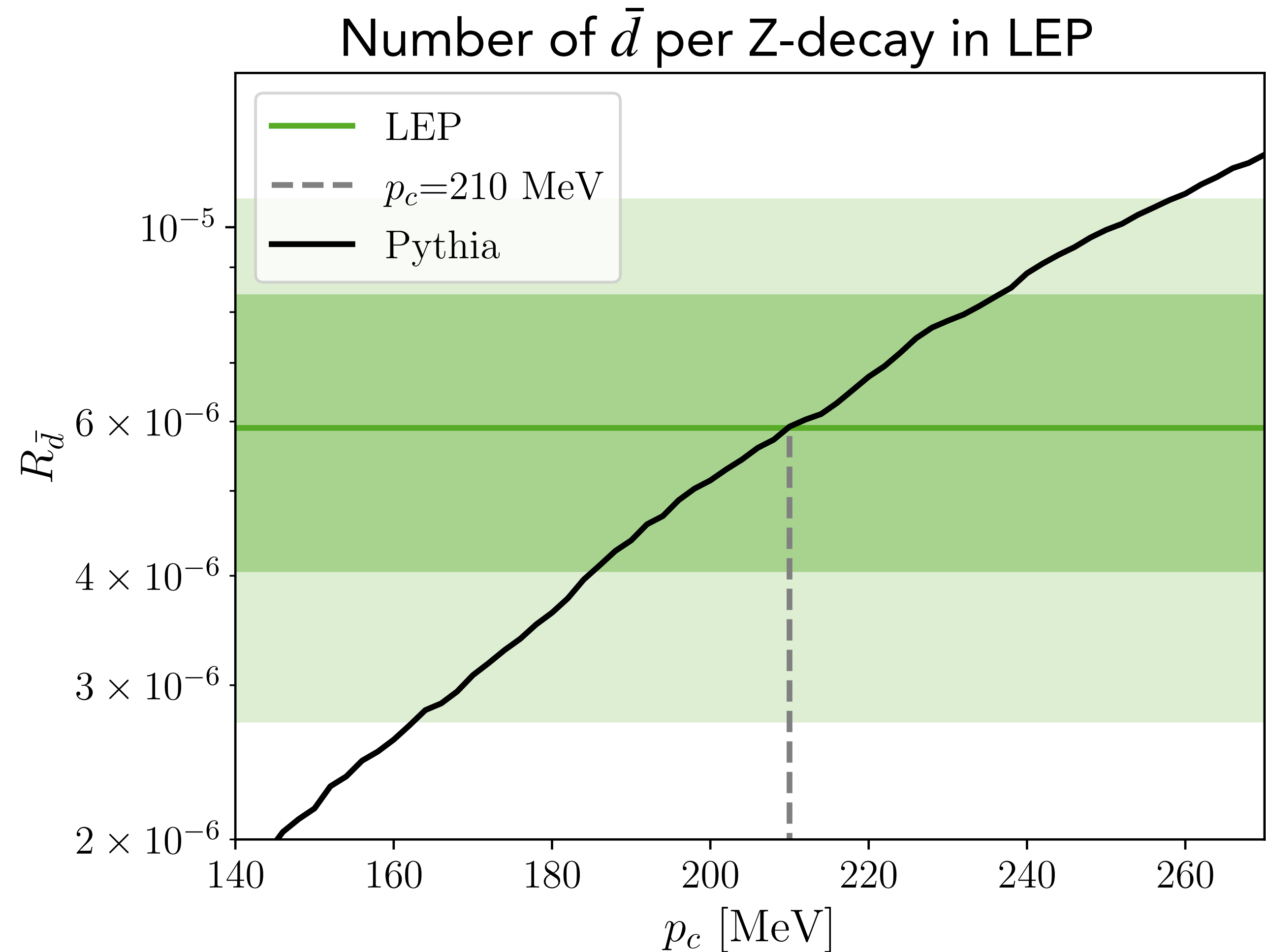


- Coalescence momentum p_c ,
determined from experiment

Antideuteron Production



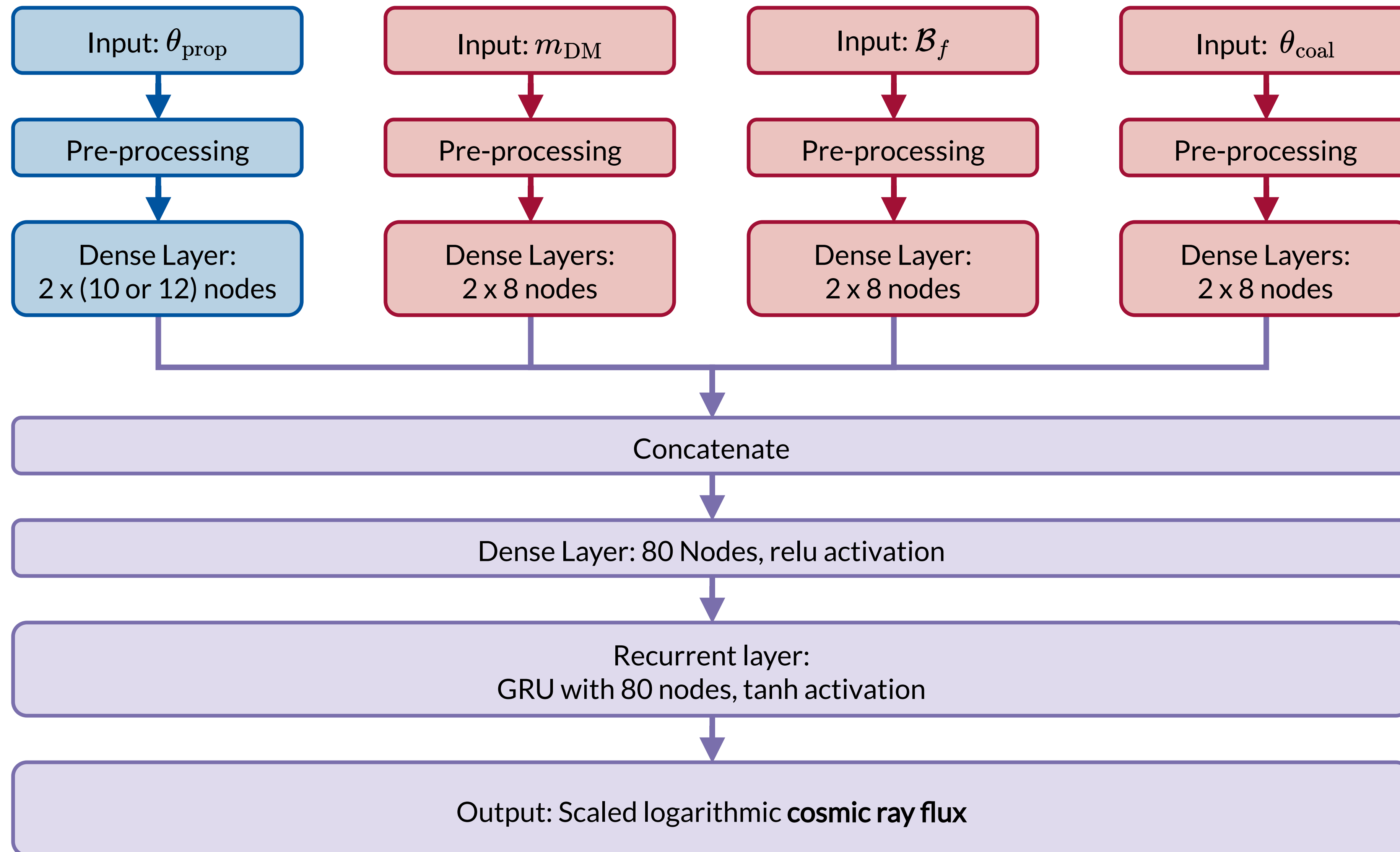
- Coalescence momentum p_c , determined from experiment
- Match number of antideuterons from simulated hadronic Z-decays to amount measured by LEP
- Spatial separation smaller than 2 fm



Propagation parameters

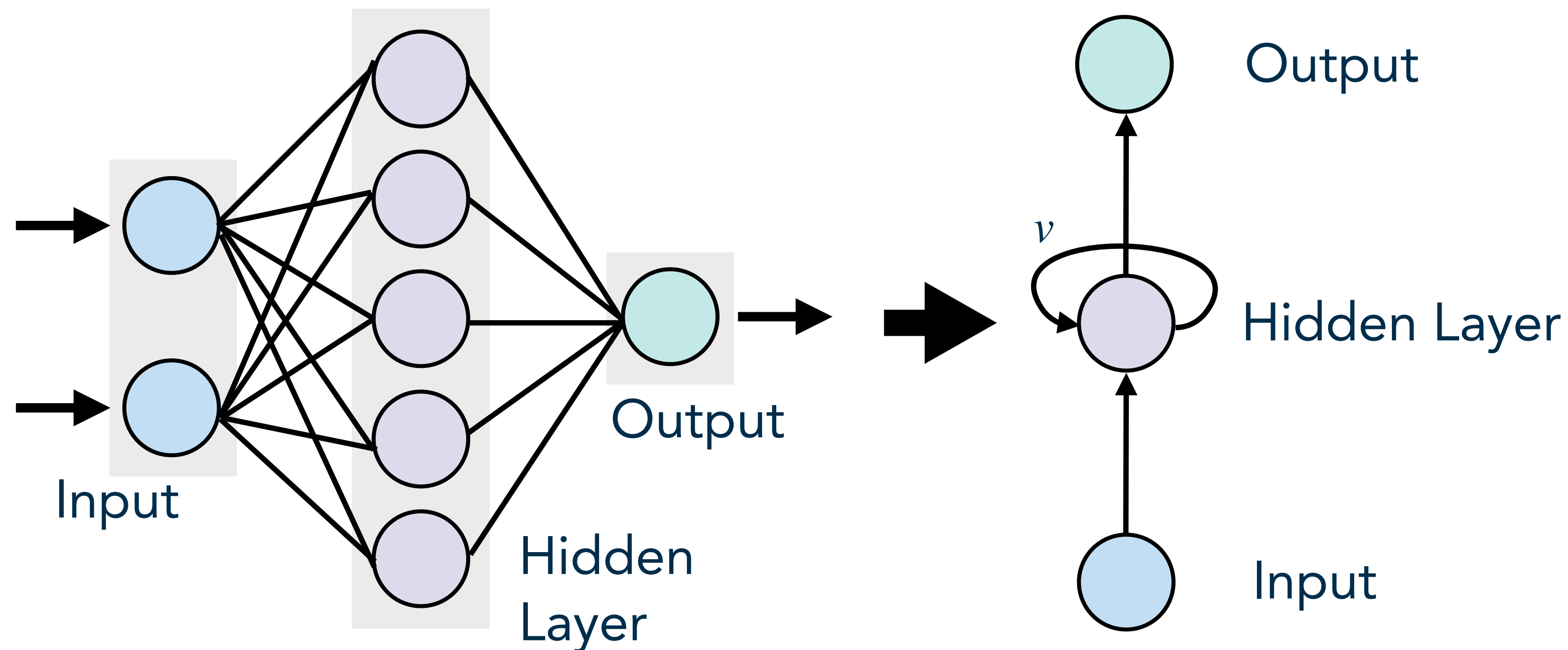
Parameters	Priors	DIFF.BRK	INJ.BRK
$\gamma_{1,p}$	1.2 – 2.1	✗	✓
γ_1	1.2 – 2.1	✗	✓
$\gamma_{2,p}$	2.1 – 2.6	✓	✓
γ_2	2.1 – 2.6	✓	✓
R_0 [GV]	1.0 – 20	✗	✓
s	0.1 – 0.7	✗	✓
D_0 [10^{28} cm ² /s]	0.5 – 10.0	✓	✓
δ_l	−1.0 – 0.5	✓	✗
δ	0.3 – 0.7	✓	✓
$\delta_h - \delta$	−0.2 – 0.0	✓	✓
$R_{D,0}$ [GV]	1.0 – 20.0	✓	✗
s_D	0.1 – 0.9	✓	✗
$R_{D,1}$ [10^3]	100 – 500	✓	✓
v_A [km/s]	0 – 30	✗	✓
$v_{0,c}$ [km/s]	0 – 60	✓	✓

Network Architecture DarkRayNet



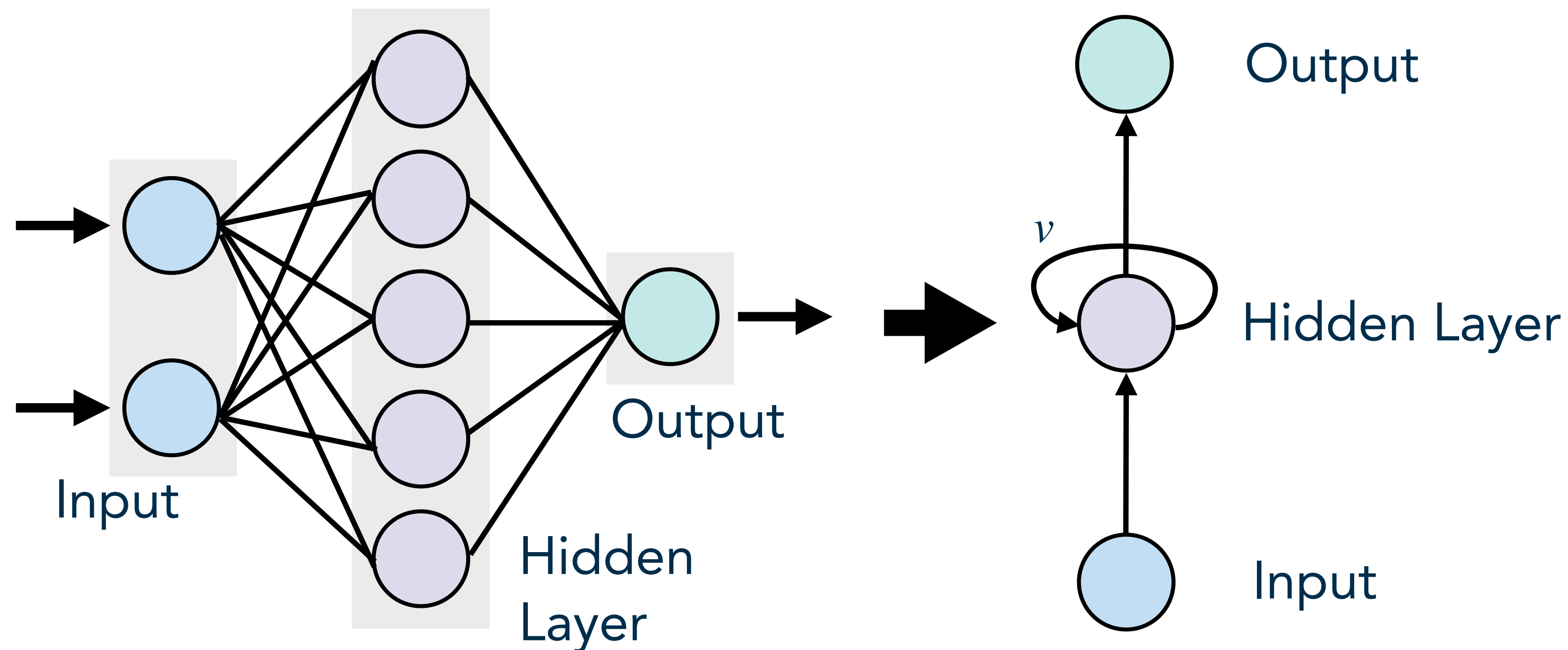
DarkRayNet

- Recurrent Neural Networks (RNN) use output of particular layer as input of the same layer → can account for correlations between energy bins
- Relative error of network $\mathcal{O}(10^{-2})$



DarkRayNet

- Recurrent Neural Networks (RNN) use output of particular layer as input of the same layer → can account for correlations between energy bins
- Relative error of network $\mathcal{O}(10^{-2})$



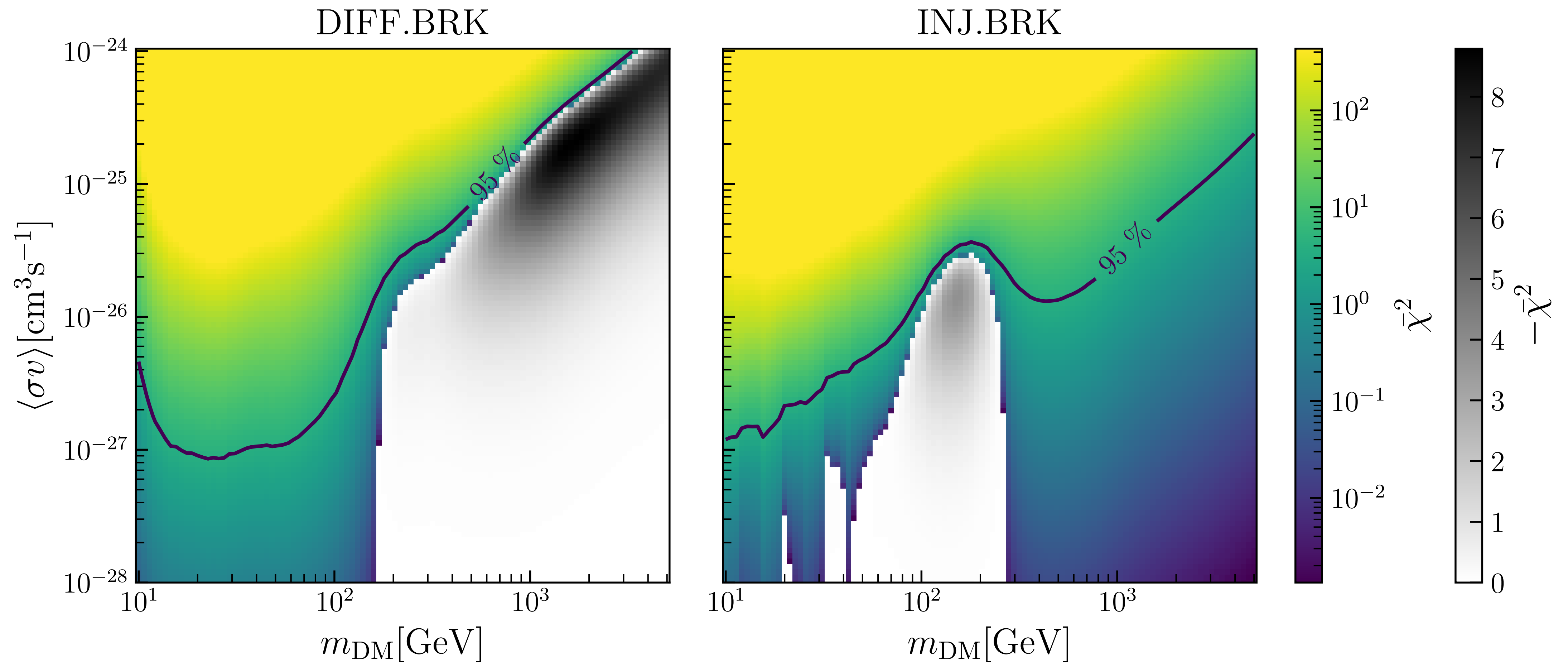
Network available in



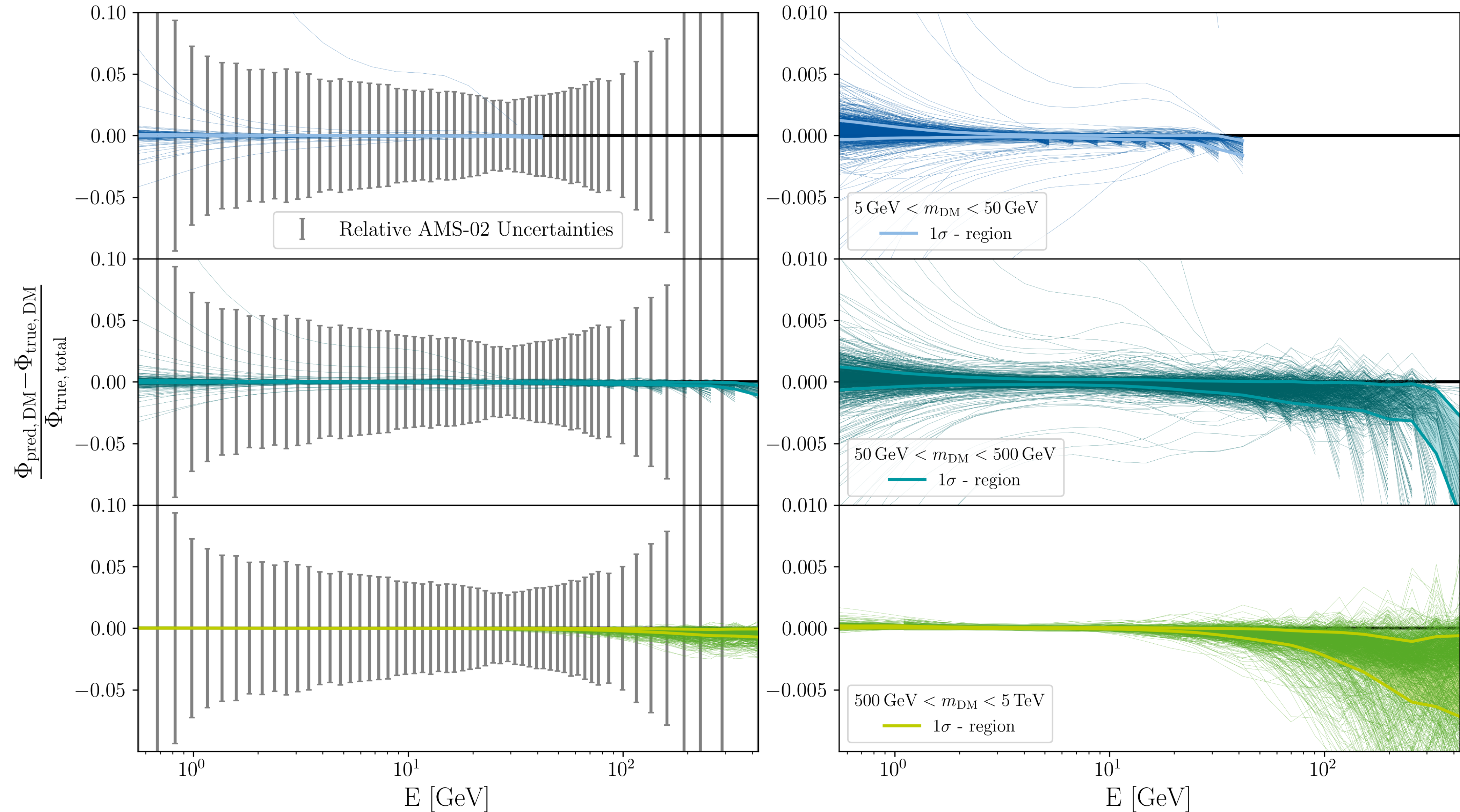
<https://github.com/kathrinnp/DarkRayNet>

Upper Limit on $\langle\sigma v\rangle$

- Upper limit using antiprotons and the DarkRayNet to marginalize over propagation parameters

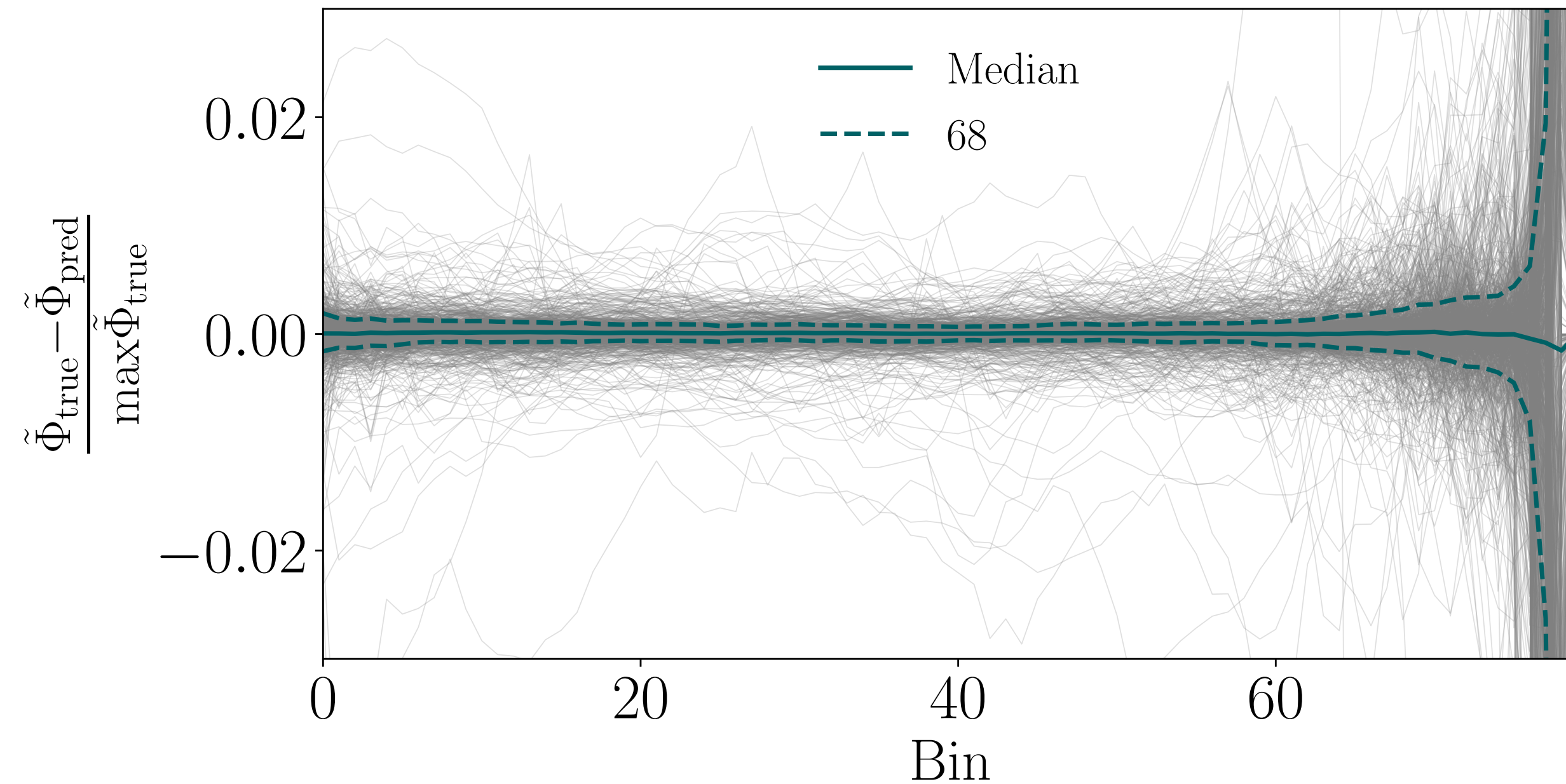


Performance DarkRayNet Antiprotons

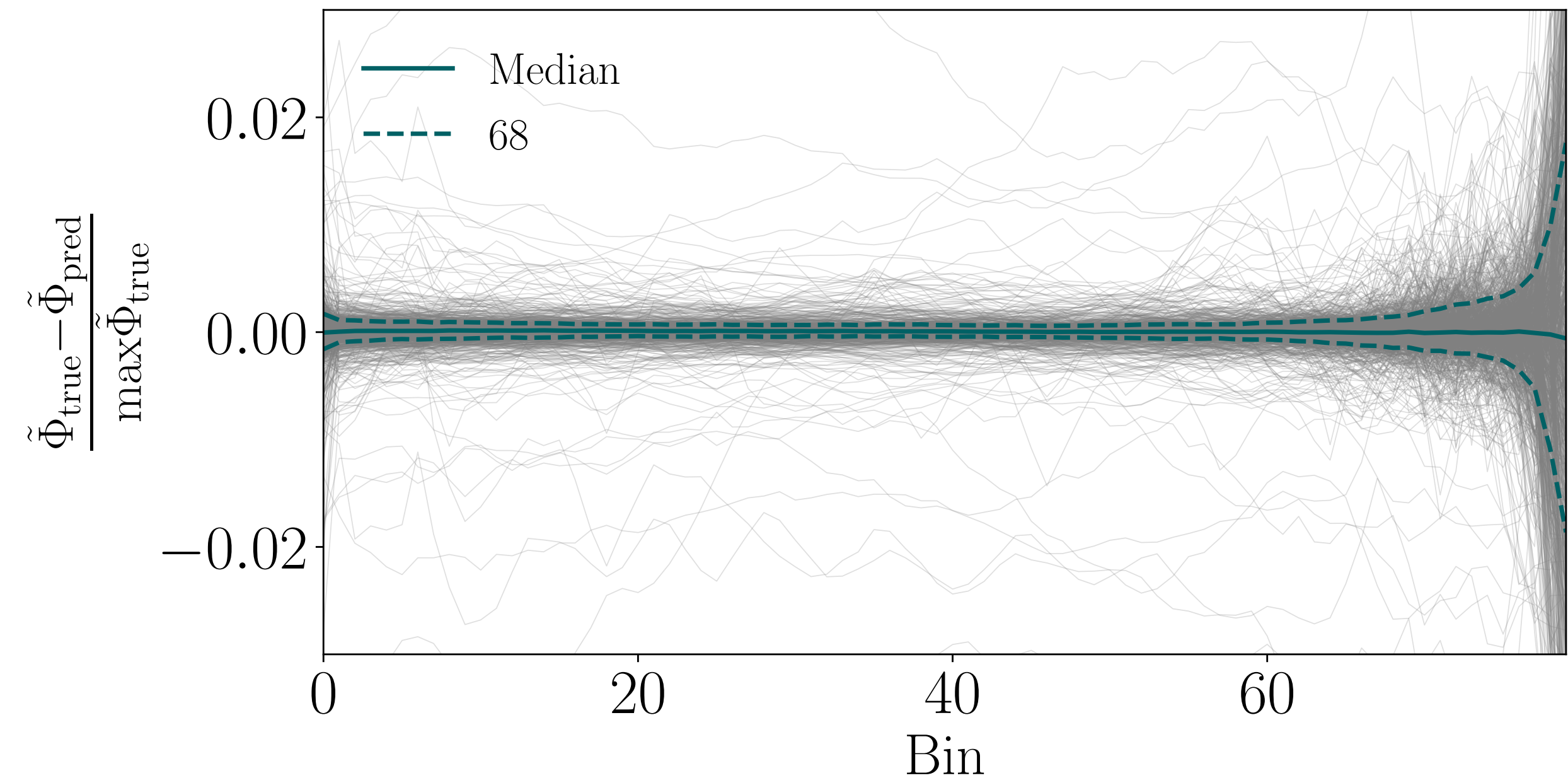


Performance DarkRayNet Antideuteron

DIFF.BRK model



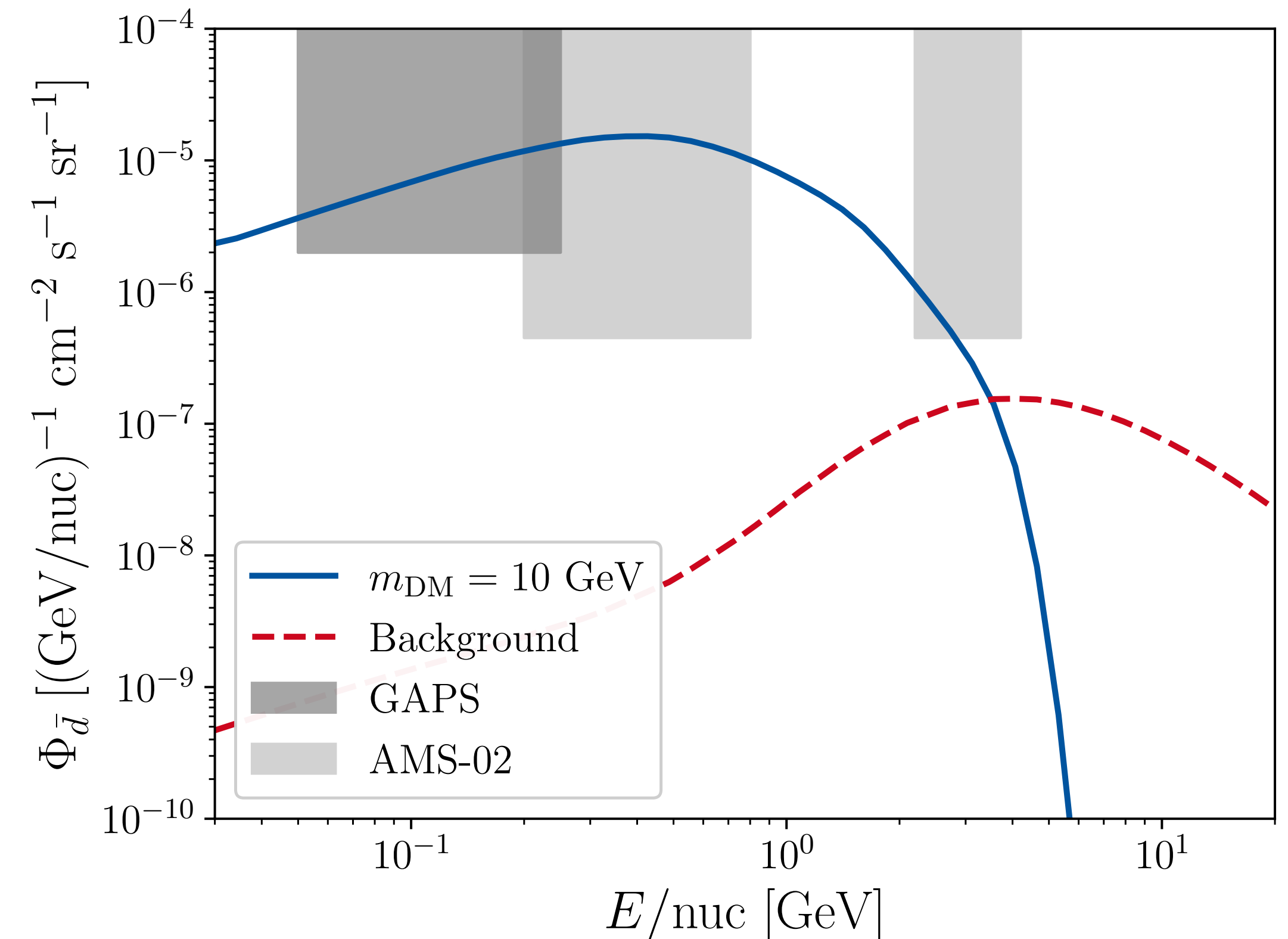
INJ.BRK model



- Relative difference of most transformed fluxes at most 6×10^{-4}
- Translates to relative error of $\mathcal{O}(10^{-2})$ in the actual flux

Prediction of Antineutron Sensitivity

- Generate fluxes for set of propagation parameters $\{\theta_{\text{prop},i}\}$ sampled from posterior of $p, \bar{p}$ and He fit
- Apply **force-field approximation** to account for solar modulation



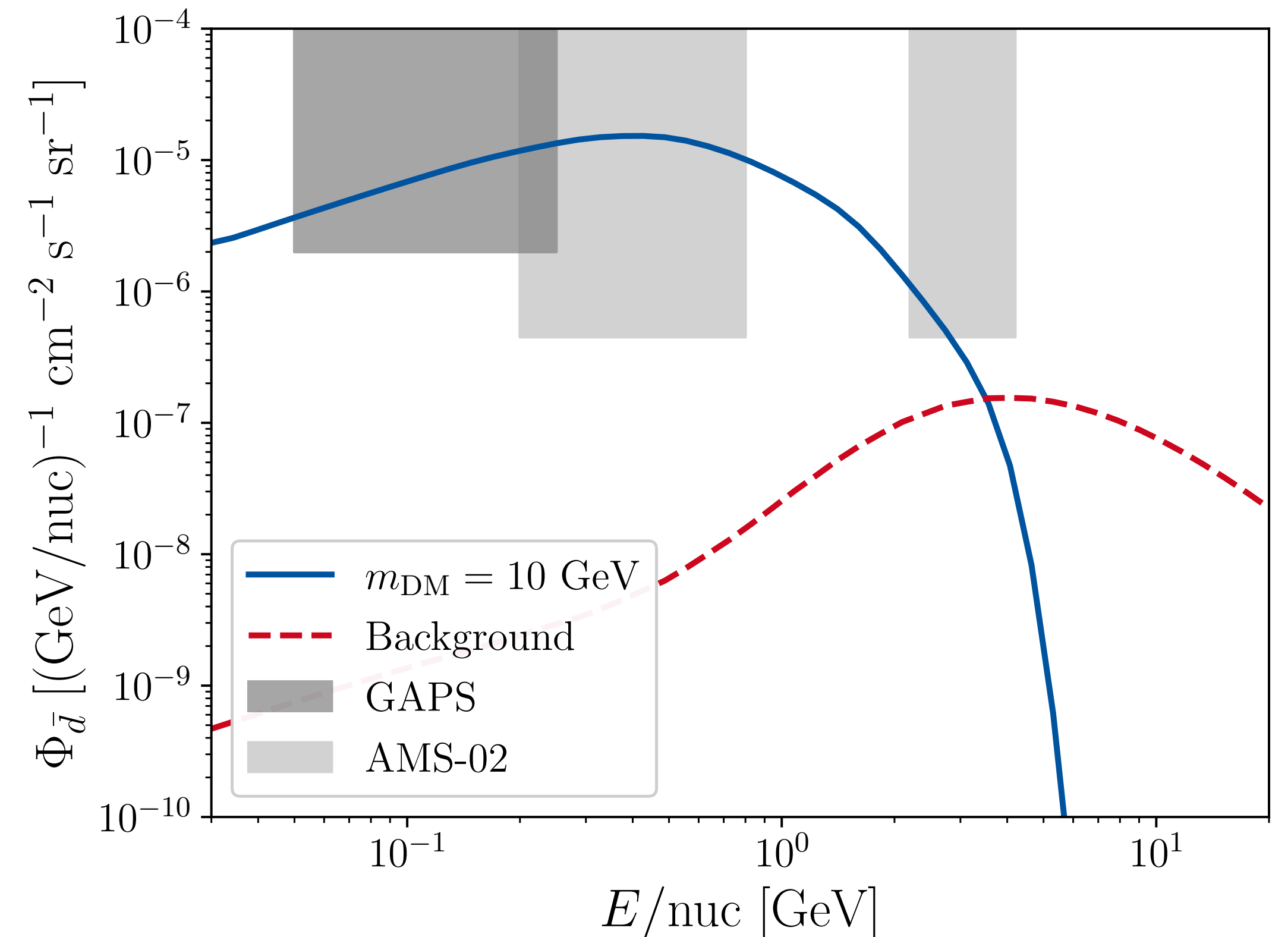
Prediction of Antineutron Sensitivity

- Generate fluxes for set of propagation parameters $\{\theta_{\text{prop},i}\}$ sampled from posterior of $p, \bar{p}$ and He fit
- Apply **force-field approximation** to account for solar modulation

- **Marginalize** over $\{\theta_{\text{prop},i}\}$:

$$\langle \Phi_{\bar{d}} \rangle = \frac{\sum_i \Phi_{\bar{d},i} \frac{\mathcal{L}_{\text{DM}}(\theta_{\text{prop},i}, x_{\text{DM}})}{\mathcal{L}(\theta_{\text{prop},i})}}{\sum_i \frac{\mathcal{L}_{\text{DM}}(\theta_{\text{prop},i}, x_{\text{DM}})}{\mathcal{L}(\theta_{\text{prop},i})}}$$

- Calculate **sensitivity factor**: $\frac{\langle \Phi_{\bar{d}} \rangle}{\Phi_{\text{exp.}}}$



Propagation Network Performance INJ.BRK

