Stockholm University

Oskar Klein centre

Constraining dark matter annihilation with cosmic-ray antiprotons using neural networks

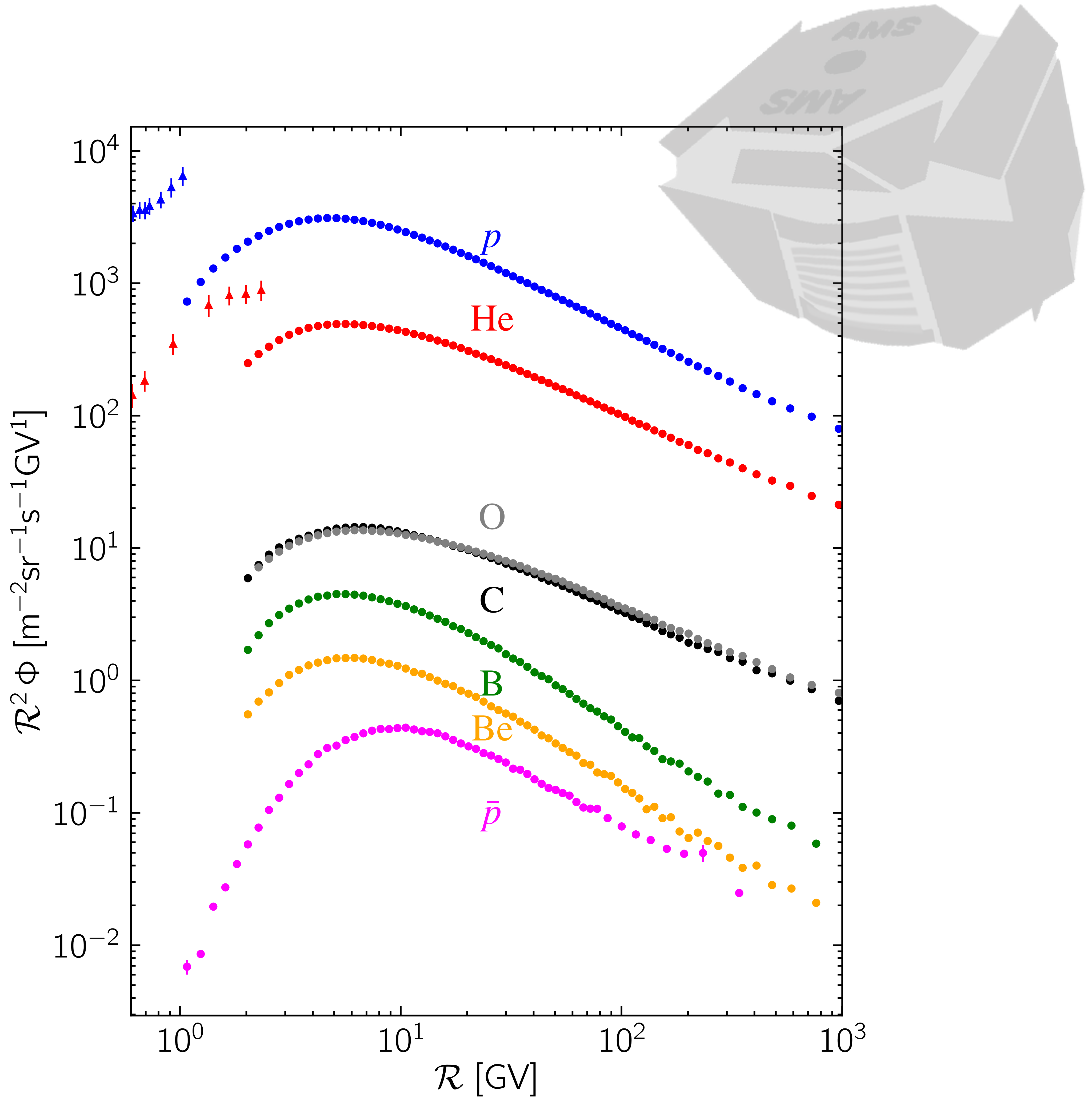
Felix Kahlhoefer, **Michael Korsmeier**, Michael Krämer,
Silvia Manconi and Kathrin Nippel
[arXiv: 2107.12395]

Partikeldagarna 2021
Chalmers Conference Centre

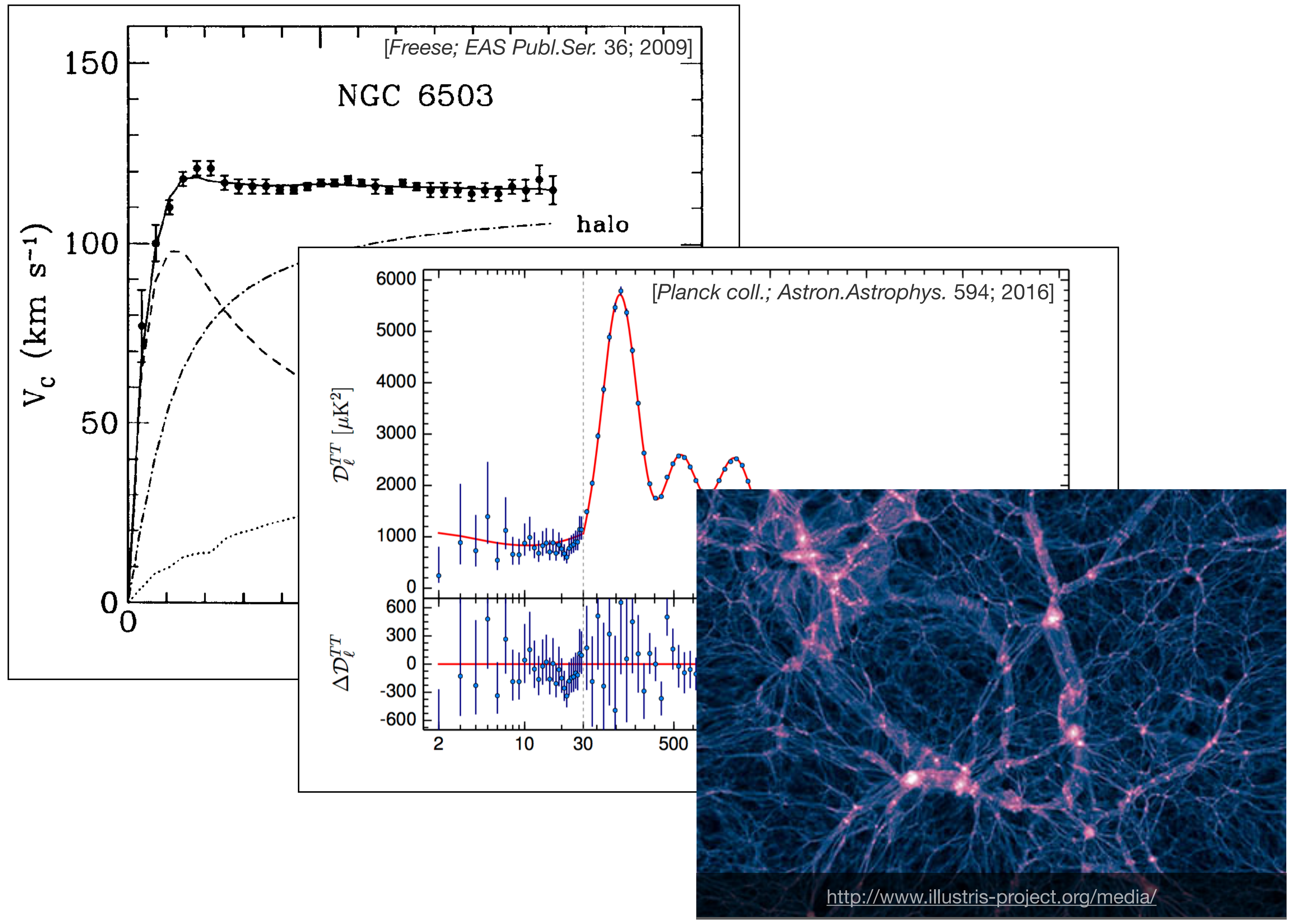
2021/11/23

Outline

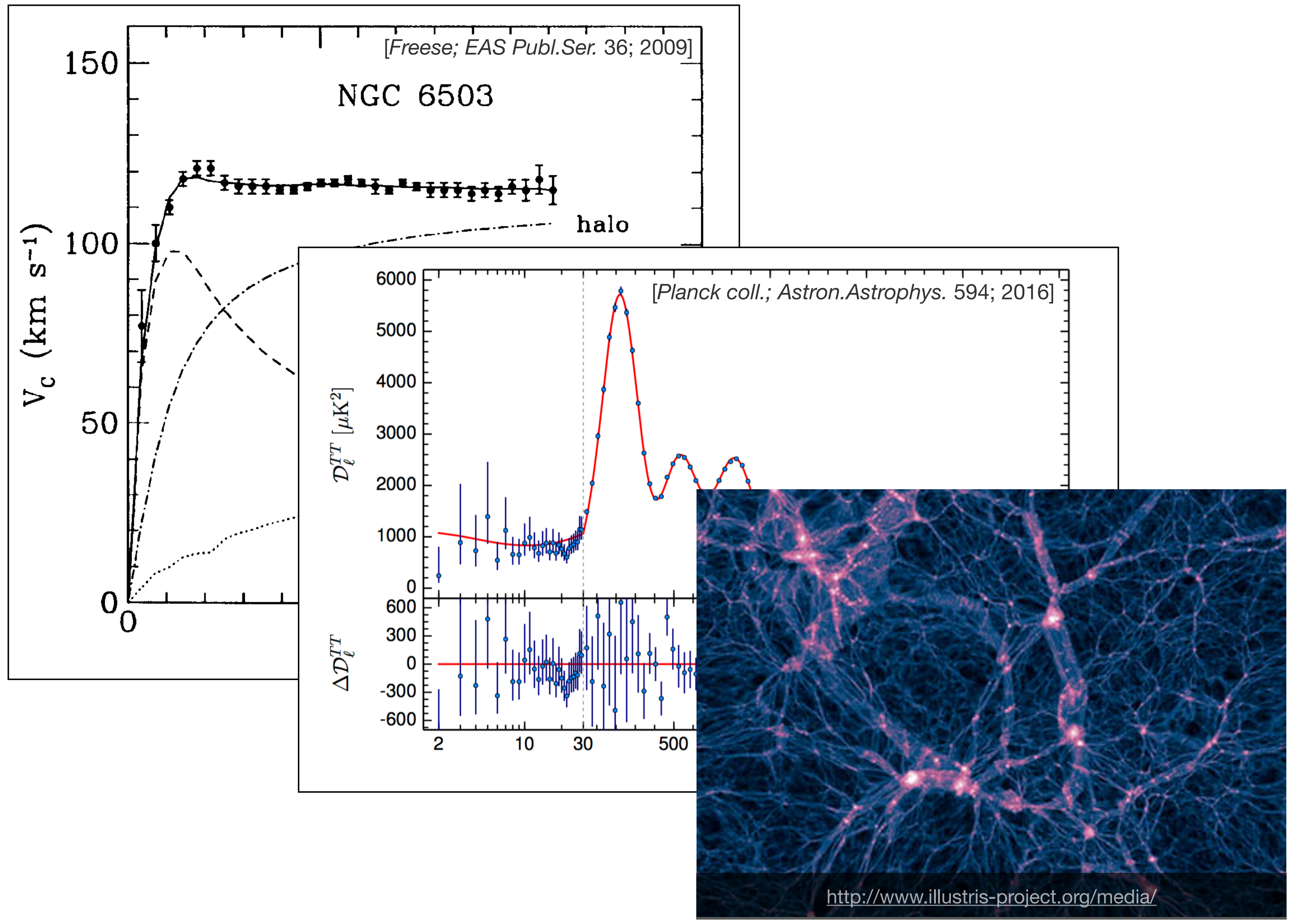
- Motivation and introduction
- Traditional approach
- New methods: RNNs and importance sampling
- Application: scalar singlet DM
- Conclusions



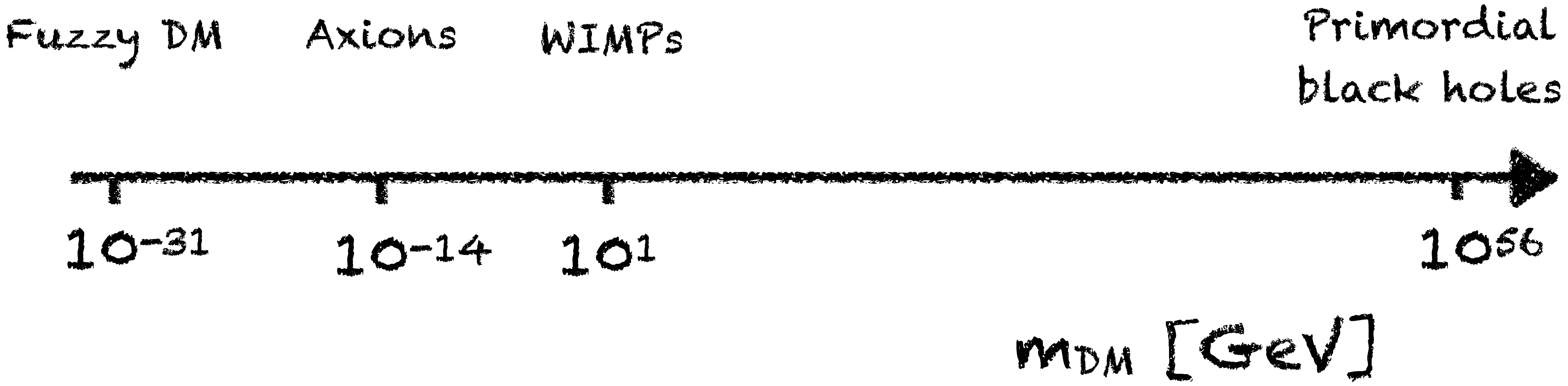
Motivation



Gravitational evidence at various scales is overwhelming.

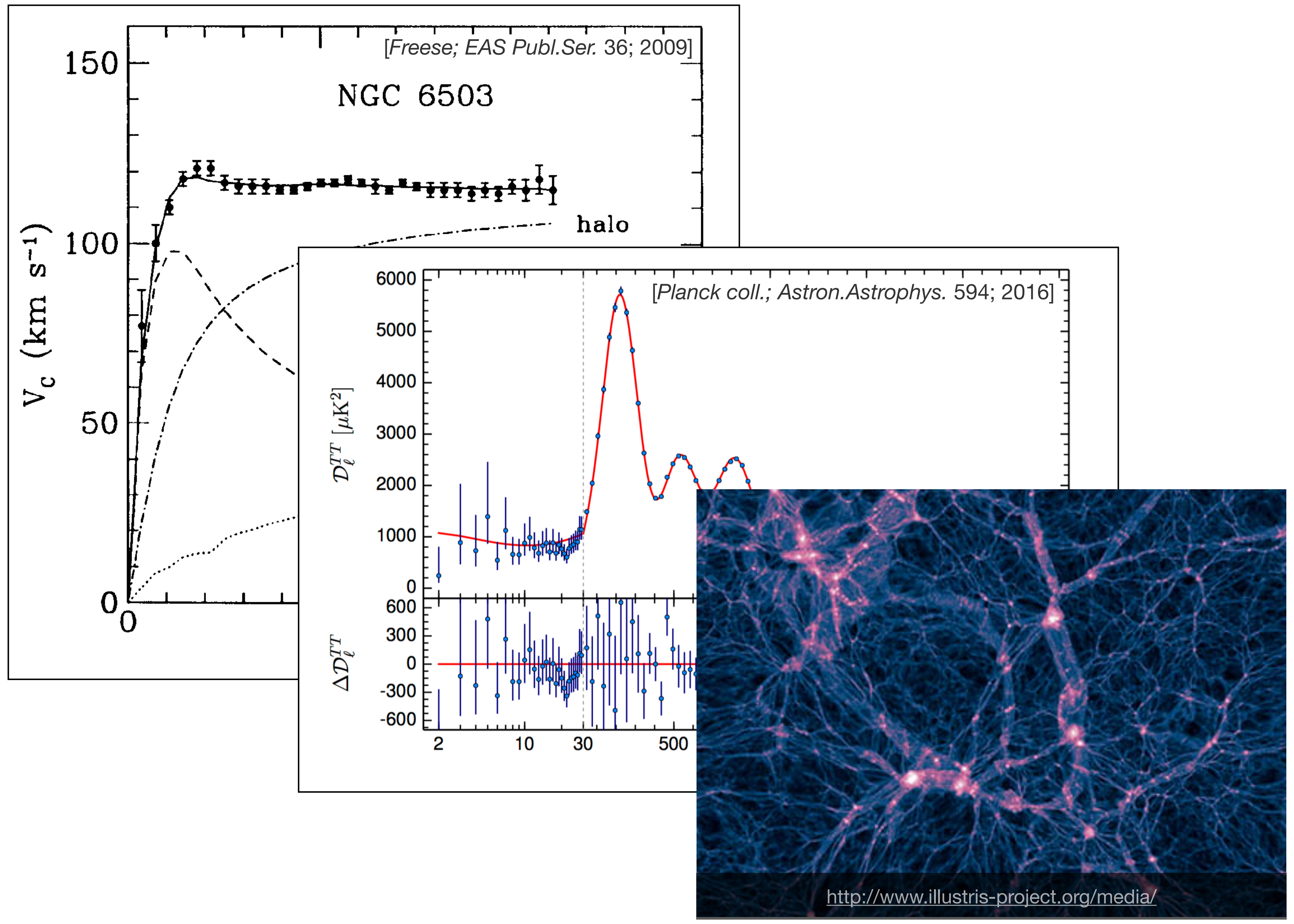


Gravitational evidence at various scales is overwhelming.

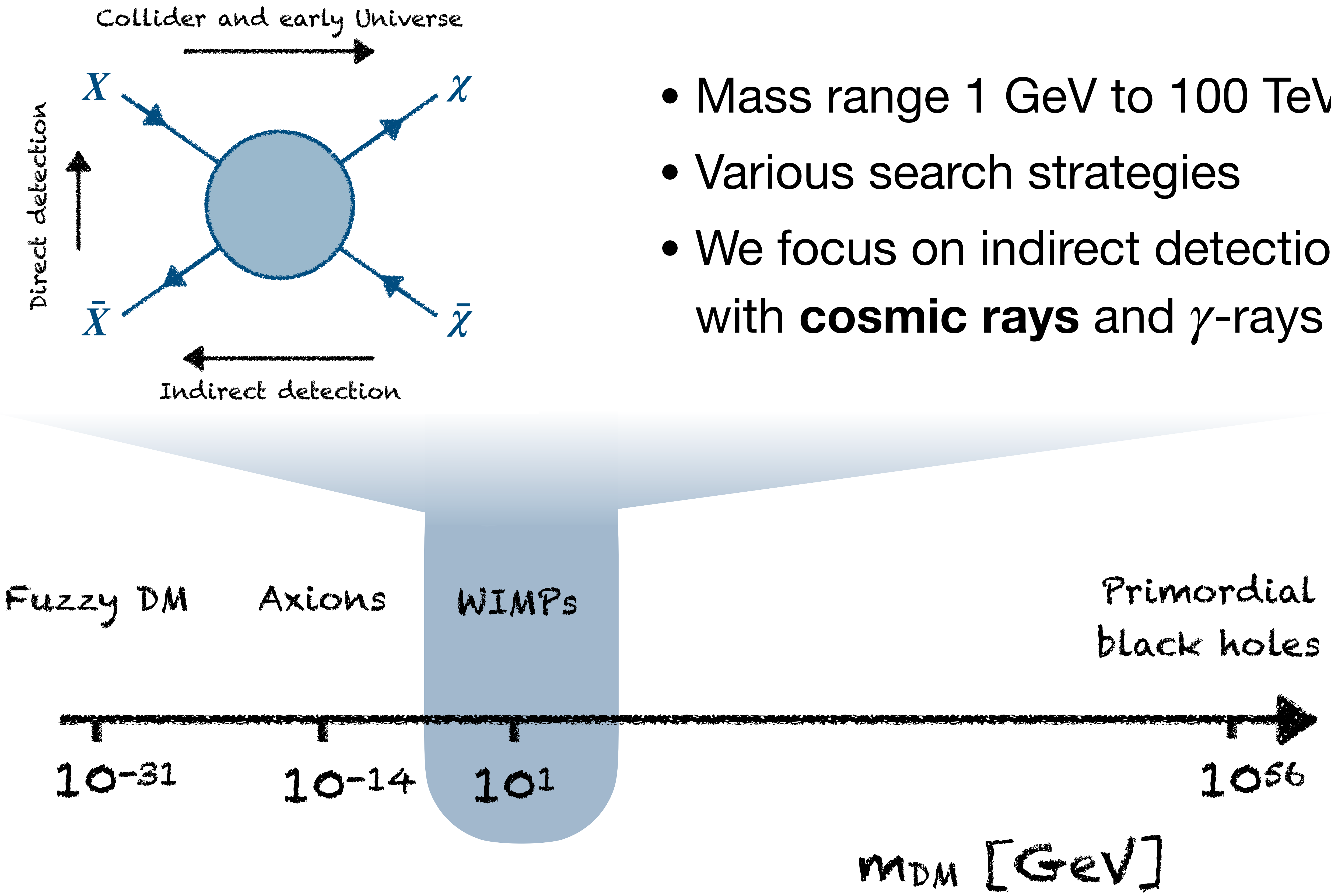


Nature of dark matter remains unknown!

Motivation



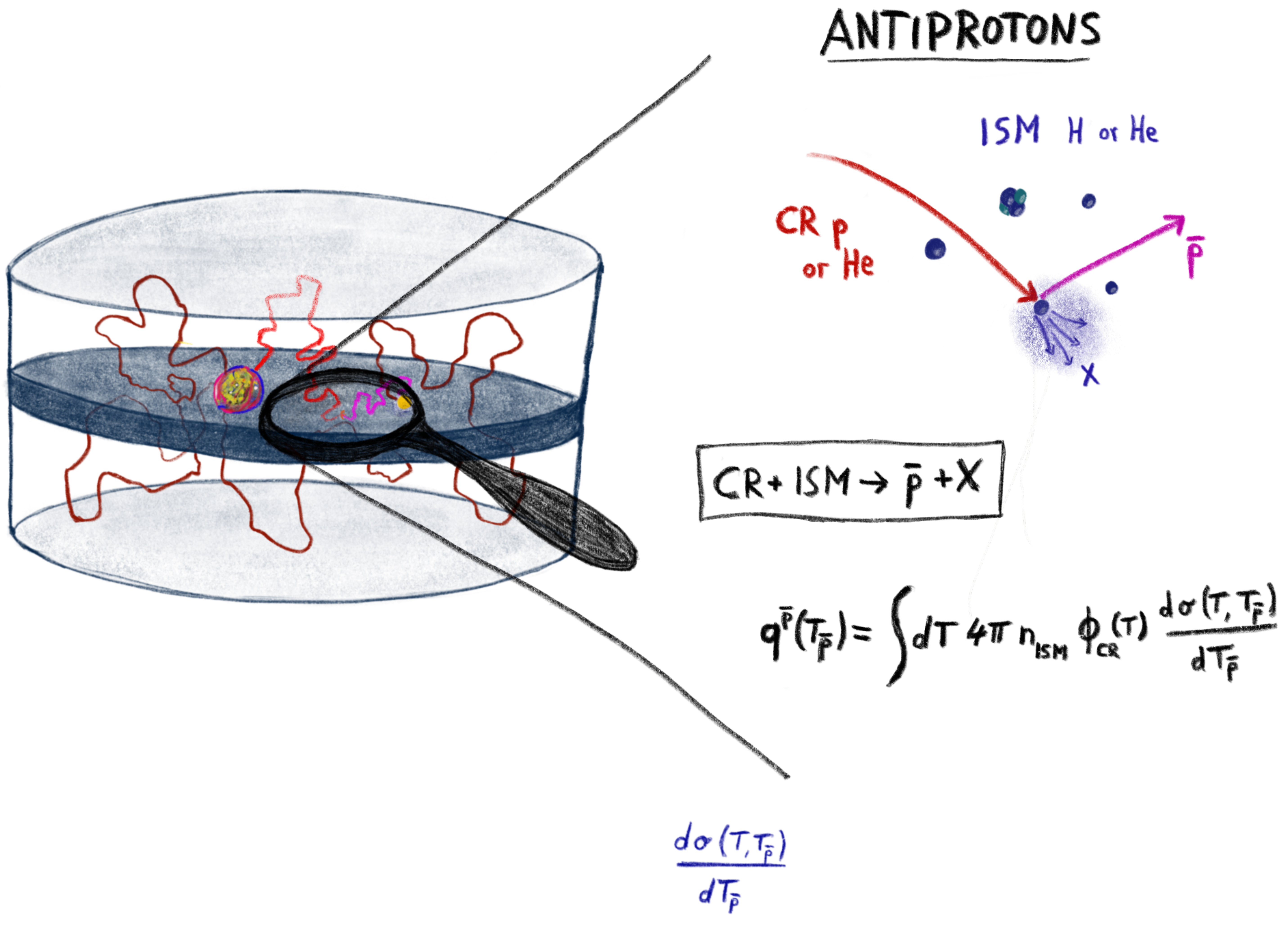
Gravitational evidence at various scales is overwhelming.



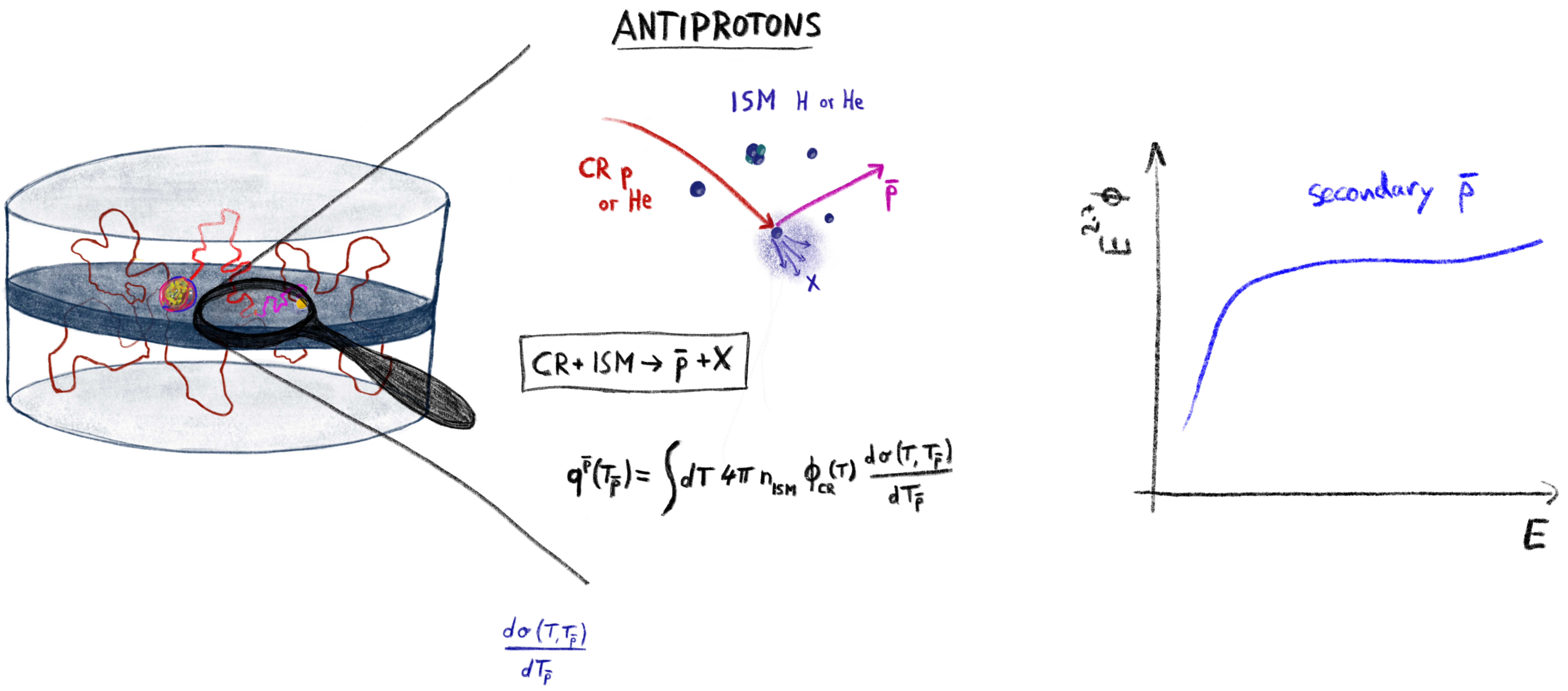
- Mass range 1 GeV to 100 TeV
- Various search strategies
- We focus on indirect detection with **cosmic rays** and γ-rays

Nature of dark matter remains unknown!

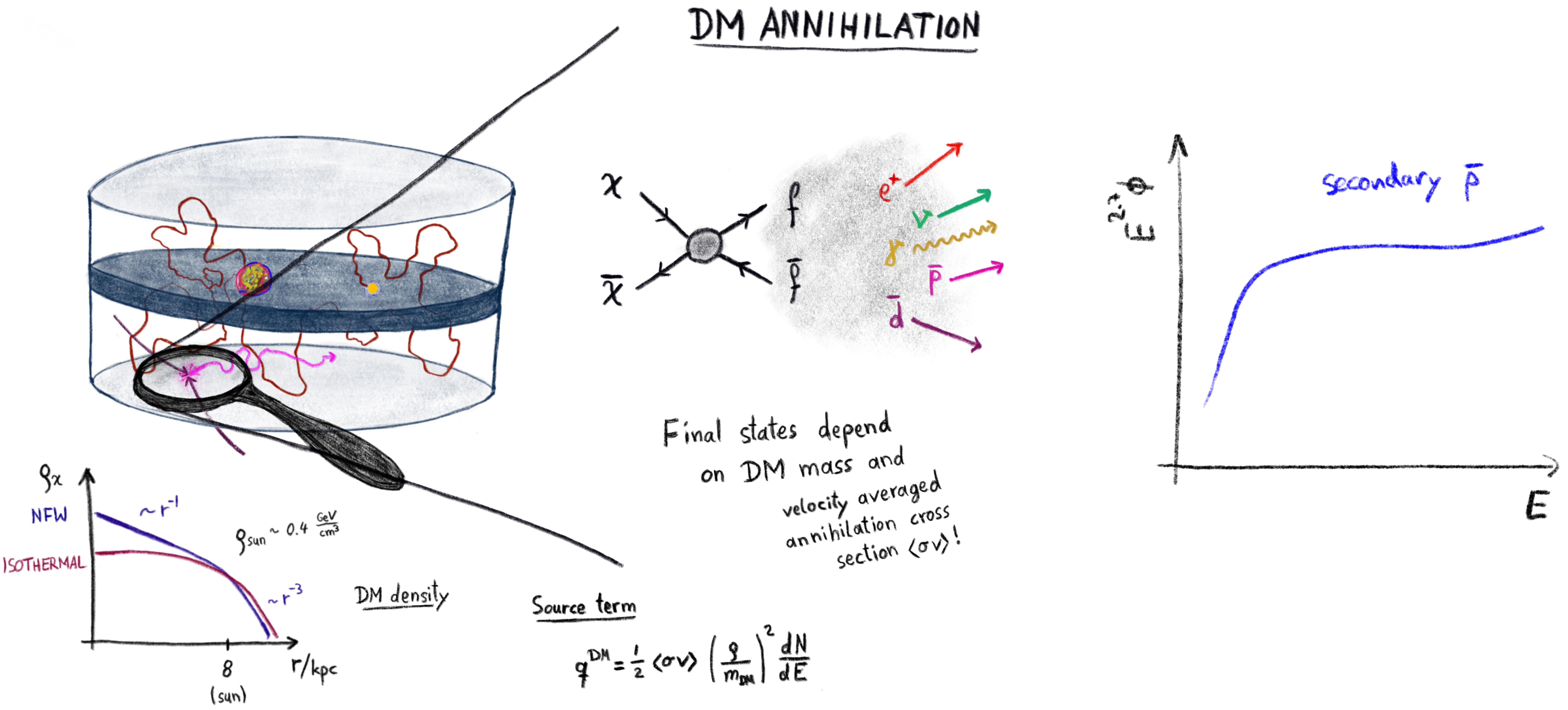
Antiprotons in cosmic rays



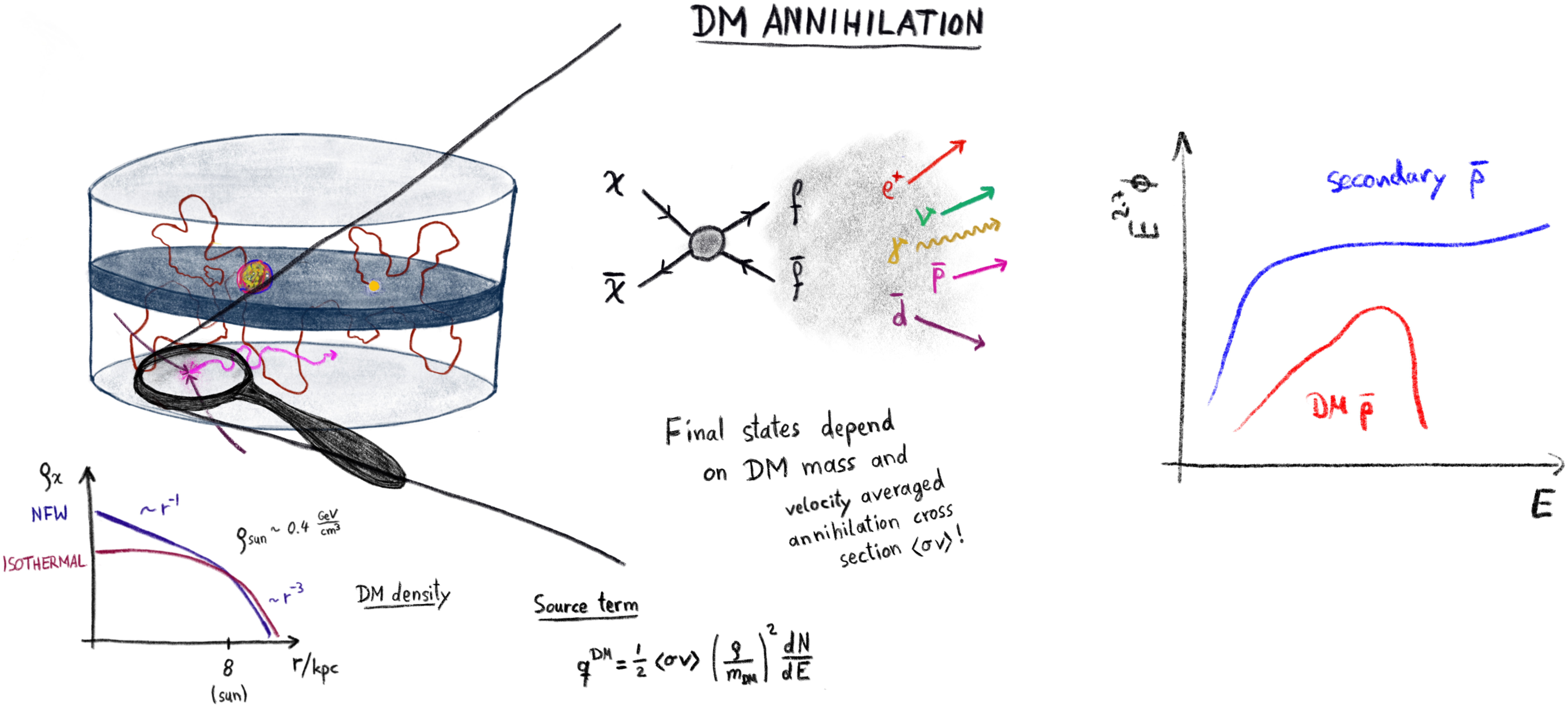
Antiprotons in cosmic rays



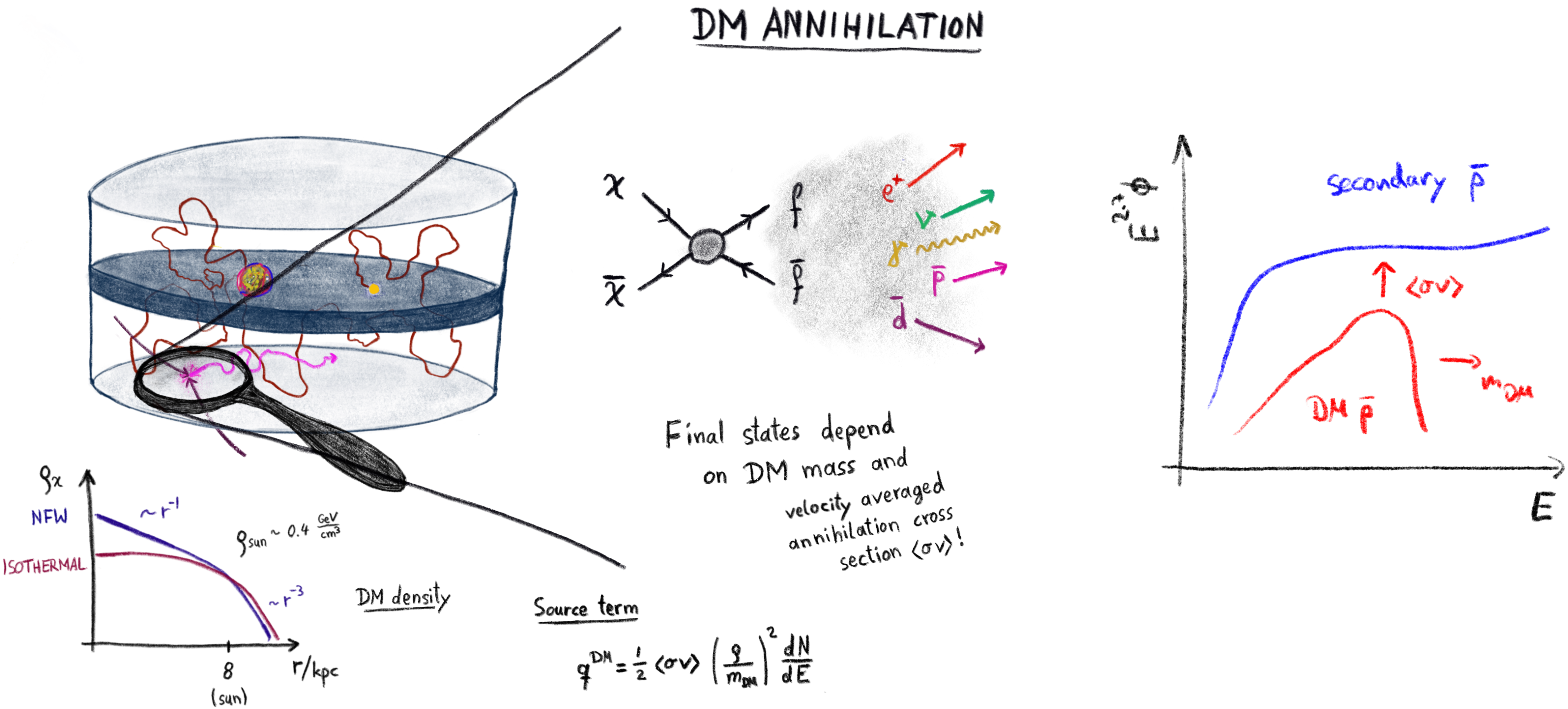
Antiprotons in cosmic rays



Antiprotons in cosmic rays



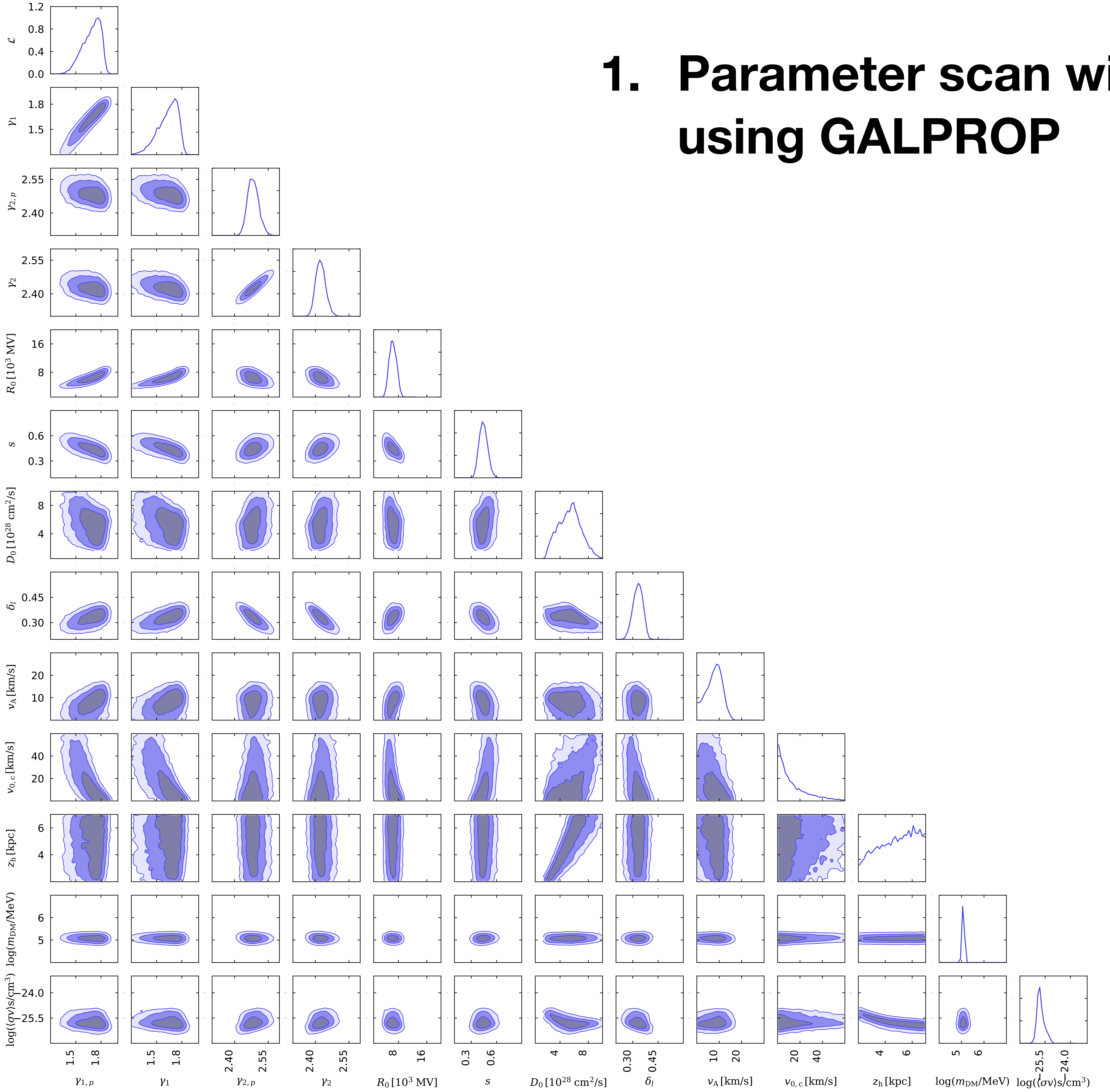
Antiprotons in cosmic rays



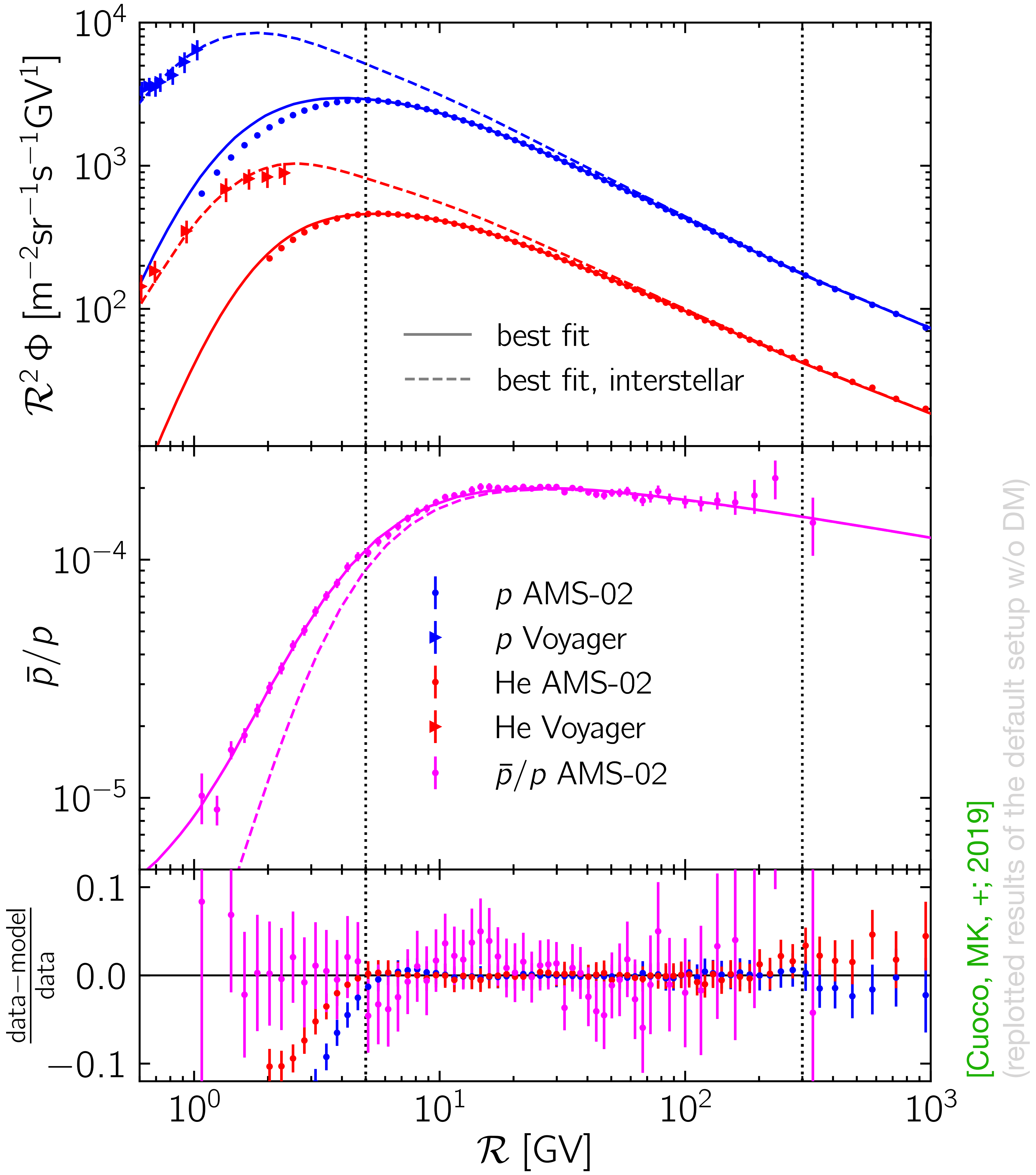
$$q^{DM} = \frac{1}{2} \langle \sigma v \rangle \left(\frac{\rho}{m_{DM}} \right)^2 \frac{dN}{dE}$$

Traditional method

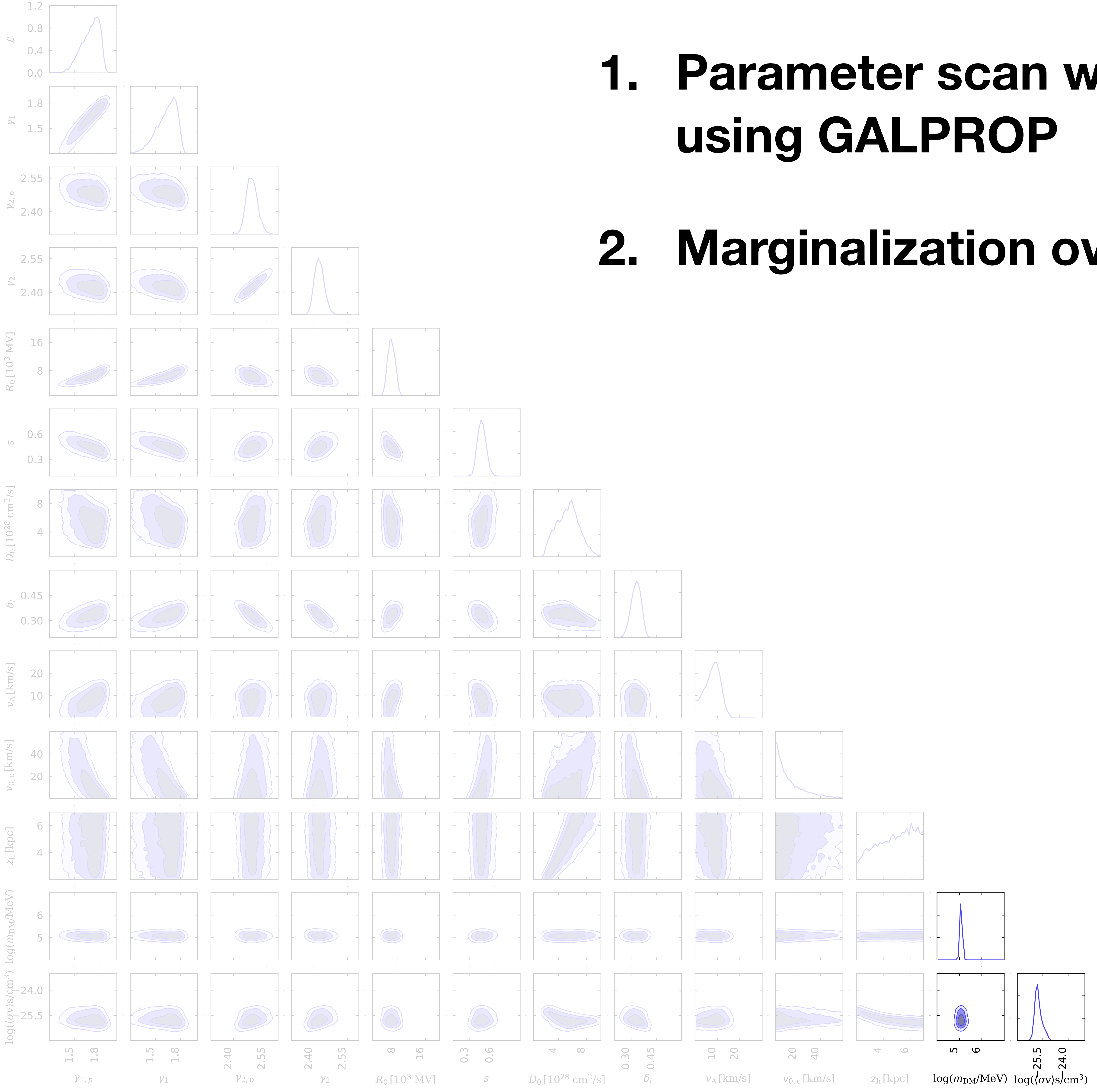
Traditional method



1. **Parameter scan with $\mathcal{O}(10^6)$ likelihood evaluations using GALPROP**



Traditional method



1. **Parameter scan with $\mathcal{O}(10^6)$ likelihood evaluations using GALPROP**

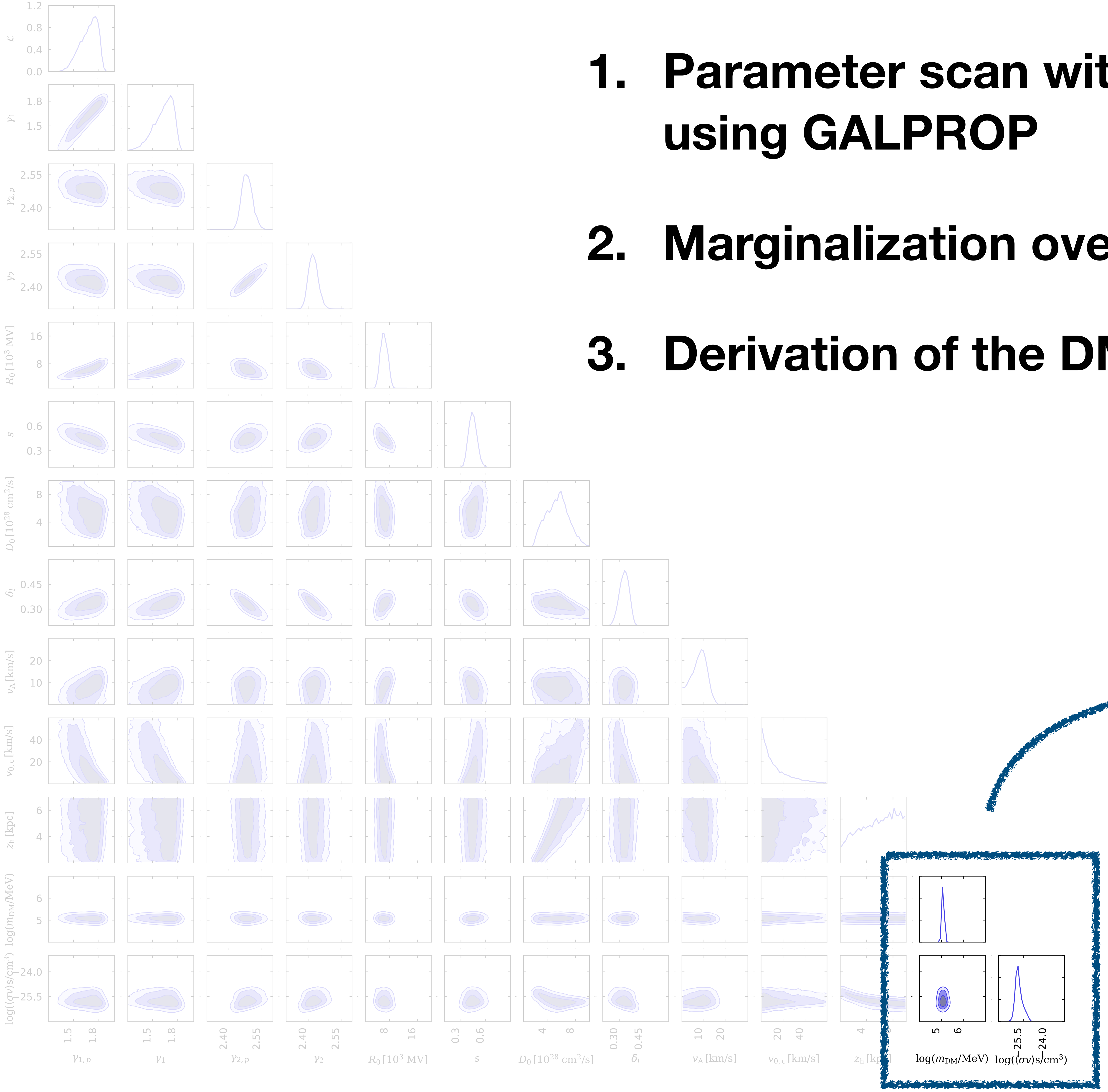
2. **Marginalization over CR parameters**

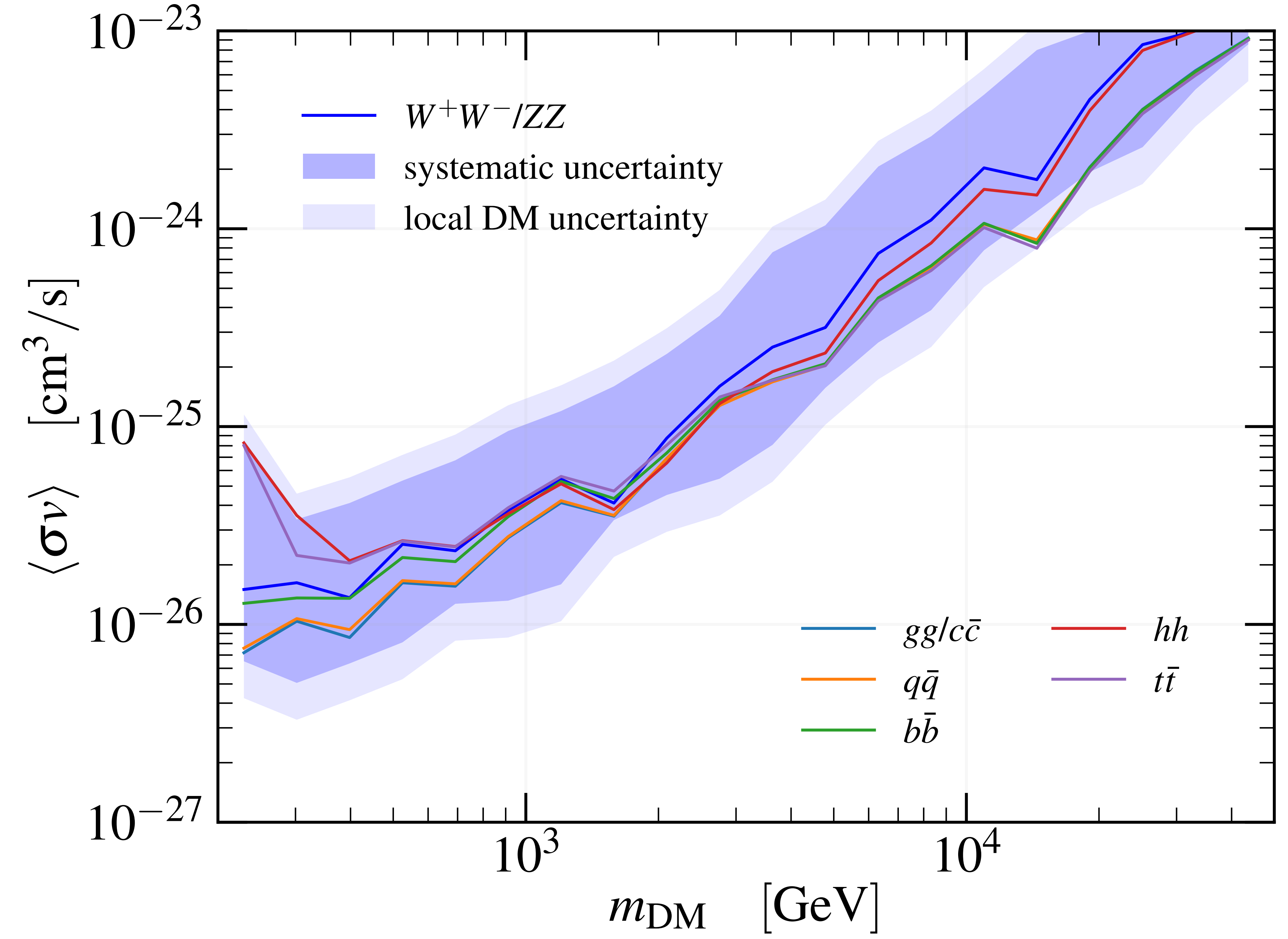
Traditional method

1. **Parameter scan with $\mathcal{O}(10^6)$ likelihood evaluations using GALPROP**

2. **Marginalization over CR parameters**

3. **Derivation of the DM limit**





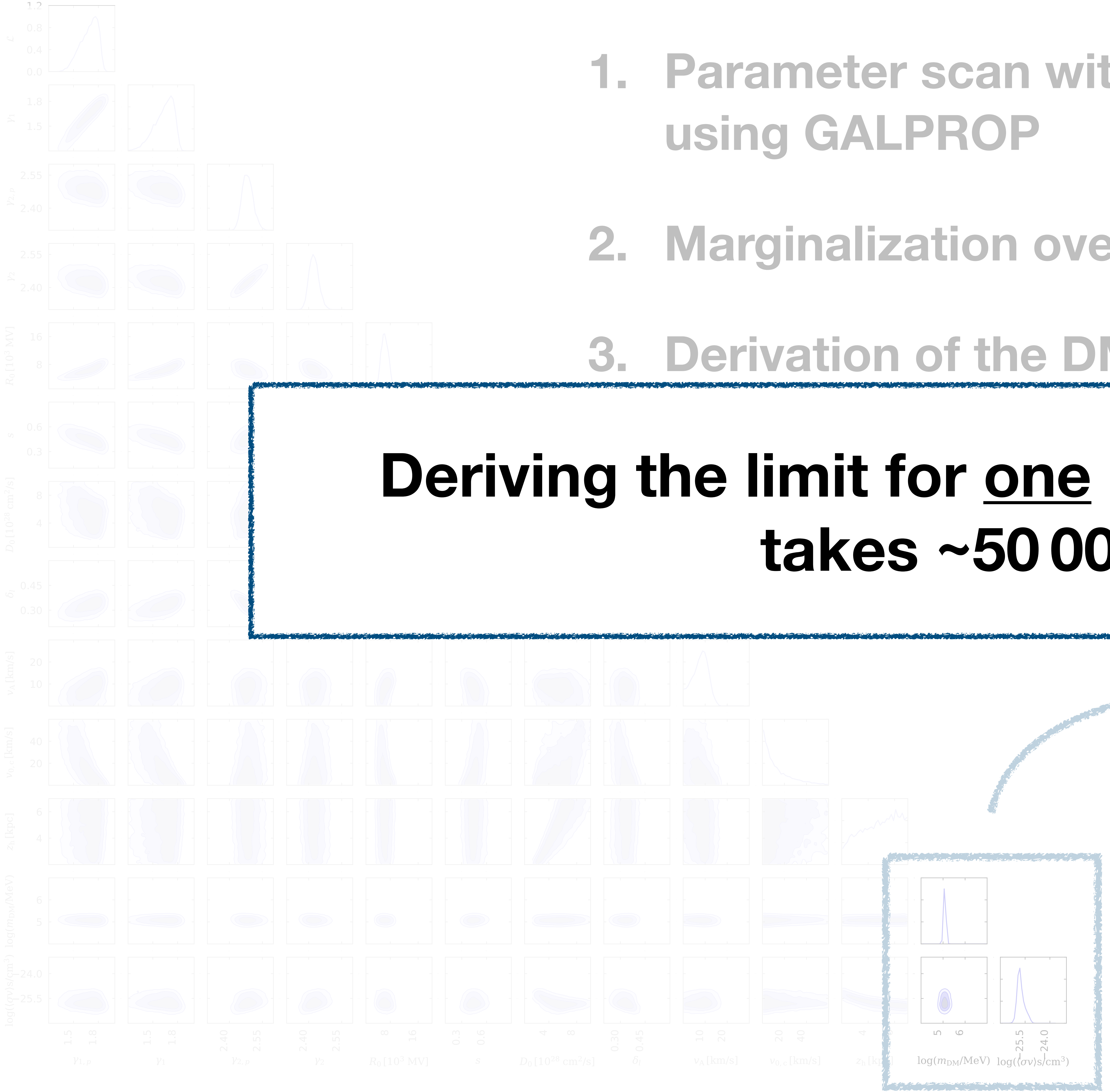
[Cuoco, Heisig, MK, Krämer; 2018]

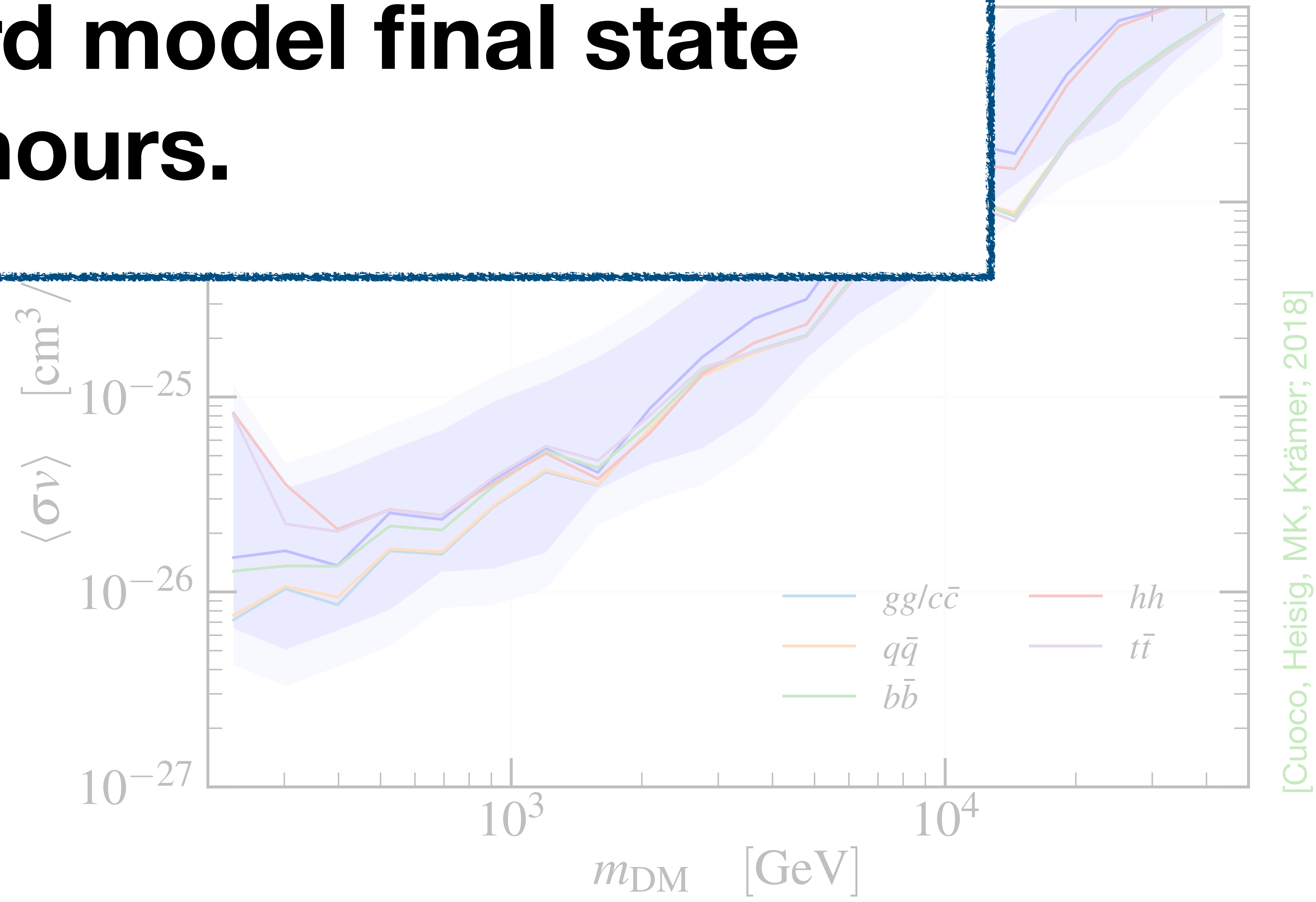
Traditional method

1. Parameter scan with $\mathcal{O}(10^6)$ likelihood evaluations using GALPROP

2. Marginalization over CR parameters

3. Derivation of the DM limit

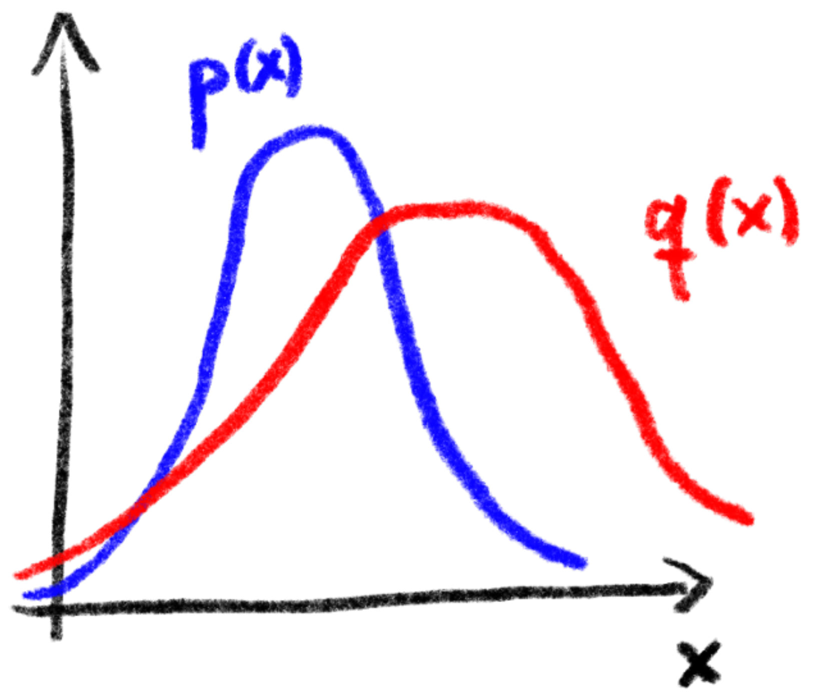




Deriving the limit for <u>one</u> standard model final state takes ~50 000 cpu-hours.

New methods to derive DM limits

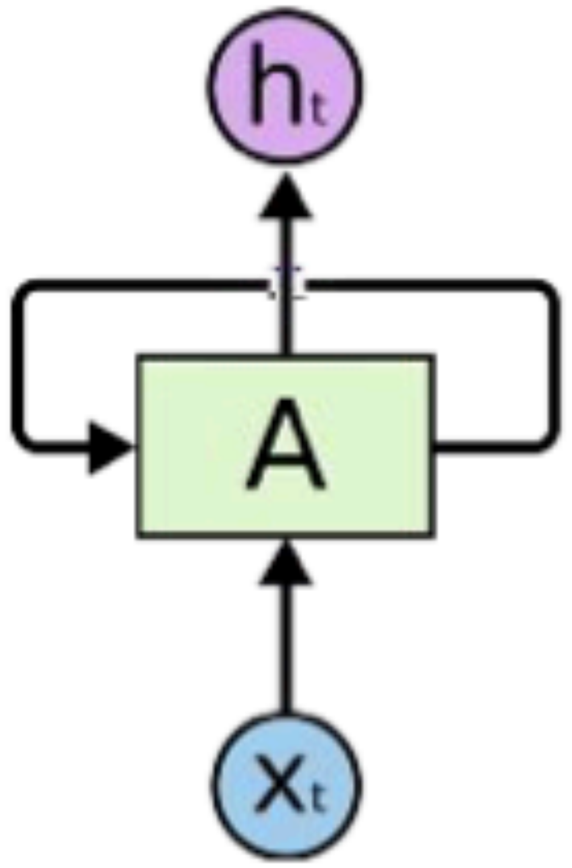
Importance Sampling to marginalize over CR propagation parameters



$$\{x_i\} \sim q(x)$$

$$E_x[p(x)] = \frac{1}{N} \sum_{i=1}^{N} x_i \frac{p(x_i)}{q(x_i)}$$

RNNs to replace GALPROP for an arbitrary DM model



Importance sampling

$\mathscr{L}_1(\boldsymbol{\theta})$ Likelihood **without DM** signal

$\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$ Likelihood **with DM** signal

$$\boldsymbol{\theta} = \{\gamma_{1,p}, \gamma_{1,\mathrm{He}}, \gamma_{2,p}, \gamma_{2,\mathrm{He}}, R_0, s, D_0, \delta, v_A, v_{0,c}, z_h, \varphi, \varphi_{\bar{p}}\}$$

$$\boldsymbol{x} = \{m_{\mathrm{DM}}, \langle\sigma v\rangle_{gg}, \langle\sigma v\rangle_{q\bar{q}}, \langle\sigma v\rangle_{c\bar{c}}, \langle\sigma v\rangle_{b\bar{b}}, \langle\sigma v\rangle_{t\bar{t}}, \langle\sigma v\rangle_{hh}, \langle\sigma v\rangle_{WW}, \langle\sigma v\rangle_{ZZ}\}$$

$$\{\boldsymbol{\theta}_i\} \sim \mathscr{L}_1(\boldsymbol{\theta})$$

$$\mathscr{L}_2(\boldsymbol{x}) = \int \mathrm{d}\theta\, \mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$$

Importance sampling

$\mathscr{L}_1(\boldsymbol{\theta})$ Likelihood **without DM** signal

$\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$ Likelihood **with DM** signal

$$\boldsymbol{\theta} = \{\gamma_{1,p}, \gamma_{1,\mathrm{He}}, \gamma_{2,p}, \gamma_{2,\mathrm{He}}, R_0, s, D_0, \delta, v_A, v_{0,c}, z_h, \varphi, \varphi_{\bar{p}}\}$$

$$\boldsymbol{x} = \{m_{\mathrm{DM}}, \langle\sigma v\rangle_{gg}, \langle\sigma v\rangle_{q\bar{q}}, \langle\sigma v\rangle_{c\bar{c}}, \langle\sigma v\rangle_{b\bar{b}}, \langle\sigma v\rangle_{t\bar{t}}, \langle\sigma v\rangle_{hh}, \langle\sigma v\rangle_{WW}, \langle\sigma v\rangle_{ZZ}\}$$

$\{\boldsymbol{\theta}_i\} \sim \mathscr{L}_1(\boldsymbol{\theta})$ ⟵ **We have this!**

$$\mathscr{L}_2(\boldsymbol{x}) = \int \mathrm{d}\theta \, \mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$$

Importance sampling

$\mathscr{L}_1(\boldsymbol{\theta})$ Likelihood **without DM** signal

$\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$ Likelihood **with DM** signal

$$\boldsymbol{\theta} = \{\gamma_{1,p}, \gamma_{1,\mathrm{He}}, \gamma_{2,p}, \gamma_{2,\mathrm{He}}, R_0, s, D_0, \delta, v_A, v_{0,c}, z_h, \varphi, \varphi_{\bar{p}}\}$$

$$\boldsymbol{x} = \{m_{\mathrm{DM}}, \langle\sigma v\rangle_{gg}, \langle\sigma v\rangle_{q\bar{q}}, \langle\sigma v\rangle_{c\bar{c}}, \langle\sigma v\rangle_{b\bar{b}}, \langle\sigma v\rangle_{t\bar{t}}, \langle\sigma v\rangle_{hh}, \langle\sigma v\rangle_{WW}, \langle\sigma v\rangle_{ZZ}\}$$

$\{\boldsymbol{\theta}_i\} \sim \mathscr{L}_1(\boldsymbol{\theta})$ ⟵ **We have this!**

$$\mathscr{L}_2(\boldsymbol{x}) = \int \mathrm{d}\theta\, \mathscr{L}_1(\boldsymbol{\theta}) \frac{\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})}{\mathscr{L}_1(\boldsymbol{\theta})} \approx \frac{1}{\tilde{N}} \sum_{i=1}^{N} \frac{\mathscr{L}_2(\boldsymbol{\theta}_i, \boldsymbol{x})}{\mathscr{L}_1(\boldsymbol{\theta}_i)}$$

Importance sampling

$\mathscr{L}_1(\boldsymbol{\theta})$ Likelihood **without DM** signal

$\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$ Likelihood **with DM** signal

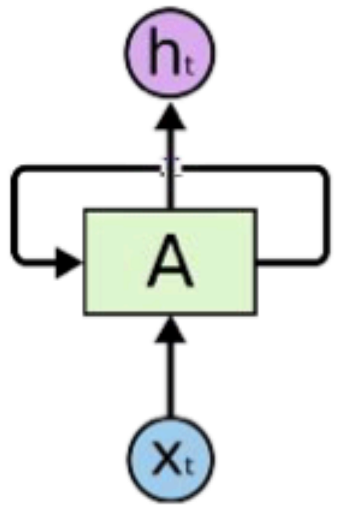
$$\boldsymbol{\theta} = \{\gamma_{1,p}, \gamma_{1,\mathrm{He}}, \gamma_{2,p}, \gamma_{2,\mathrm{He}}, R_0, s, D_0, \delta, v_A, v_{0,c}, z_h, \varphi, \varphi_{\bar{p}}\}$$

$$\boldsymbol{x} = \{m_{\mathrm{DM}}, \langle\sigma v\rangle_{gg}, \langle\sigma v\rangle_{q\bar{q}}, \langle\sigma v\rangle_{c\bar{c}}, \langle\sigma v\rangle_{b\bar{b}}, \langle\sigma v\rangle_{t\bar{t}}, \langle\sigma v\rangle_{hh}, \langle\sigma v\rangle_{WW}, \langle\sigma v\rangle_{ZZ}\}$$

$\{\boldsymbol{\theta}_i\} \sim \mathscr{L}_1(\boldsymbol{\theta})$ ⟵ **We have this!**

$$\mathscr{L}_2(\boldsymbol{x}) = \int \mathrm{d}\theta\, \mathscr{L}_1(\boldsymbol{\theta}) \frac{\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})}{\mathscr{L}_1(\boldsymbol{\theta})} \approx \frac{1}{\tilde{N}} \sum_{i=1}^{N} \frac{\mathscr{L}_2(\boldsymbol{\theta}_i, \boldsymbol{x})}{\mathscr{L}_1(\boldsymbol{\theta}_i)}$$

⟵ **We make it fast!**



Importance sampling

$\mathscr{L}_1(\boldsymbol{\theta})$ Likelihood **without DM** signal

$\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})$ Likelihood **with DM** signal

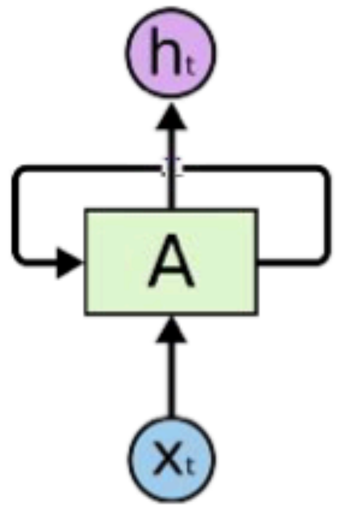
$$\boldsymbol{\theta} = \{\gamma_{1,p}, \gamma_{1,\mathrm{He}}, \gamma_{2,p}, \gamma_{2,\mathrm{He}}, R_0, s, D_0, \delta, v_A, v_{0,c}, z_h, \varphi, \varphi_{\bar{p}}\}$$

$$\boldsymbol{x} = \{m_{\mathrm{DM}}, \langle\sigma v\rangle_{gg}, \langle\sigma v\rangle_{q\bar{q}}, \langle\sigma v\rangle_{c\bar{c}}, \langle\sigma v\rangle_{b\bar{b}}, \langle\sigma v\rangle_{t\bar{t}}, \langle\sigma v\rangle_{hh}, \langle\sigma v\rangle_{WW}, \langle\sigma v\rangle_{ZZ}\}$$

$$\{\boldsymbol{\theta}_i\} \sim \mathscr{L}_1(\boldsymbol{\theta})$$ ← **We have this!**

Note of caution: We assume that the full parameter space (θ, x) is sufficiently covered by our sample.

$$\mathscr{L}_2(\boldsymbol{x}) = \int \mathrm{d}\theta \, \mathscr{L}_1(\boldsymbol{\theta}) \frac{\mathscr{L}_2(\boldsymbol{\theta}, \boldsymbol{x})}{\mathscr{L}_1(\boldsymbol{\theta})} \approx \frac{1}{\tilde{N}} \sum_{i=1}^{N} \frac{\mathscr{L}_2(\boldsymbol{\theta}_i, \boldsymbol{x})}{\mathscr{L}_1(\boldsymbol{\theta}_i)}$$ ← **We make it fast!**

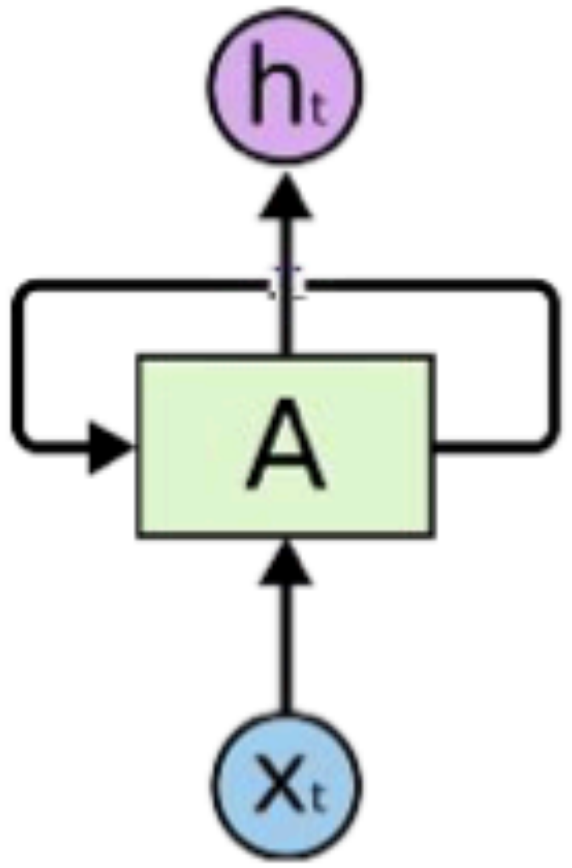


Importance Sampling to marginalize over CR propagation parameters

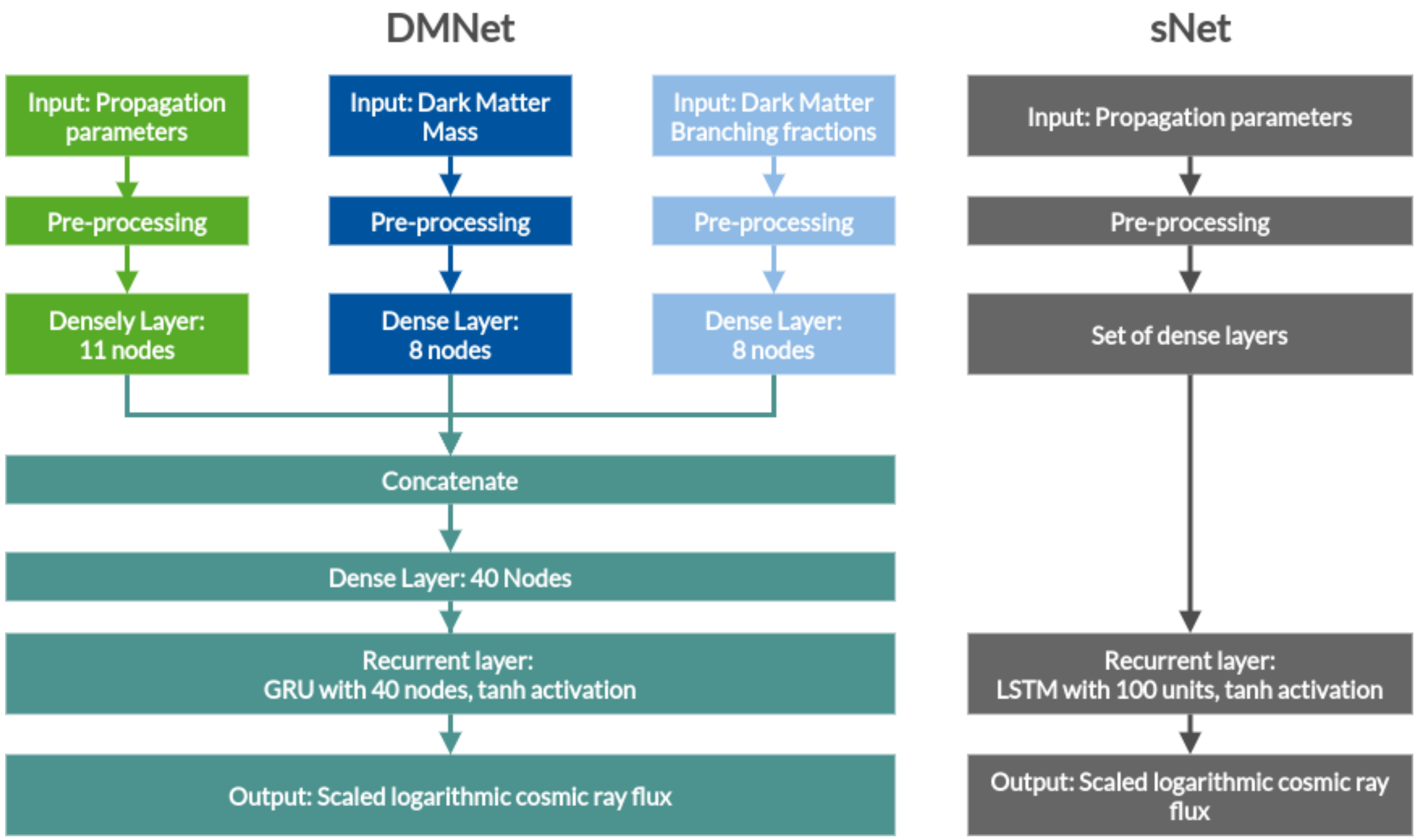
$$\{x_i\} \sim q(x)$$

$$E_x[p(x)] = \frac{1}{N}\sum_{i=1}^{N} x_i \frac{p(x_i)}{q(x_i)}$$

RNNs to replace GALPROP for an arbitrary DM model



Architecture and training



$$\tilde{\phi}_s(E) = \log_{10}(E^{2.7}\phi(E)) \qquad \tilde{\phi}_{DM}(x) = \log_{10}(m_{DM}^3 x \phi(E))$$

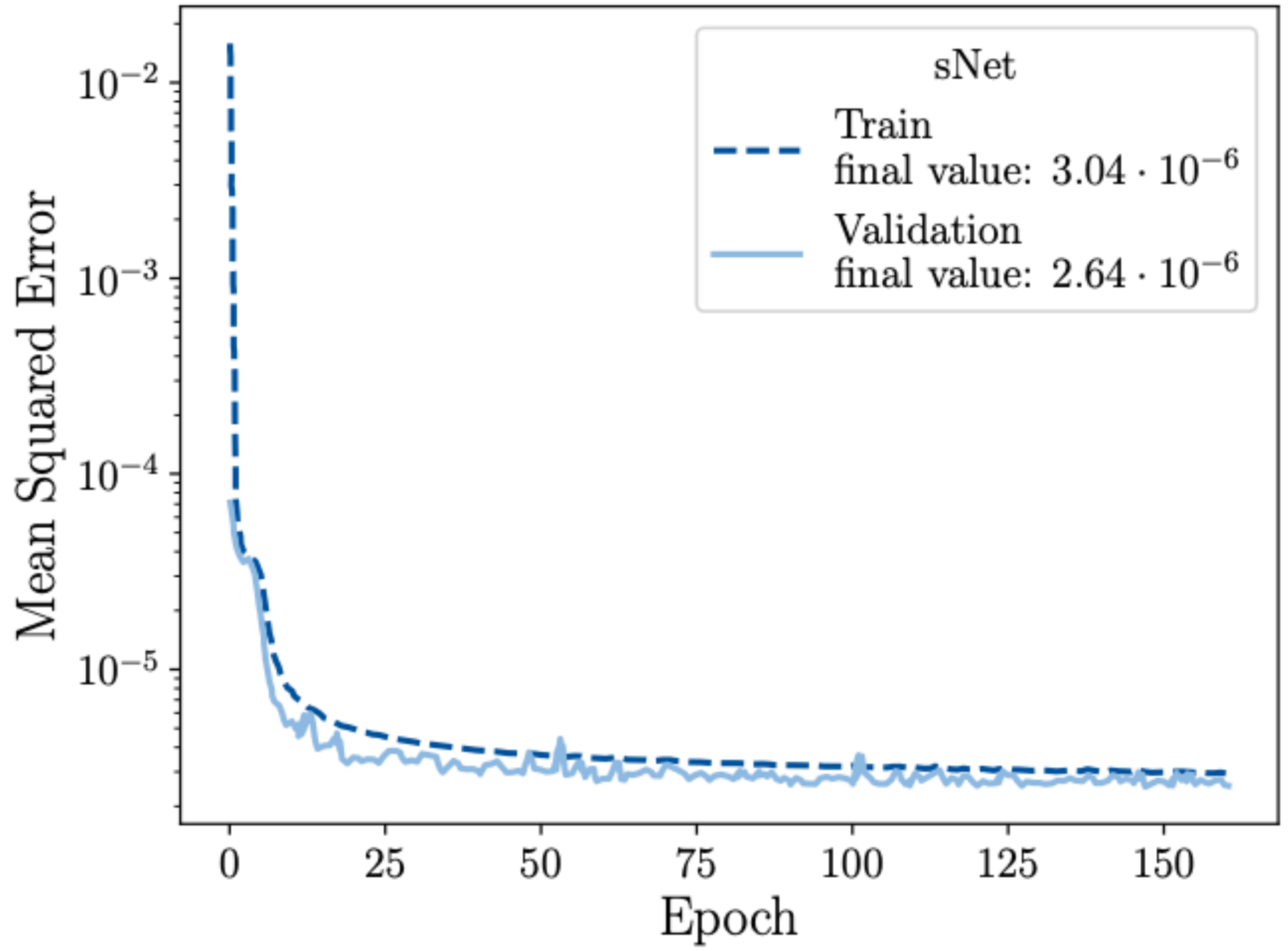
Training Data: Chain of a MultiNest fit

RNNs efficiently learn smooth spectra

Architecture and training

DMNet sNet

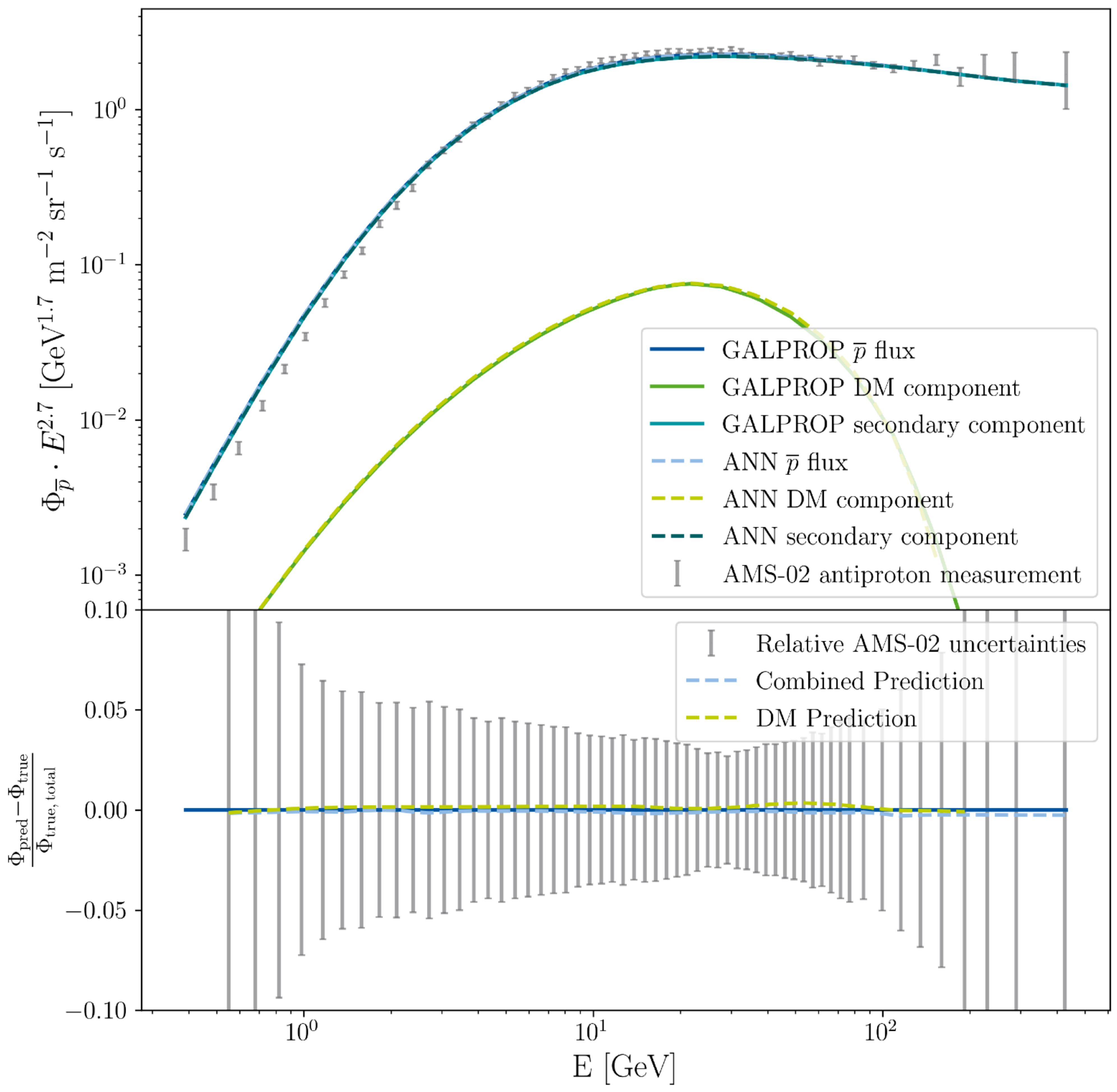
Input: Propagation Input: Dark Matter Input: Dark Matter

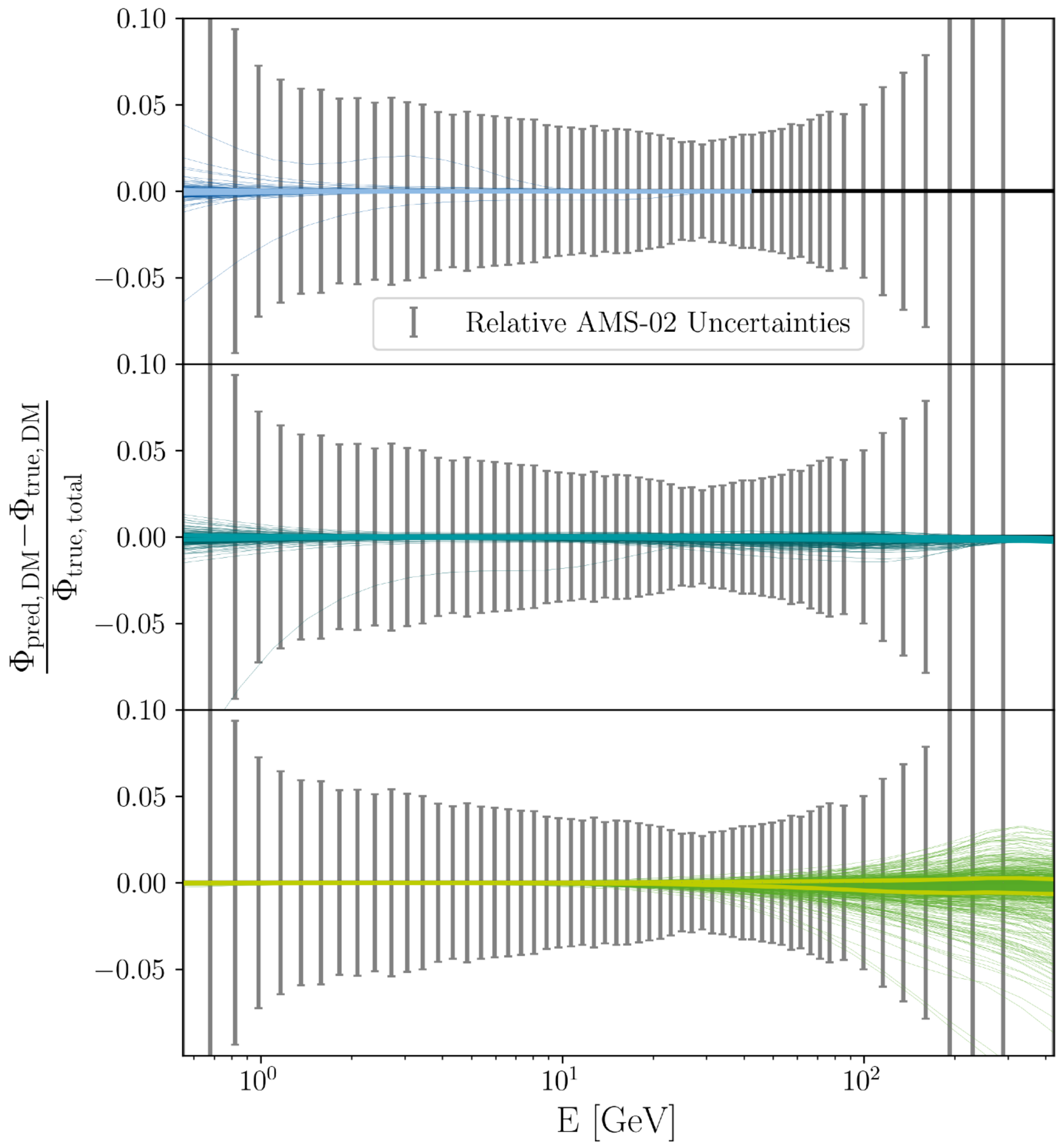


Hyperparameters	
Activation	ReLU
Dropout fraction	0.1 %
Optimizer	Adam, learning rate scheduling $l \in [10^{-2}, 10^{-5}]$, patience 10 epochs
Loss	Mean squared error (MSE)
Batch size	500
Validation split	20 %
Early stopping	Monitor val. loss, patience = 40

$\tilde{\phi}_{\mathrm{s}}(E) = \log_{10}(E^{2.7}\phi(E))$ $\tilde{\phi}_{\mathrm{DM}}(x) = \log_{10}(m_{\mathrm{DM}}^3 x \phi(E))$

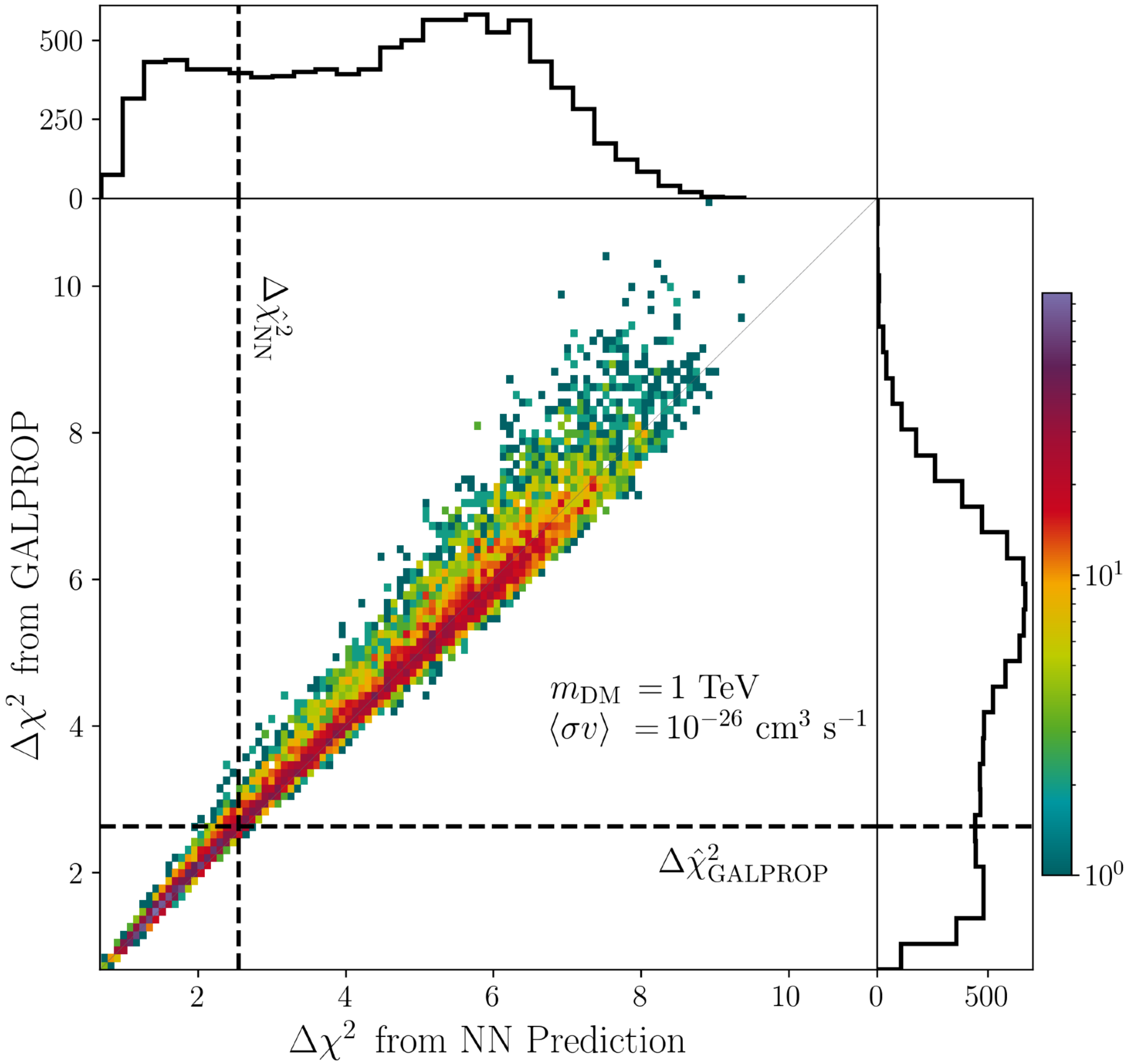
Validation





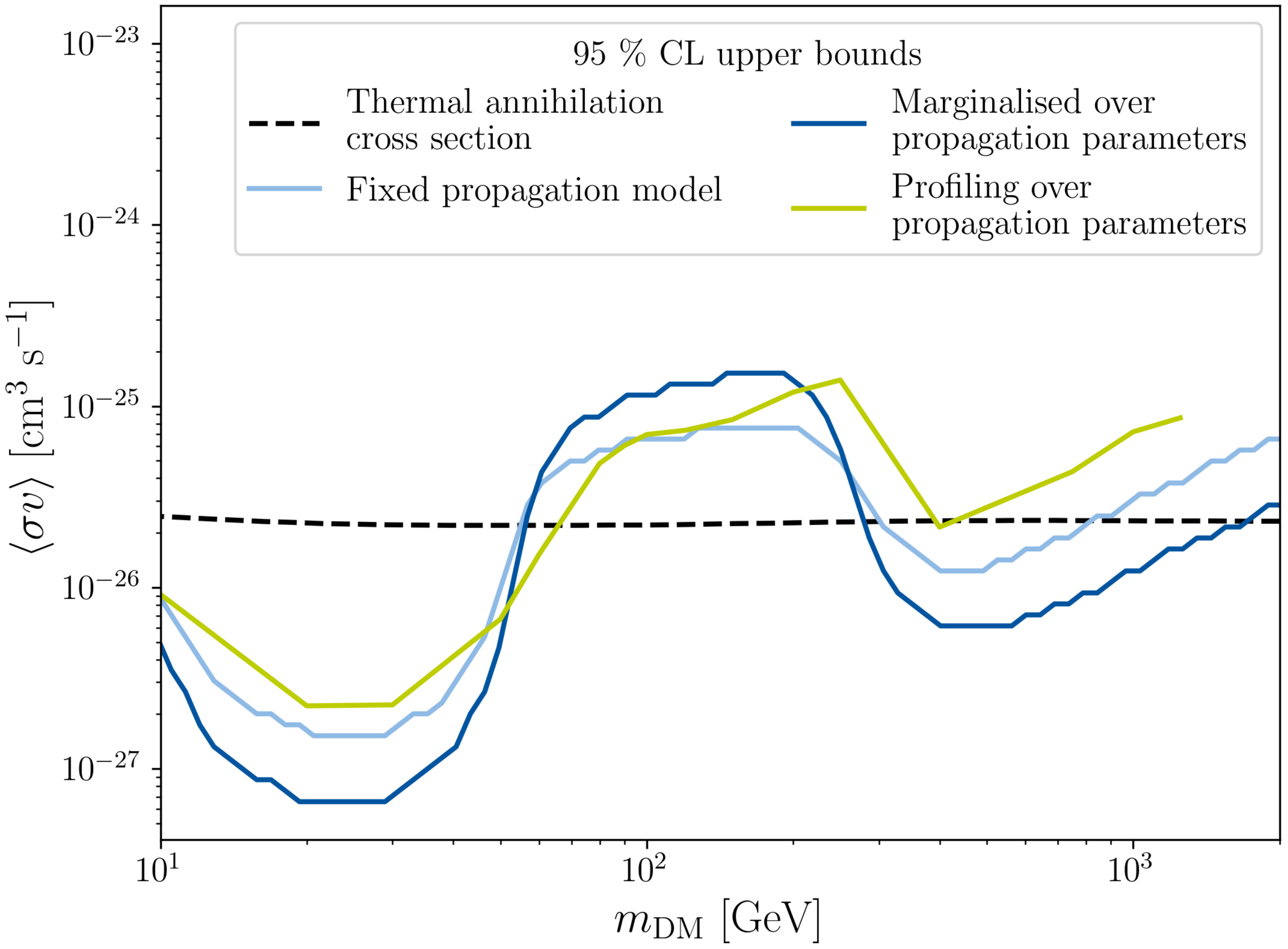
Uncertainties from the NN prediction are almost negligible compared to AMS-02 measurement uncertainties.

Validation

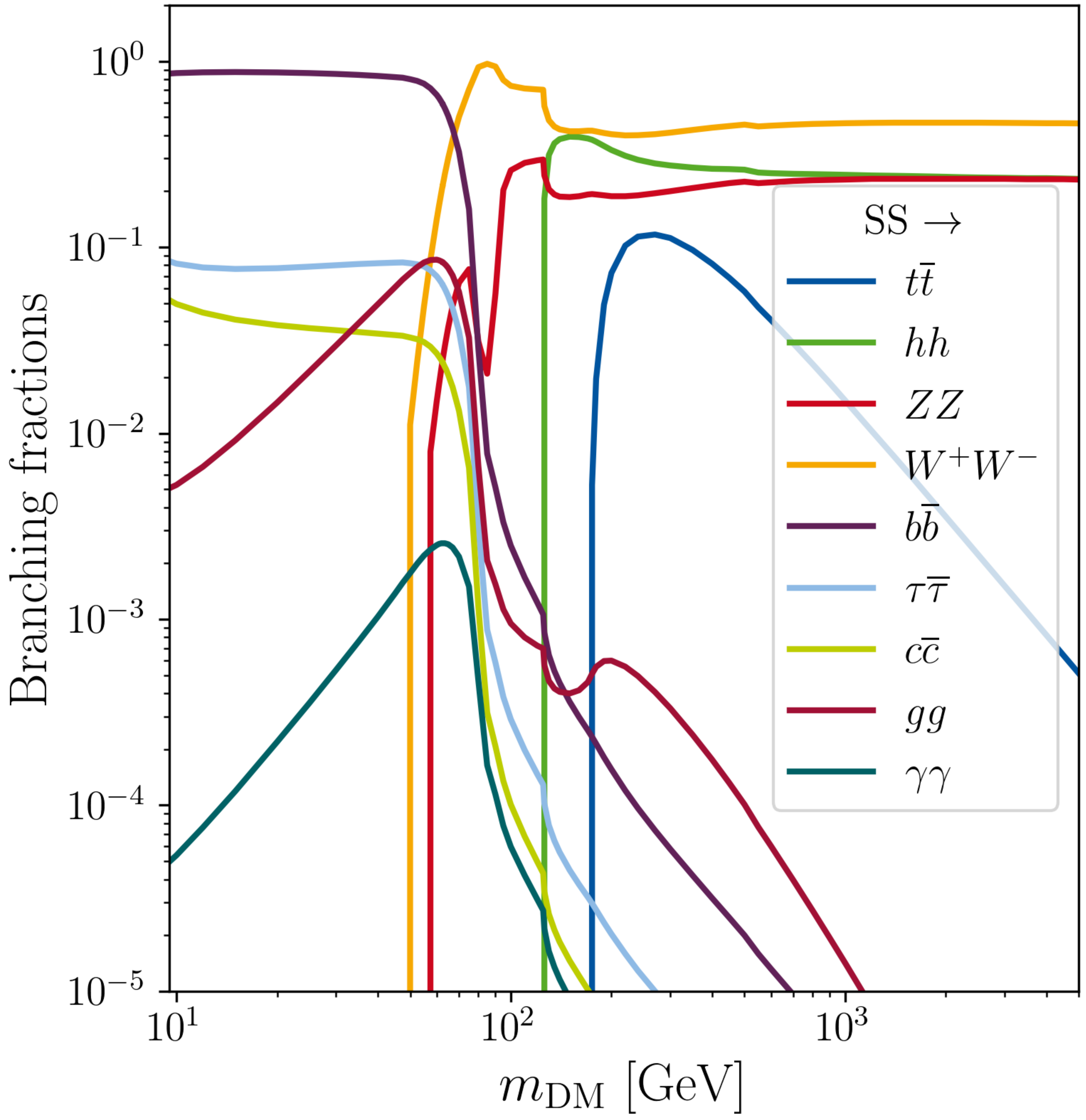


Comparison of the $\Delta\chi^2$ evaluation between the RNN and GALPROP for a large validation set shows very good agreement!

DM limits (example $b\bar{b}$)

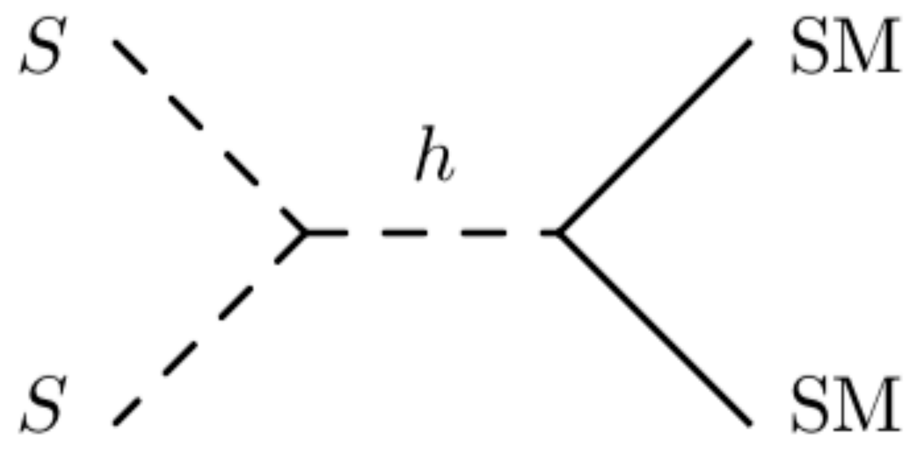


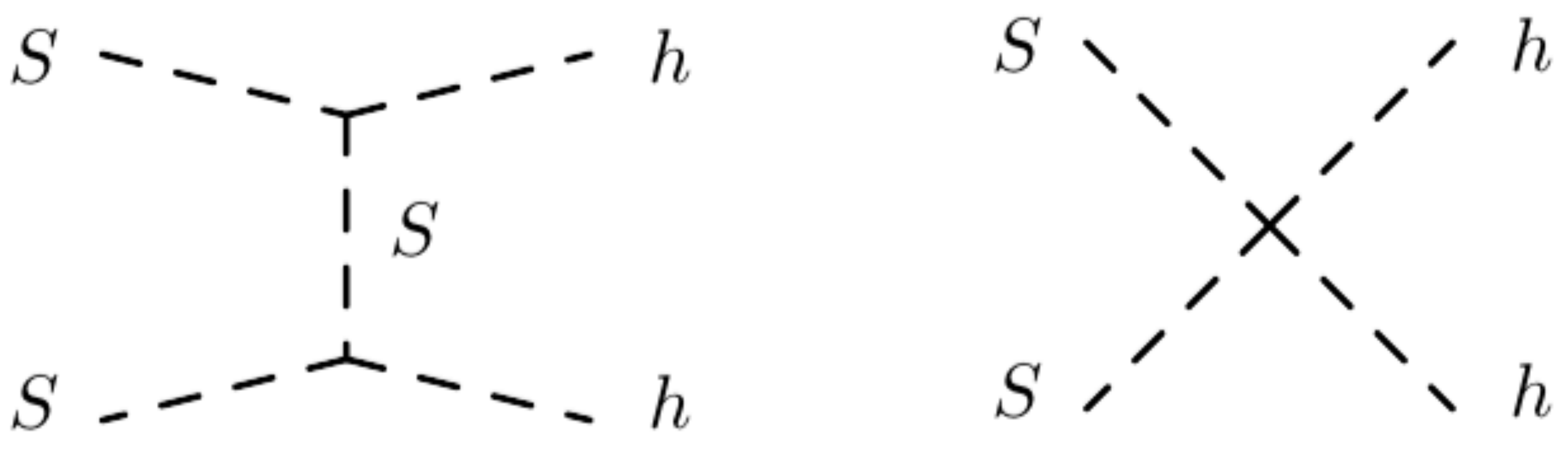
Scalar singlet dark matter



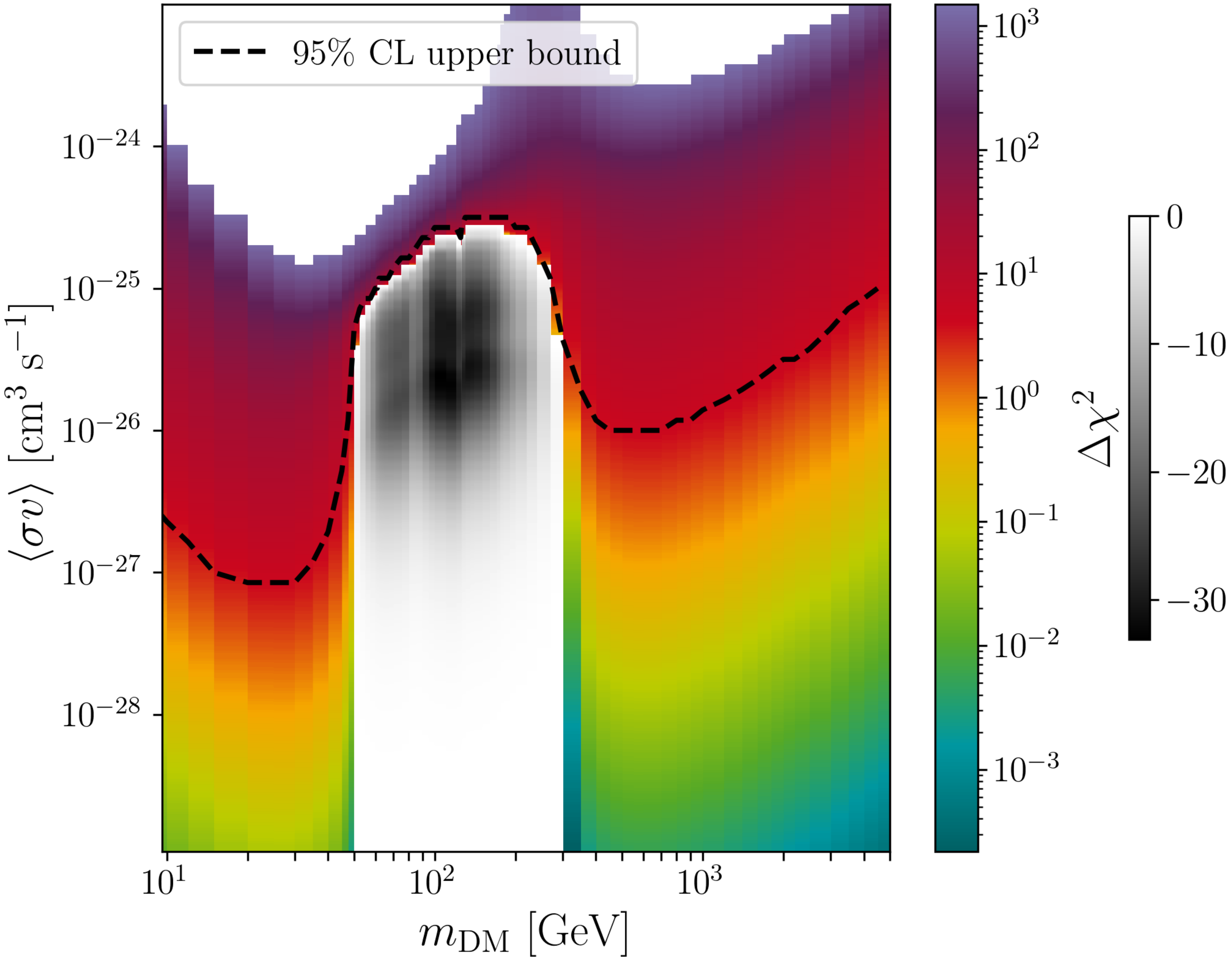
After EW symmetry breaking:

$$\mathscr{L} \supset -\frac{1}{2}m_S^2 S^2 - \frac{1}{4}\lambda_S S^4 - \frac{1}{4}\lambda_{HS} h^2 S^2 - \frac{1}{2}\lambda_{HS} vhS^2$$

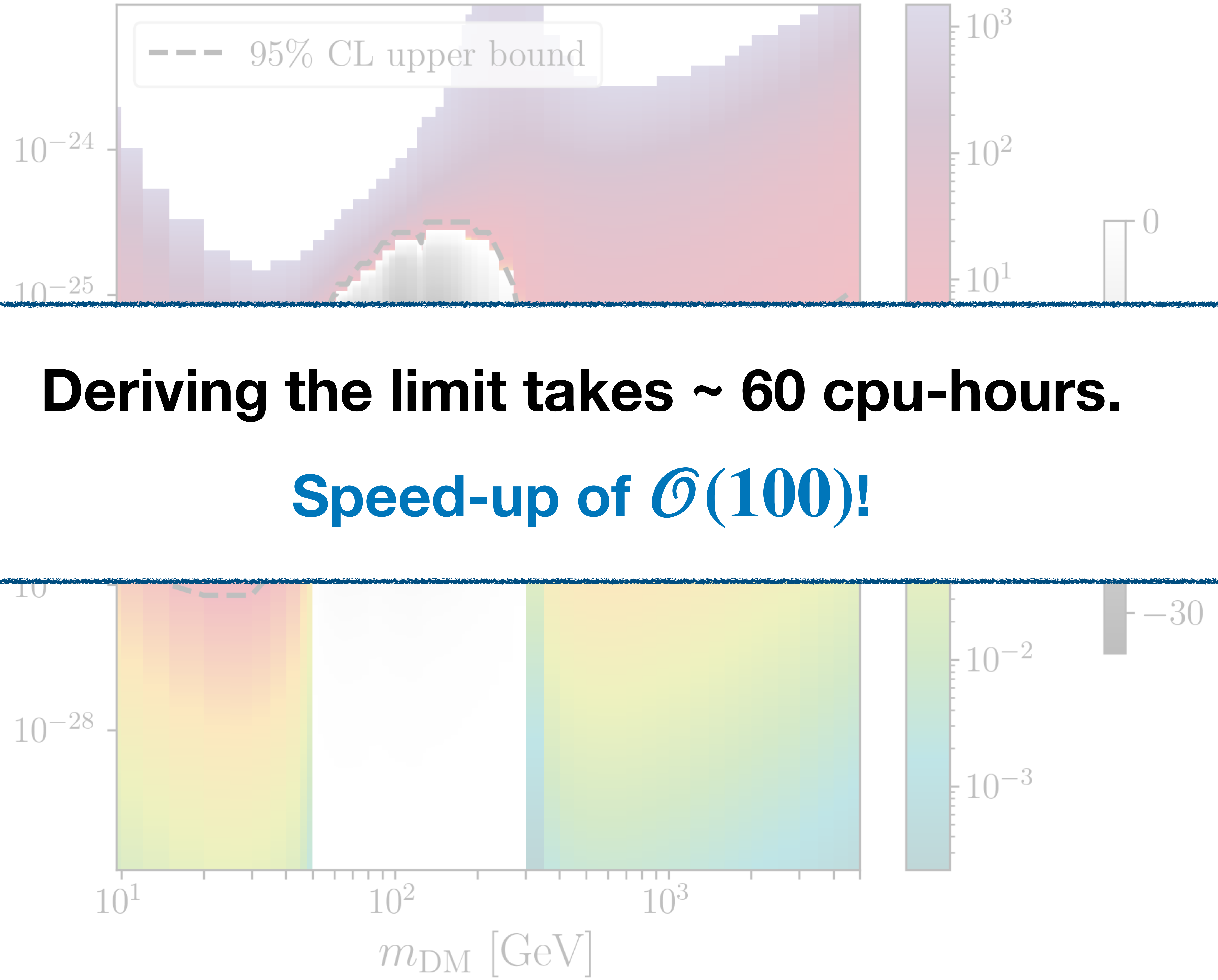




Scalar singlet dark matter



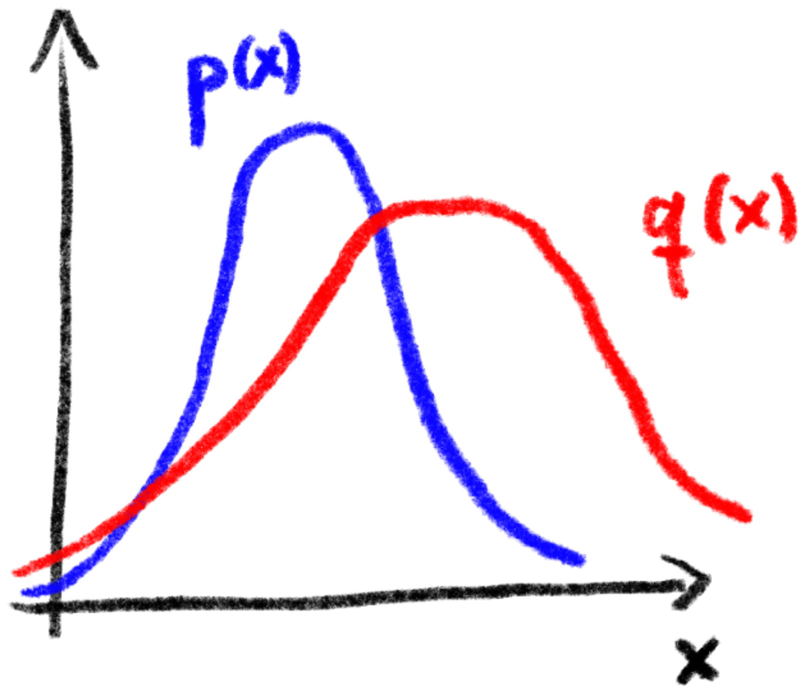
Scalar singlet dark matter



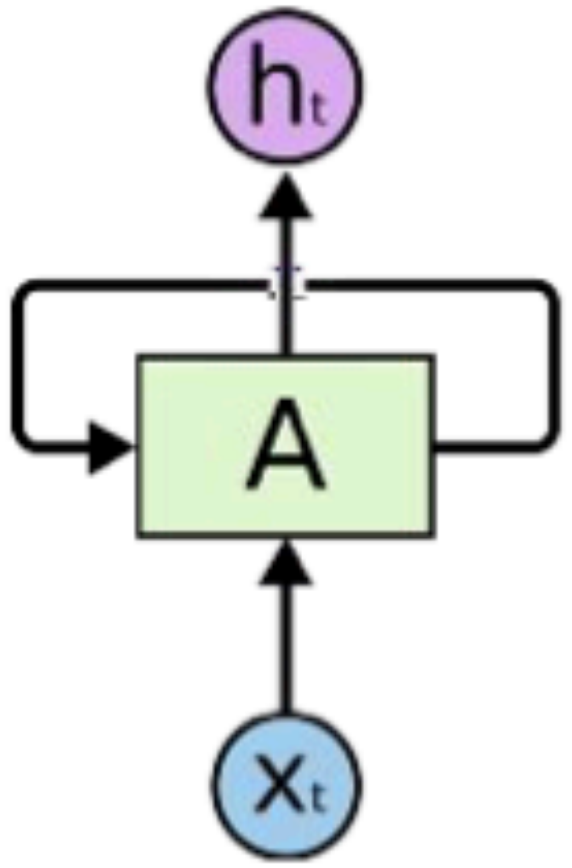
Deriving the limit takes ~ 60 cpu-hours.

Speed-up of $\mathcal{O}(100)$!

Conclusions



We have developed tools to quickly derive DM limits for a large number of DM models



RNNs are particularly well suited to predict CR spectra both for DM and the astrophysical background

The networks are published.
Try it yourself on GitHub !