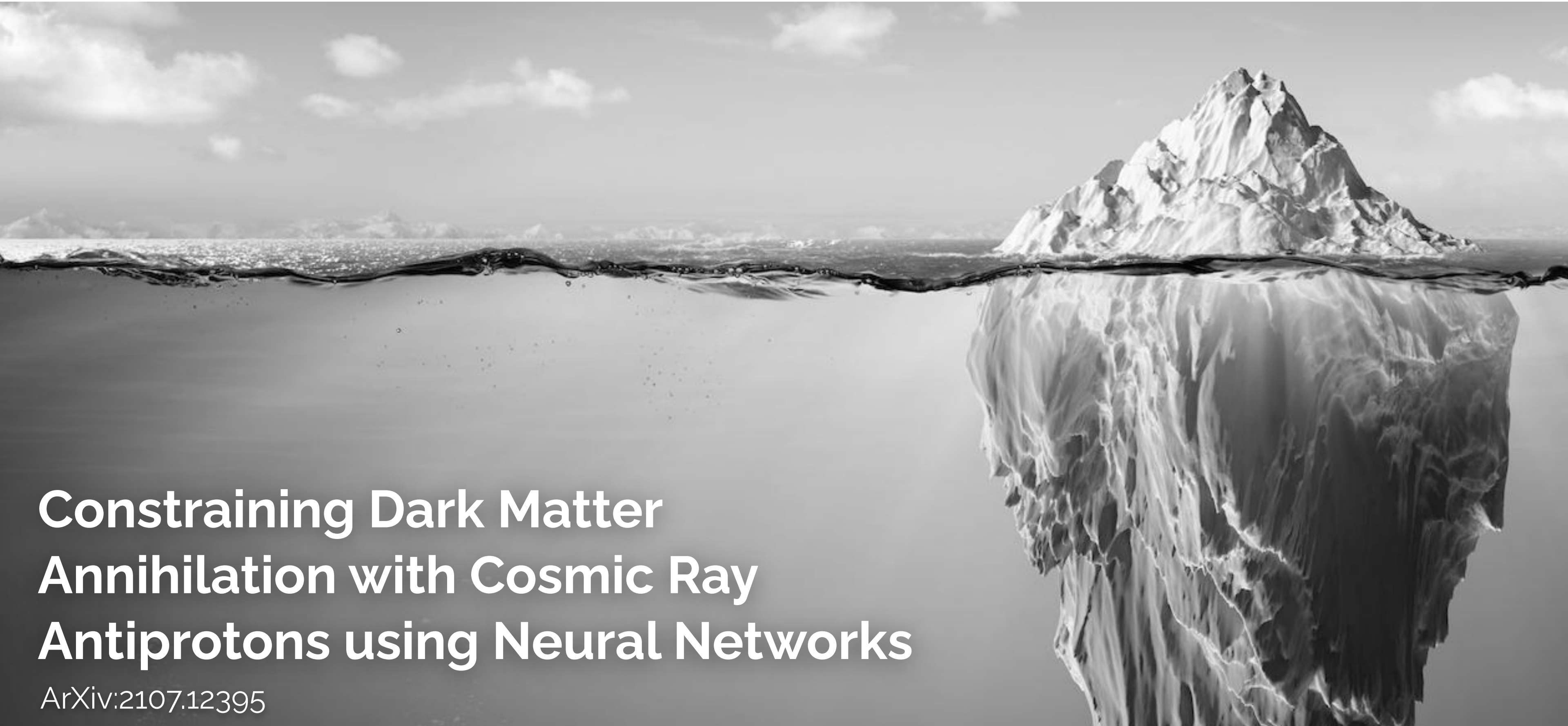
TeV Particle Astroparticle 2021 // 29.10.2021

Kathrin Nippel, Felix Kahlhöfer, Michael Korsmeier,
Michael Krämer, Silvia Manconi

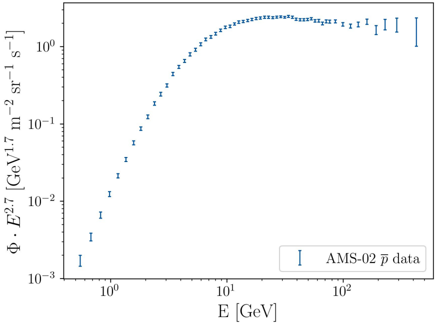


Constraining Dark Matter Annihilation with Cosmic Ray Antiprotons using Neural Networks

ArXiv:2107.12395

Antiprotons at AMS-02 for Indirect DM Searches

AMS-02 Measurement
→ Phys. Rept. 894 (2021) 1–116.

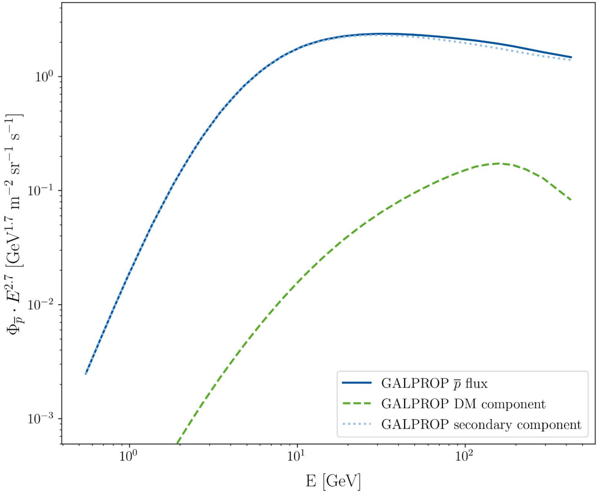


Compare using χ^2 statistic

Model // Simulations

Dark Matter Annihilation

Cosmic-Ray Propagation



See [1610.03071], [1610.03840], [1712.00002], [1903.02549], [1903.01472]

29.10.2021
Kathrin Nippel

Antiprotons at AMS-02 for Indirect DM Searches

Model Parameters

Dark Matter Annihilation

m_{DM}

$\langle \sigma v \rangle_{\mathrm{ann}}$

$\chi\chi \rightarrow q\bar{q},\ c\bar{c},\ b\bar{b},\ t\bar{t},\ W^+W^-,\ ZZ,\ gg, hh$

Cosmic Ray Propagation

z_h D_0 δ

v_a $v_{0,c}$ R_0

s γ_1 γ_2

$\gamma_{1,p}$ $\gamma_{2,p}$

Profiled Parameters

$\varphi_{\text{Solar modulation, AMS-02}}$

(Force Field Approximation: Astrophys. J. 206 (1976) 333–341)

$A_{\bar{p},\text{AMS-02}}$

Simulation

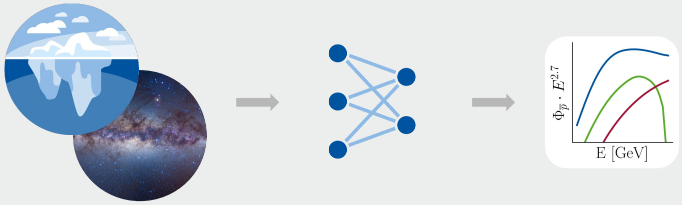
GALPROP

(Astrophys. J. 537 (2000) 763–784)

→ Computationally expensive!

$\mathcal{O}(10^5)$ cpu hours scan for this number of parameters

Outline



1. Artificial Neural Network Development
   - 1.1. Training set
   - 1.2. Architecture & training
   - 1.3. Validation
2. Application to ID Analysis
   - 2.1. Marginalization using importance sampling
   - 2.2. Constraining $\chi\chi \rightarrow b\bar{b}$
   - 2.3. Constraining scalar singlet DM
3. Conclusion

Training Set

Dark Matter Annihilation

WIMP, 5 GeV - 5 TeV
Injection spectra following “PPPC4DMID”
[1012.4515]

Choose branching fractions randomly

Cosmic Ray Propagation

MultiNest fit to AMS-02 data (assuming no DM)

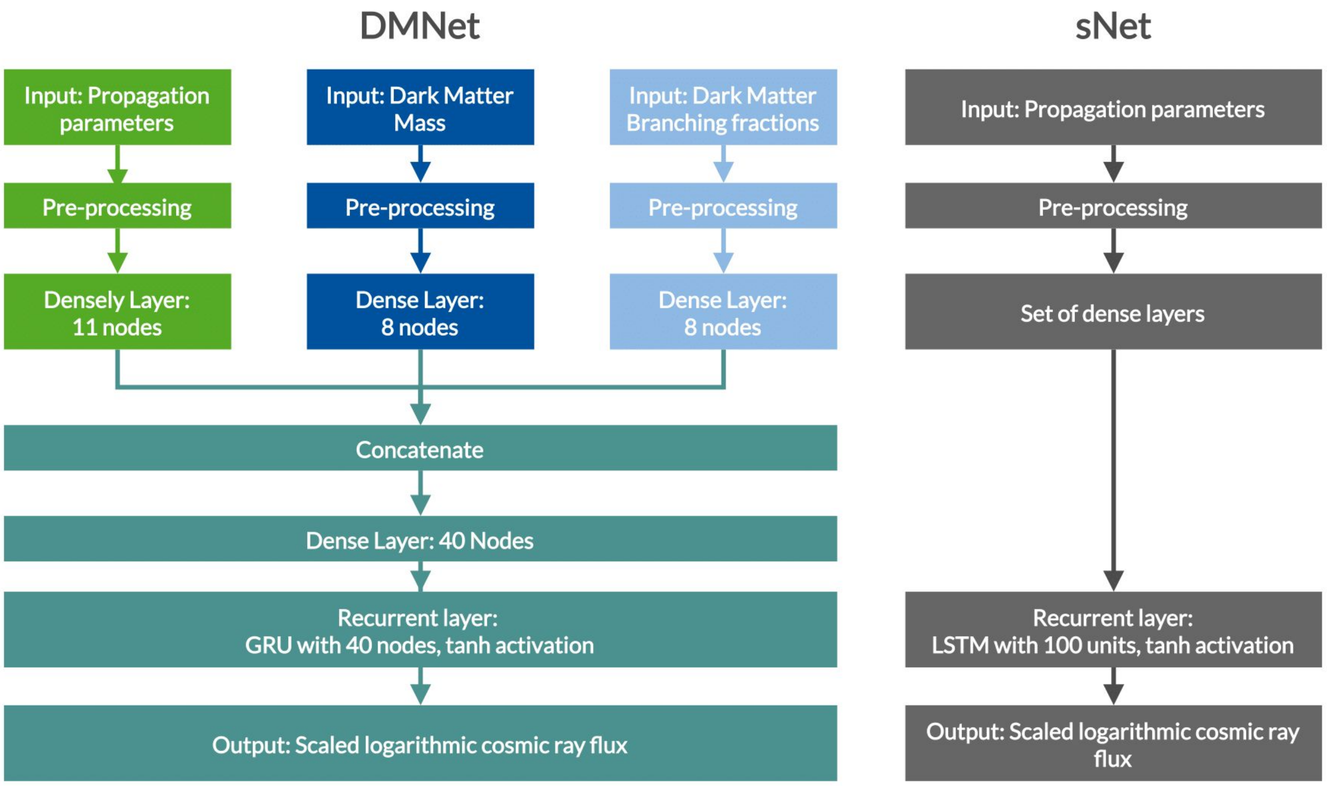
→Understand relevant parameter space

see e.g. [1903.01472]

Extensive training set based on physical reasoning

Goal: Neural network is suited to well represent relevant parameter space

Architecture & Training



- Inputs normalized
- Outputs scaled with

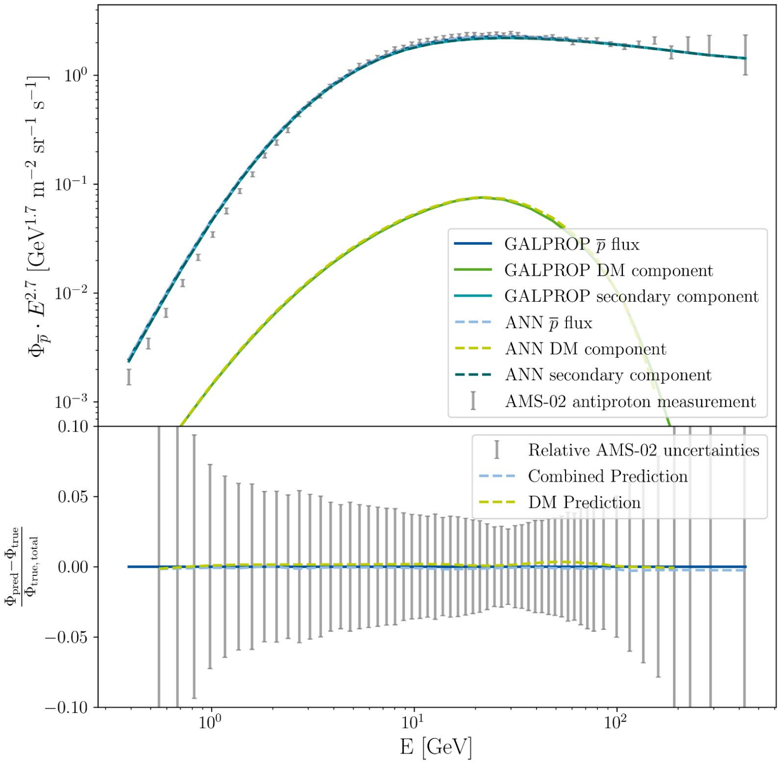
$$\tilde{\Phi}_{\mathrm{DM}}(x) = \log_{10}\left(m_{\mathrm{DM}}^3\, x\, \Phi(E)\right)$$

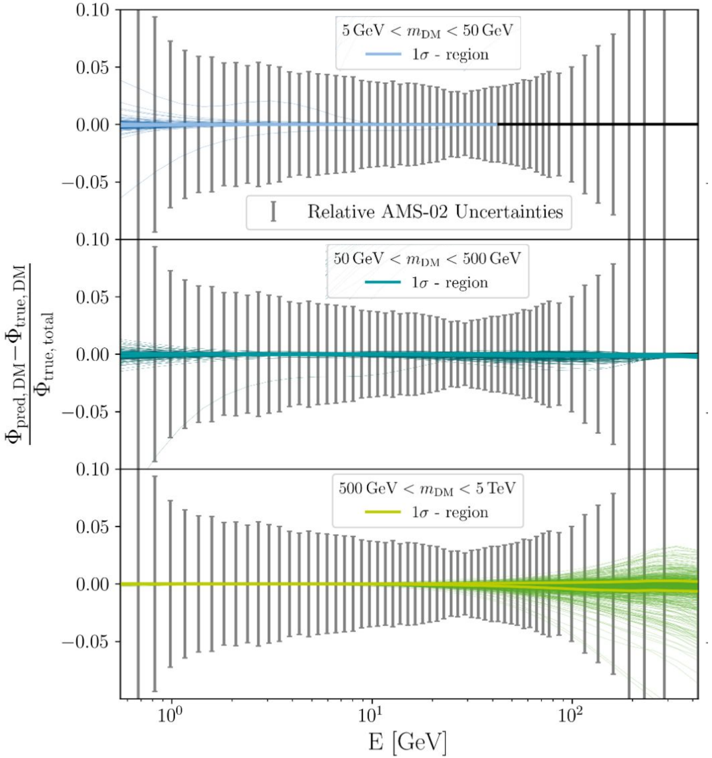
$$\tilde{\Phi}_{\mathrm{s}}(E) = \log_{10}\left(\Phi(E)\, E^{2.7}\right)$$

- Recurrent layer used to learn smooth spectrum

→ Exact input to output mapping

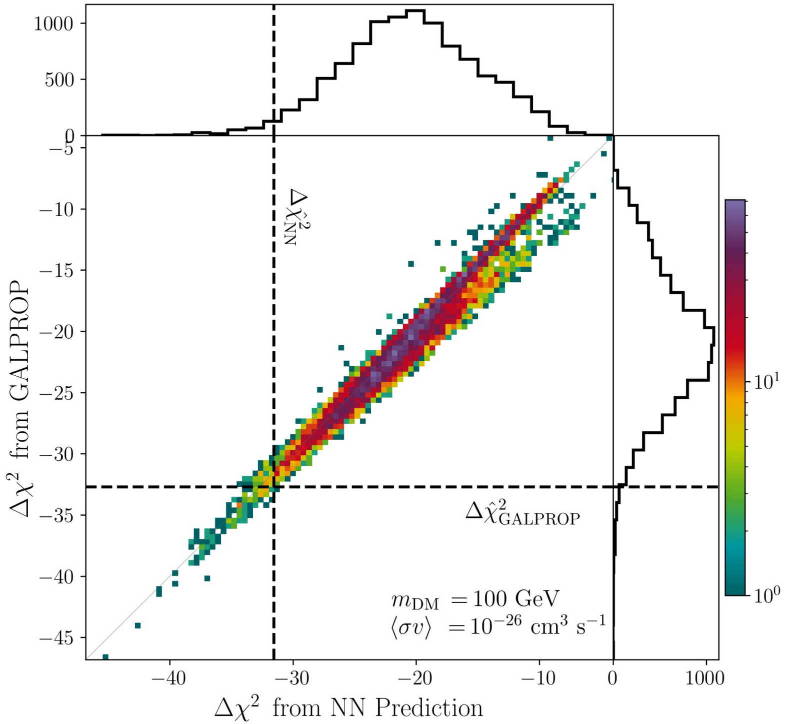
Validation
Direct Comparison of Spectra

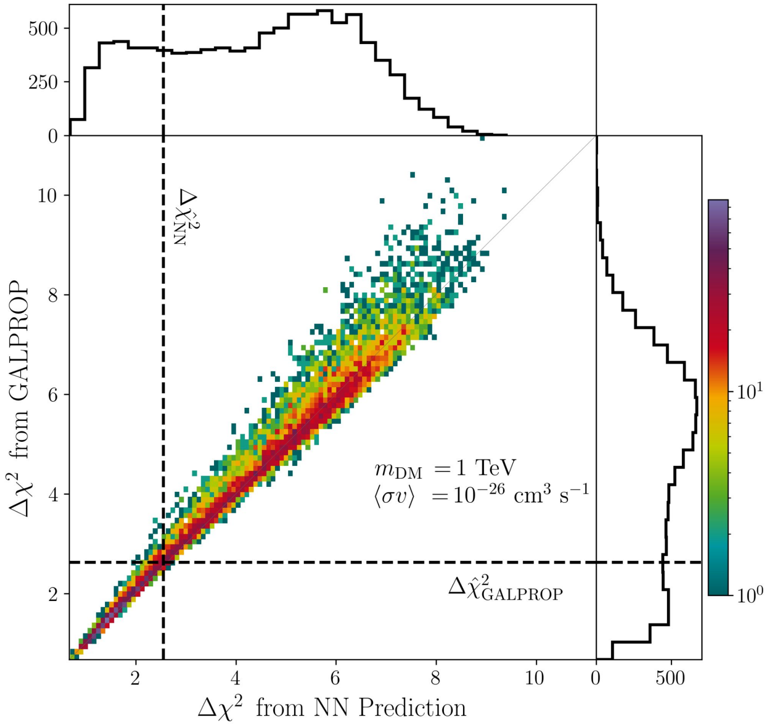




→ Differences between network predictions and simulations of spectra sufficiently small compared to measurement uncertainties

Validation
Comparison of χ^2





→ χ^2 distributions agree for individual DM models and large sample of propagation parameter points

Marginalization using Importance Sampling

Importance sampling
see e.g. [2102.05407]

$$\bar{\mathcal{L}}(\mathbf{x}_{\mathrm{DM}}) = \int \mathcal{L}(\mathbf{x}_{\mathrm{DM}}, \boldsymbol{\theta}_{\mathrm{prop}}) p(\boldsymbol{\theta}_{\mathrm{prop}}) \mathrm{d}\boldsymbol{\theta}_{\mathrm{prop}}$$

$$\bar{\mathcal{L}}(\mathbf{x}_{\mathrm{DM}}) \approx \frac{\sum_{i=1}^{N} \mathcal{L}(\mathbf{x}_{\mathrm{DM}}, \boldsymbol{\theta}_i) \frac{p(\boldsymbol{\theta}_i)}{q(\boldsymbol{\theta}_i)}}{\sum_{i=1}^{N} \frac{p(\boldsymbol{\theta}_i)}{q(\boldsymbol{\theta}_i)}}$$

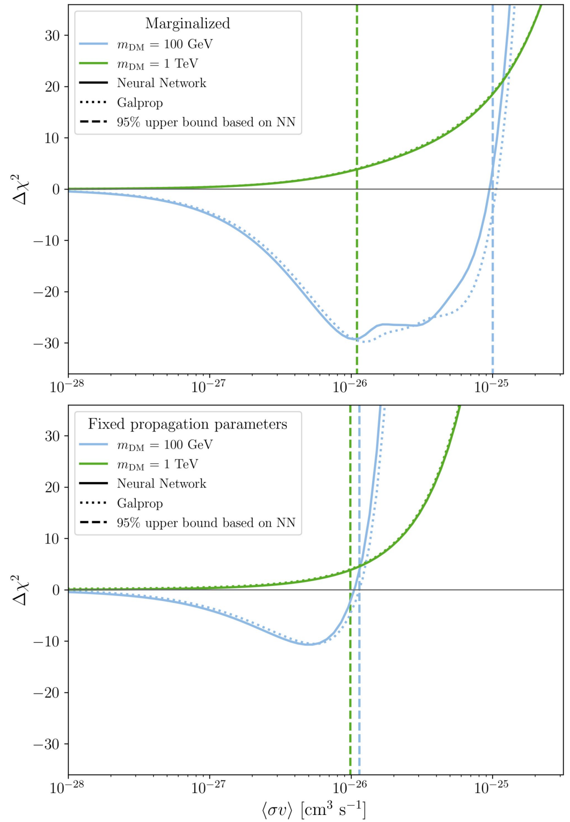
Posterior probability in absence of DM signal

$$p(\boldsymbol{\theta}_i)/q(\boldsymbol{\theta}_i) \propto 1/\mathcal{L}_0(\theta_i)$$

$\chi^2 \propto -2 \log \mathcal{L}$

$$\Delta\bar{\chi}^2(\mathbf{x}_{\mathrm{DM}}) = -2 \log \frac{\sum_{i=1}^{N} \exp\left(-\frac{\Delta\chi^2(\mathbf{x}_{\mathrm{DM}}, \boldsymbol{\theta}_{\mathrm{prop}})}{2}\right)}{\sum_{i=1}^{N} \exp\left(\frac{\chi_0^2(\boldsymbol{\theta}_{\mathrm{prop}})}{2}\right)} - \bar{\chi}_0^2$$

Resulting $\chi^2\langle\sigma v\rangle$ distribution
(before and after marginalization)

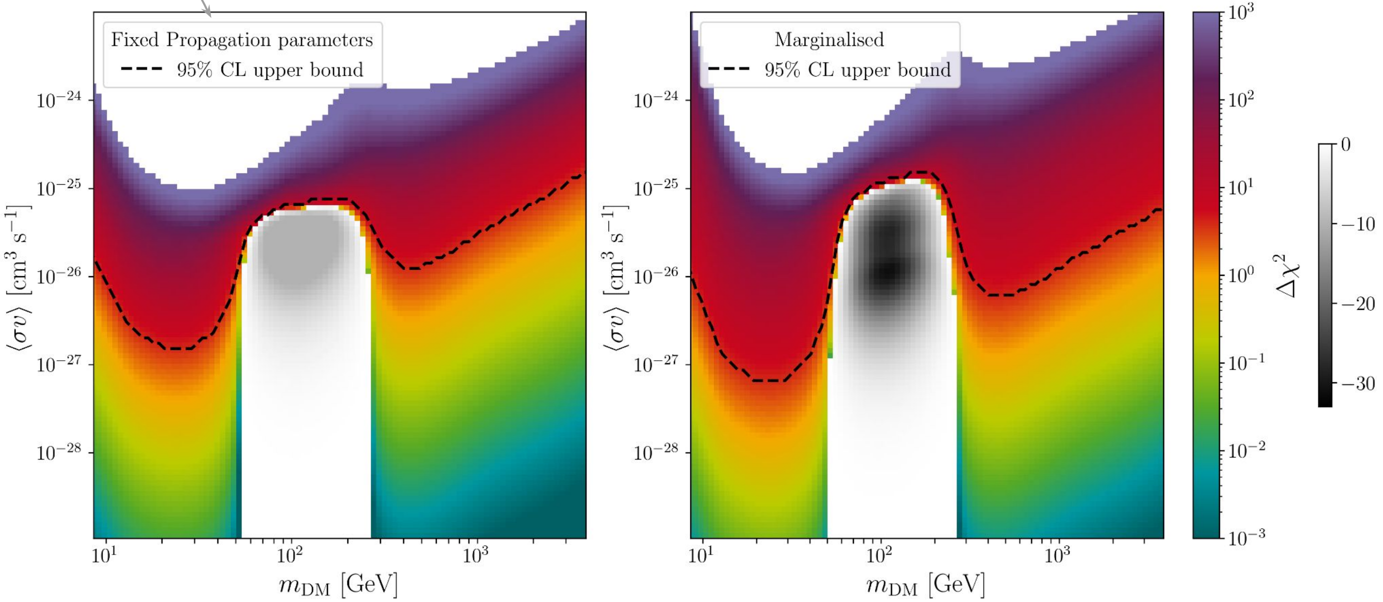


→ Calculation of 95% CL upper bounds on $\langle\sigma v\rangle$

Constraining
$\chi\chi \rightarrow b\bar{b}$

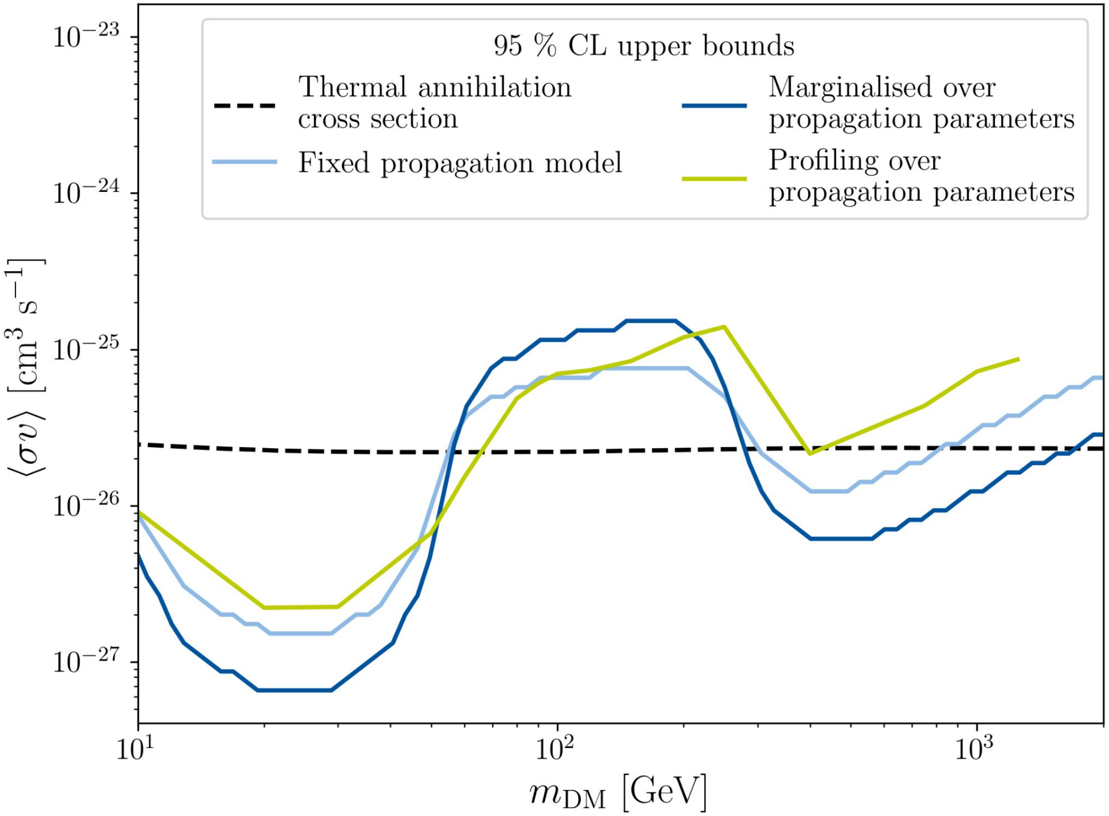
Resulting $\Delta\chi^2$ distribution for m_{DM} - $\langle\sigma v\rangle$ parameter space + 95% CL exclusion bounds

Best fit value
(without DM)



→ Preference for DM with masses in the range of ~ 50–250 GeV
Marginalisation leads to relaxed exclusion bounds

Constraining
$\chi\chi \to b\bar{b}$



- ➜ Differences between bounds are result of the respective methods
- ➜ Significant speed-up achieved by implementing ANN
  - individual data points: > $\mathcal{O}(10^3)$
  - marginalization: $\mathcal{O}(10^3)$
  - profiling: $\mathcal{O}(10^2)$

Profiling over propagation parameters:

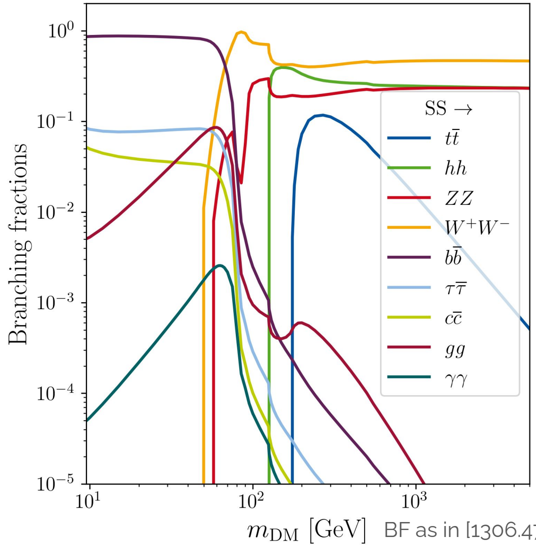
- MultiNest fit based on NN spectra
- Need to ensure no convergence towards untrained parameter regions

Constraining
Scalar Singlet Dark Matter

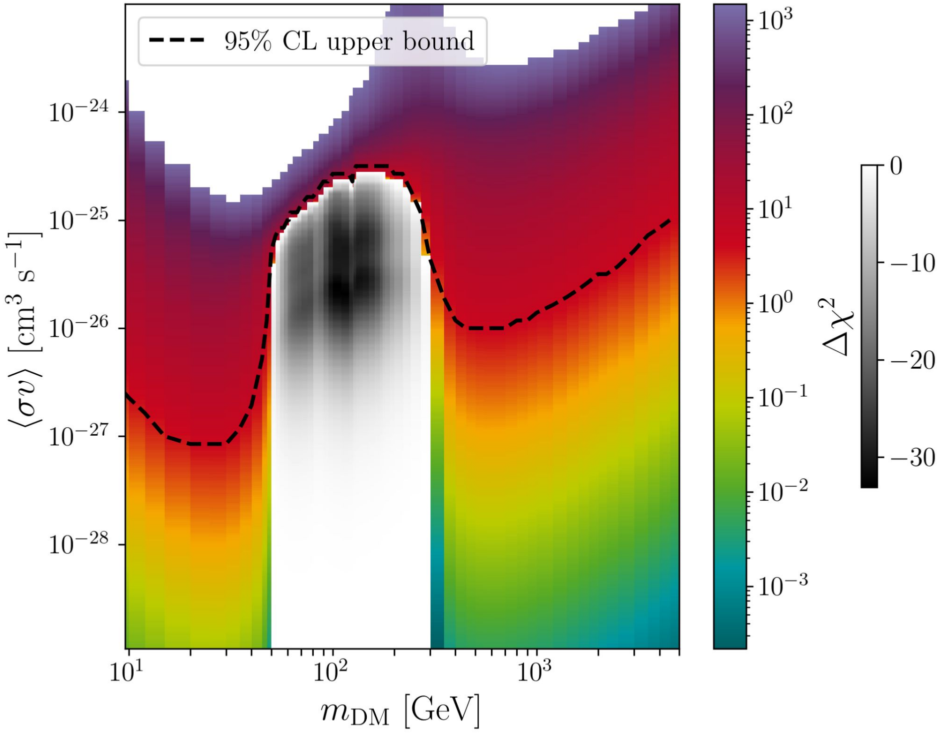
$$\mathcal{L} \supset \frac{1}{2}\partial_\mu S \partial^\mu S - \frac{1}{2}m_{S,0}^2 S^2 - \frac{1}{4}\lambda_S S^4 - \frac{1}{2}\lambda_{hs} S^2 H^\dagger H$$

→ see e.g. [hep-ph/0011335]

$$m_S = \sqrt{m_{S,0}^2 + \frac{1}{2}\lambda_{hs} v_0^2}$$

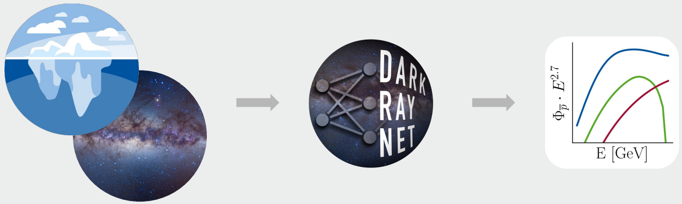


BF as in [1306.4710]



➡ Neural network approach well suited for models with rich branching fraction structure

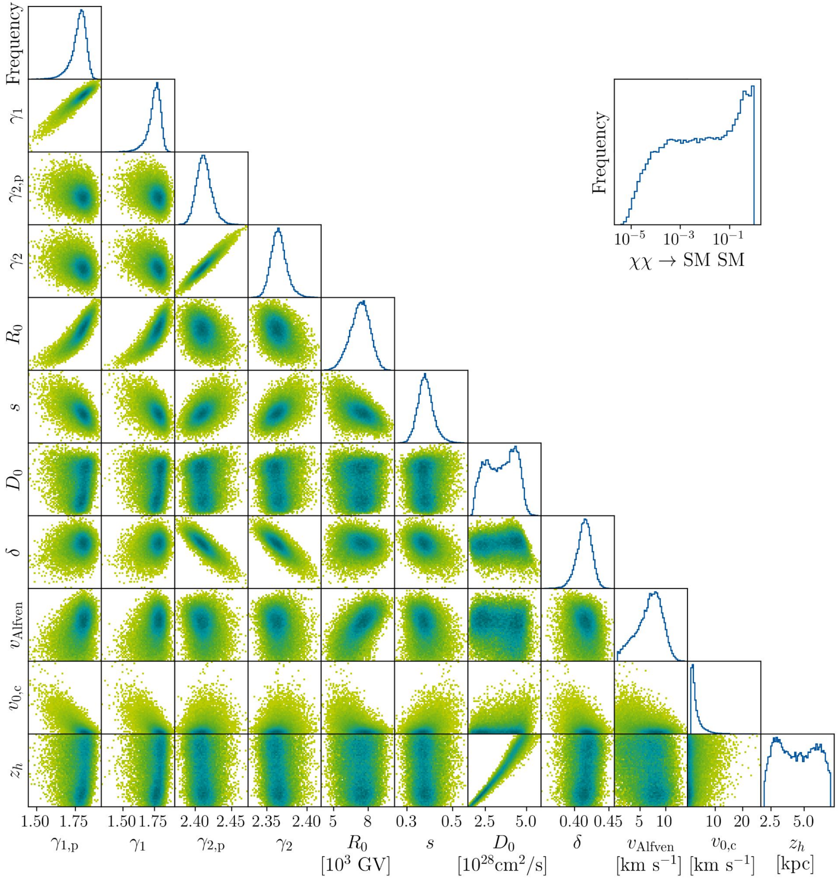
Conclusions



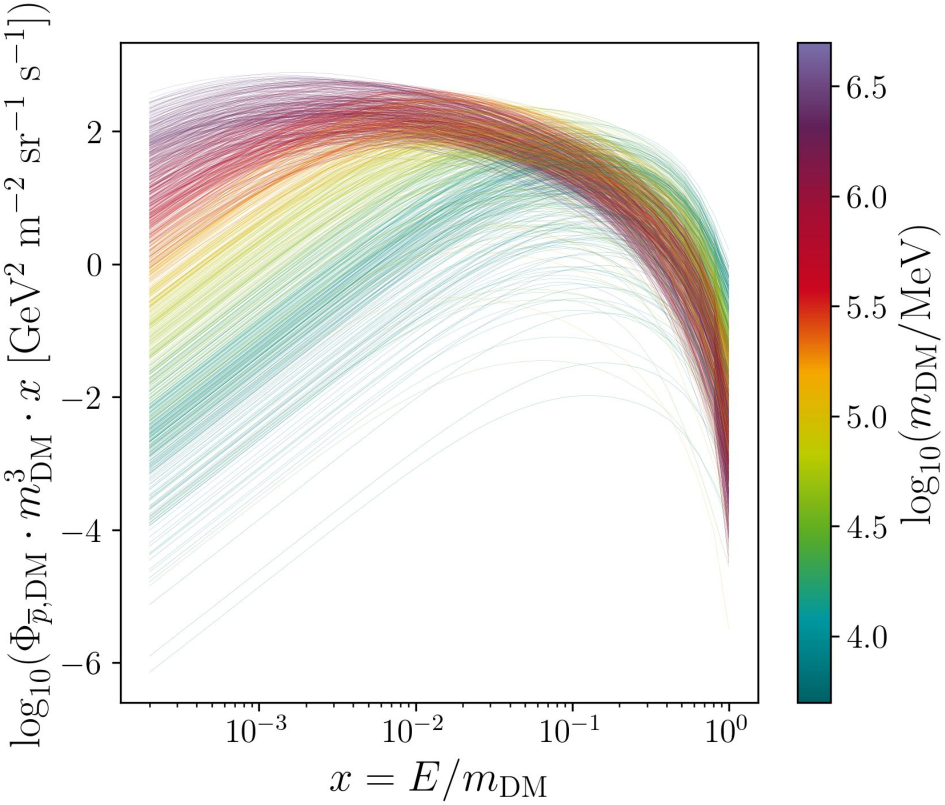
github.com/kathrinnp/DarkRayNet

- Replace GALPROP with Neural Network in indirect DM analysis
  - Based on AMS-02 antiproton measurement
  - Physically motivated training set
  - Recurrent network well suited
  - Multiple checks of network accuracy
- Constrain $\chi\chi \rightarrow b\bar{b}$ and SSDM
  - Use advantage of previous fit of CR propagation for importance sampling
  - Speed up of two to three orders of magnitude
  - Network and method can be applied to further models
- Future Direction:
  - Bayesian Neural Network
  - Include different models for CR propagation

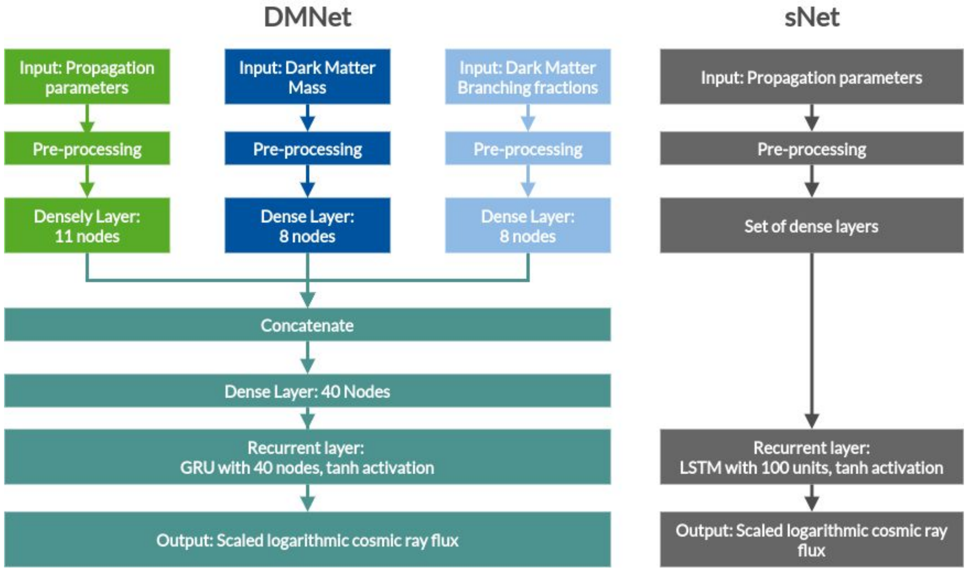
Backup: Propagation Parameter Space



Backup:
Scaling of DM Spectra



Backup: Neural Network Hyperparameters



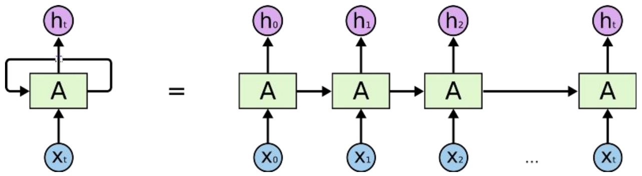
Hyperparameters	
Activation	ReLU
Dropout fraction	0.1 %
Optimizer	Adam, learning rate scheduling $l \in [10^{-2}, 10^{-5}]$, patience 10 epochs
Loss	Mean squared error (MSE)
Batch size	500
Validation split	20 %
Early stopping	Monitor val. loss, patience = 40

Backup:
Recurrent Neural Networks

Idea:

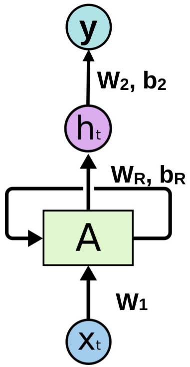
- Neural Network with connections with nodes
- Optimization: 'Back propagation through time'
- Sequence processing

Basic Concept:



$$\boldsymbol{h}_t = f(\boldsymbol{h}_{t-1}, \boldsymbol{x}_t)$$
$$= g(\boldsymbol{x}_t, \boldsymbol{x}_{t-1}, \boldsymbol{x}_{t-2}, ..., \boldsymbol{x}_0)$$

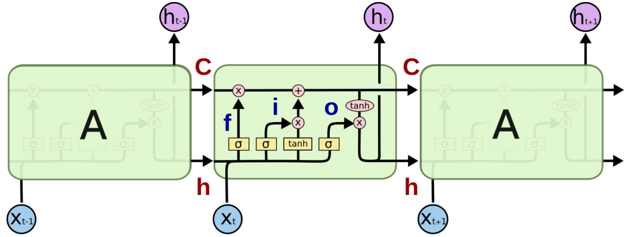
recursive formula



$$h_t = \tanh(\boldsymbol{W}_R h_{t-1} + \boldsymbol{W}_1 x_t + \boldsymbol{b}_R)$$
$$\boldsymbol{y} = \sigma(\boldsymbol{W}_2 \boldsymbol{h} + \boldsymbol{b}_2)$$

Variations:

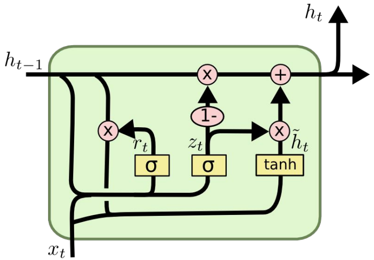
LSTM:



C - cell state
h - hidden / normal state

f - forget gate
i - input gate
o - output gate

GRU:



Update Gate:

$$z_t = \sigma(\boldsymbol{W}_z h_{t-1} + \boldsymbol{U}_z x_t)$$

Reset gate:

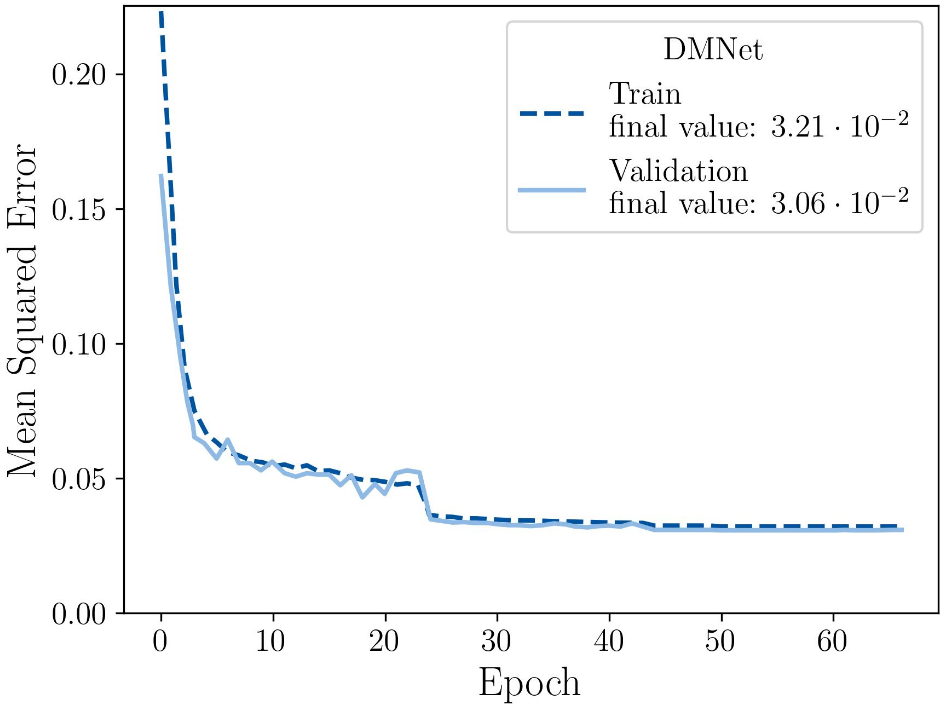
$$\boldsymbol{r}_t = \sigma(\boldsymbol{W}_r h_{t-1} + \boldsymbol{U}_r x_t)$$

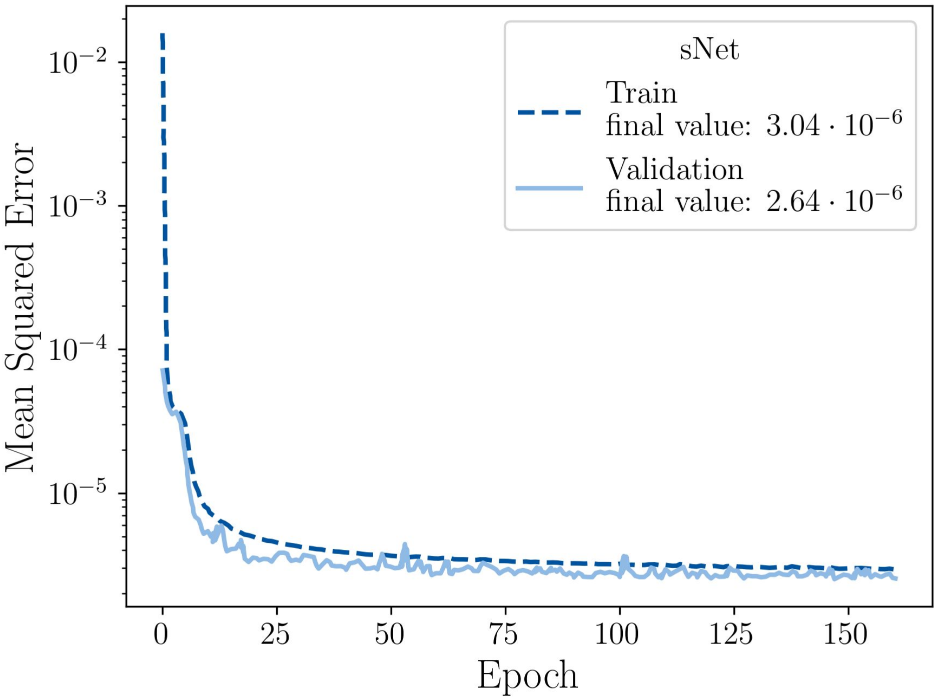
Hidden state:

$$\tilde{h}_t = \tanh(r_t \boldsymbol{W}_h h_{t-1} + \boldsymbol{U}_h x_t)$$
$$h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t$$

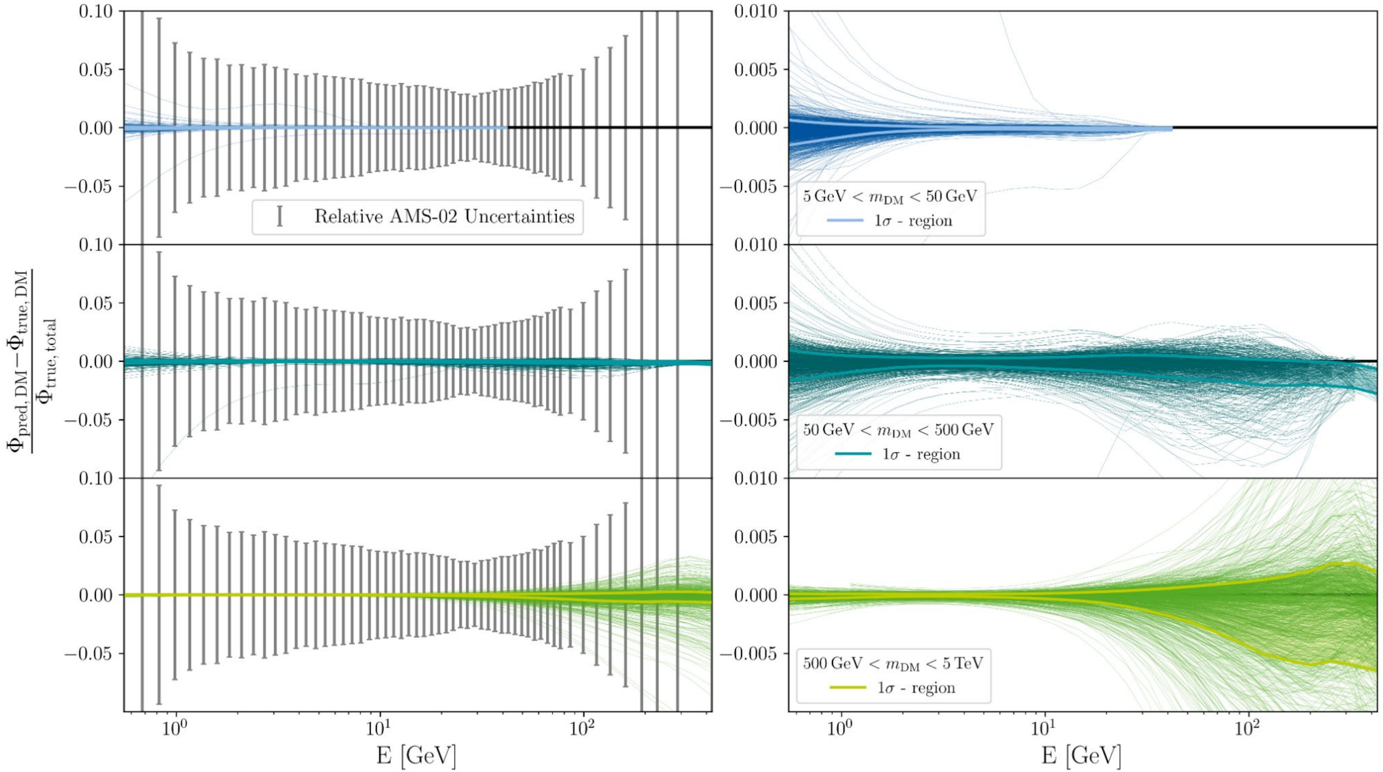
29.10.2021
Kathrin Nippel

Backup: Loss Curves



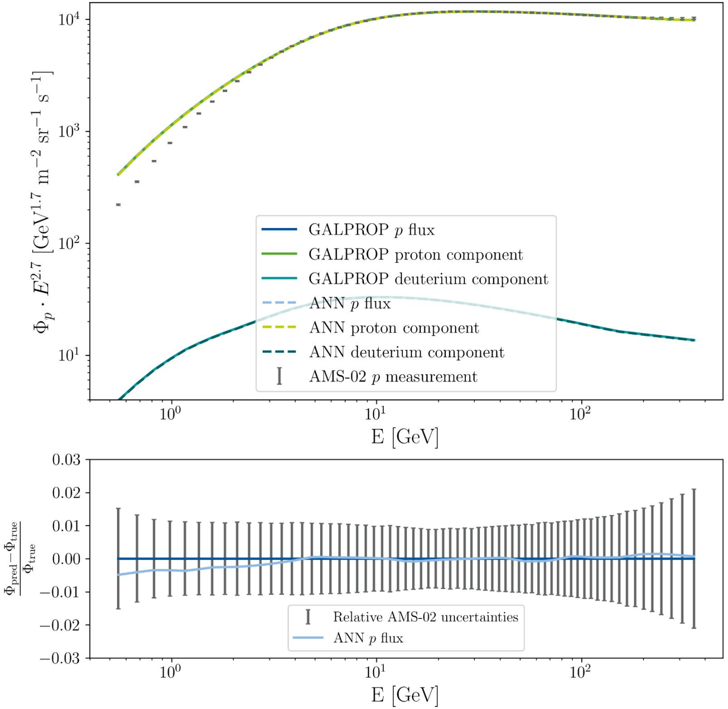


Backup:
Zoomed-in Spectra Validation



29.10.2021
Kathrin Nippel

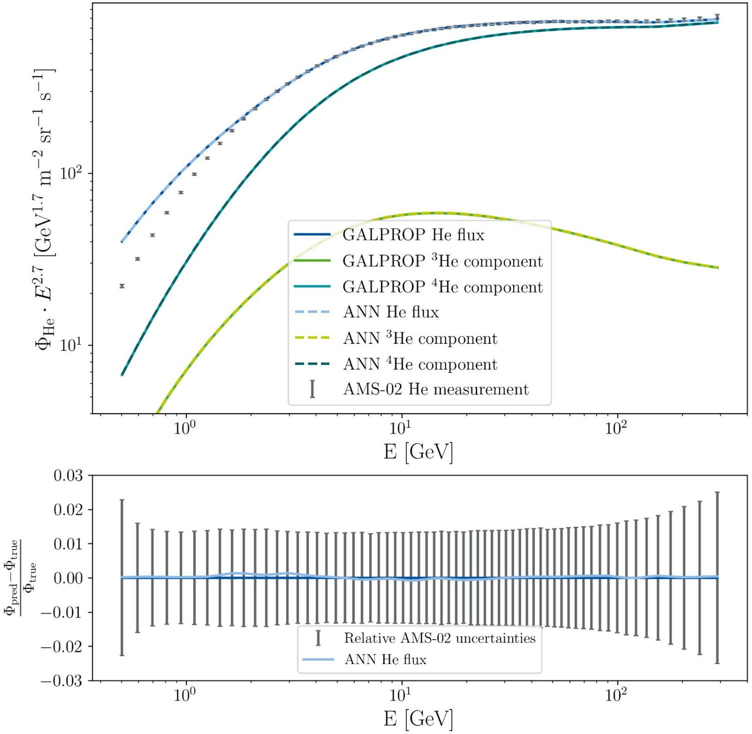
Backup:
Proton and Helium Networks



Propagation parameters:
$\gamma_{1,\mathrm{p}} = 1.79$
$\gamma_1 = 1.77$
$\gamma_{2,\mathrm{p}} = 2.42$
$\gamma_2 = 2.37$
$R_0 = 7417.87$ MV
$s = 0.35$
$D_0 = 2.52\mathrm{e}{+28}$ cm^2/s
$\delta = 0.41$
$v_{\mathrm{Alfven}} = 6.65$ km s^{-1}
$v_{0,c} = 0.86$ km s^{-1}
$z_h = 3.19$ kpc

Profiled parameters:
GALPROP:
$\varphi_{\mathrm{AMS-02}} = 0.25$ GV
$A_{\mathrm{AMS-02}} = 1.18$

ANN:
$\varphi_{\mathrm{AMS-02}} = 0.25$ GV
$A_{\mathrm{AMS-02}} = 1.18$



Propagation parameters:
$\gamma_{1,\mathrm{p}} = 1.87$
$\gamma_1 = 1.85$
$\gamma_{2,\mathrm{p}} = 2.41$
$\gamma_2 = 2.36$
$R_0 = 9285.81$ MV
$s = 0.34$
$D_0 = 3.64\mathrm{e}{+28}$ cm^2/s
$\delta = 0.41$
$v_{\mathrm{Alfven}} = 10.44$ km s^{-1}
$v_{0,c} = 1.13$ km s^{-1}
$z_h = 4.47$ kpc

Profiled parameters:
GALPROP:
$\varphi_{\mathrm{AMS-02}} = 0.29$ GV
$A_{\mathrm{AMS-02}} = 1.23$

ANN:
$\varphi_{\mathrm{AMS-02}} = 0.29$ GV
$A_{\mathrm{AMS-02}} = 1.23$