Characterizing the *Fermi*-LAT high-latitude sky with simulation-based inference

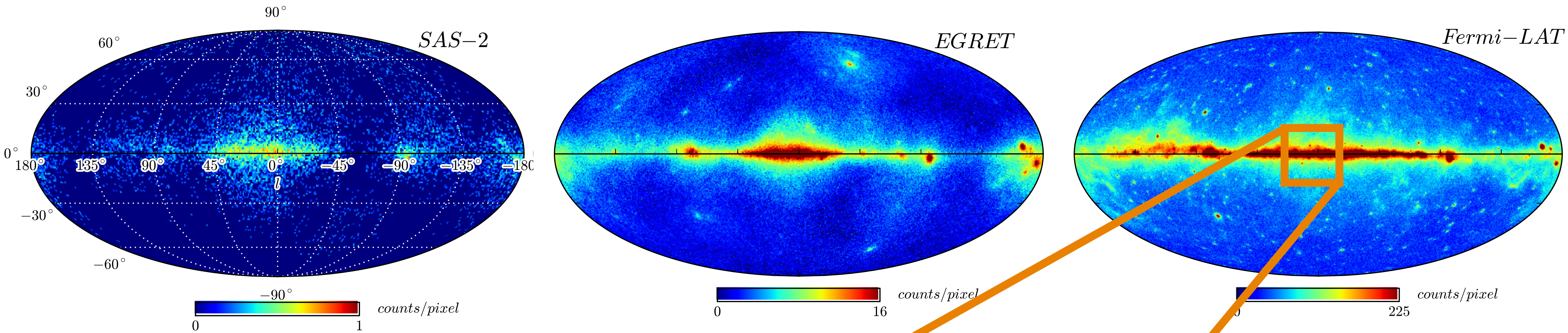
Christopher Eckner (christopher.eckner@ung.si)
SMASH post-doctoral fellow
University of Nova Gorica, Center for Astrophysics and Cosmology, Slovenia

1st SMASHing Workshop 2024
7th - 11th of October 2024 | University of Nova Gorica, Vipava

Fundamental physics with gamma rays is hard

The high-energy gamma-ray sky seen over the decades (space-borne telescopes).

1972-73 **1991-2000** **2008-now**

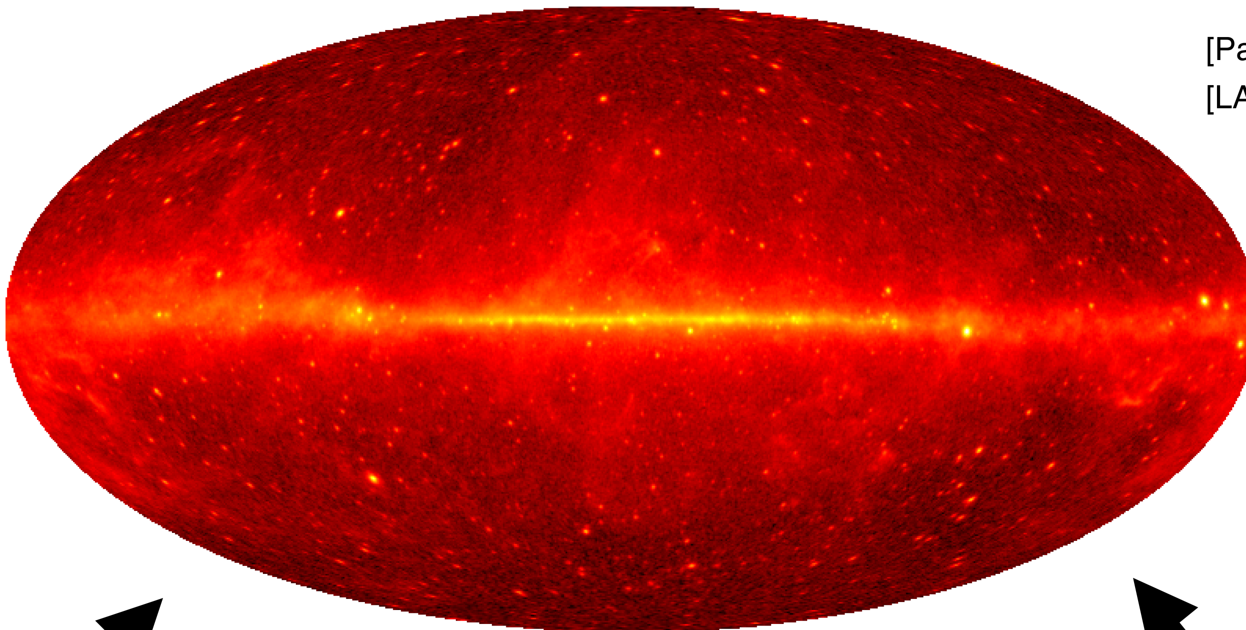


[Fermi-LAT collaboration, ApJS 223 (2016) 2]

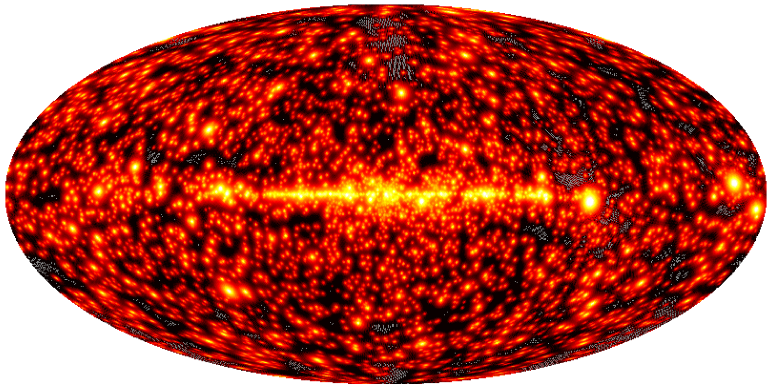
[Bertone, Tait, *Nature* 562 (2018) 7725]

Signatures of fundamental physics are potentially hiding there!
→ How to deal with the **complexity** of all the astrophysics?
There is a lot to model …

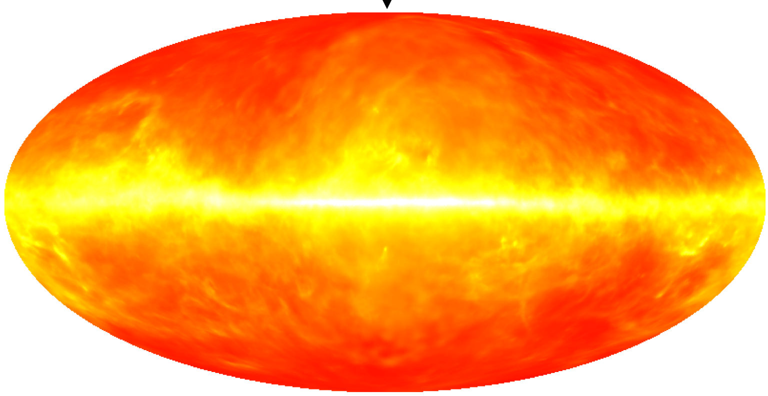
Understanding the gamma-ray sky



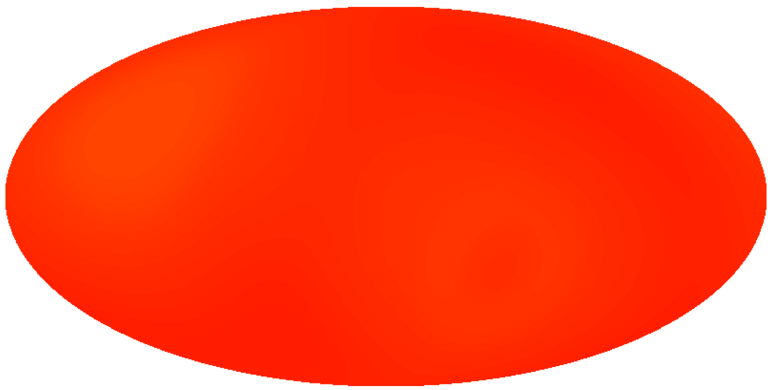
[Pass8 (R3), ULTRACLEANVETO, FRONT+BACK]
[LAT data (9.5 years), $E \in [0.5, 500]$ GeV]



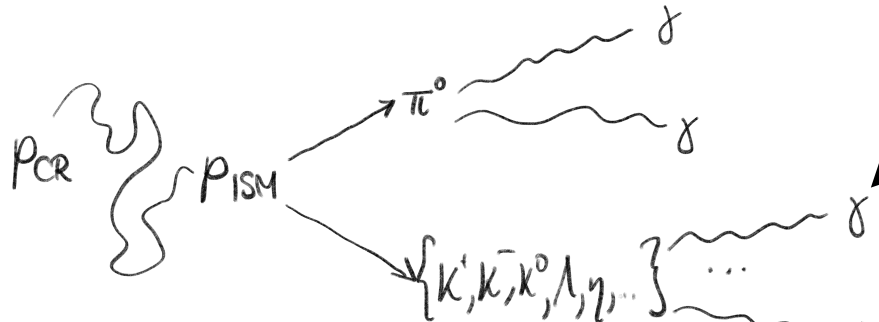
localised sources



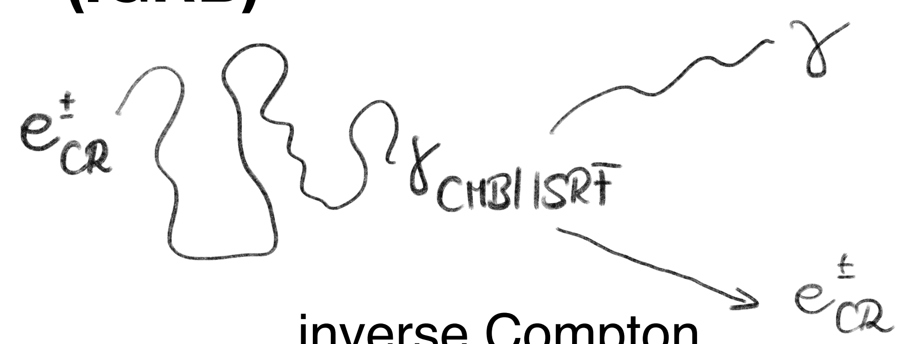
Galactic diffuse emission



isotropic γ-ray background (IGRB)



π^0-decay + Bremsstrahlung



inverse Compton

A word about the diffuse emission

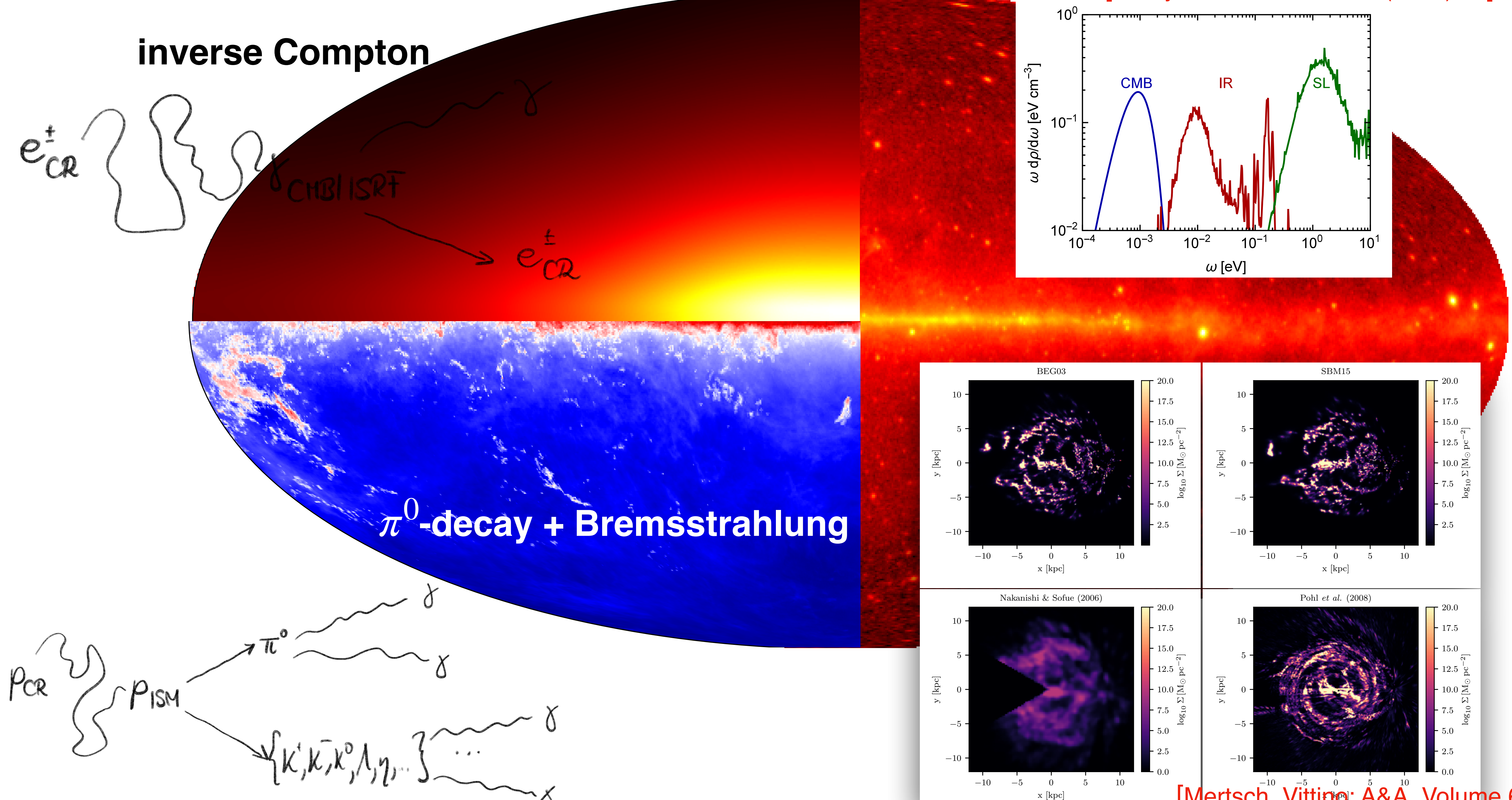
Product of charged cosmic rays interactions within the Milky Way:

— primary cosmic rays $(p, e^{\pm})$ accelerated and injected at source site

— propagate through the Milky Way (diffusion, convection, diffusive re-acceleration, popular solvers: GALPROP, DRAGON)

— **interactions with gas** (hadronic processes, Bremsstrahlung) and **radiation fields** (inverse Compton)

[Dobrynina et al.; PRD 95 (2017) 10]



[Mertsch, Vittino; A&A, Volume 655, A64]

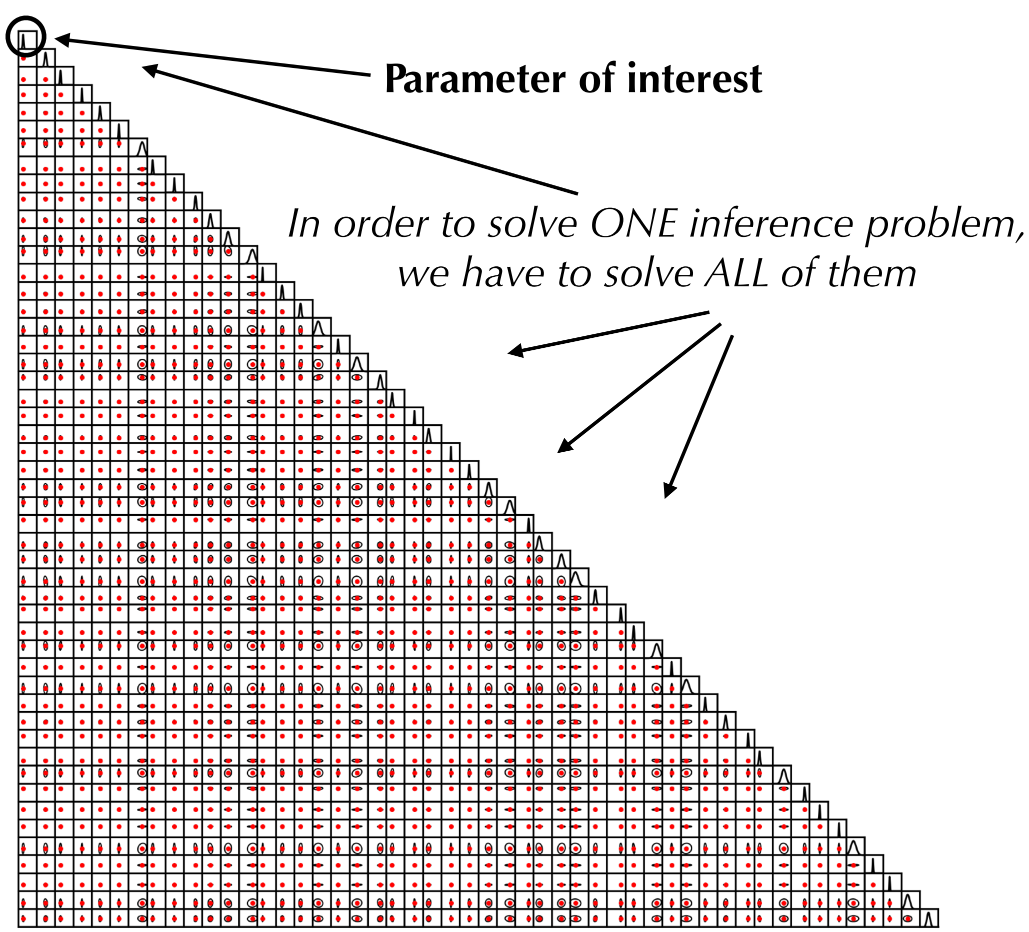
Striving for model complexity is expensive

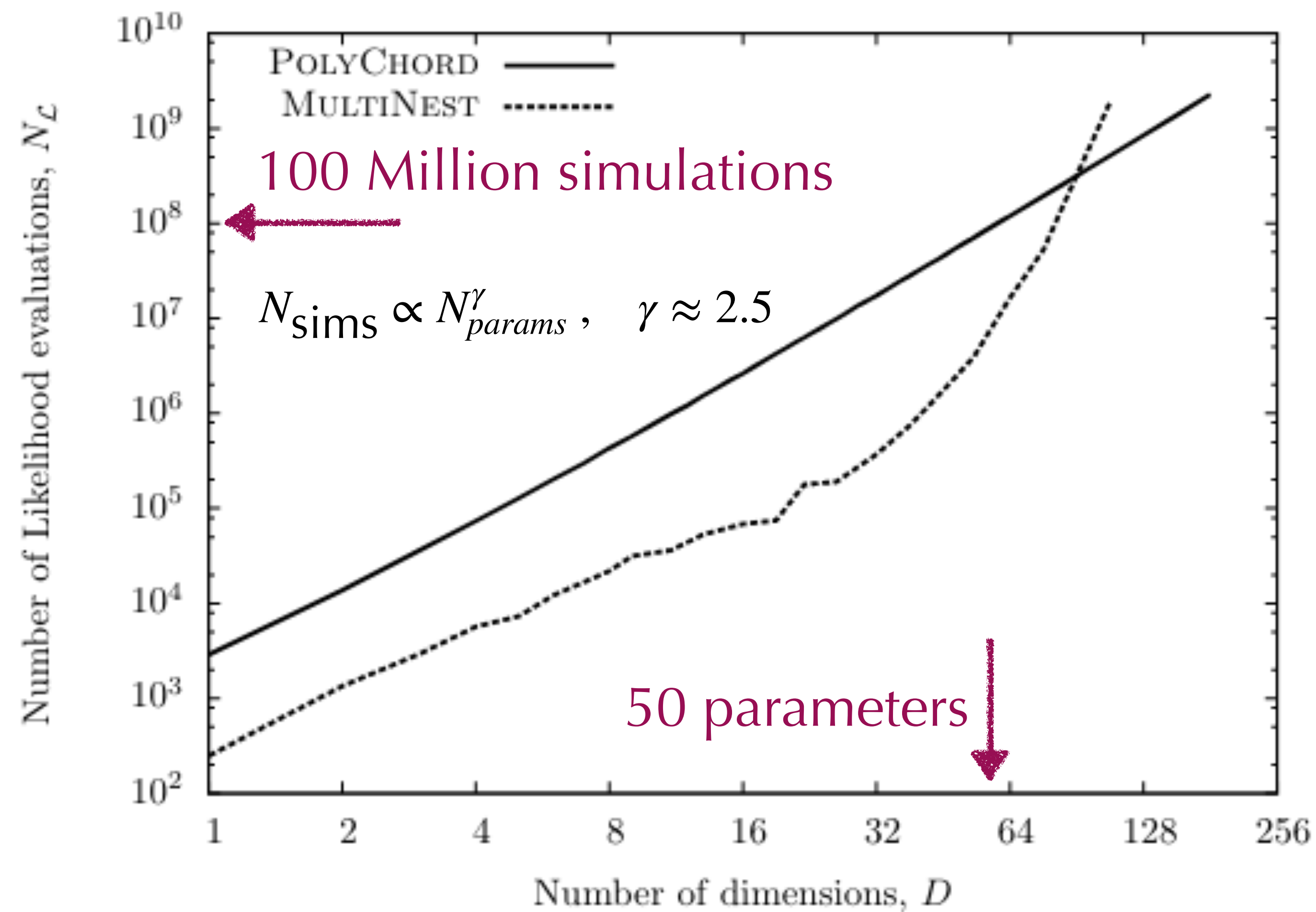
The curse of dimensionality in Bayesian inference problems:

posterior — likelihood — prior

Bayes' Theorem $$p(Z|X) = \frac{p(X|Z)\,p(Z)}{p(X)}$$

parameters Z**, data** X





[credit: Christoph Weniger]

[F. Feroz et al., MNRAS 398 (2009)]
[W. J. Handley et al., MNRAS 450 (2015) 1]

Simulation-based inference (*swyft*) to the rescue!

Simulation-based inference (SBI) breaks this curse:

[B. Miller et al., J. Open Source Softw. 7 (2022) 75]

Use machine learning to learn the posterior-to-prior ratio r with a binary classifier to reconstruct posteriors! **Neural Ratio Estimation (NRE)**

swyft

Truncated
Marginal
Neural
Ratio
Estimation

$$r(X;Z) = \frac{p(Z|X)}{p(Z)} = \frac{p(X,Z)}{p(X)p(Z)}$$

See also talk by F. List!

joint sample

Class 1: Matching (data, parameter) pairs

(😽, cat) (🐕, dog) (🏠, house) (⭐, star) (🦍, monkey)

$X, Z \sim P(X, Z)$

marginal sample

Class 0: Scrambled (data, parameter) pairs

(😽, dog) (🐕, cat) (⭐, star) (🏠, monkey) (🦍, house)

$X, Z \sim P(X)P(Z)$

In a nutshell: We train a neural network as a binary classifier to tell us if in a pair (X, Z) Z generated X which **only requires a forward model of the physics involved**.

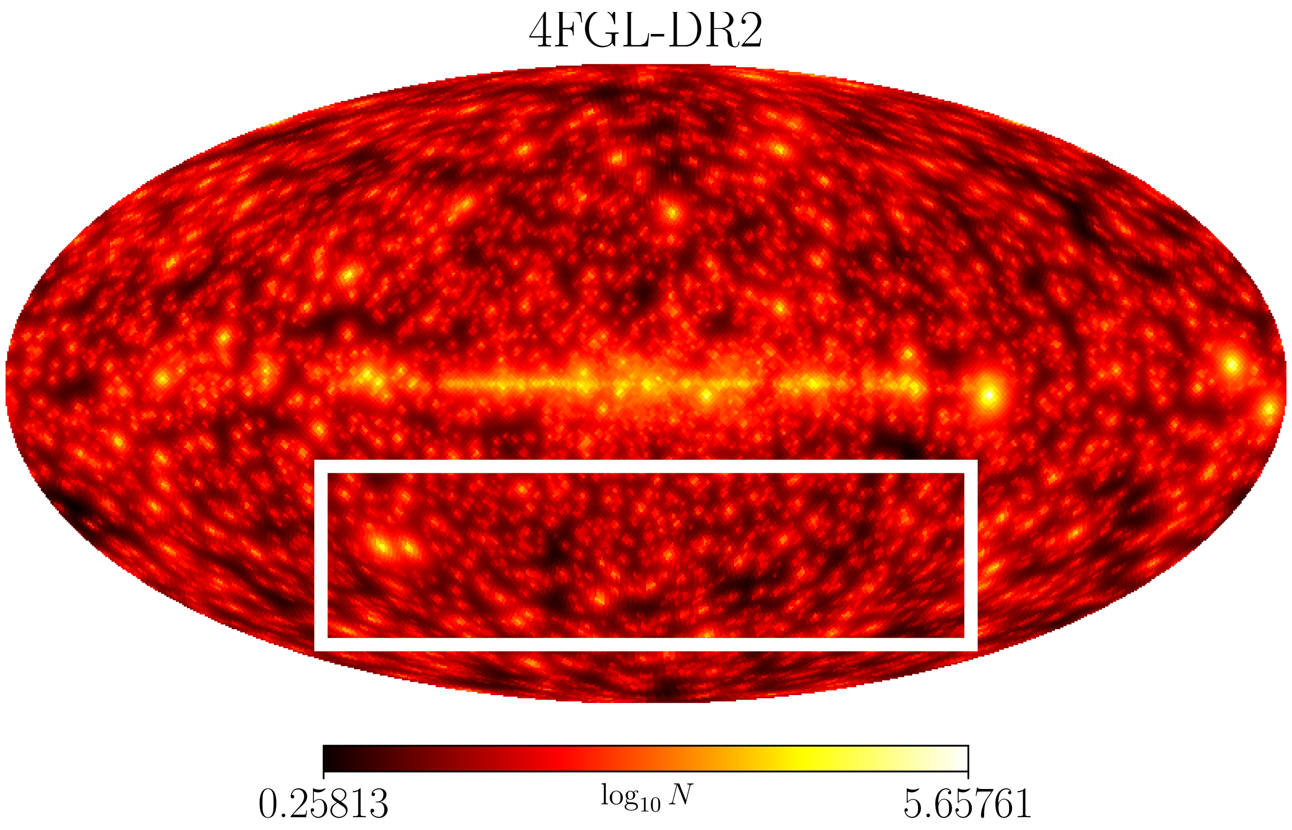
Dissecting the high-latitude GeV gamma-ray sky

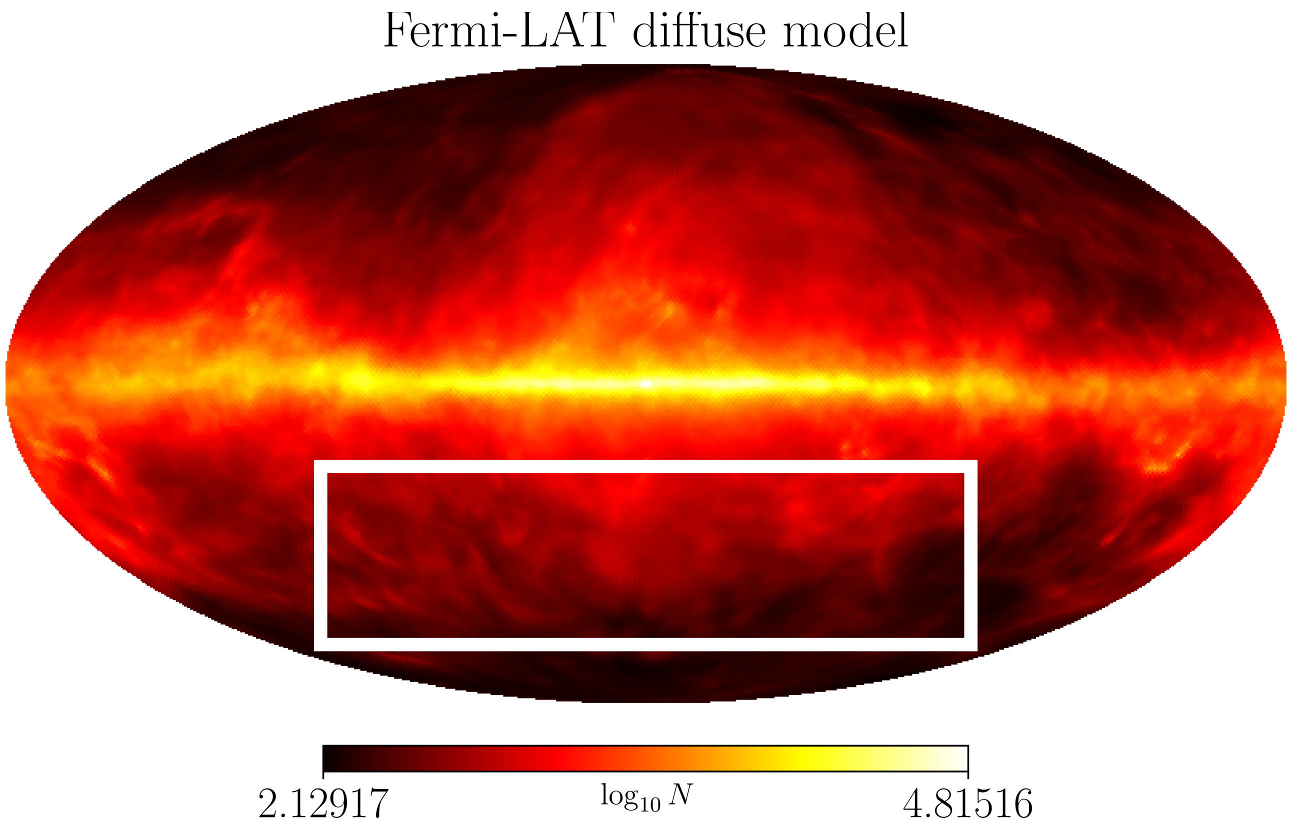
Our approach: Verify and benchmark the performance of our SBI approach on well-known terrain before addressing more complex questions about the LAT sky!

Optimal science case: Exploring the properties of the high-latitude gamma-ray sky

Why?

1. Much less affected by Galactic diffuse emission than, e.g., the Galactic center.
2. Limited number of gamma-ray source classes present (majority of extragalactic origin).
3. Well-tested science case: Opportunity for performance of cross-checks!
4. Science case: Composition of the IGRB → Contribution of astrophysical/exotic source classes.





Dissecting the high-latitude GeV gamma-ray sky

Objective: Infer the **source-count distribution** of high-latitude sources and the astrophysical diffuse gamma-ray emission and **localise the bright part of the population (detection)**.

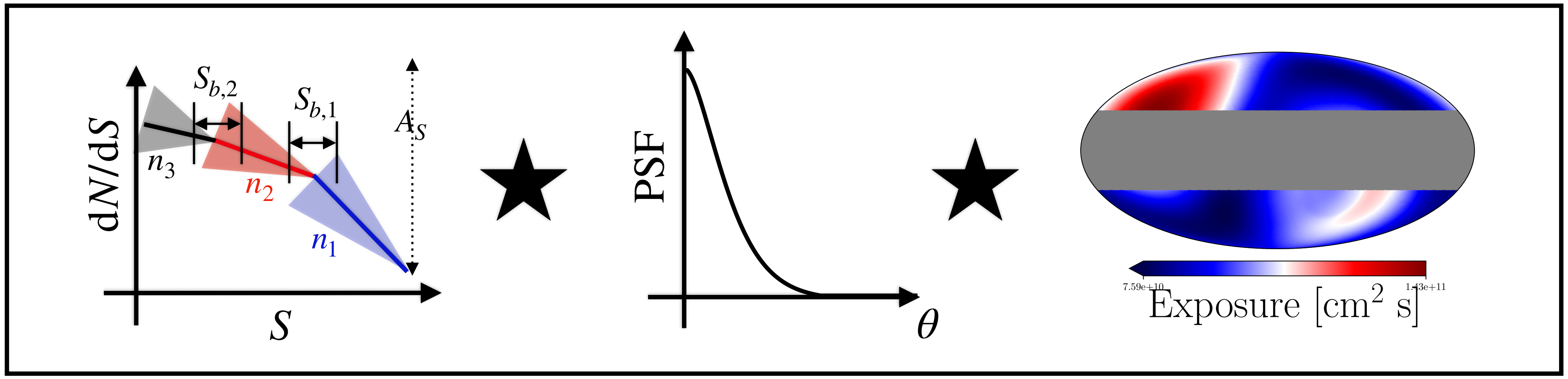
Source-count distribution $\mathrm{d}N/\mathrm{d}S$**:** # of sources N per $\mathrm{d}\Omega$ with integral flux in $(S, S + \mathrm{d}S)$.

Dissecting the high-latitude GeV gamma-ray sky

Objective: Infer the **source-count distribution** of high-latitude sources and the astrophysical diffuse gamma-ray emission and **localise the bright part of the population (detection)**.

Source-count distribution $\mathrm{d}N/\mathrm{d}S$**:** # of sources N per $\mathrm{d}\Omega$ with integral flux in $(S, S+\mathrm{d}S)$.

Forward simulator: $\mathrm{d}N/\mathrm{d}S$ as multiply broken power law (norm, break positions and slopes)



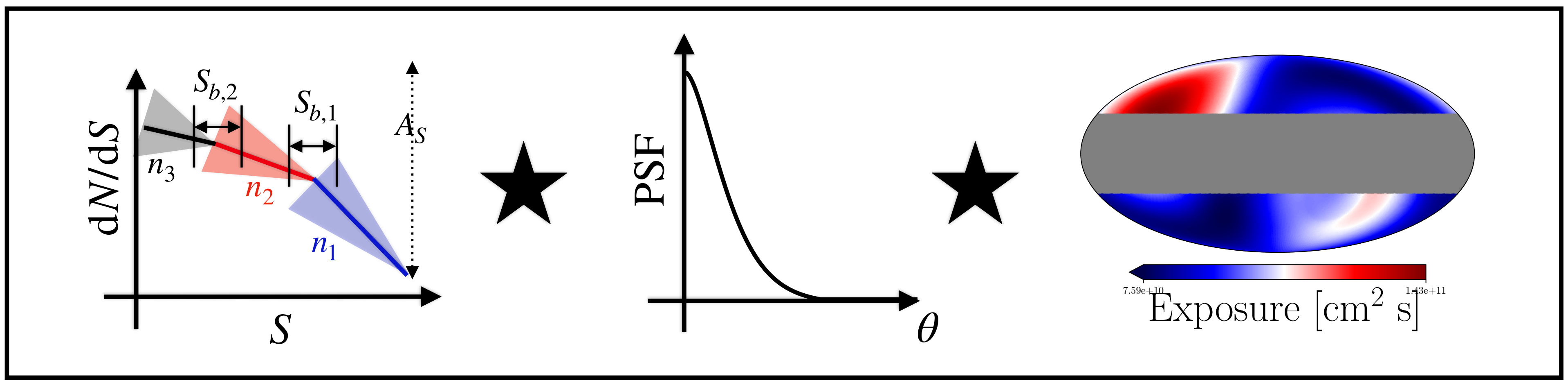
- Flux S for single energy bin from 1 GeV to 10 GeV.
- Correction for PSF effects using effective PSF derived from data (`gtpsf`).
- Uses *Fermi*-LAT non-uniform exposure.

Dissecting the high-latitude GeV gamma-ray sky

Objective: Infer the **source-count distribution** of high-latitude sources and the astrophysical diffuse gamma-ray emission and **localise the bright part of the population (detection)**.

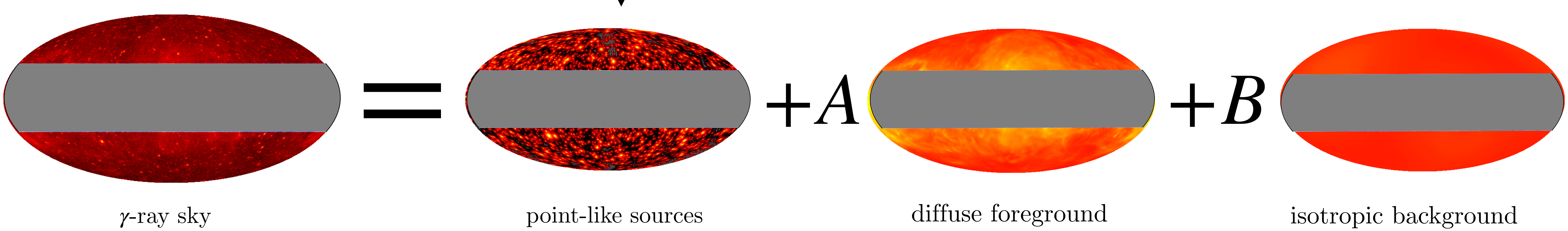
Source-count distribution $\mathrm{d}N/\mathrm{d}S$**:** # of sources N per $\mathrm{d}\Omega$ with integral flux in $(S, S + \mathrm{d}S)$.

Forward simulator: $\mathrm{d}N/\mathrm{d}S$ as multiply broken power law (norm, break positions and slopes)

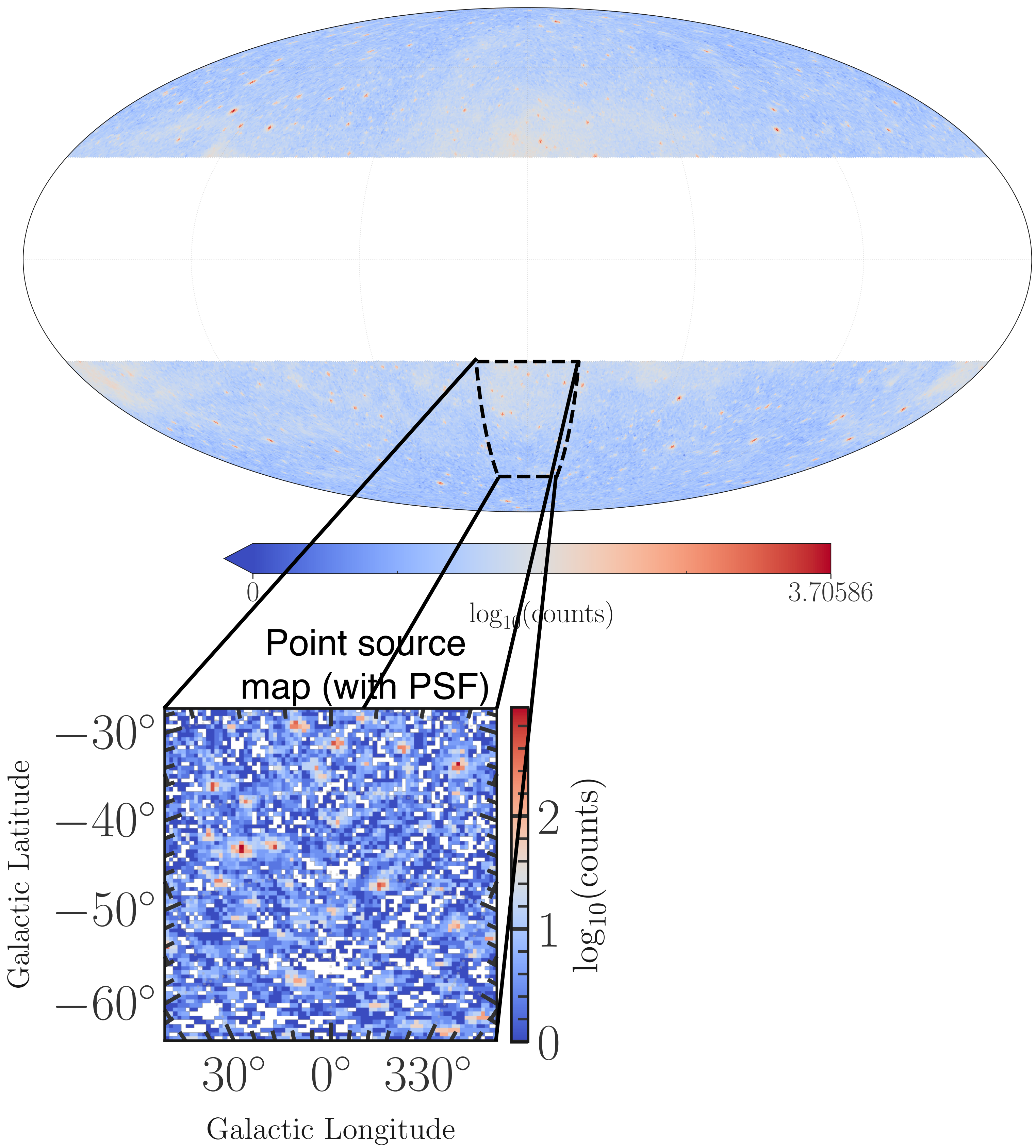


- Flux S for single energy bin from 1 GeV to 10 GeV.
- Correction for PSF effects using effective PSF derived from data (`gtpsf`).
- Uses *Fermi*-LAT non-uniform exposure.

isotropic distribution in the sky



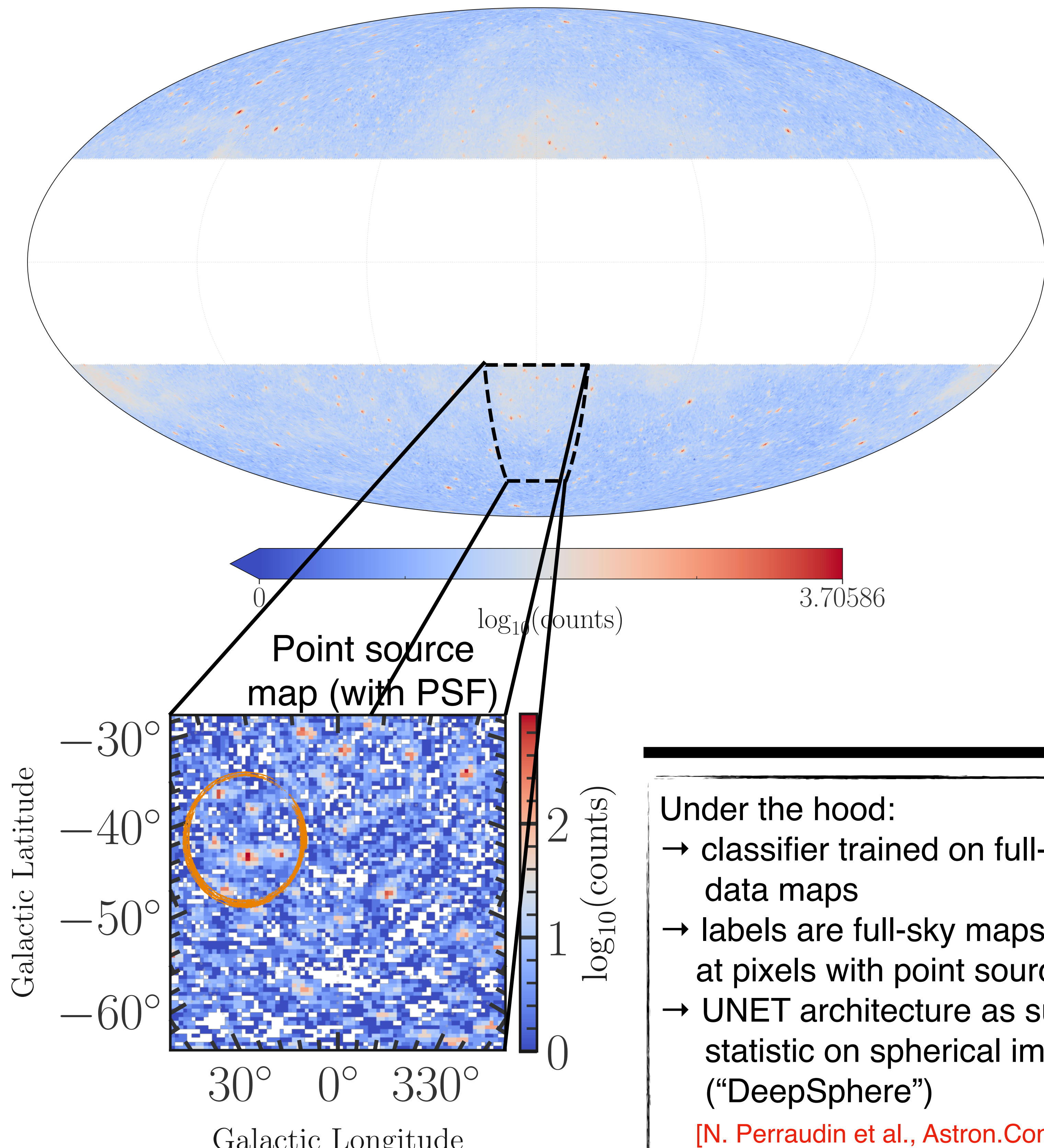
Source Detection using SBI — Method



[N. Anau Montel & C. Weniger, arXiv:2211.04291]

Source detection in SBI language:
Given the actual observed sky, what is the probability of observing a source at a certain position with flux S exceeding a certain threshold S_{th}?

$$r(\Omega, S_{\text{th}}; x) = \frac{p(\mathbb{1}_x(S \geq S_{\text{th}}) = 1, \Omega \,|\, x)}{p(\mathbb{1}_x(S \geq S_{\text{th}}) = 1, \Omega)}$$



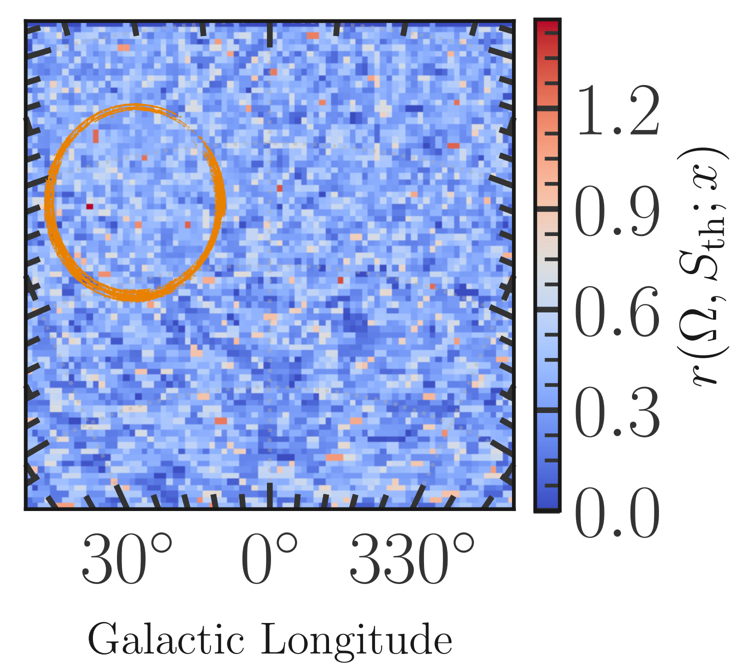
[N. Anau Montel & C. Weniger, arXiv:2211.04291]

Source detection in SBI language:
Given the actual observed sky, what is the probability of observing a source at a certain position with flux S exceeding a certain threshold S_{th}?

$$r(\Omega, S_{\mathrm{th}}; x) = \frac{p(\mathbb{I}_x(S \geq S_{\mathrm{th}}) = 1, \Omega \,|\, x)}{p(\mathbb{I}_x(S \geq S_{\mathrm{th}}) = 1, \Omega)}$$

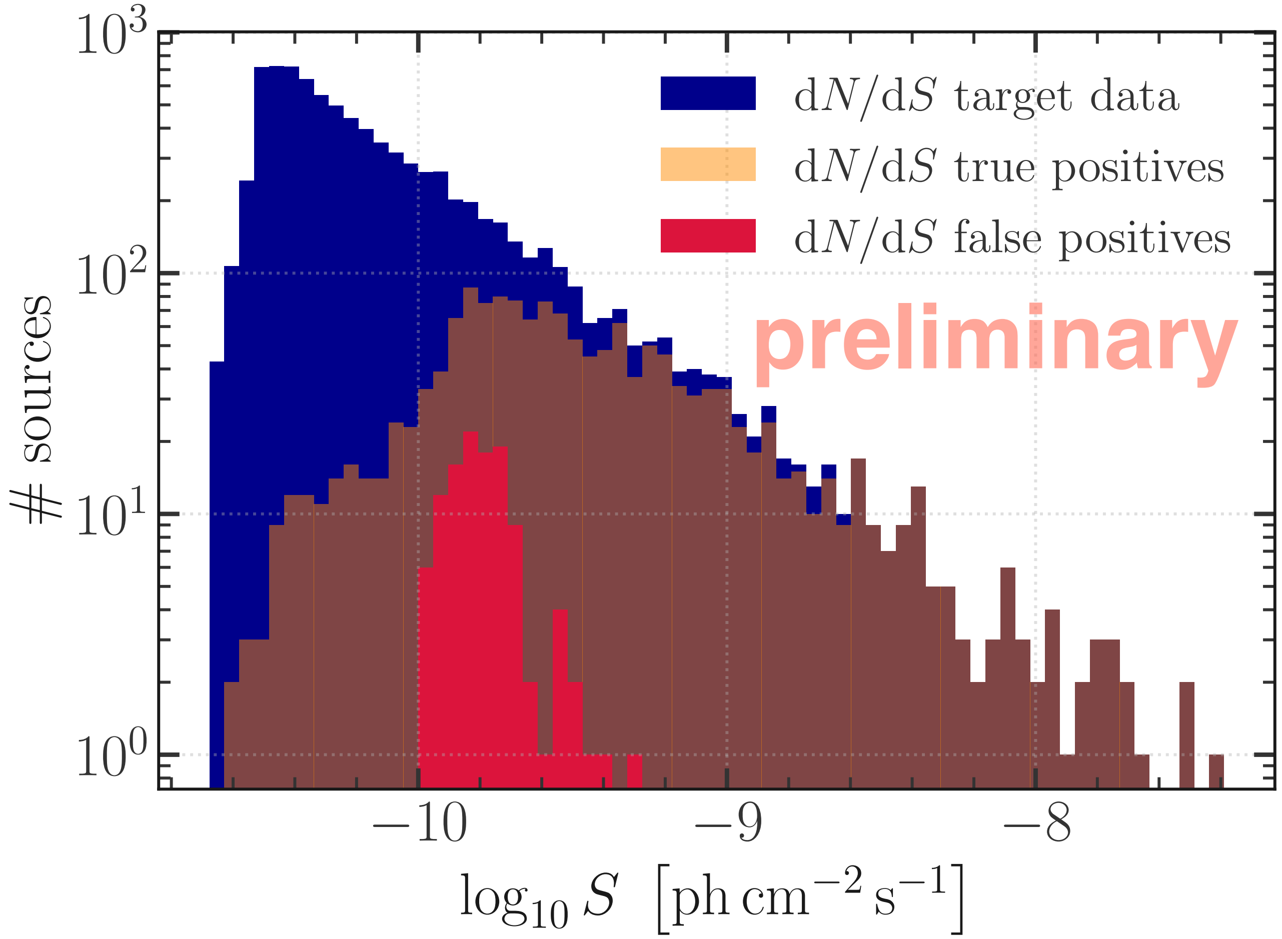
Under the hood:
→ classifier trained on full-sky data maps
→ labels are full-sky maps firing at pixels with point sources
→ UNET architecture as summary statistic on spherical image data (“DeepSphere”)

[N. Perraudin et al., Astron.Comput. 27 (2019)]



Source Detection using SBI — Results

The **shown results** concern **simulated data** with a double-broken $\mathrm{d}N/\mathrm{d}S$ using the best fit parameters derived in [Zechlin et al., Astrophys.J.Suppl. 225 (2016) 2, 18]!

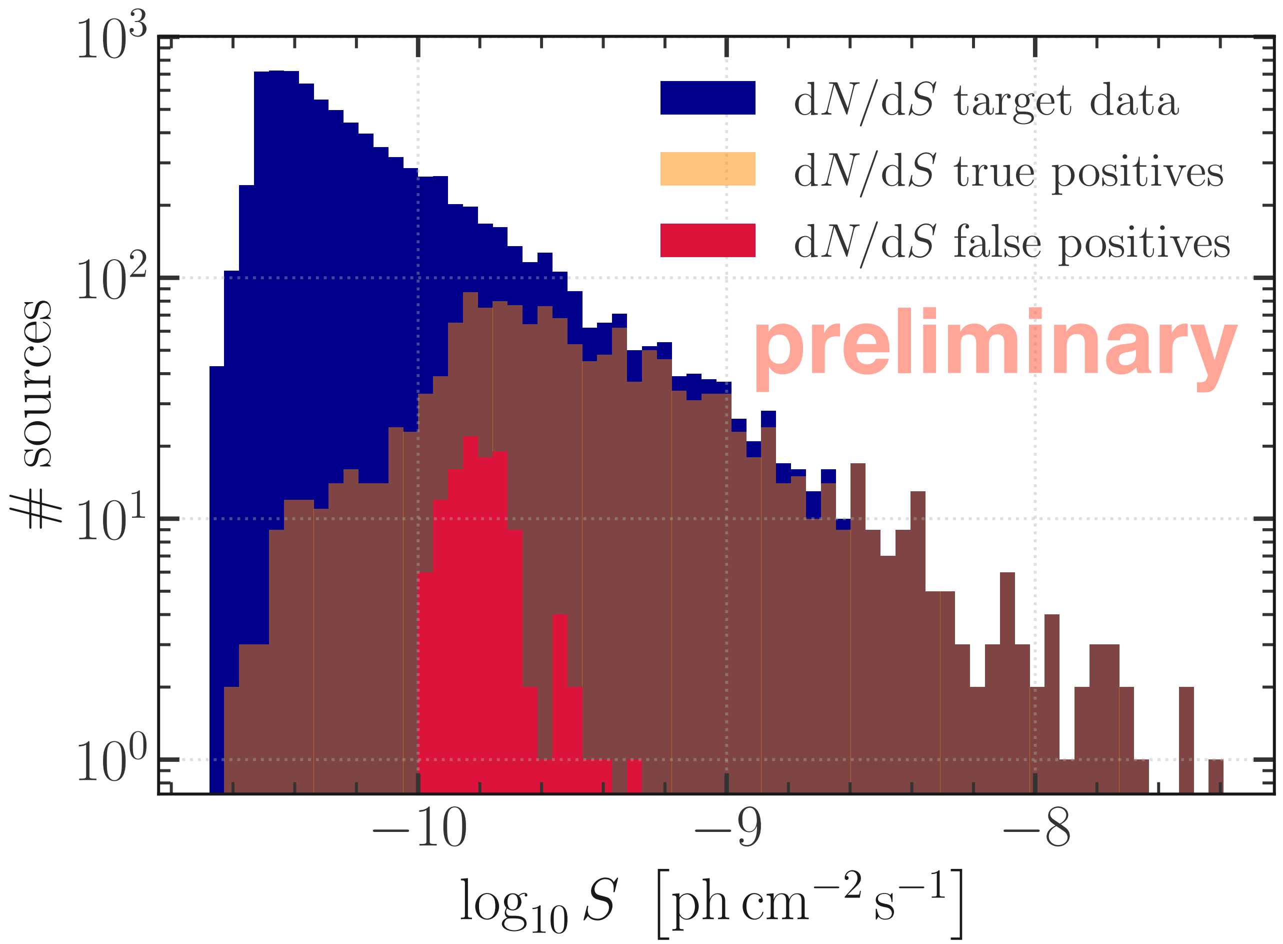


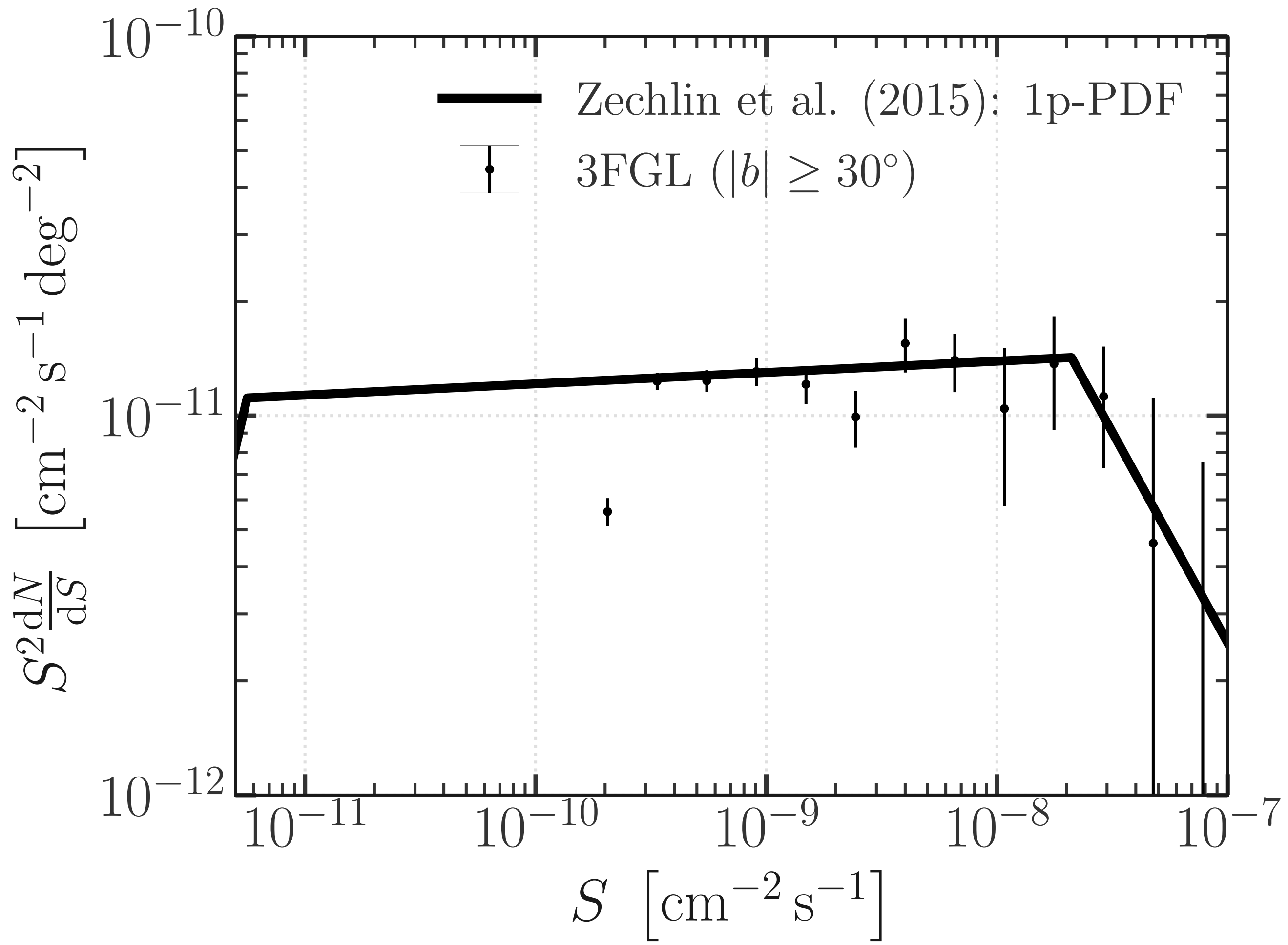
Blue: true $\mathrm{d}N/\mathrm{d}S$ of simulated target data
Orange: detected true sources using a cut on $r(\Omega, S_{\mathrm{th}}; x)$
Red: false positives (misclassified background fluctuations)

→ Overall false positive rate here: 7.5% of total detections.

Source Detection using SBI — Results

The **shown results** concern **simulated data** with a double-broken $\mathrm{d}N/\mathrm{d}S$ using the best fit parameters derived in [Zechlin et al., Astrophys.J.Suppl. 225 (2016) 2, 18]!





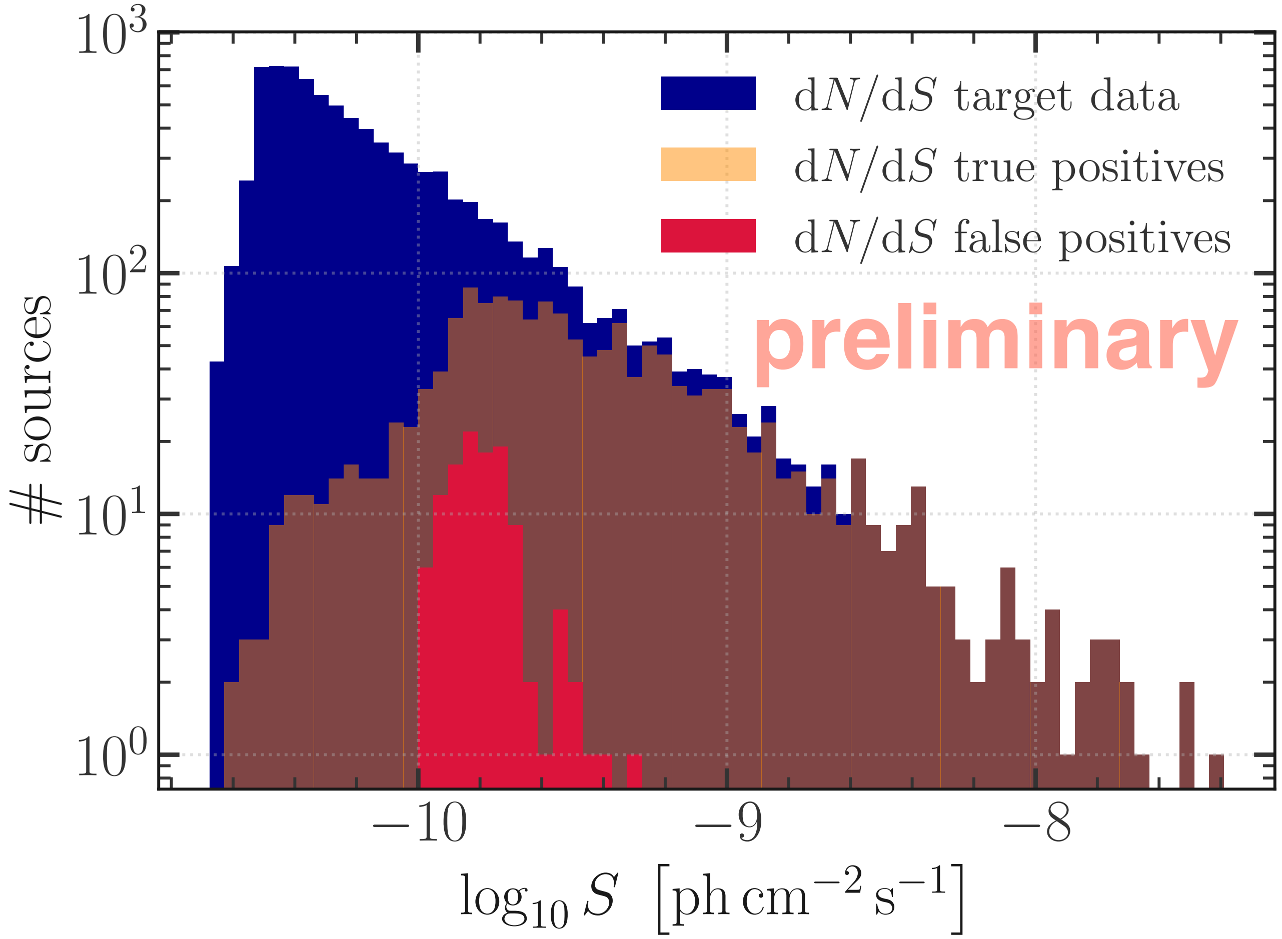
Blue: true $\mathrm{d}N/\mathrm{d}S$ of simulated target data
Orange: detected true sources using a cut on $r(\Omega, S_{\mathrm{th}}; x)$
Red: false positives (misclassified background fluctuations)

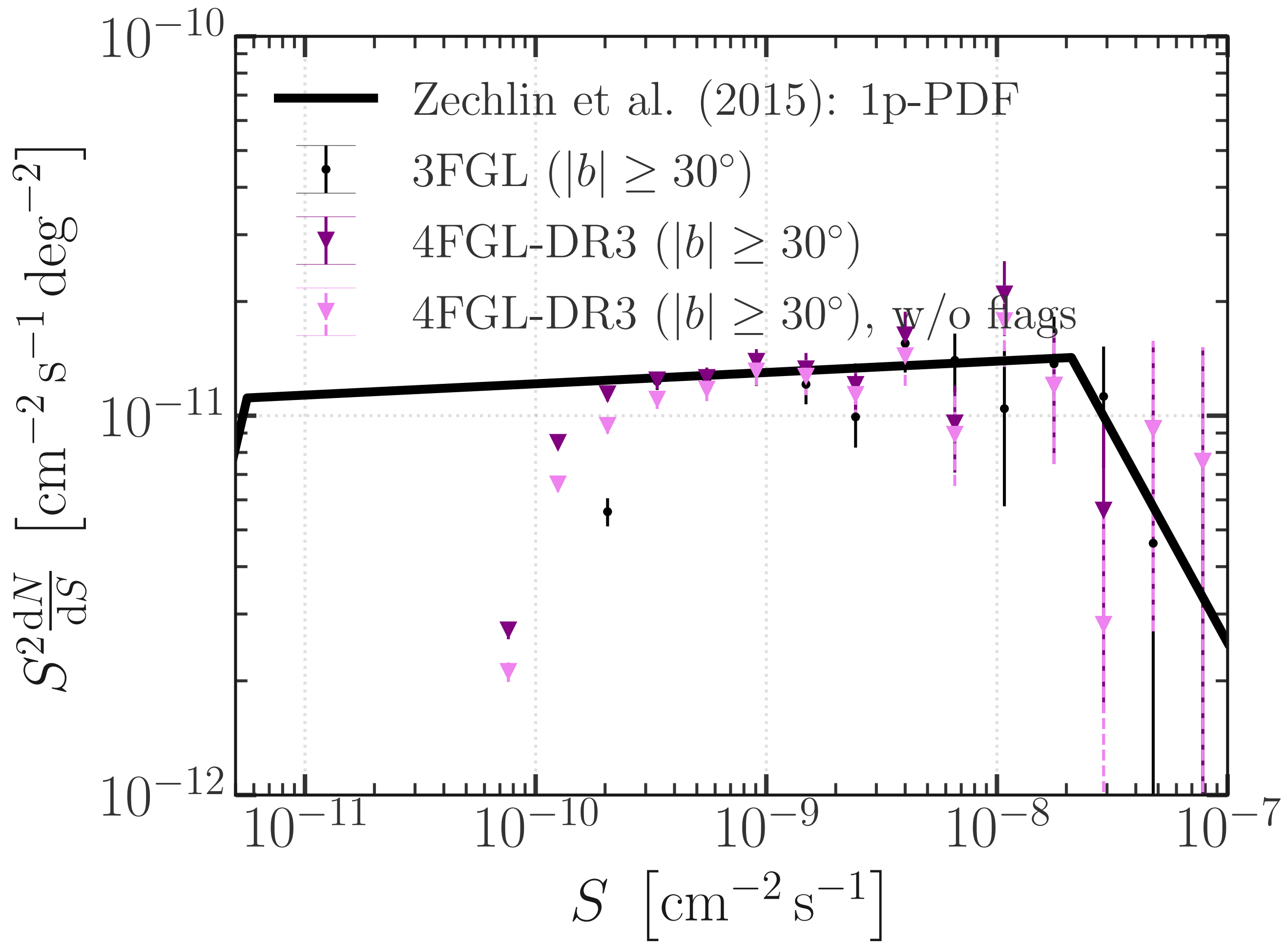
→ Overall false positive rate here: 7.5% of total detections.

$\mathrm{d}N/\mathrm{d}S$ and corresponding *Fermi*-LAT catalog to which our simulated data correspond
→ *Source positions and fluxes are different in our target data!*

Source Detection using SBI — Results

The **shown results** concern **simulated data** with a double-broken $\mathrm{d}N/\mathrm{d}S$ using the best fit parameters derived in [Zechlin et al., Astrophys.J.Suppl. 225 (2016) 2, 18]!





Blue: true $\mathrm{d}N/\mathrm{d}S$ of simulated target data
Orange: detected true sources using a cut on $r(\Omega, S_{\mathrm{th}}; x)$
Red: false positives (misclassified background fluctuations)

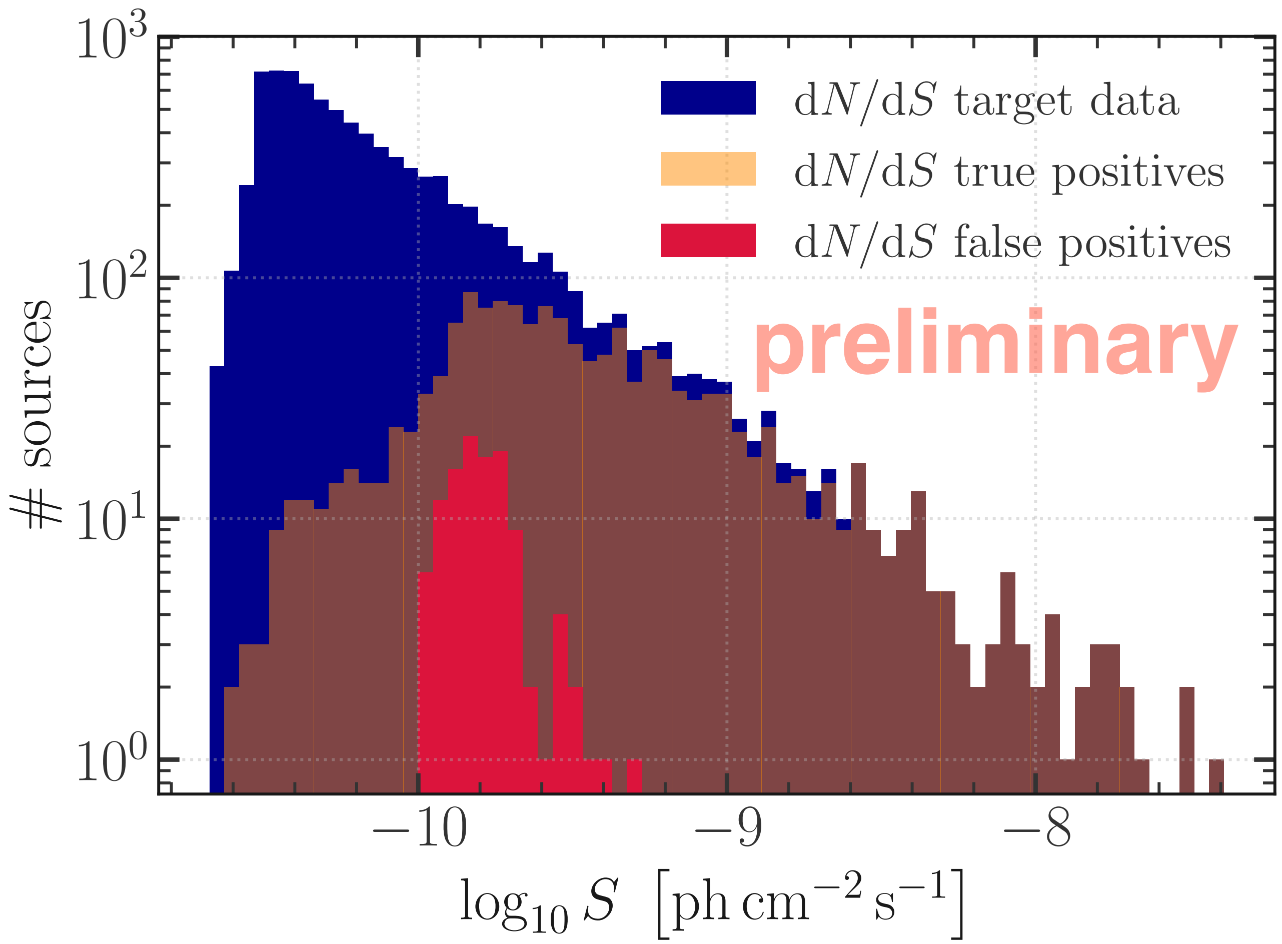
→ Overall false positive rate here: 7.5% of total detections.

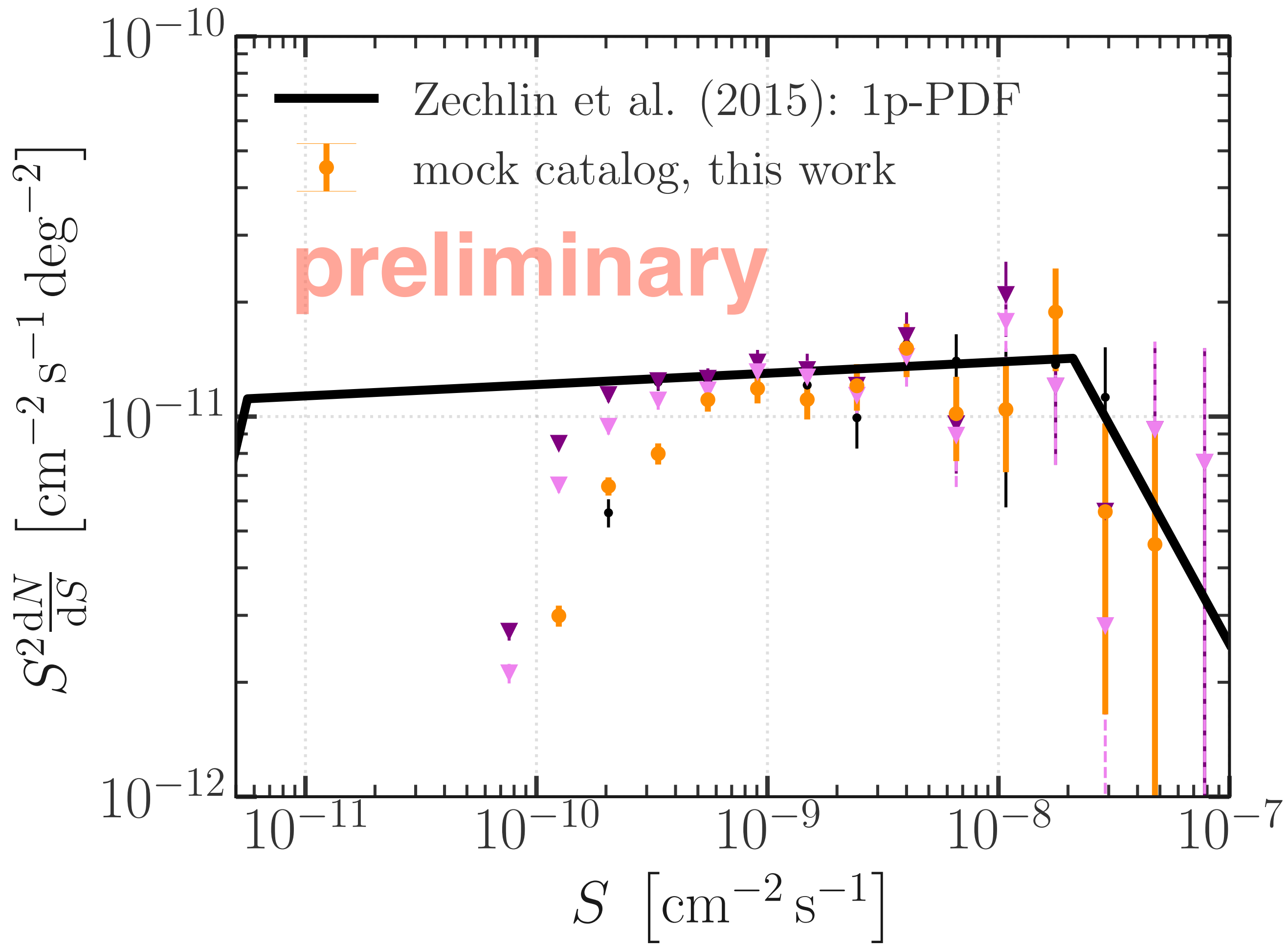
Improvement in number of detected sources with 12 years (4FGL-DR3) instead of 4 years.

→ Sources are flagged if the background is very bright at their position or they could be false positives.

Source Detection using SBI — Results

The **shown results** concern **simulated data** with a double-broken $\mathrm{d}N/\mathrm{d}S$ using the best fit parameters derived in [Zechlin et al., Astrophys.J.Suppl. 225 (2016) 2, 18]!





Blue: true $\mathrm{d}N/\mathrm{d}S$ of simulated target data
Orange: detected true sources using a cut on $r(\Omega, S_{\mathrm{th}}; x)$
Red: false positives (misclassified background fluctuations)

→ Overall false positive rate here: 7.5% of total detections.

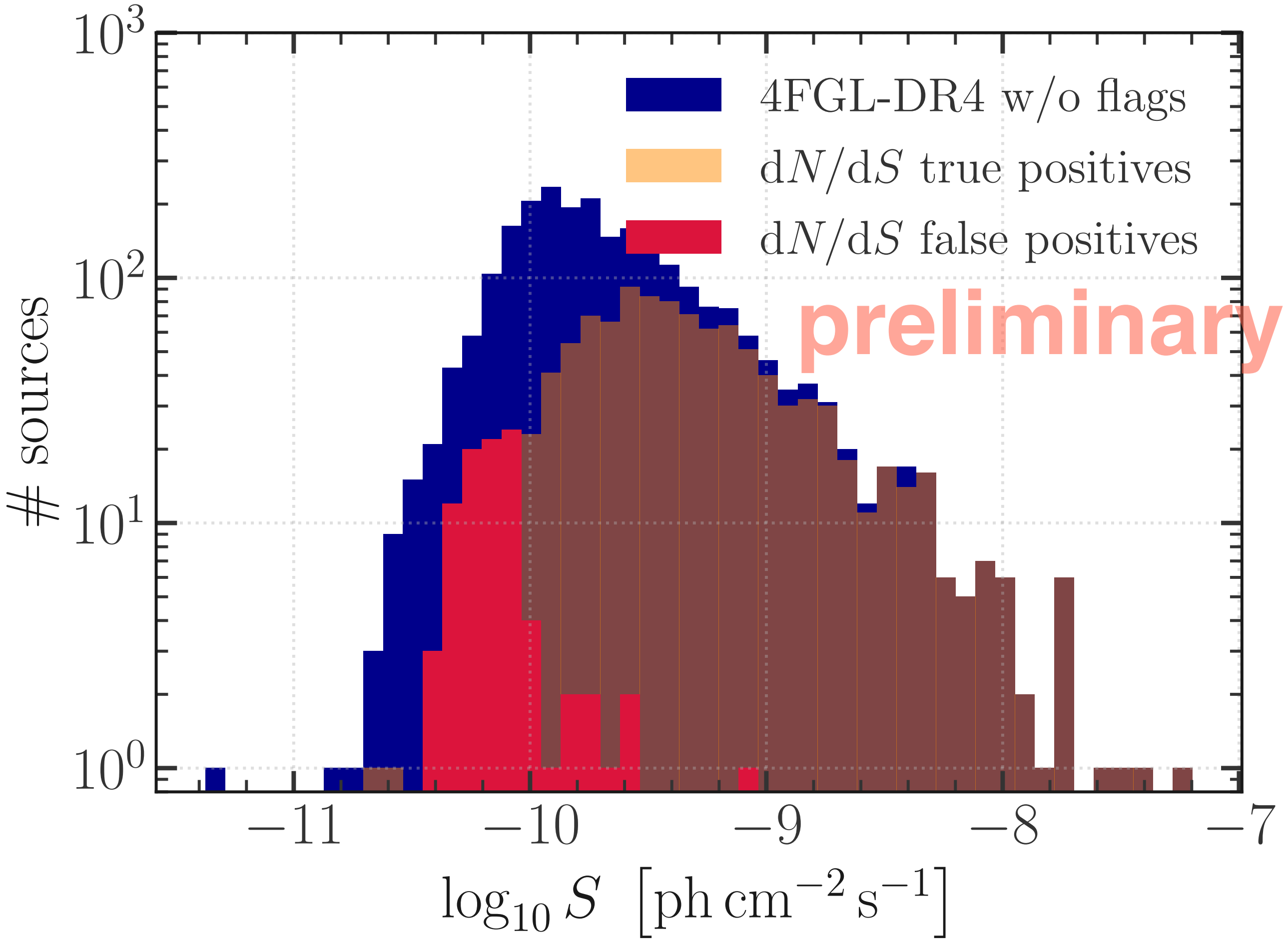
We use the same exposure as 4FGL-DR3.

→ **Our catalog loses efficiency of ~100% around $S = 10^{-9}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, comparable to flagged 4FGL-DR3.**

→ **In the dim-source regime, it performs like 3FGL.**

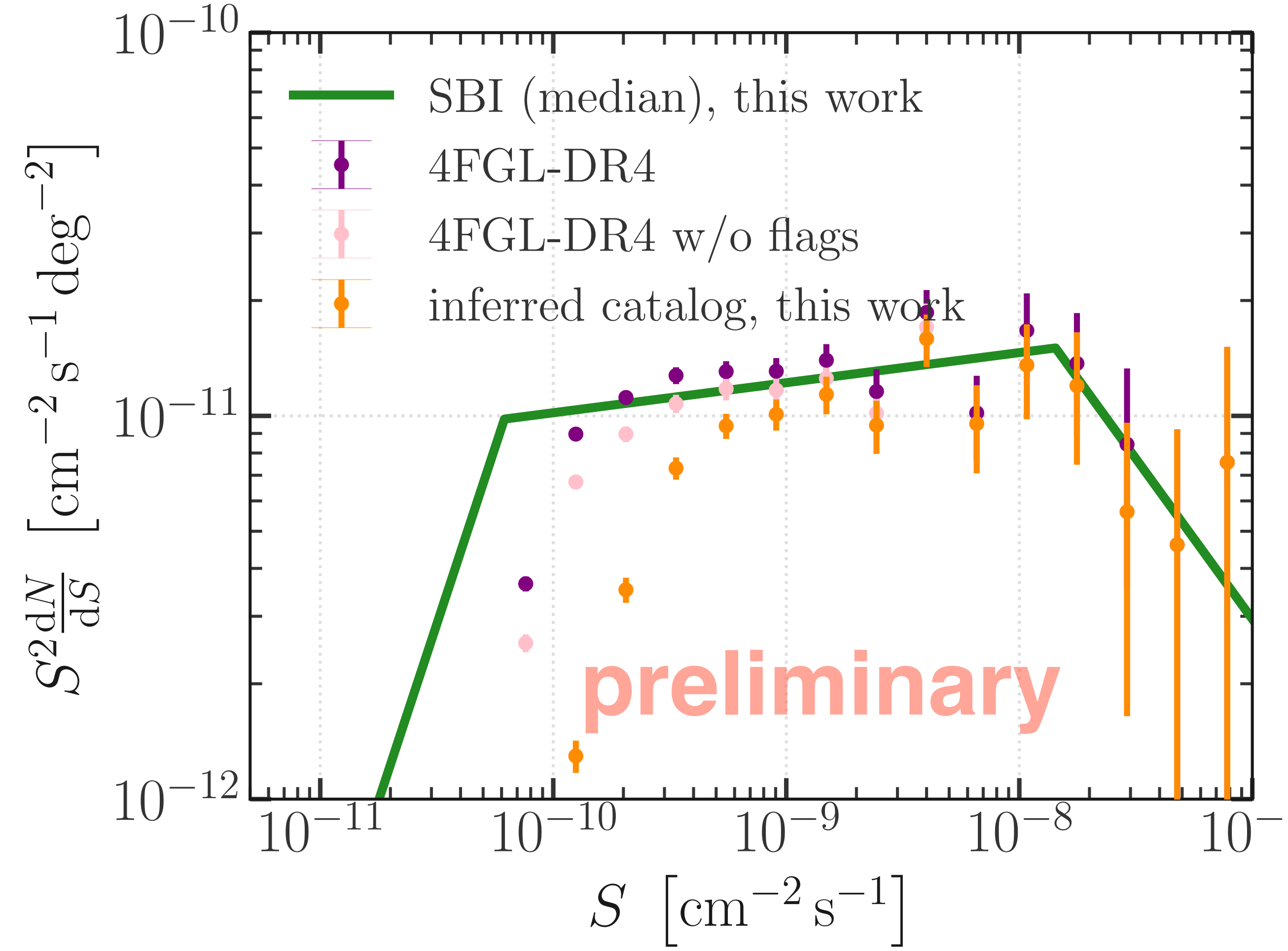
Source Detection using SBI — Results

Now we exchange the target data for the **real *Fermi*-LAT sky**; 12 years of data and, for the experts, `SOURCEVETO` event class with `FRONT` converted photons (1 GeV to 10 GeV).



Blue: true $\mathrm{d}N/\mathrm{d}S$ of simulated target data
Orange: detected true sources using a cut on $r(\Omega, S_{\mathrm{th}}; x)$
Red: false positives (misclassified background fluctuations *or new sources*?)

→ Overall false positive rate here: 9% of total detections (catalog-dependent statement; some might be actual sources)



Comparison with 4FGL-DR4 (14 years) for maximal amount of sources.

→ **Our framework detects quite robustly those point-like sources that have low background contamination (w/o flags).**

→ **In the dim-source regime, it performs like 3FGL.**

Summary and outlook

The road so far:

- We presented an SBI scheme that is features a **realistic simulator of *Fermi*-LAT data.** It is currently able to **localise bright sources** in binned all-sky gamma-ray data and to **infer the underlying parameters** of the model components. **Obtaining at the same time catalog and population parameters is a novelty!**

- The efficiency of **our SBI source catalog** does not reach the one of 4FGL-DR3 but it is **compatible once flagged sources are removed** while we achieve 3FGL efficiency for dim sources.

- The source detection on real LAT data traces the bright part of 4FGL-DR4 while showing **better performance on the clean (w/o flag) sample** than the full gamma-ray catalogue!

Future prospects:

- Using **non-parametric source-count distributions** as simulation framework.

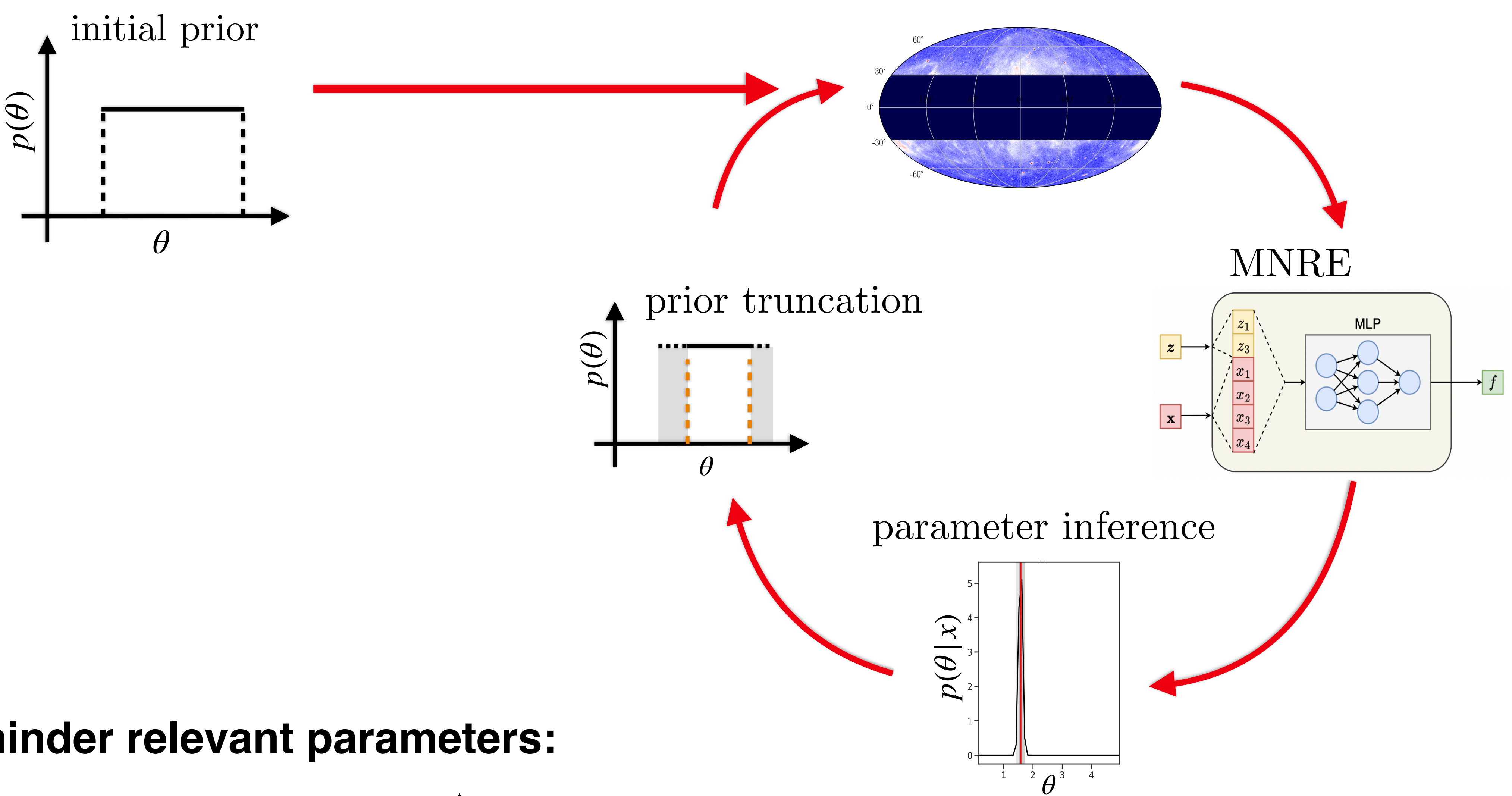
- Extending the framework to **multiple energy bins** and consequently multiple source classes with characteristic spectral shapes.

- On-the-fly **sampling** of diffuse Milky Way foreground from **uncertainty of gas structure.**
  [A. Ramírez et al., arXiv:2407.02410]

Backup slides

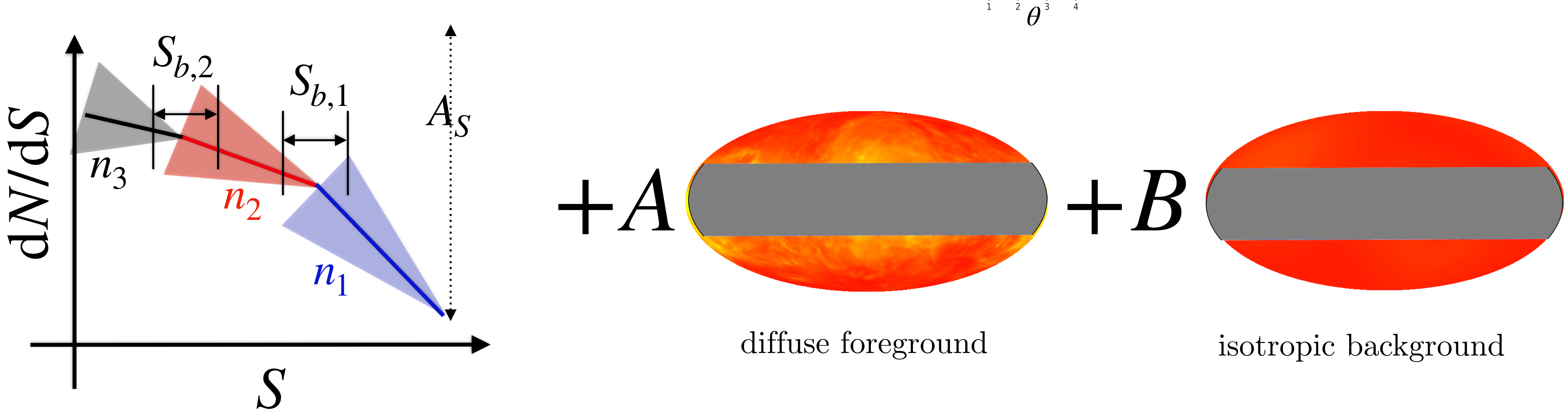
Parameter inference in our SBI framework

The **parameter inference** scheme of our SBI framework allows to perform **sequential inference** in multiple training rounds **based on the results of the previous round**.



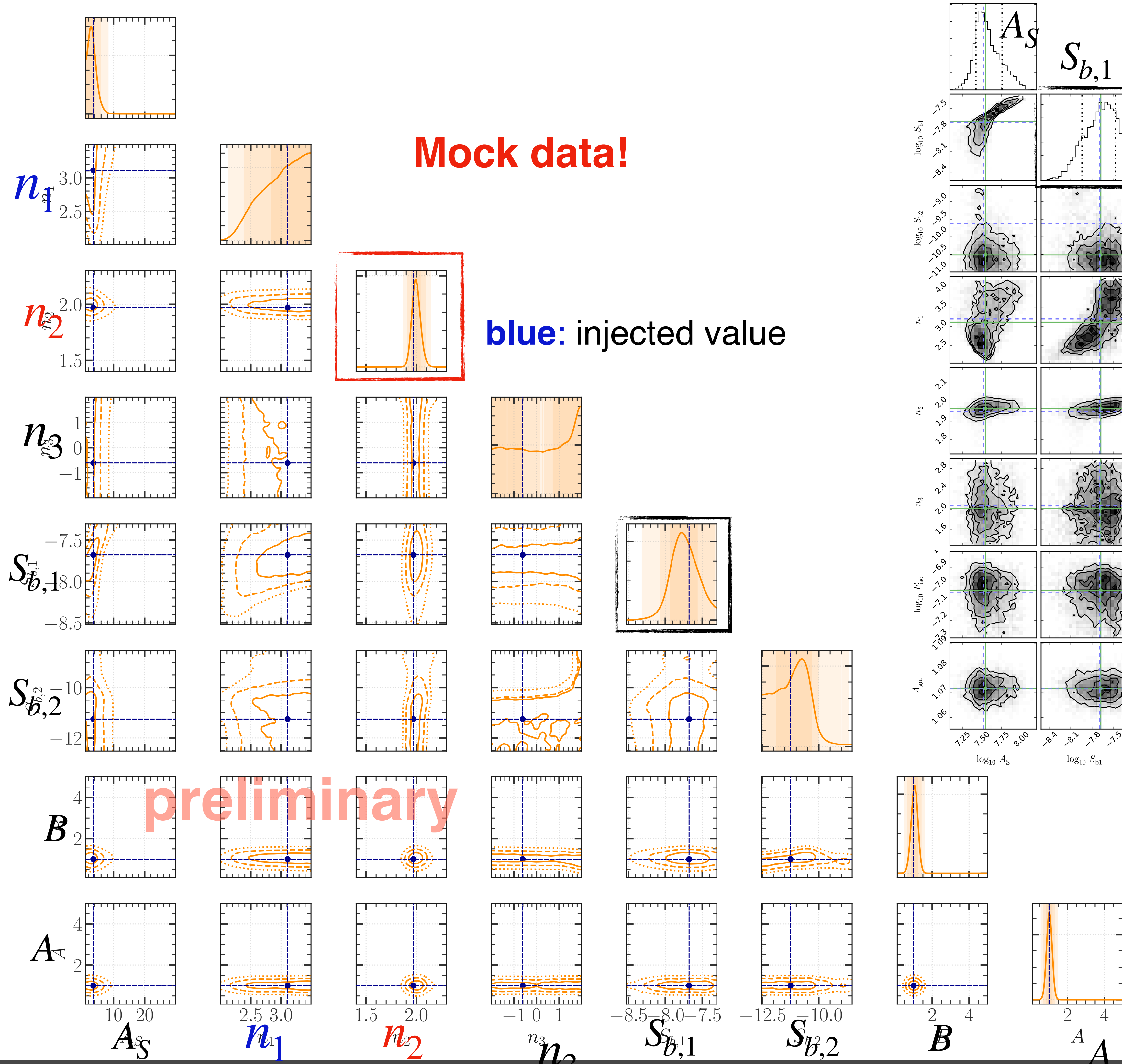
Involves analytic prescription or neural networks to compress data and derive a summary statistic.

Reminder relevant parameters:



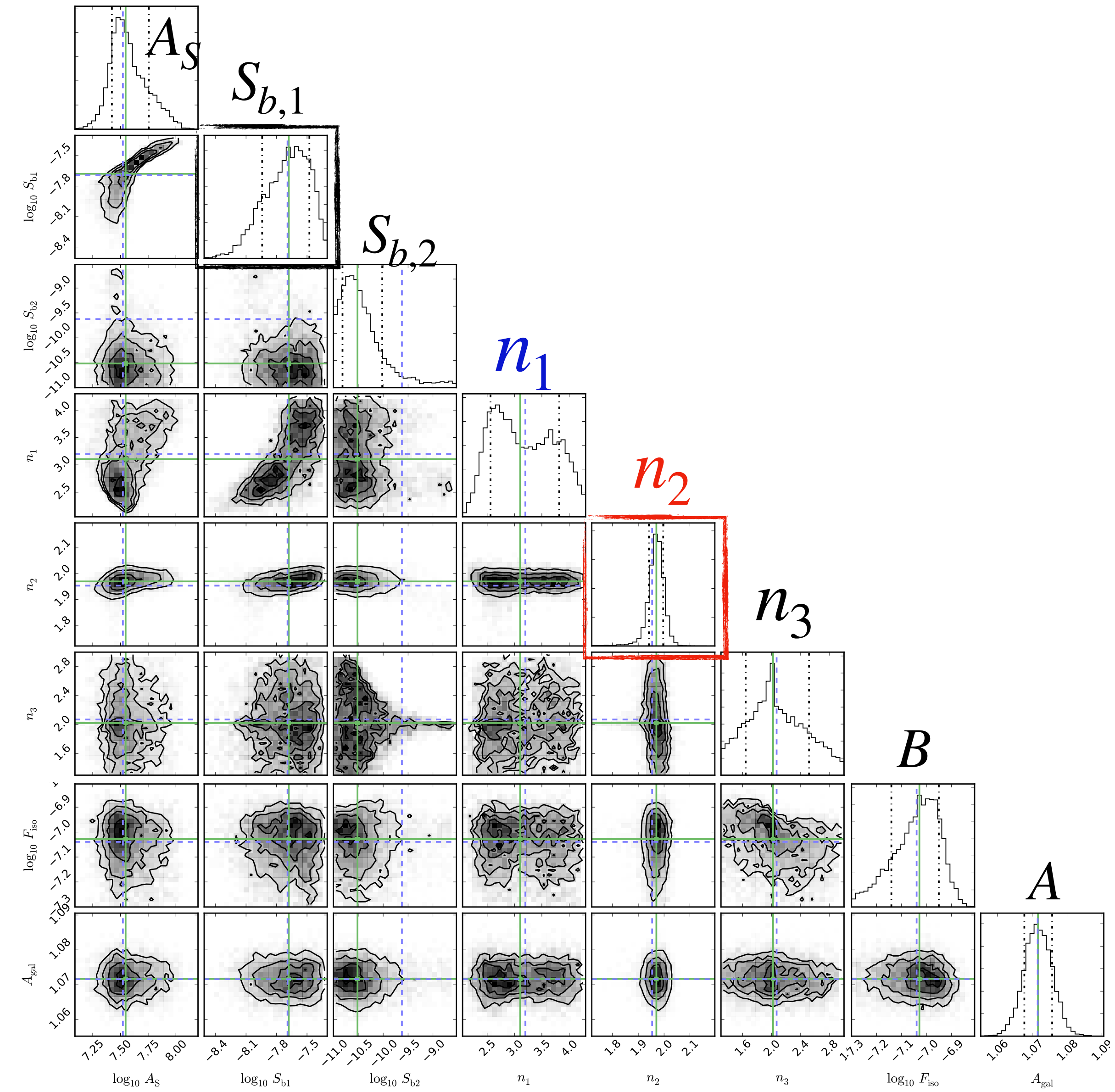
Parameter inference in our SBI framework — Results

1st round: Amortised information (universally applicable to any target data set) with parameter correlations; summary statistic: convolutional neural network



Mock data!

blue: injected value



[Zechlin et al., Astrophys.J.Suppl. 225 (2016) 2, 18]

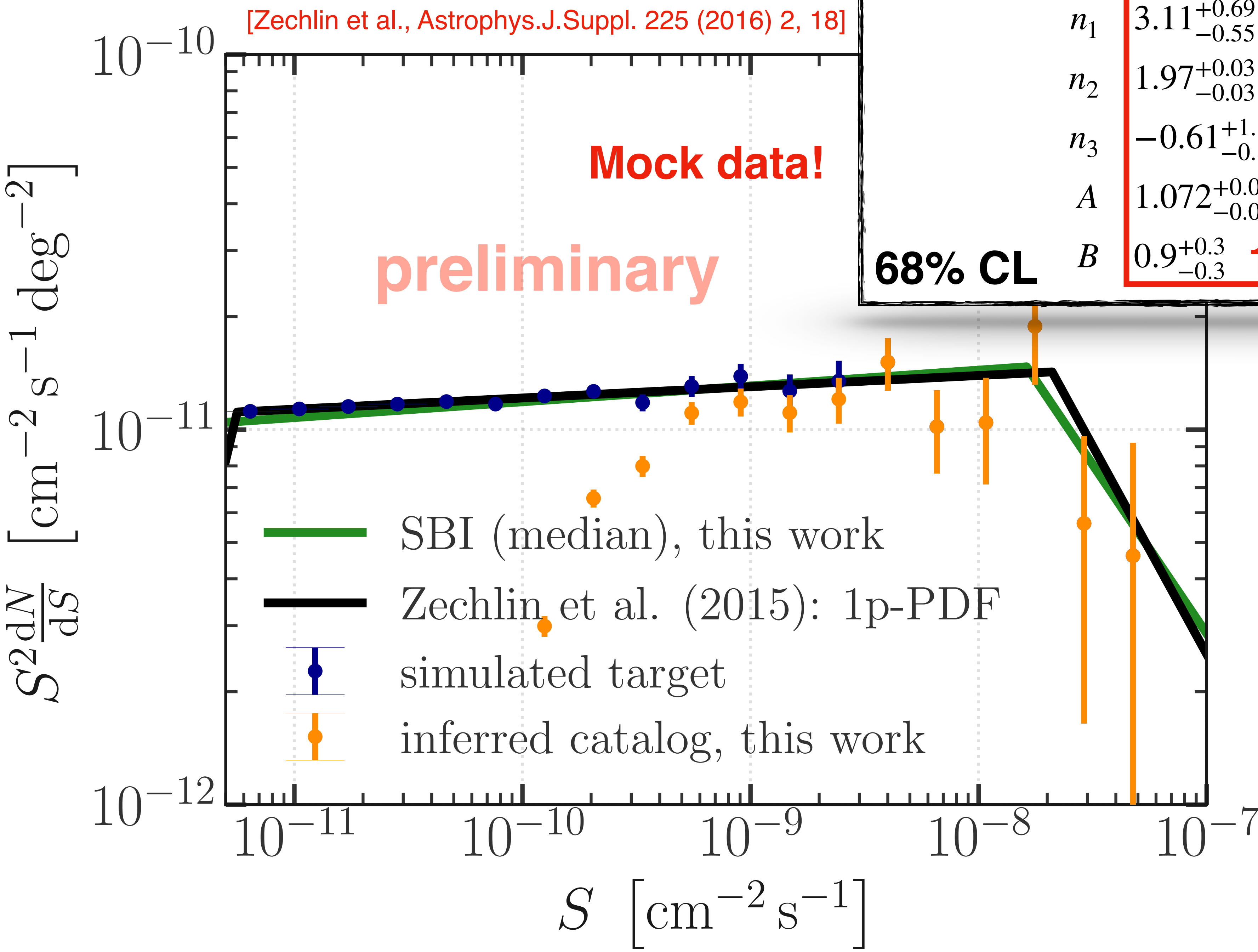
Parameter correlations and posteriors profiles recovered compared to previous work.

Parameter inference in our SBI framework — Results

5th round: Only valid with respect to target data; summary statistic: convolutional neural network

	1p-PDF	this work
A_S [s cm^2 sr^{-1}]	$3.5^{+1.6}_{-1.0}$	$3.3^{+0.6}_{-1.3}$
$S_{b,1}$ [s^{-1} cm^{-2}]	$2.1^{+0.9}_{-1.2} \times 10^{-8}$	$1.65^{+1.0}_{-0.6} \times 10^{-8}$
$S_{b,2}$ [s^{-1} cm^{-2}]	$5.6^{+5.6}_{-5.1} \times 10^{-12}$	$4.8^{+0.2}_{-4.5} \times 10^{-12}$
n_1	$3.11^{+0.69}_{-0.55}$	$2.88^{+0.24}_{-0.39}$
n_2	$1.97^{+0.03}_{-0.03}$	$1.96^{+0.03}_{-0.03}$
n_3	$-0.61^{+1.13}_{-0.89}$	$0.24^{+1.74}_{-0.35}$
A	$1.072^{+0.004}_{-0.004}$	$1.010^{+0.010}_{-0.007}$
B	$0.9^{+0.3}_{-0.3}$	$0.9^{+0.1}_{-0.1}$

68% CL



- Despite being derived from different target data, the reconstructed 68% credible intervals are consistent with each other, SBI may even perform better on some parameters.
- The input $\mathrm{d}N/\mathrm{d}S$ is almost perfectly reconstructed.