The impact of model realism on interpretations of the Galactic Center Excess

based on arXiv:2211.09796

Christopher Eckner (LAPP, CNRS)

—15th February 2023—

—Cosmo Coffee Seminar Series, CERN TH—

Outline

1. **The Facts:** The Fermi-LAT GeV gamma-ray excess (GCE), backgrounds and properties.

2. **Its Origin:** Exciting New Physics or mundane?

3. **Hypothesis Testing:** How can we discern different hypotheses based on the GCE's characteristics?

4. **Current Status:** A brief summary of recent results and controversies.

5. **Mind the Gap**: Why background uncertainties need to be overcome to progress on the GCE.

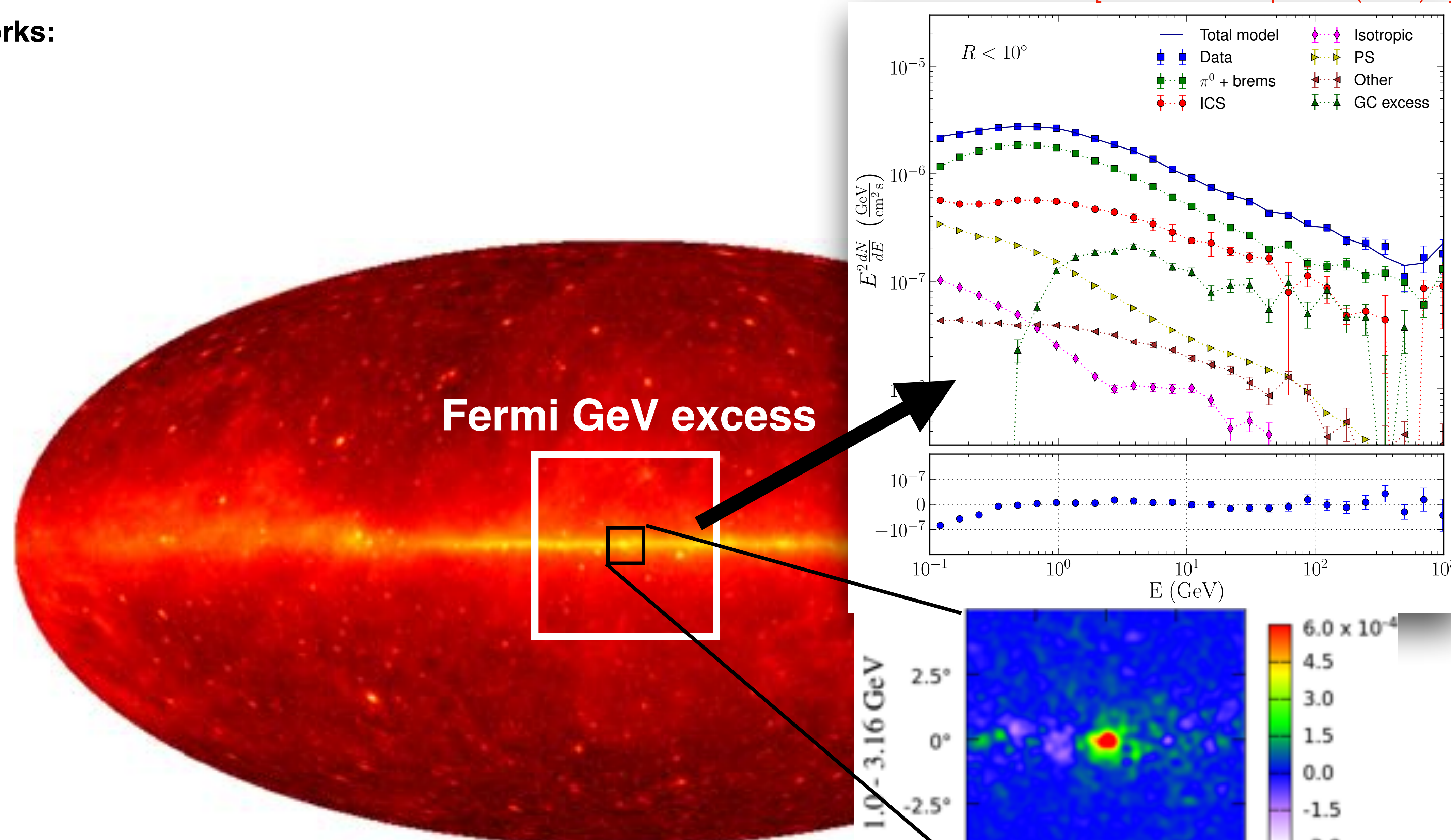
6. **How to proceed?** My humble opinion.

The Facts: The Fermi-LAT GeV gamma-ray excess (GCE), backgrounds and properties.

After more than 10 years, the Fermi GeV excess …

We all agree: There is a significant excess of GeV gamma rays (GCE) toward the Galactic centre measured by Fermi-LAT **above known astrophysical backgrounds**.

[Fermi collab. ApJ 840 (2017) 1]



[Daylan et al., Phys.Dark Univ. 12 (2016)]

An incomplete list of works:

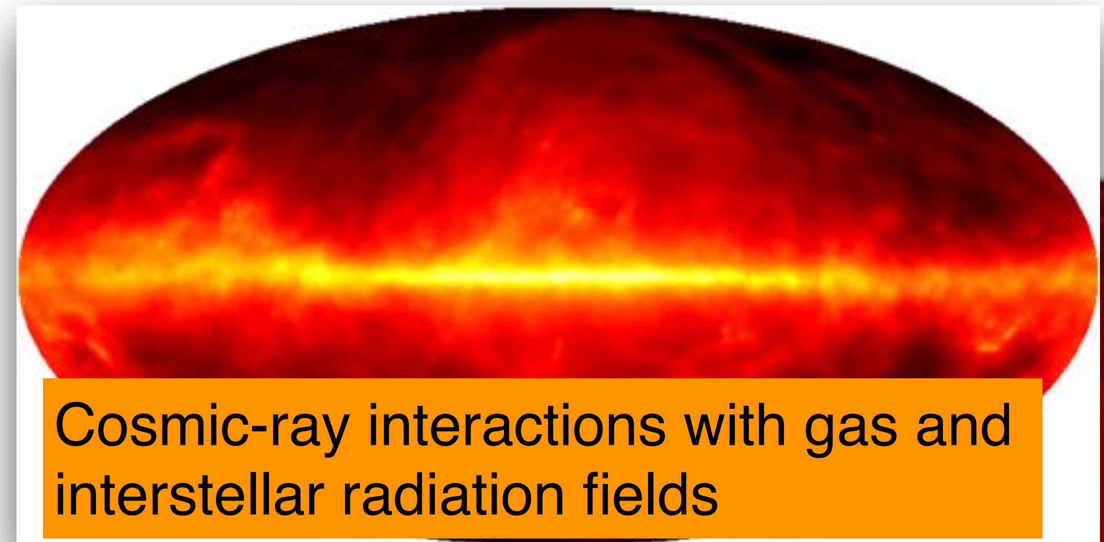
Goodenough & Hooper (2009)
Vitale & Morselli (2009)
Hooper & Goodenough (2011)
Hooper & Linden (2011)
Boyarsky et al (2011)
Abazajian & Kaplinghat (2012)
Gordon & Macias (2013)
Macias & Gordon (2014)
Abazajian et al (2014, 2015)
Calore et al (2014)
Daylan et al (2014)
Selig et al (2015)
Huang et al (2015)
Gaggero et al (2015)
Carlson et al (2015, 2016)
de Boer et al (2016)
Yang & Aharonian (2016)
Fermi Coll. (2016)
Horiuchi et al (2016)
Linden et al (2016)
Ackermann et al (2017)
Macias et al (2018)
Bartels et al (2018)
Balaji et al (2018)
Zhong et al (2019)
Macias et al (2019)
Chang et al (2020)
Buschmann et al (2020)
Leane & Slatyer (2020)
Abazajian et al (2020)
List et L (2020)
Di Mauro (2020)
Burns et al (2020)
Cholis et al (2022)
Pohl, Macias+ (2022)
…

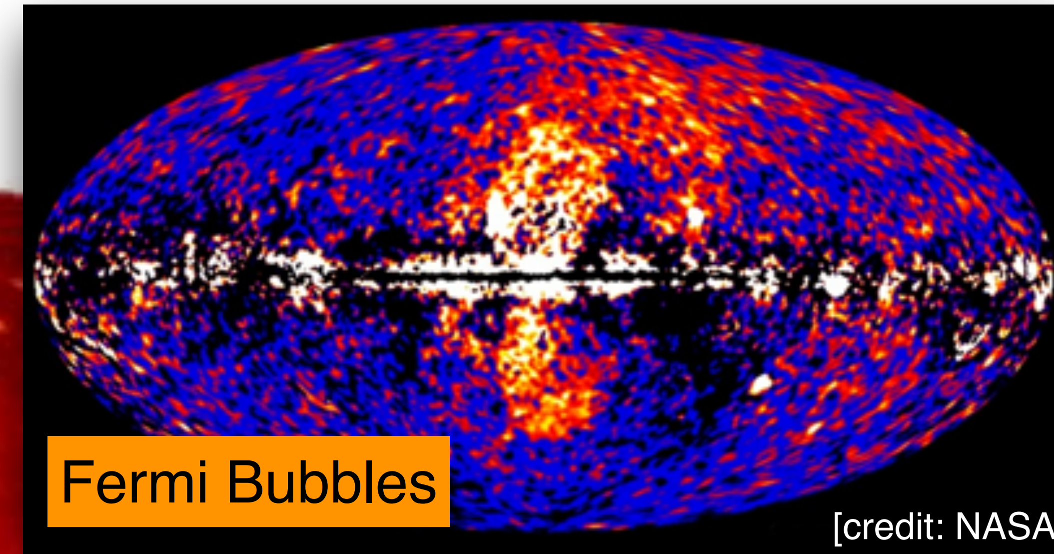
Where we do not agree:

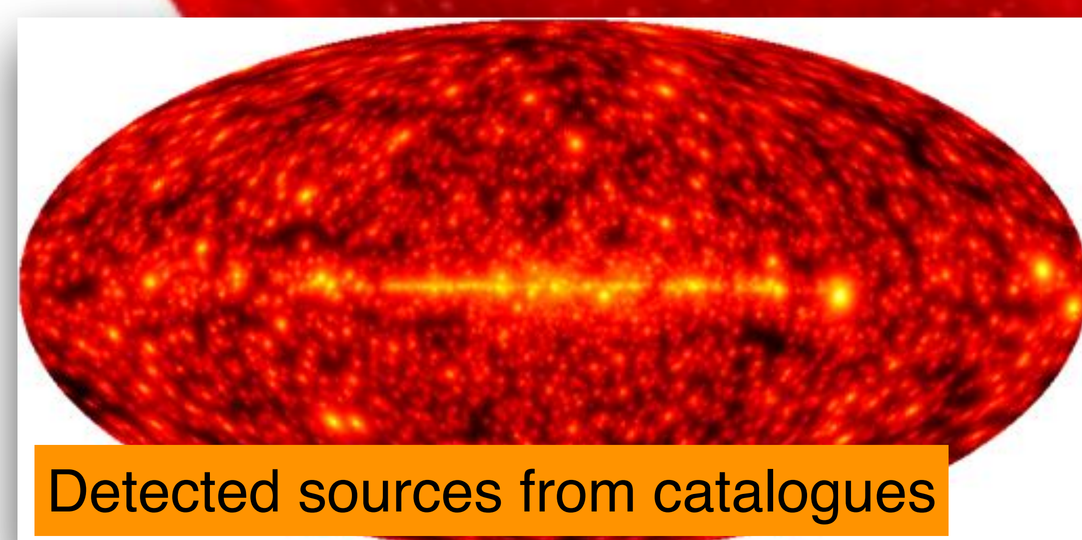
1. What is the preferred spatial morphology of the excess?
2. What is causing the Fermi GeV excess?

The known* astrophysical backgrounds

Main ingredients: diffuse emission + detected, localised gamma-ray sources







Additional components: Sun/Moon, isotropic gamma-ray background, Loop I, etc.

*modulo known unknowns: contribution of faint/unresolved sources, level of systematic uncertainties in models

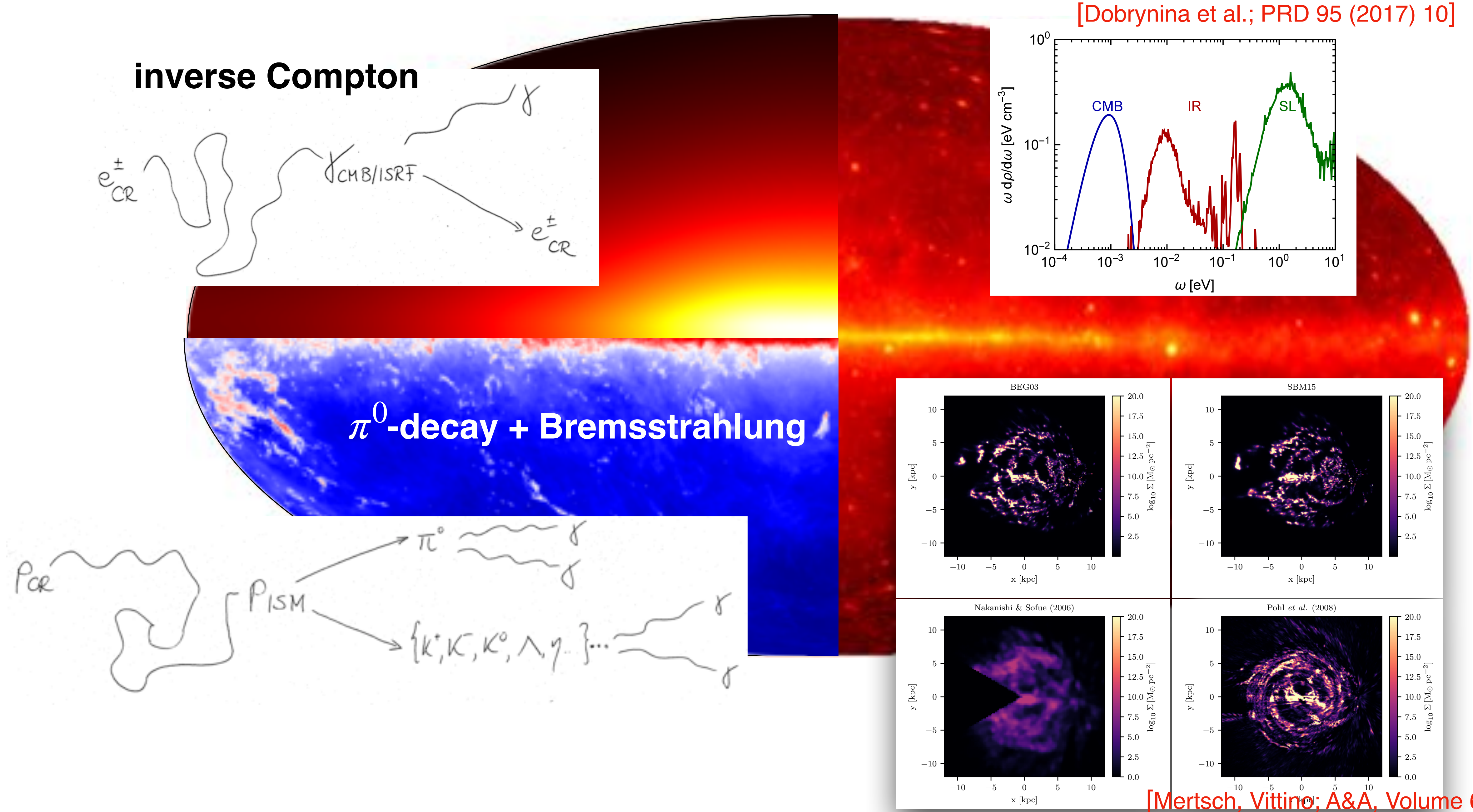
A word about the diffuse emission

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)



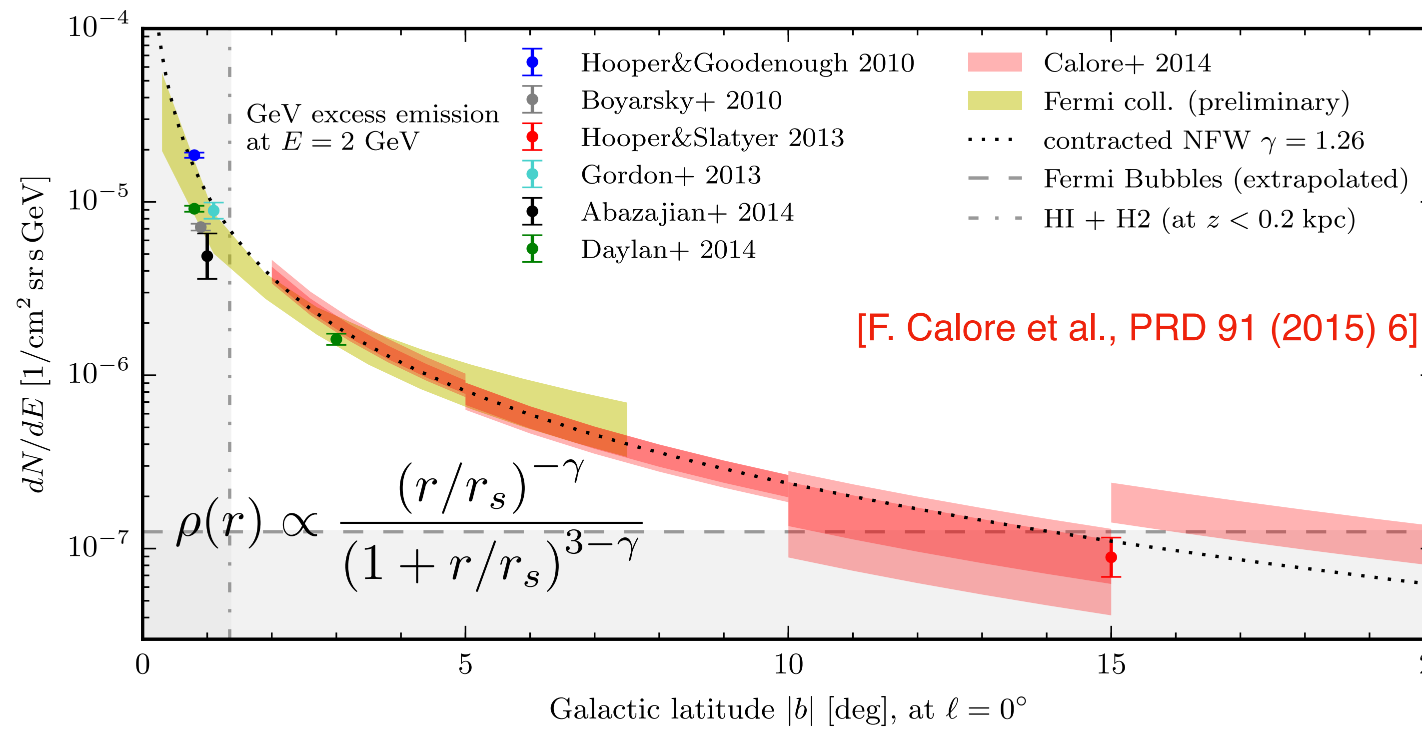
[Dobrynina et al.; PRD 95 (2017) 10]

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

The GCE's characteristics in a nutshell

Morphology:

- Appears to be spherically symmetric (a review on this later) with flux that falls as $\sim r^{-2.4}$ out to at least $\sim 10°$.
- Consistent with a generalised NFW profile with inner slope parameter $\gamma \sim 1.2$.



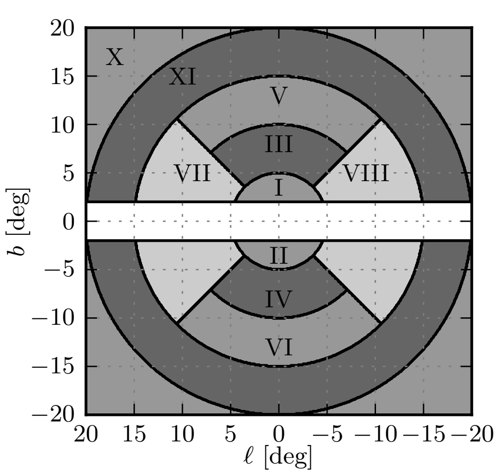
[F. Calore et al., PRD 91 (2015) 6]

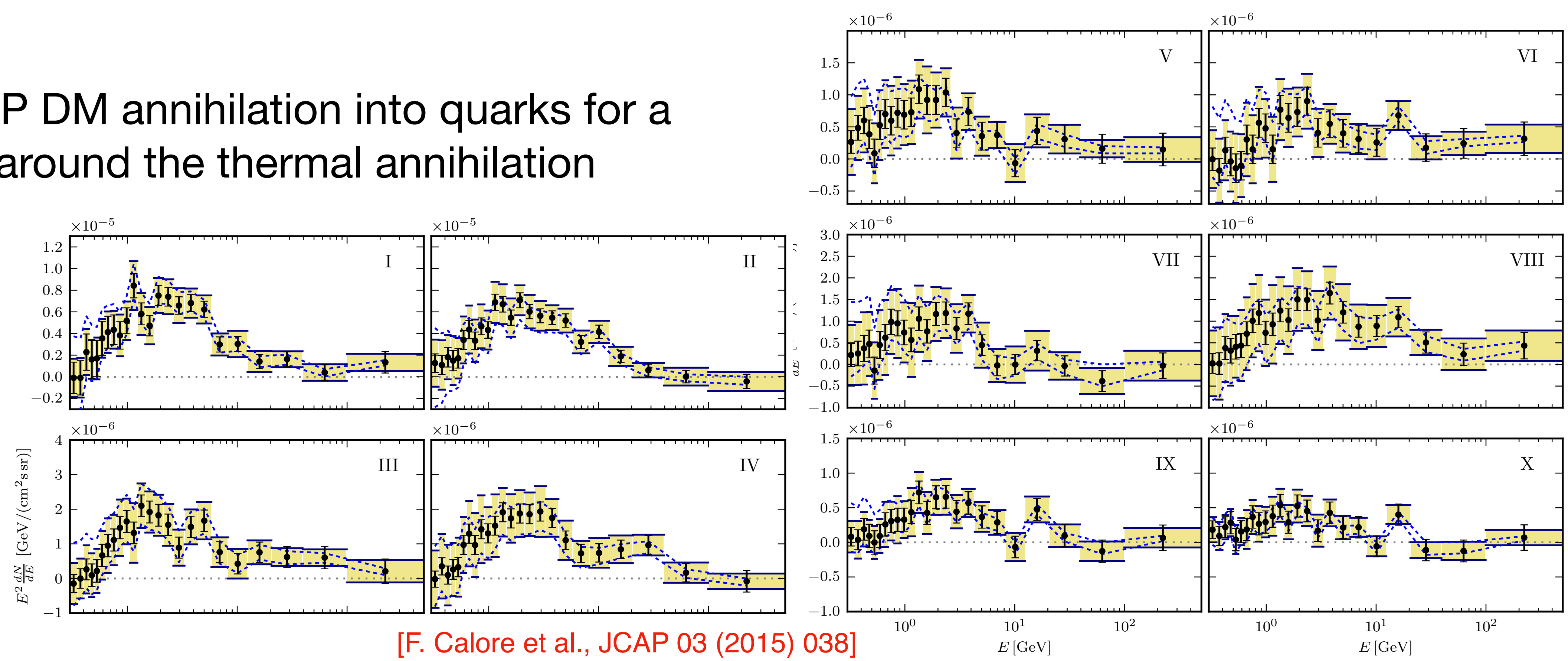
Spectrum:

- Peaked around 1 - 3 GeV, presence above ~10 GeV less robust
- Noticeable asymmetries (North-South, West-East)

Intensity:

- Consistent with WIMP DM annihilation into quarks for a ~40 - 60 GeV particle around the thermal annihilation cross-section.

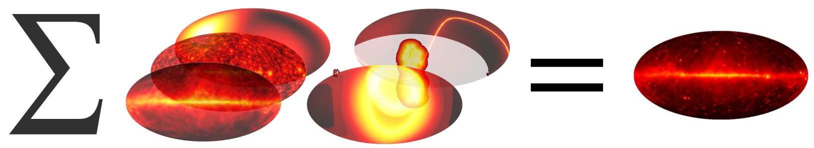




[F. Calore et al., JCAP 03 (2015) 038]

Methodology to analyse the GCE — A summary

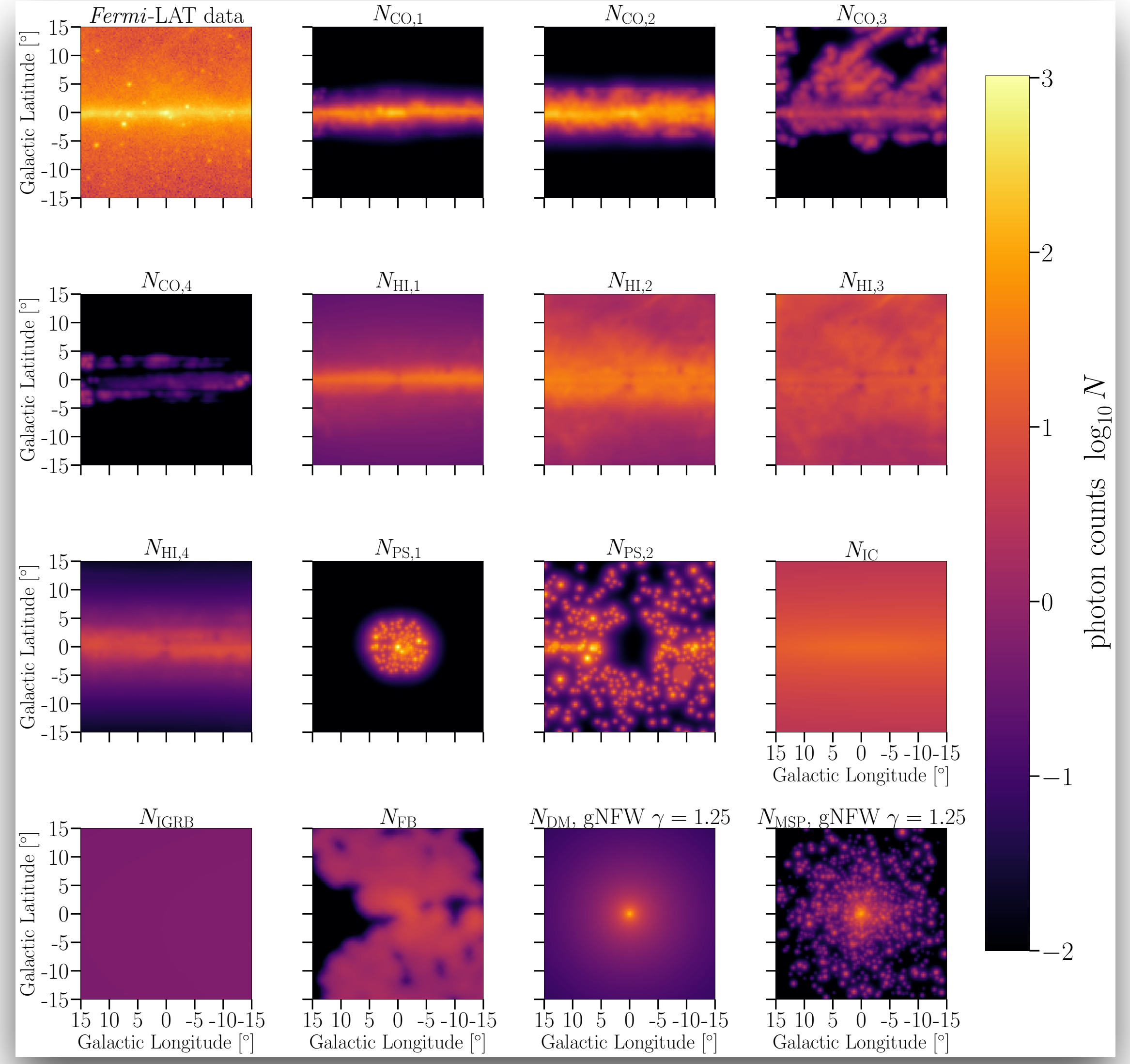
Template-based approaches



- Template-fitting (Poisson likelihood) (e.g. [F. Calore et al., PRD 91 (2015) 6], [Fermi-LAT collab., ApJ 840 (2017) 1], [Pohl et al., ApJ 929 (2022) 2], [Cholis et al., PRD 105 (2022) 10])
- Template-fitting (weighted Poisson likelihood) (e.g. [M. di Mauro, PRD 103 (2021) 6], [Fermi-LAT collab., AjS 247 (2020) 1])
- Adaptive template-fitting (SkyFACT)
  [E. Storm et al. JCAP 08 (2017) 022]
  [R. Bartels et al. Nature Astron. 2 (2018) 10]
- Photon-count statistics:
  1) 1pPDF
  [F. Calore & S. Manconi, PRL 127 (2021) 16]
  2) non-Poissonian template fitting (NPTF)
  (e.g. [S. Lee et al., PRL 116 (2016)])
- Machine learning:
  [S. Caron et al., JCAP 05 (2018) 058]
  [F. List et al. PRL 125 (2020) 241102]
  [F. List et al. PRD 104 (2021) 12]
  [S. Mishra-Sharma and K. Cranmer, PRD 105 (2022) 6]
  [S. Caron, C. Eckner et al., arXiv:2211.09796]

Other methods

- Wavelet transform
  [R. Bartels et al., PRL 116 (2016) 5]
  [Y. Zhong et al., PRL 124 (2020) 23]
- Spectral fits, D3P0
  [X. Huang et al., JCAP 04 (2016) 030]

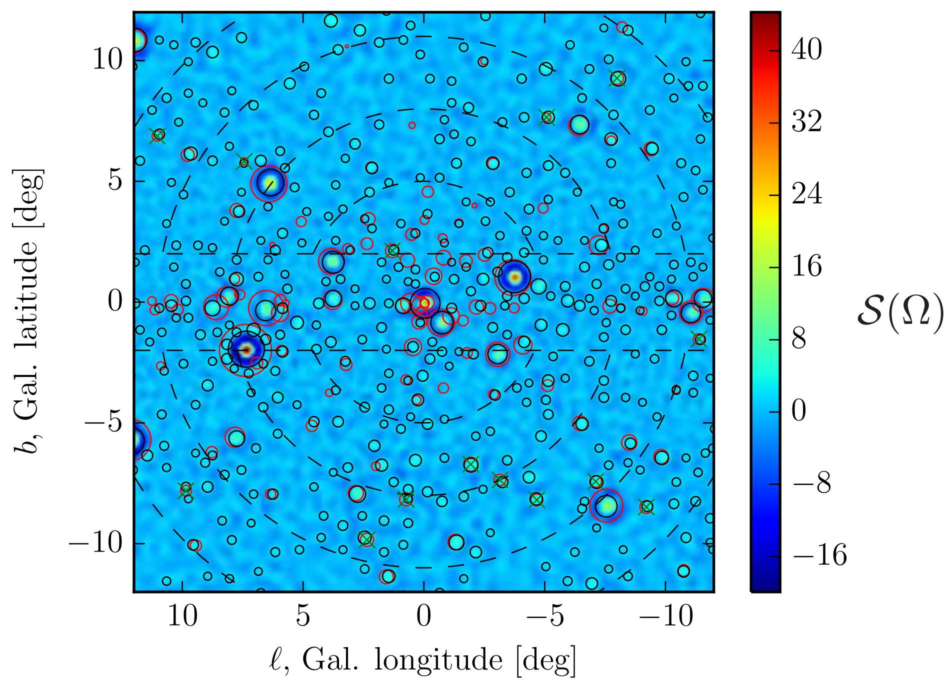


What about the other techniques?

Wavelet analyses: [R. Bartels et al., PRL 116 (2016) 5]

Convolve photon counts with wavelet kernel (Mexican hat wavelet family)

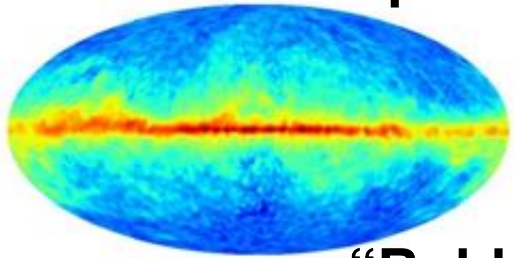
$$\mathcal{F}_{\mathcal{W}}[\mathcal{C}](\Omega) \equiv \int d\Omega' \, \mathcal{W}(\Omega - \Omega')\mathcal{C}(\Omega') \qquad \mathcal{S}(\Omega) \equiv \frac{\mathcal{F}_{\mathcal{W}}[\mathcal{C}](\Omega)}{\sqrt{\mathcal{F}_{\mathcal{W}^2}[\mathcal{C}](\Omega)}}$$



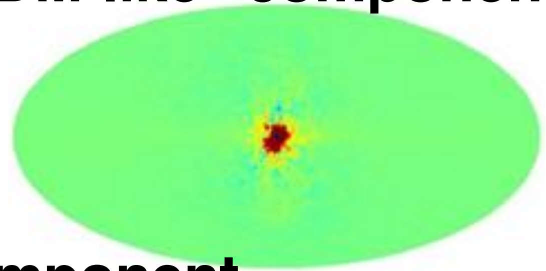
Results:
1. Clustering of photons compatible with GCE detected at 10σ.
2. Small-scale diffuse mis-modelling may be another source of such clustering.
3. A more recent study does not find such evidence anymore. ([Y. Zhong et al., PRL 124 (2020) 23]).

Spectral decomposition:

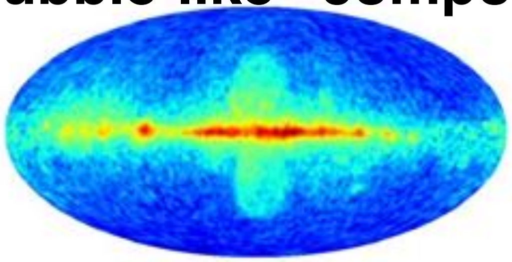
“Cloud-like” component



“DM-like” component



“Bubble-like” component



[X. Huang et al., JCAP 04 (2016) 030]

Pixel-by-pixel spectral decomposition:

$$\frac{dN}{dE} = \alpha_1 \left.\frac{dN}{dE}\right|_{\mathrm{Bu}} + \alpha_2 \left.\frac{dN}{dE}\right|_{\mathrm{Cl}} + \alpha_3 \left.\frac{dN}{dE}\right|_{b\bar{b}} + \mathrm{PSC}$$

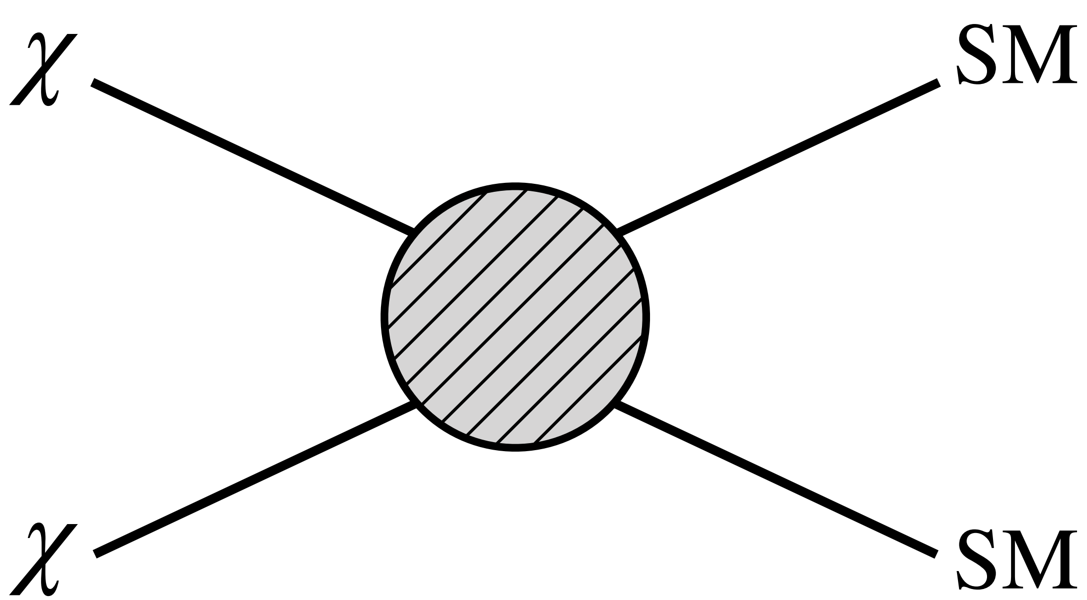
Results:
1. DM-like component favoured in innermost part.
2. Disfavoured by off-plane regions (steeper drop of emission than expected from N-body simulations).
3. Other spectra lead to different results (tested in [W. De Boer et al., arXiv:1610.08926]).

Its Origin: Exciting New Physics or mundane?

What produces the excess?

The excess is tantalising since it coincides well with the expectations for the sought-after signal of **thermal dark matter pair-annihilating** in the Galactic centre. However, **unresolved populations of gamma-ray sources** are a strong contender! [See D. Hooper, arXiv:2209.14370 for a different view.]

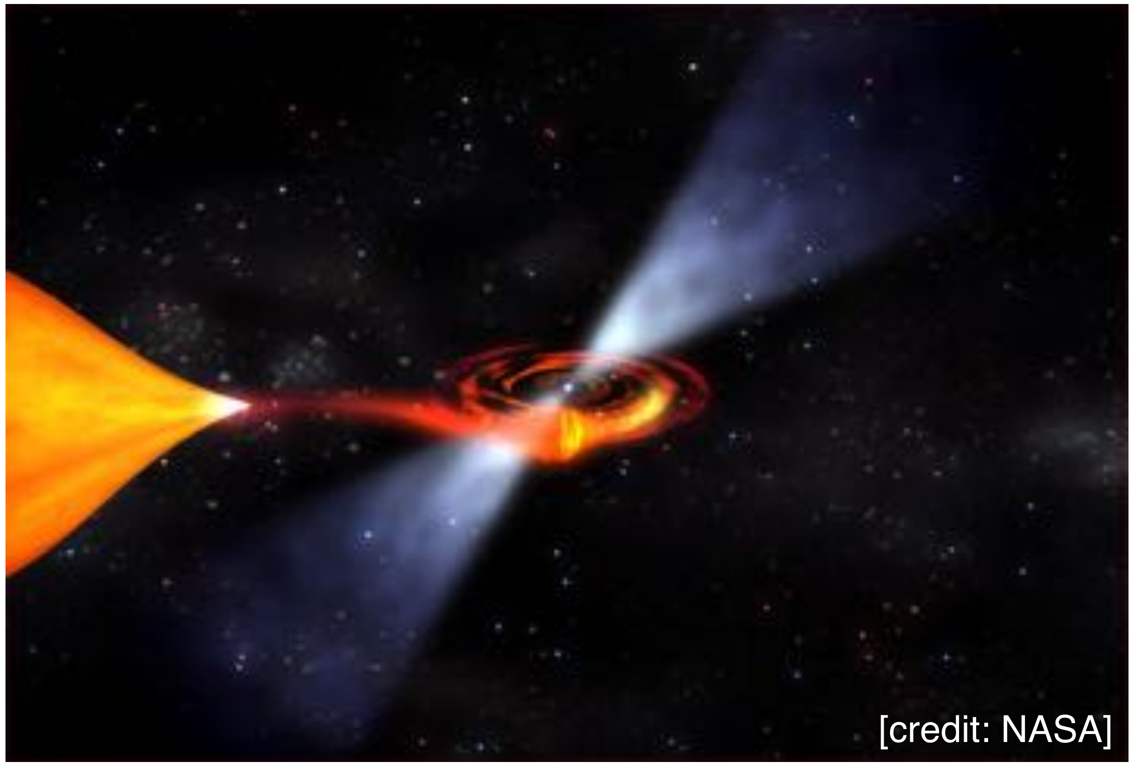
Thermal dark matter



supported by (incomplete collection):
[Fermi collab. ApJ 840 (2017) 1];
[R. K. Leane and T. R. Slatyer, PRL 123 (2019) 24];
[M. di Mauro, PRD 103 (2021) 6]; [I. Cholis et al., PRD 105 (2022) 10];
[S. D. McDermott et al., arXiv:2209.00006]

VS.

Unresolved Galactic source population (here: millisecond pulsars [MSPs])



supported by (incomplete collection):
[R. Bartels et al., PRL 116 (2016) 5];
[C. Eckner et al., ApJ 862 (2018) 1];
[R. Bartels et al., Nature Astron. 2 (2018) 10];
[O. Macias et al., JCAP 09 (2019) 042];
[F. Calore et al., PRL 127 (2021) 16]

Other interpretations are cosmic-ray based, e.g., a past enhanced star formation/leptonic burst in the Galactic centre [E. Carlsom, S. Profumo; PRD 90 (2014) 2][J. Petrovic et al.; JCAP 10 (2014) 052][D. Gaggero et al., JCAP 12 (2015) 056].

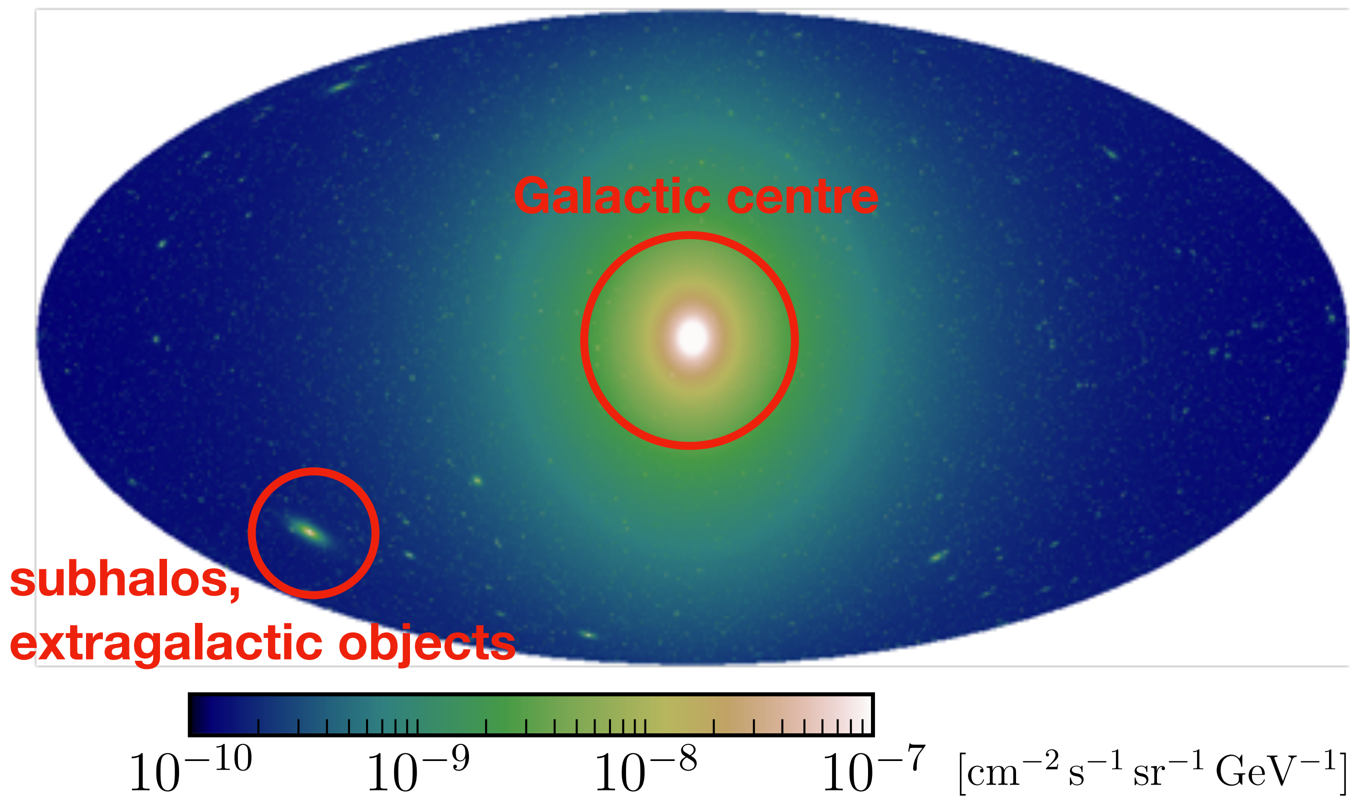
The (WIMP) dark matter hypothesis

Dark matter signal shape:

$$\frac{d\Phi_\gamma}{d\Omega\, dE_\gamma}(E_\gamma, \psi) = \frac{1}{4\pi} \boxed{\int_{\mathrm{l.o.s}} d\ell(\psi)\rho_\chi^2(\mathbf{r})} \boxed{\left(\frac{\langle \sigma v \rangle_{\mathrm{ann}}}{2 S_\chi m_\chi^2} \sum_f B_f \frac{dN_\gamma^f}{dE_\gamma} \right)}$$

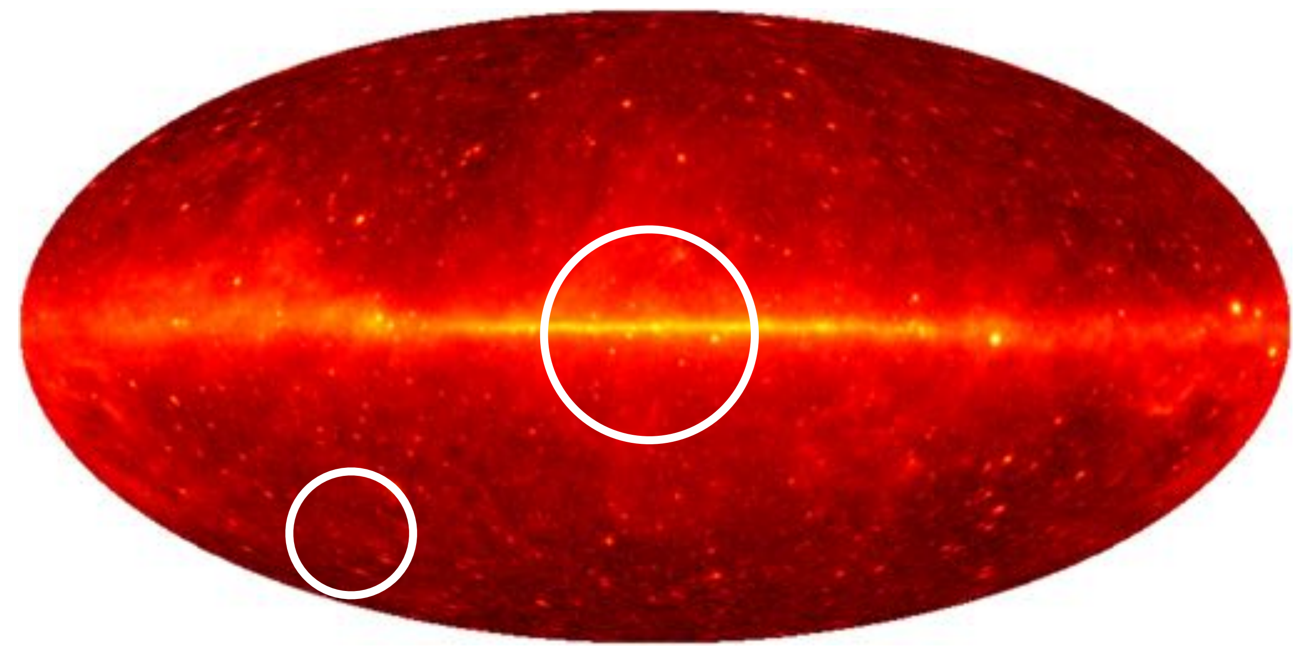
cosmology/astrophysics particle physics

Dark matter gamma-ray diff. intensity (MW-like)



[M. Hütten et al., JCAP 1609 (2016) 047]

Fermi-LAT gamma-ray sky



strong signal
high background

Galactic centre

...

moderate signal
moderate background

LMC, Andromeda galaxy

...

moderate/low signal
low background

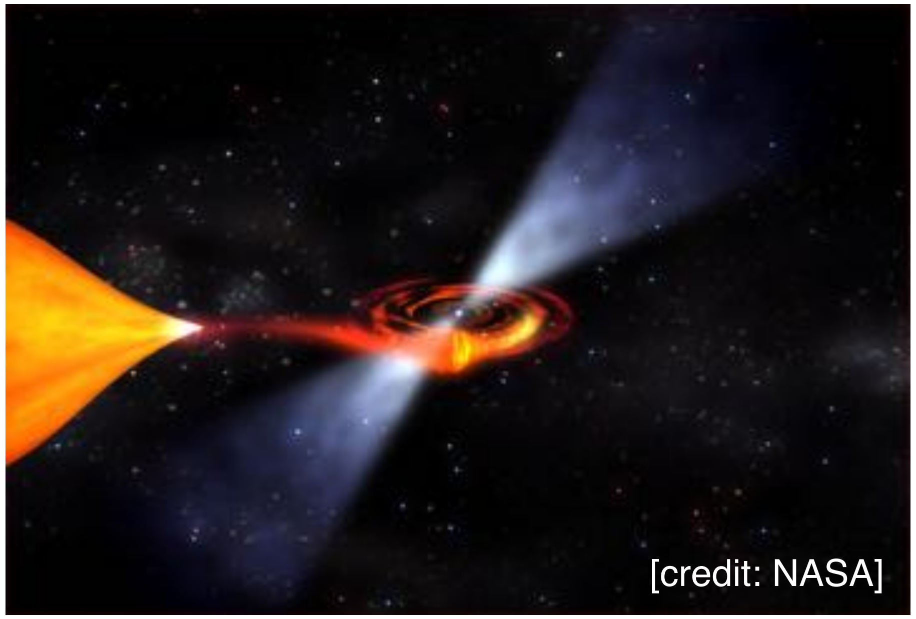
Milky Way dwarf satellites

— Sub-class of pulsars, rapidly spinning neutron stars that gradually convert rotational kinetic energy into gamma rays and radio emission.

— In lifetime of a pulsar, starts with periods ~1 s and slows down over time to become faint after ~$10^6 - 10^8$ y.

— Millisecond pulsars are old almost dead pulsars that accrete from a companion star and spin-up to even shorter periods than young pulsars (~1.5 ms).

— Rather low magnetic fields (~$10^8 - 10^9$ G), which results in a less pronounced spin down leading to lifetimes $> 10^9$ y.



[credit: NASA]

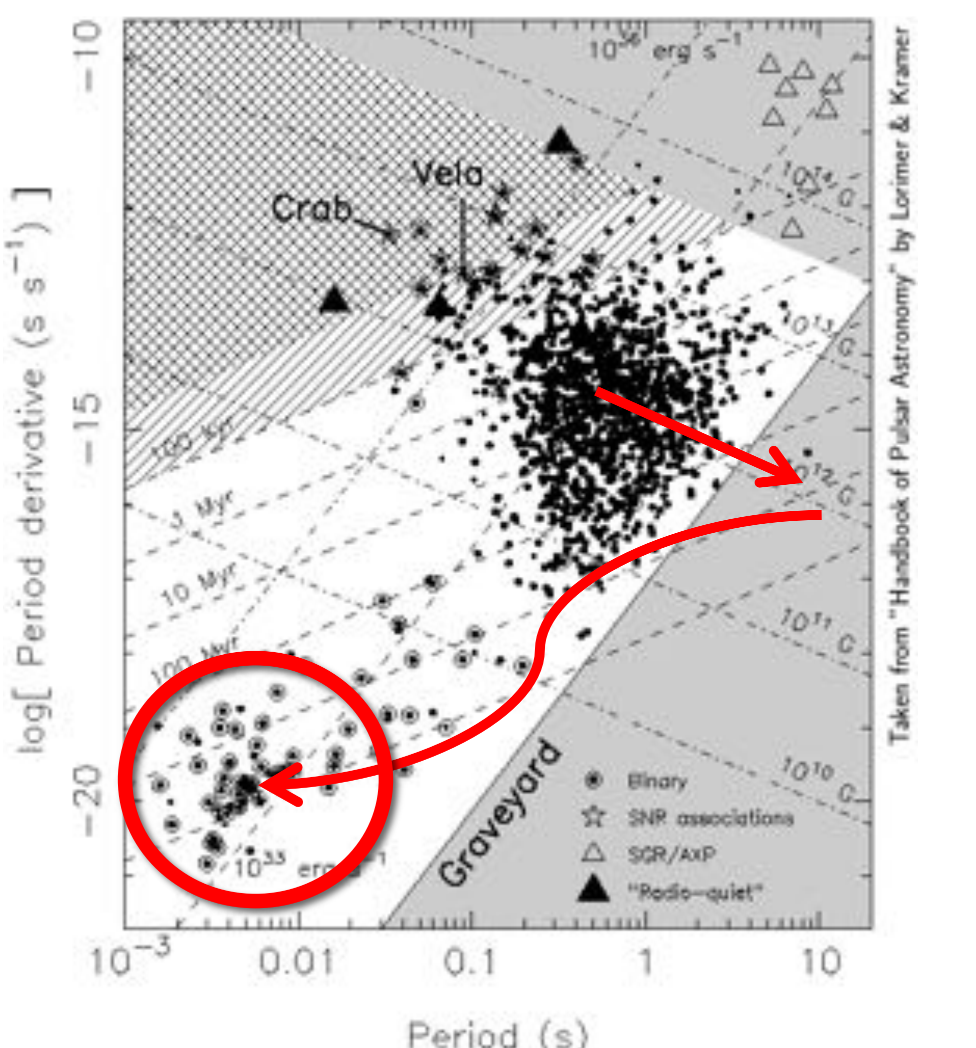
Where do these MSPs in the Galactic centre come from?

— primordial formation (from binaries born in the GC)

— dynamical formation (binary formation via stellar encounters)

— deposition by disrupted globular clusters [O. Gnedin et al., ApJ 785 (2014) 71] [T.D. Brandt & B. Kocsis, ApJ 812 (2015) 1]

— accretion-induced collapse of binaries containing white dwarfs [A. Gautam et al., Nature Astron. 6 (2022) 6]



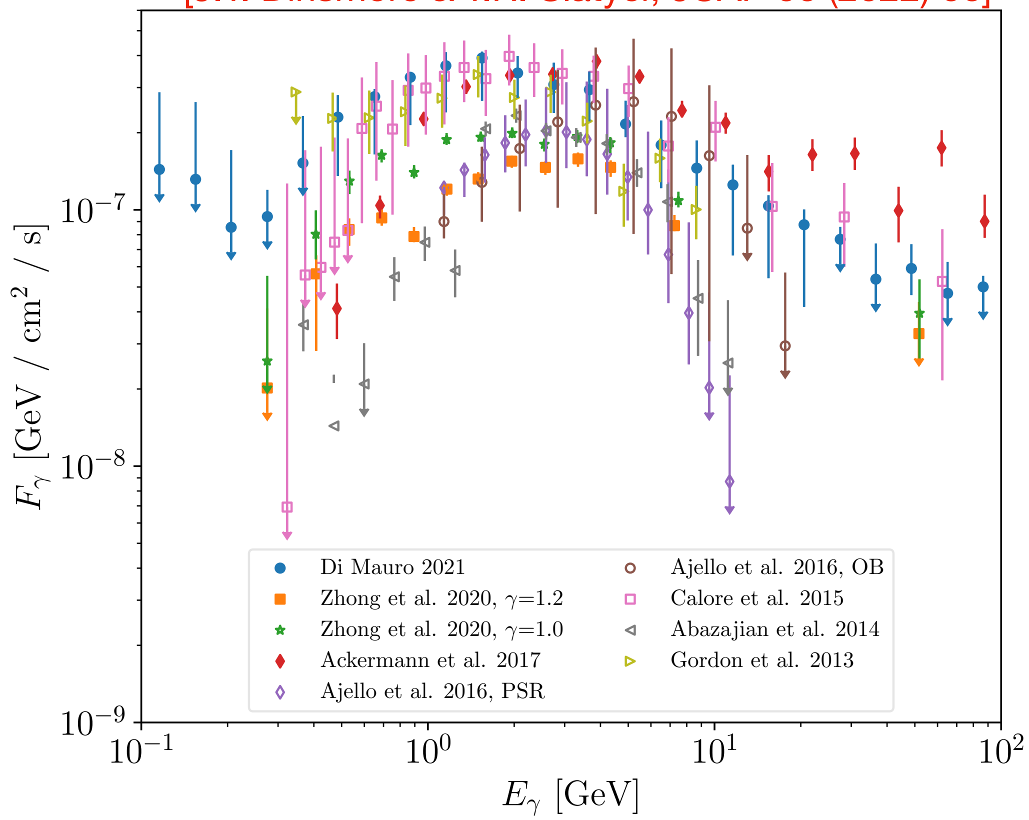
[credit: D. Hooper, IDM 2021]

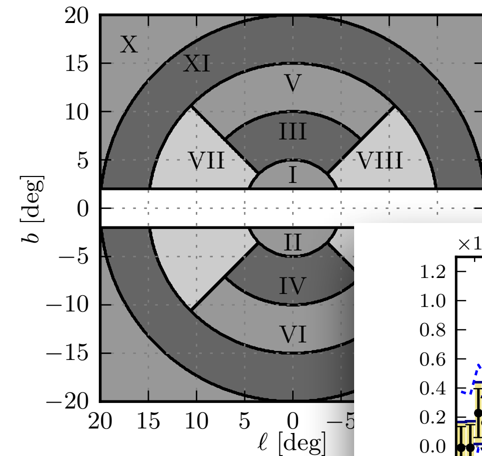
Hypothesis Testing: How can we discern different hypotheses based on the GCE's characteristics?

Discrimination via spectral shape

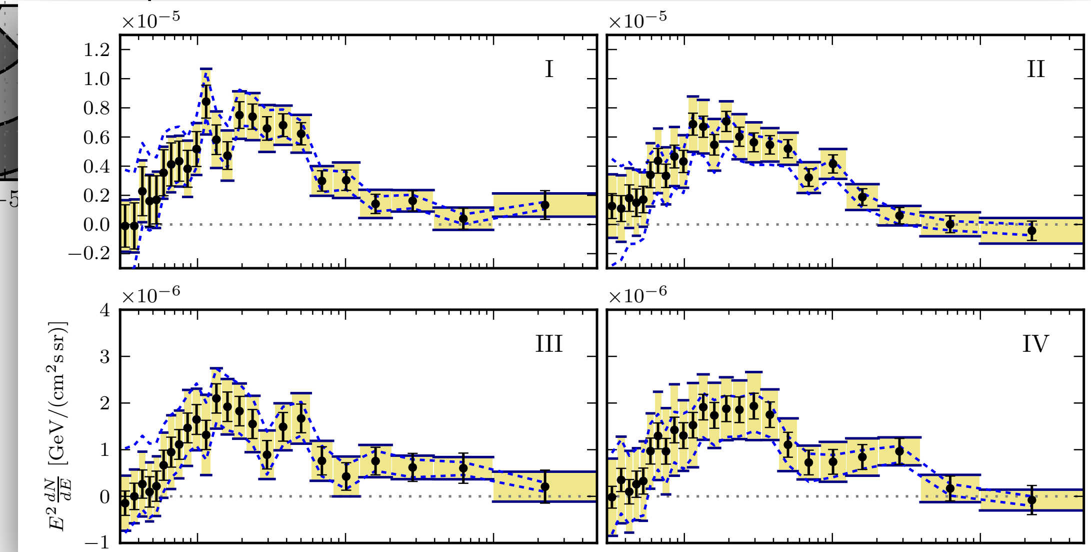
GCE spectral reconstructions:

[J.T. Dinsmore & T.R. Slatyer, JCAP 06 (2022) 06]

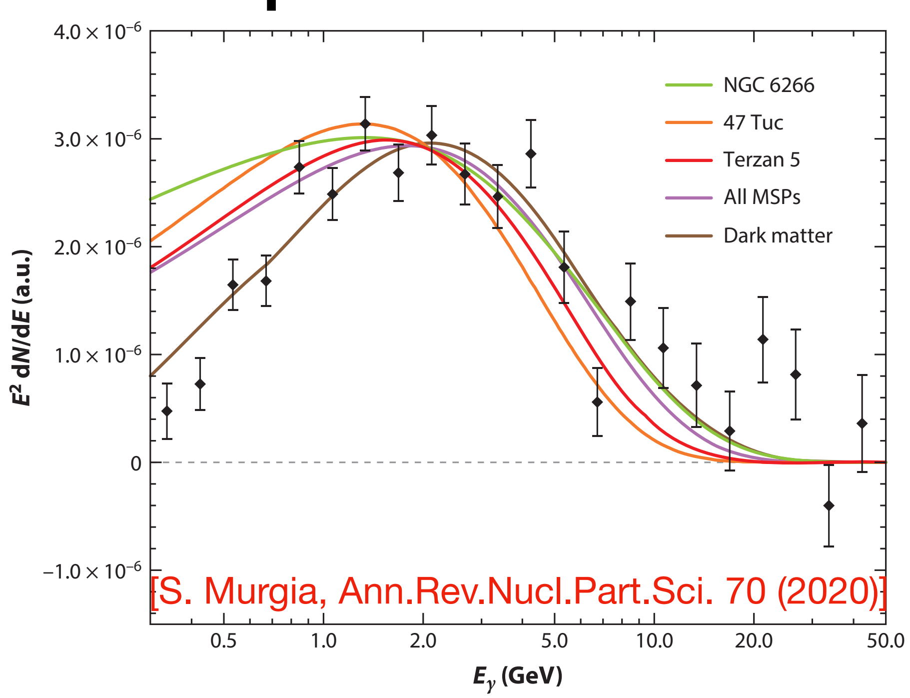




[F. Calore et al., JCAP 03 (2015) 038]



Potential spectra:



[S. Murgia, Ann.Rev.Nucl.Part.Sci. 70 (2020)]

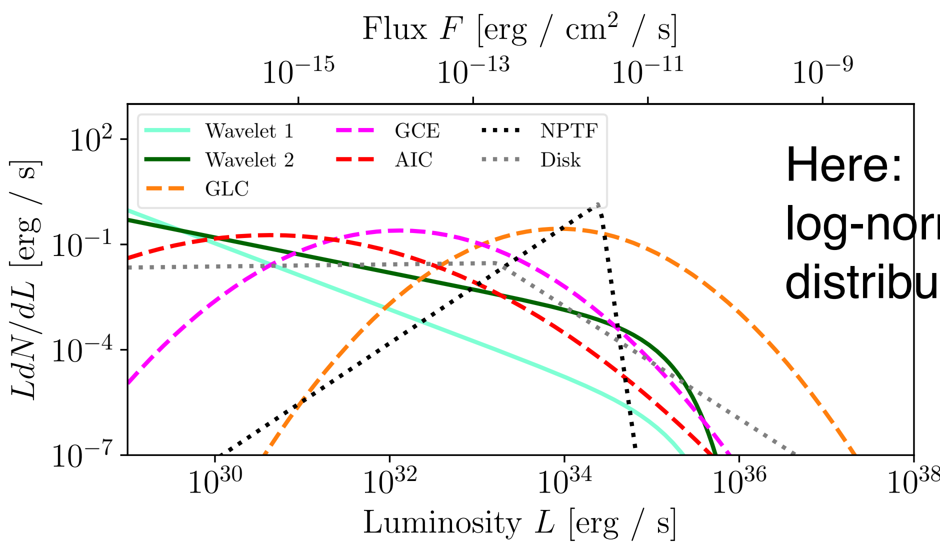
Discriminating power:

- dark matter and MSP emission share similar spectra
- spectral tails > 10 GeV may be decisive, however not robust against backgrounds

Verdict: Possible avenue but not convincing

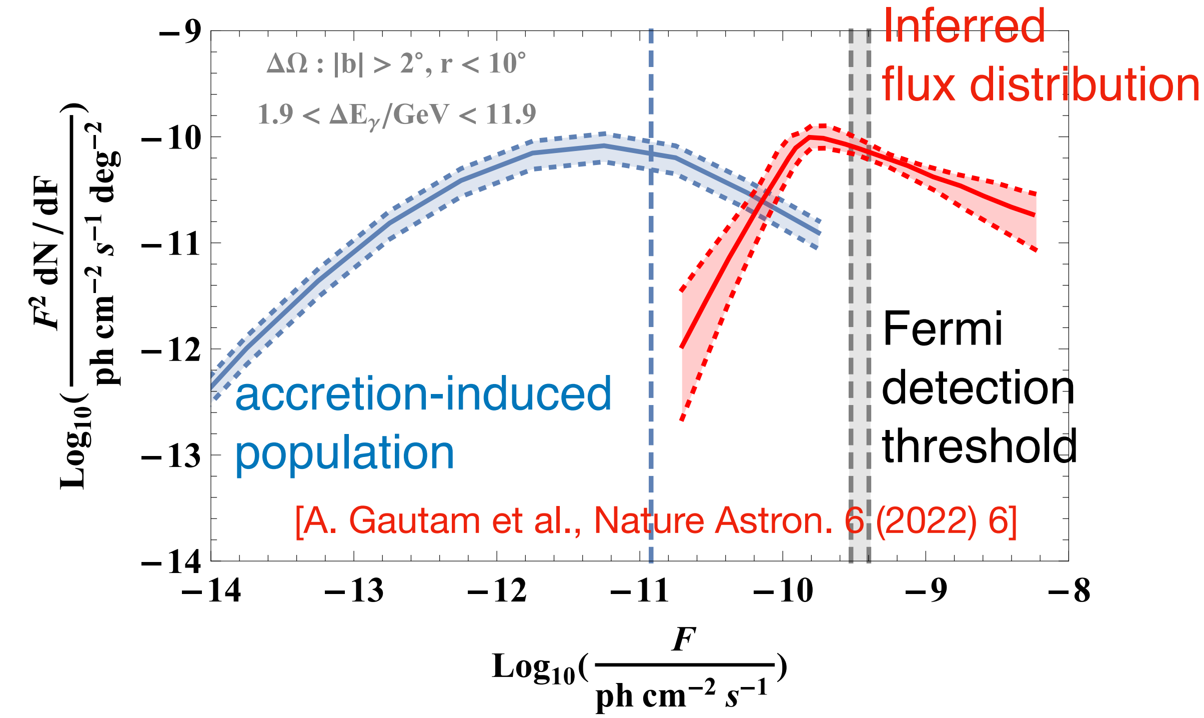
Discrimination via intensity

Luminosity of MSP population:

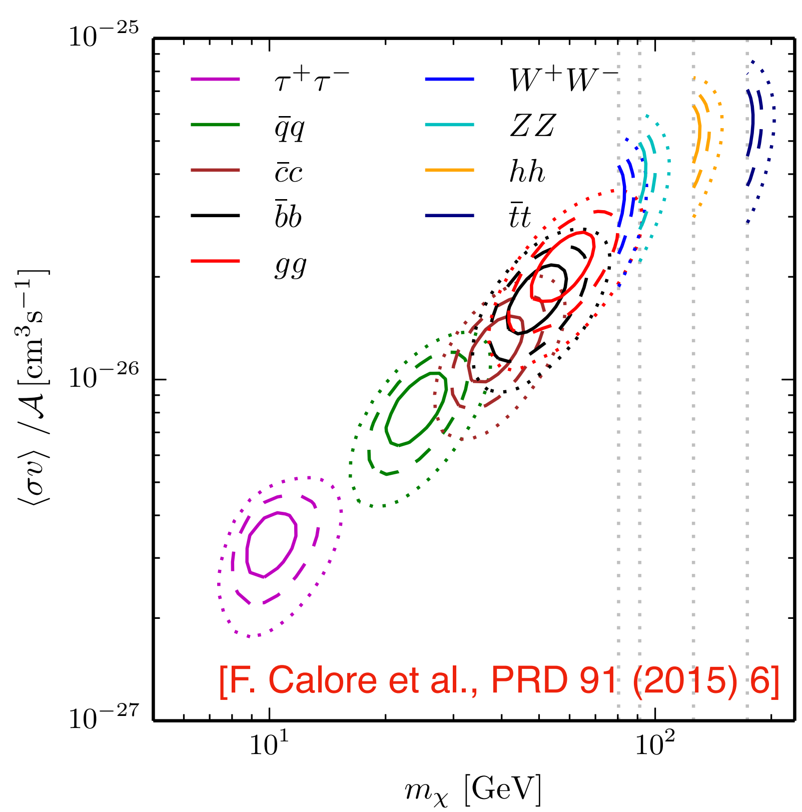


Here: log-normal distribution of L

[J.T. Dinsmore & T.R. Slatyer, JCAP 06 (2022) 06]



Best-fits of WIMP-y DM:



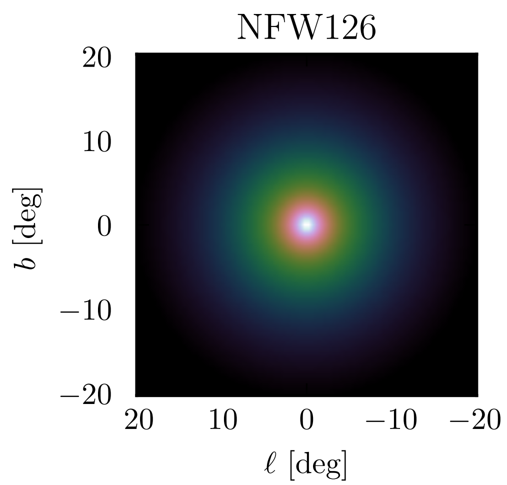
Discriminating power:

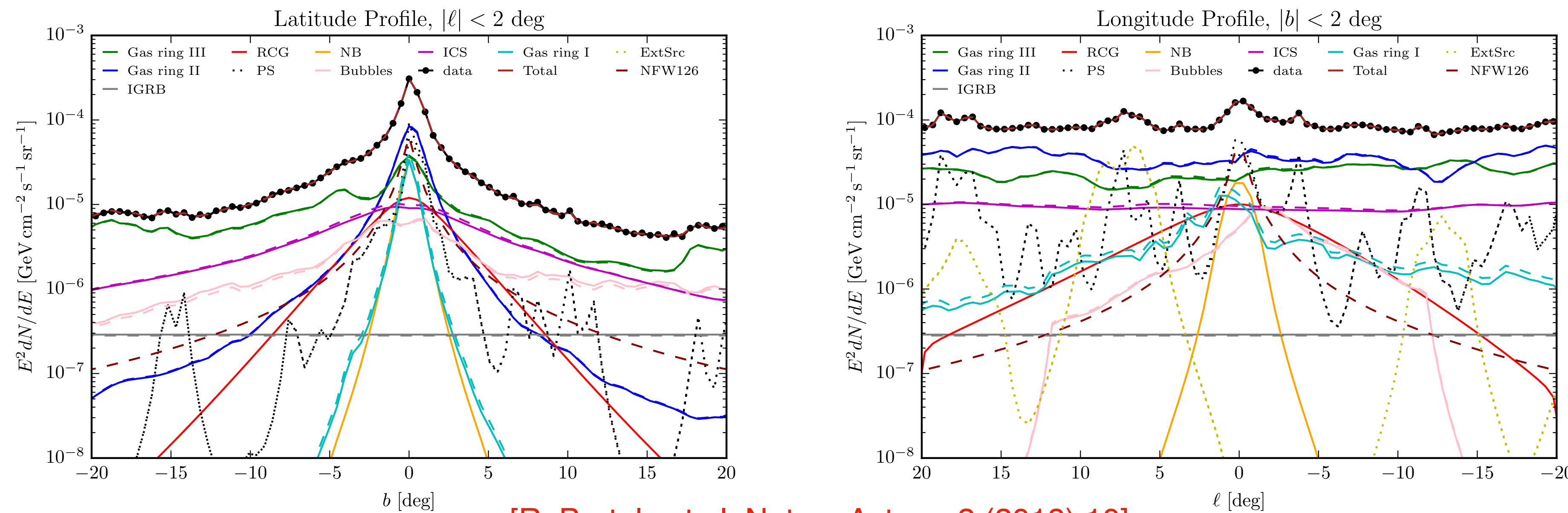
- Vanilla WIMP models fit the intensity very well around the thermal cross-section
- MSP population parameters uncertain, large variety of luminosity functions can explain reconstructed intensity
- MSP Formation/deposition mechanism unclear.
- Even claims that observed MSPs are not sufficient [Y. Zhong et al., PRL 124 (2020) 23]

Verdict: More MSP research is necessary.

Discrimination via spatial morphology

DM spatial distribution:

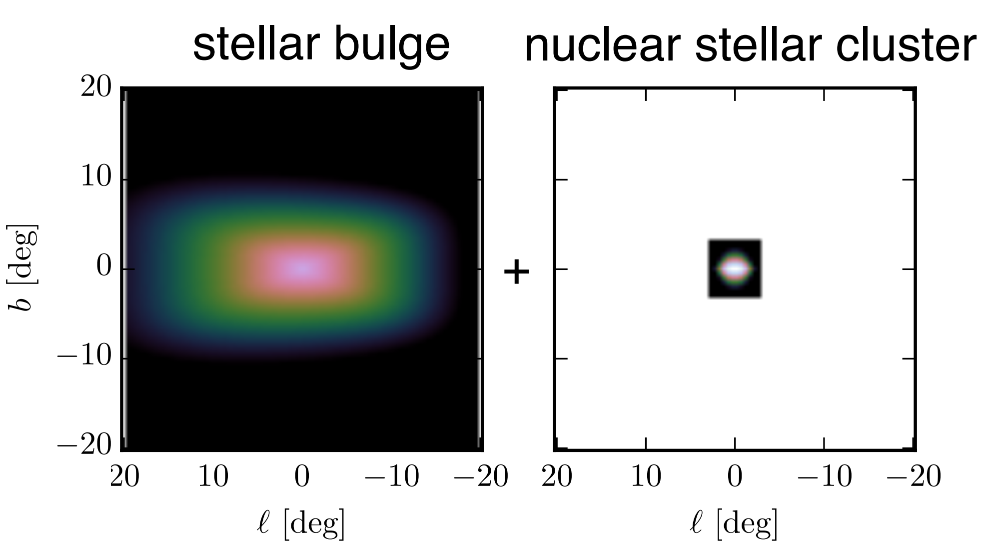




[R. Bartels et al. Nature Astron. 2 (2018) 10]

Spherical symmetry found in earlier works.

Stellar density profile in the Galactic bulge:



Pile-up of evidence for GCE tracing the stellar density: [C. Eckner et al., ApJ 862 (2018) 1]
[R. Bartels et al. Nature Astron. 2 (2018) 10][Pohl et al., ApJ 929 (2022) 2]

Discriminating power:

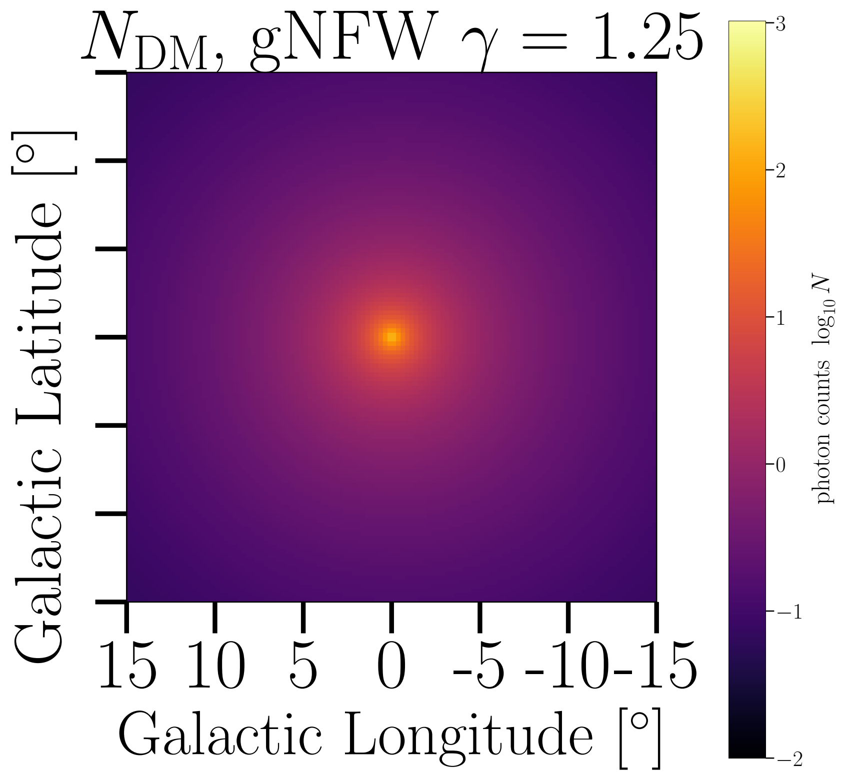
- Data seems to exhibit asymmetries in both directions.
- Standard assumptions for DM-/MSP-like morphology are not fully degenerate.
- **Recent debate about the reconstructed morphology of the GCE (more later!)**
- Baryonic feedback on DM density profile not fully understood.

Verdict: Very promising!

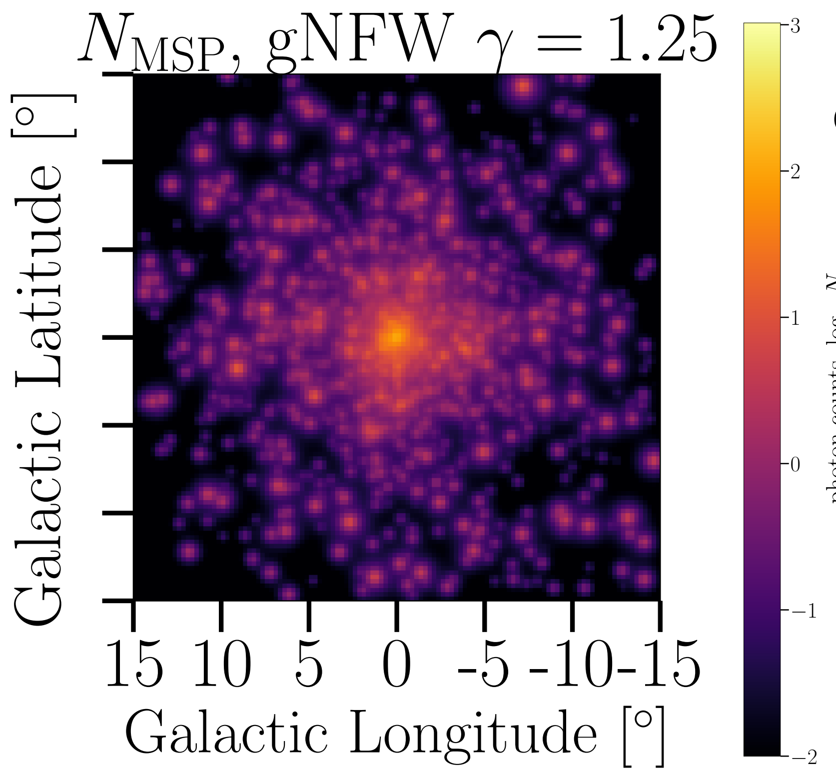
Discrimination via small-scale photon clustering

A decisive feature of the GeV excess is its **photon clustering behaviour**, spectrally they can be almost identical.

DM annihilation
(smooth morphology,
Poisson-distributed photon events)



Faint millisecond pulsar population
(photon clustering on small scales,
non-Poissonian noise component)



1 - 2 GeV energy band

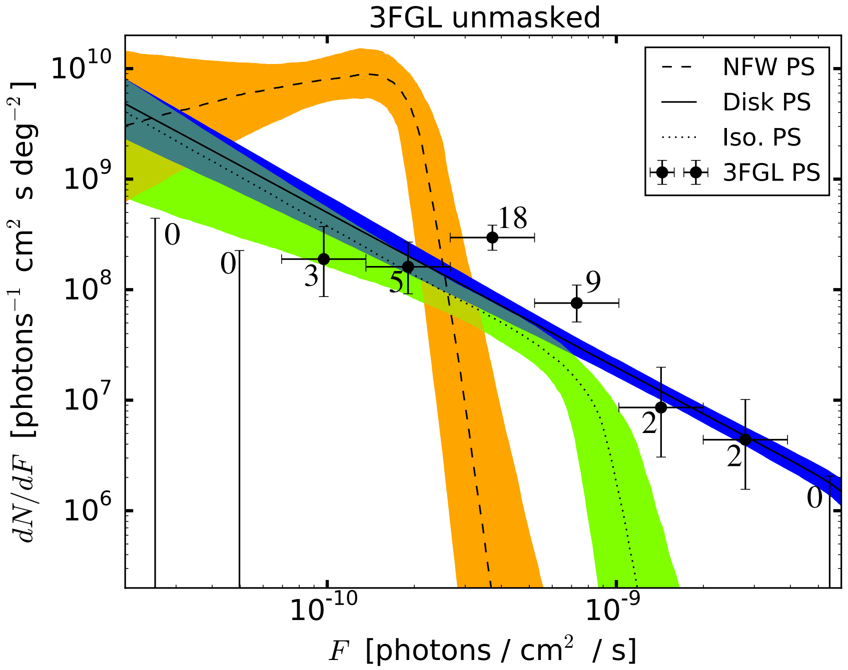
- Traditional likelihood methods cannot explore this difference in any practical way (probabilistic nature of point source locations and fluxes!)
- Effective methods have been proposed: **1p-PDF** (seminal work: [D. Malyshev & D.W. Hogg, ApJ 738 (2011) 181]), **non-Poissonian template fitting, wavelet analysis**.
  These approaches seem to prefer an excess due to MSPs (e.g. [M. Buschmann et al. PRD 102 (2020) 2]; [R. Bartels et al., PRL 116 (2016) 5]).
- **Machine learning with convolutional networks could generalise over point source distribution as a generic feature and include uncertainties in astrophysical background modelling!**

Current Status: A brief summary of recent results and controversies.

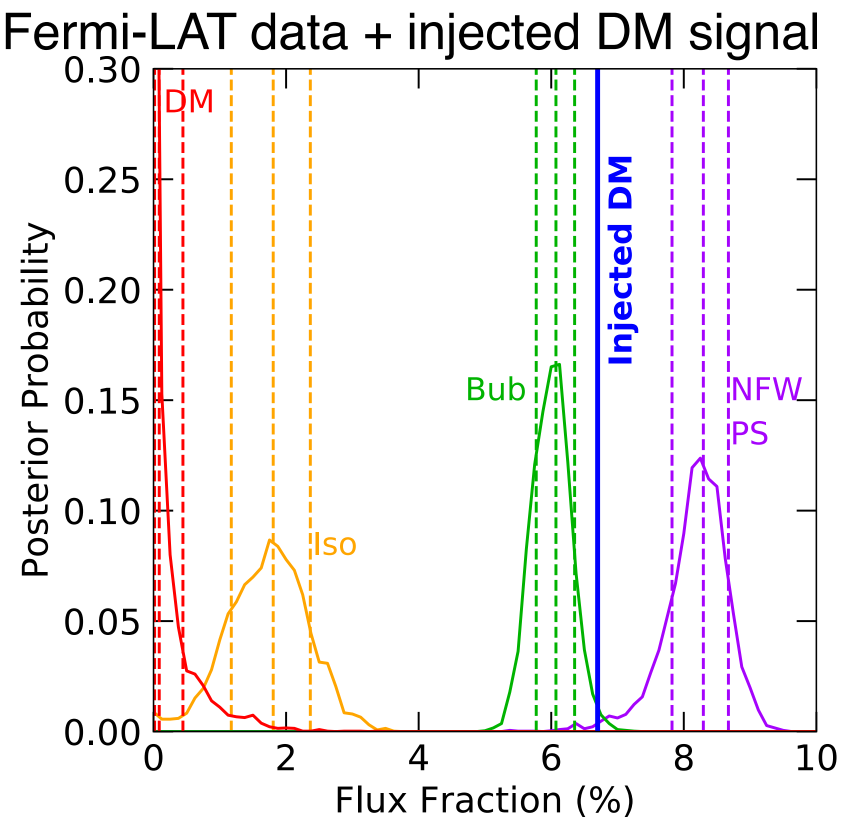
1. What do photon-count statistics approaches tell us?
2. Recent advances based on machine learning
3. Multi-wavelength picture of the GCE
4. The GCE morphology debate

And where mis-modelling of backgrounds enters in connection with template-based approaches!

The situation of non-Poissonian template fitting



[S. Lee et al., PRL 116 (2016)]



[R. K. Leane and T. R. Slatyer, PRL 123 (2019) 24]

Answers: Is the GCE smooth (DM-like) or granular (point-source-like)? (Point sources follow DM spatial distribution)

2016: First application of NPTF to inner Galaxy; strong evidence for a 100% point-source origin of the GCE.
[S. Lee et al., PRL 116 (2016)]

2019: Rebuttal of 2016 results; NPTF approach mis-attributes injected DM-like signal to point sources
[R. K. Leane and T. R. Slatyer, PRL 123 (2019) 24]

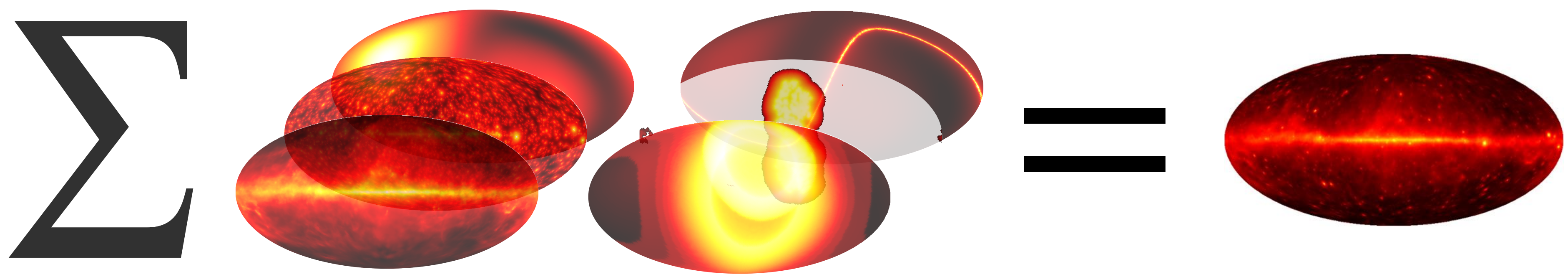
2019 & 2020: Robustness check of NPTF method by independent group: Mis-modelling of backgrounds led to reported inconsistency; GCE at least partially made of point-source population.
[M. Buschmann et al., PRD 102 (2020) 2] [L.J. Chang et al., PRD 101 (2020) 2]

2020: Evidence for point-source origin may be driven by spurious sources due to un-modelled North-South asymmetry of the GCE in earlier works.
[R. K. Leane and T. R. Slatyer, PRL 125 (2020) 12] [R. K. Leane and T. R. Slatyer, PRD 102 (2020) 6]

Status: Background model uncertainties inhibit progress.

How well can templates model the inner Galaxy?

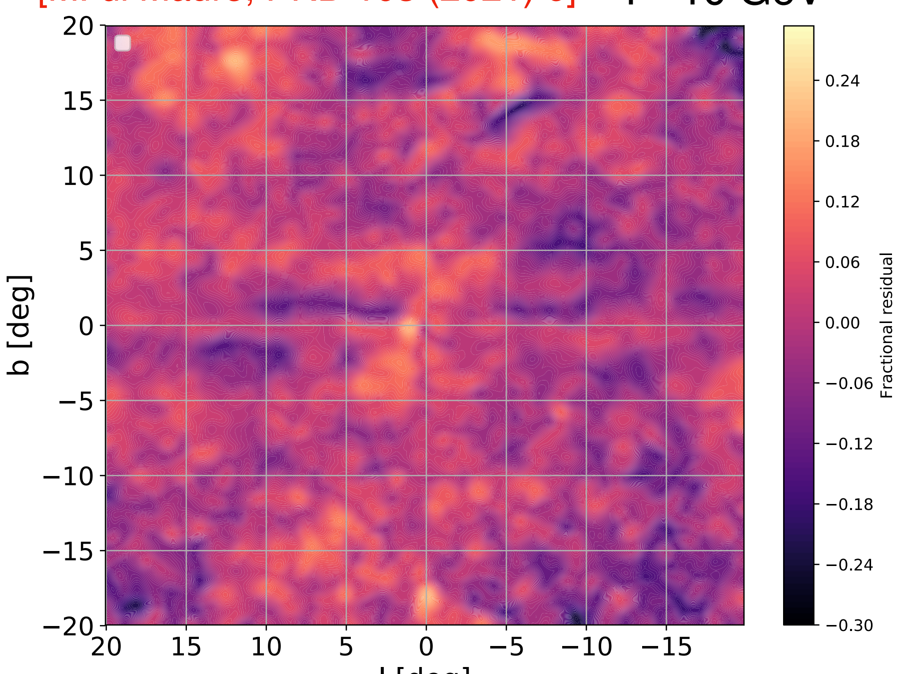
Reminder: When templates are involved we do something along the lines of:



more or less rigid spatial morphology + spectral freedom

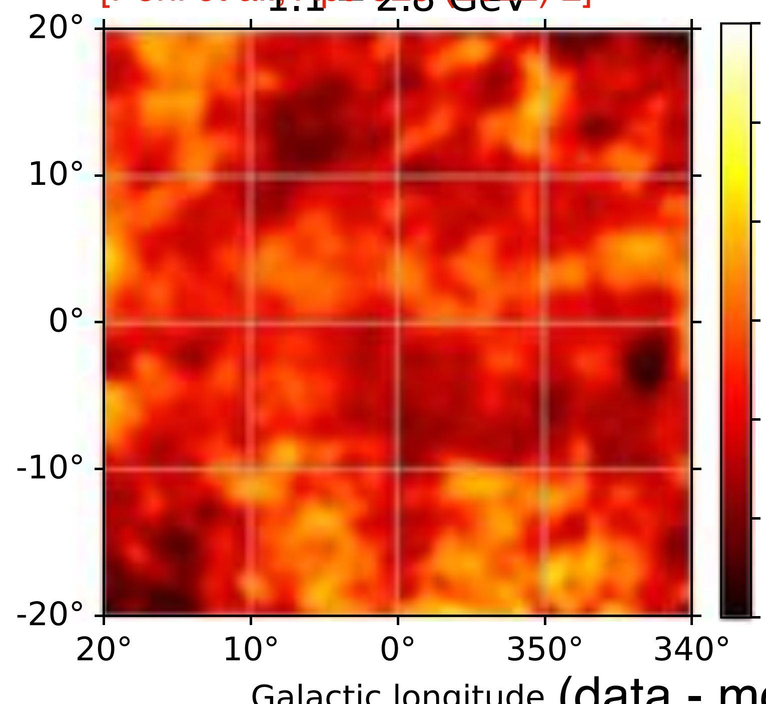
1. Residuals of best-fitting models can still reach ~30% and exhibit "some structure".
2. Trade-off between masking complex regions and having physically motivated/realistic models.
3. Mis-modelling typically impacts small-scales: See spurious sources due to North-South asymmetry reported in [R. K. Leane and T. R. Slatyer, PRL 125 (2020) 12] [C. Karwin et al., arXiv:2206.02809]

[M. di Mauro, PRD 103 (2021) 6] 1 - 10 GeV

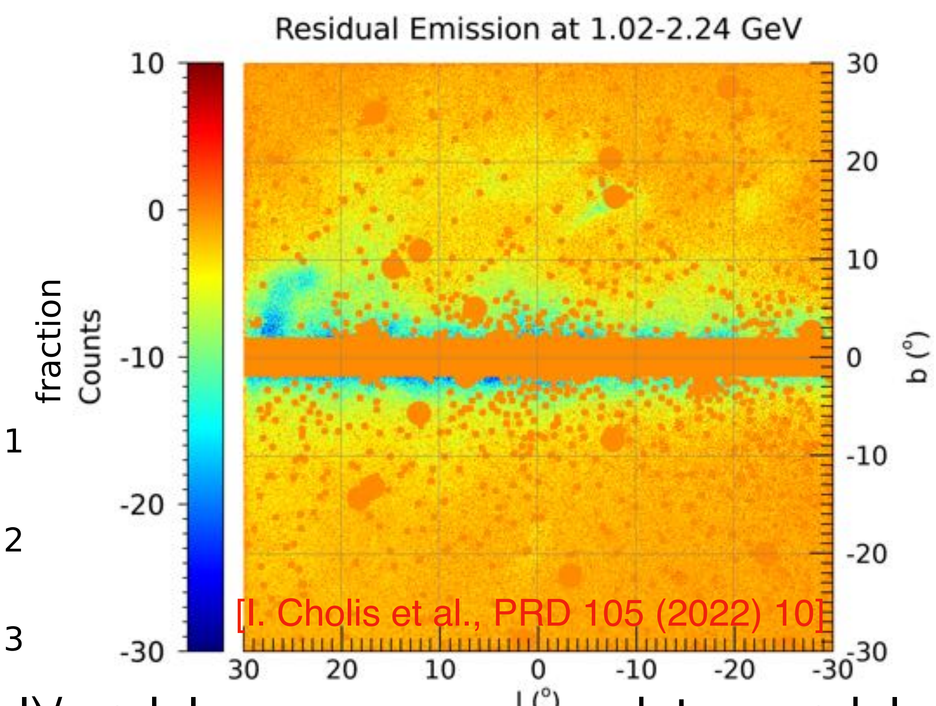


(data - model)/(data+model)

[Pohl et al., ApJ 929 (2022) 2]



(data - model)/model

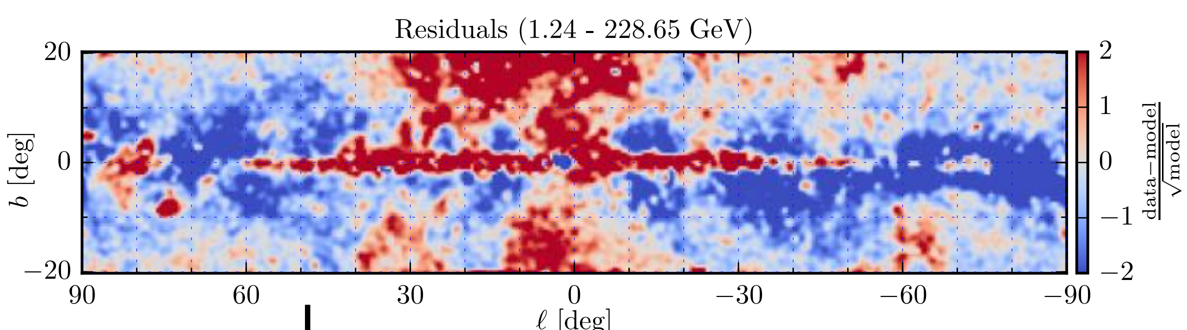


data - model

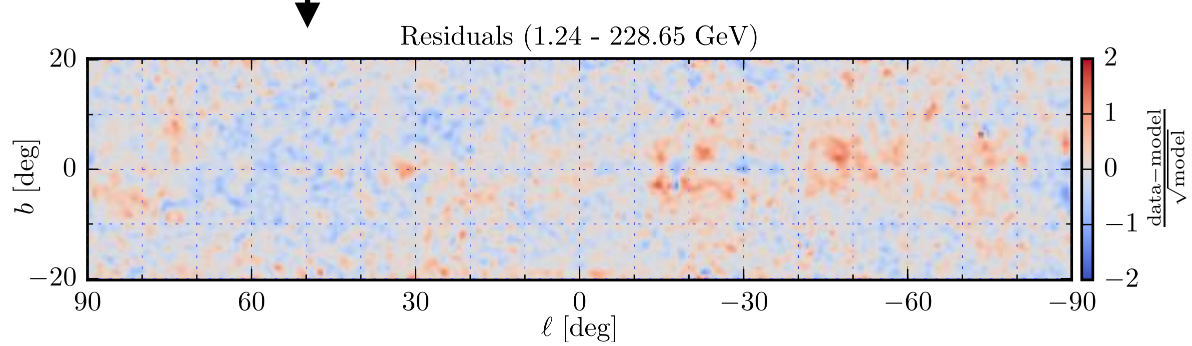
How people tackled the problem of mis-modelling

Data-driven approaches:

- Spherical harmonic marginalisation (—> improvement of diffuse templates by allowing for nuisance parameters on a range of angular scales)
  [M. Buschmann et al., PRD 102 (2020) 2]
- Gaussian processes
  [S. Mishra-Sharma and K. Cranmer, PRD 105 (2022) 6]
- Adaptive template-fitting: SkyFACT
  [E. Storm et al. JCAP 08 (2017) 022]

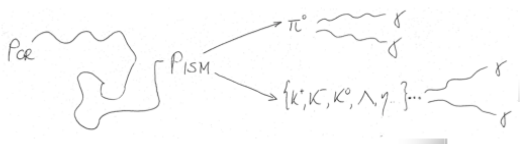


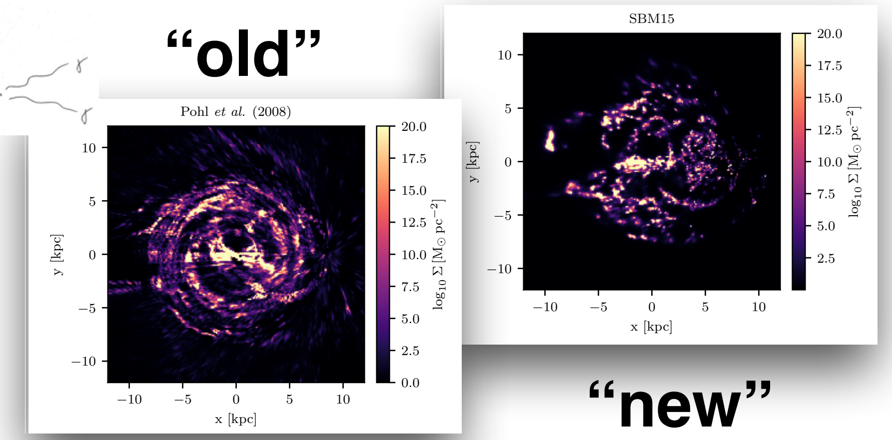
from vanilla model to model including GCE



Model improvements:

Main target: Improving our model for gas (HI, HII, H_2) in the Galaxy via spectral lines, dispersion measures, etc.
—> reduced small-scale mis-modelling

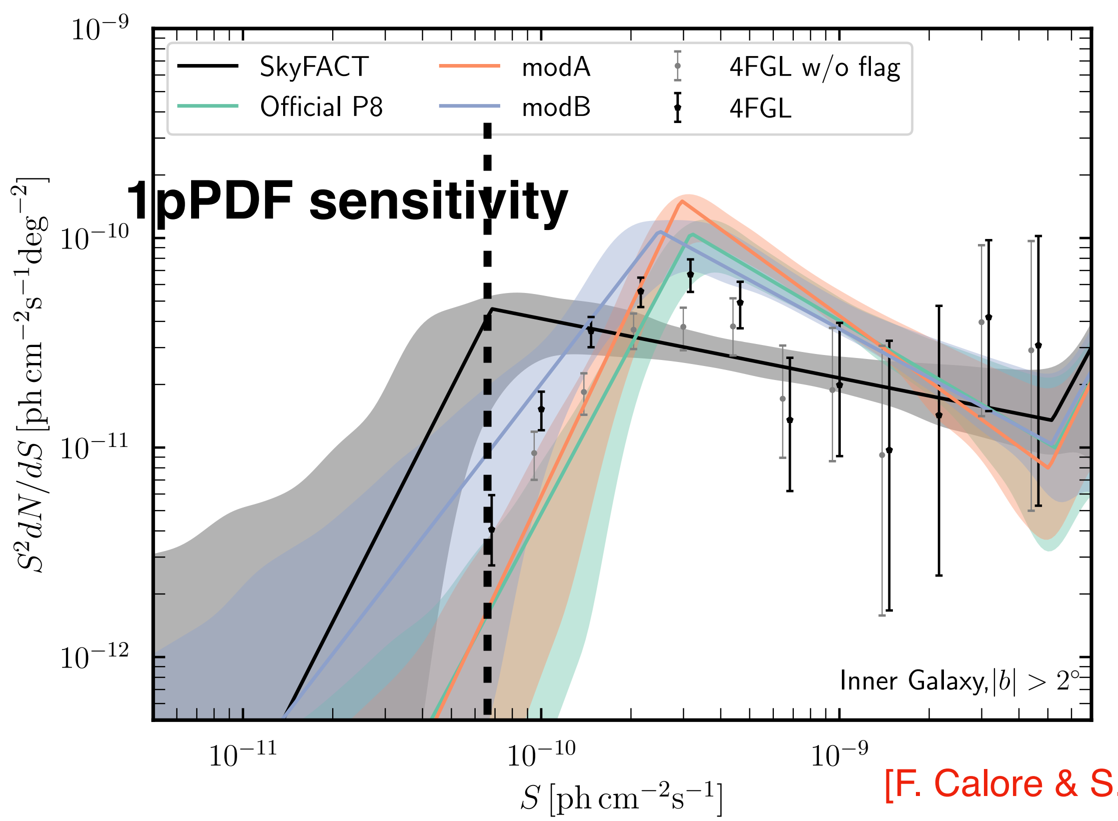


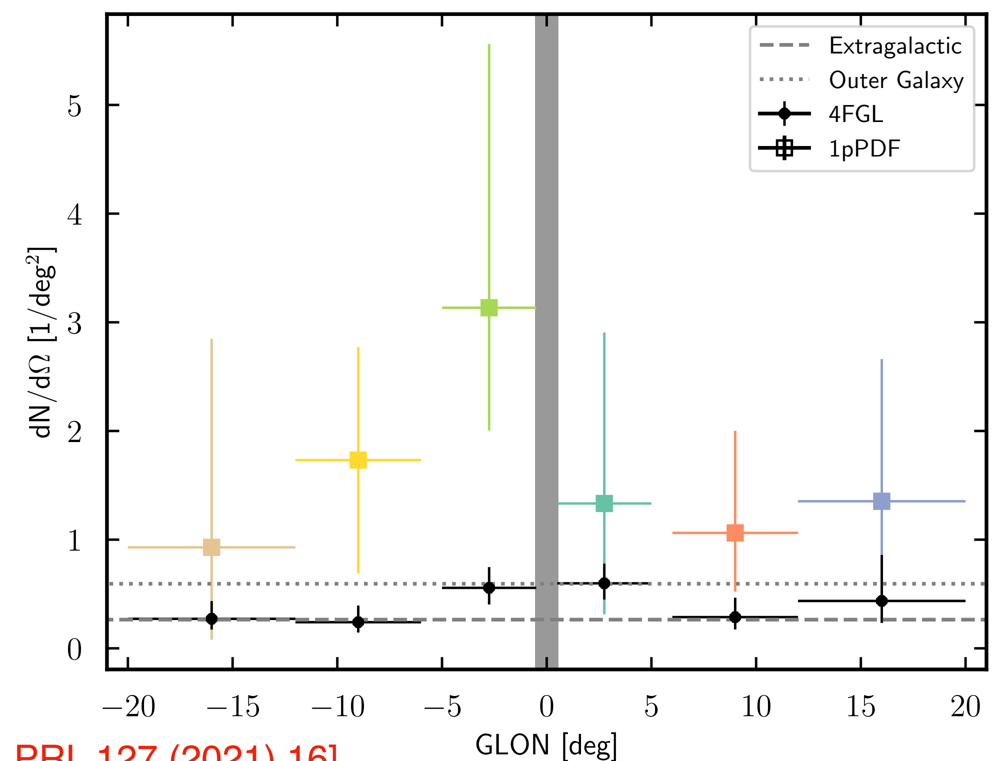


- new atomic HI reconstr, with radiation model of emission & absorption
  [A. Shmakov et al., arXiv:2206.02819]
- convolutional neural nets to fill gaps in molecular H_2-tracers like CO
  [C. Karwin et al., arXiv:2206.02809][A. Shmakov et al., arXiv:2206.02819]
- Bayesian inference of CO and HI maps
  [P. Mertsch, Vittino; A&A, Volume 655, A64] [P. Mertsch, Vo Hong Minh Phan; arXiv:2202.02341]

SkyFACT and photon-count statistics

Idea:
- SkyFACT: Derive optimised diffuse template (while injecting GCE template: DM- or MSP-like).
- Model faint sources with 1pPDF method (after reducing the residuals).





[F. Calore & S. Manconi, PRL 127 (2021) 16]

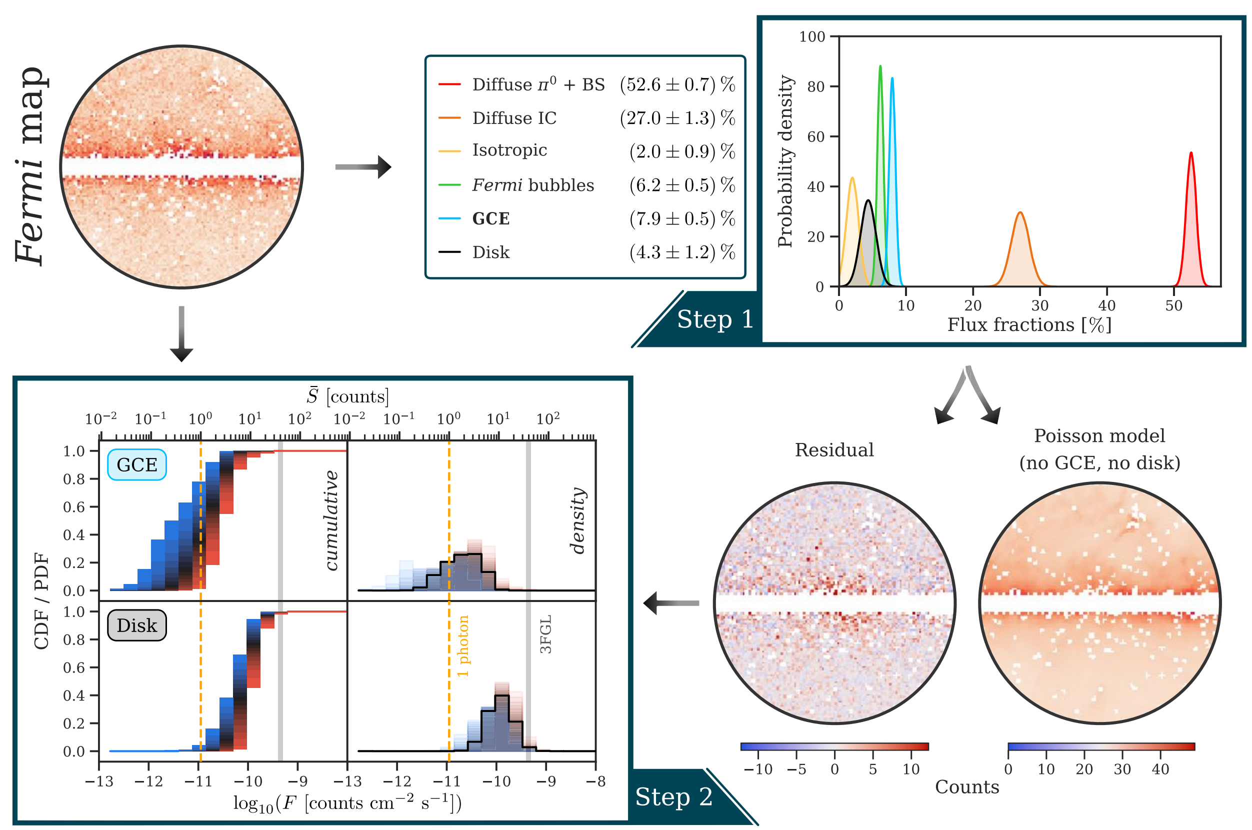
- Reconstruction of faint sources strongly depends on diffuse model.
- Still some spurious sources in 4FGL catalogue.
- Stellar bulge morphology preferred over spherically symmetric NFW profile.
- Faint sources not completely symmetric in longitude.

Optimising the diffuse model corroborates the hypothesis that the GCE is (at least) partially due to unresolved source populations.

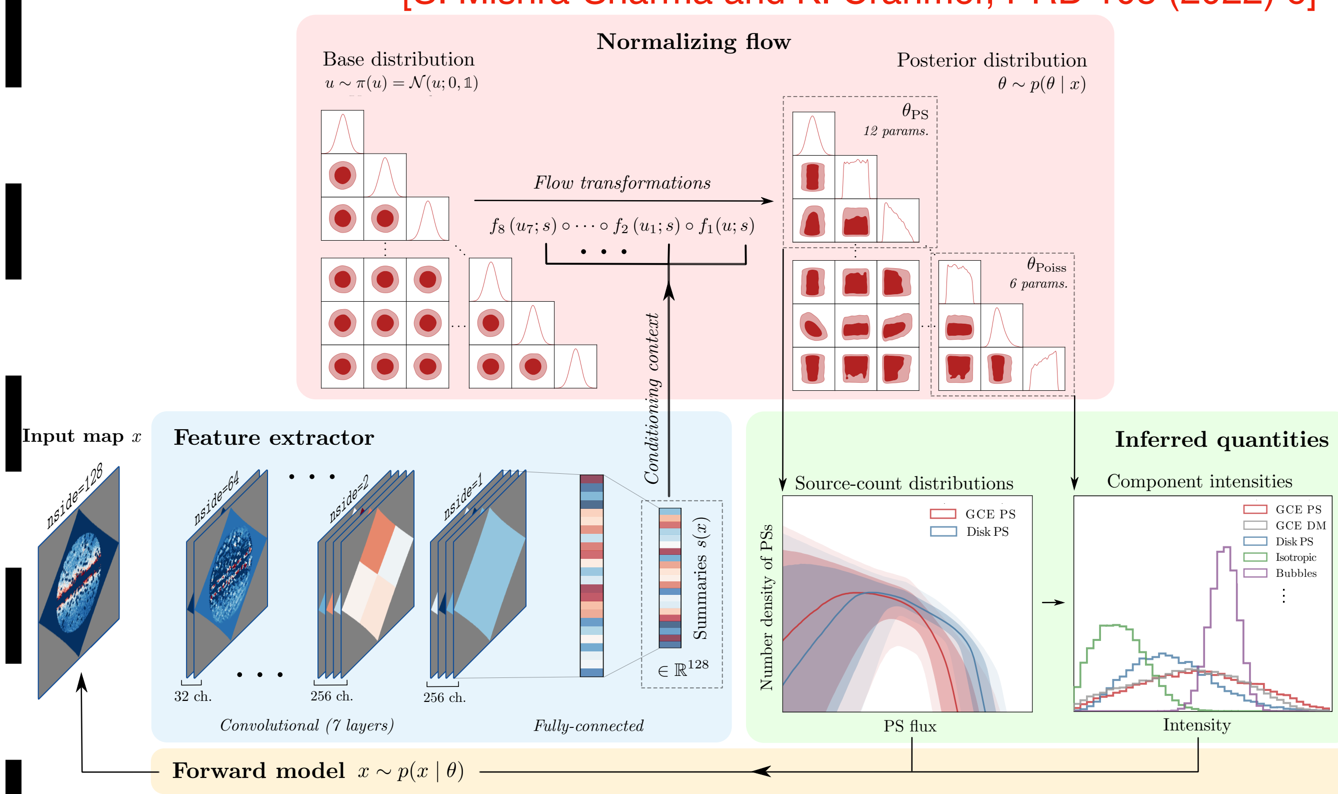
Machine vision on the GCE

Disclaimer: All machine learning works so far assumed DM and point-source-like GCE to follow the same spatial profile (generalised NFW).

[F. List et al. PRL 125 (2020) 241102] [F. List et al. PRD 104 (2021) 12]



[S. Mishra-Sharma and K. Cranmer, PRD 105 (2022) 6]



Combination of decomposing the gamma-ray sky with convolutional neural networks (CCN) and histogramme regression.

Point-source contribution to GCE < 66% and peaked around $4 \times 10^{-11}\ \mathrm{ph\,cm^{-2}\,s^{-1}}$**.**

Simulation-based inference with CCN to derive summary statistics and normalising flows for posterior estimation.

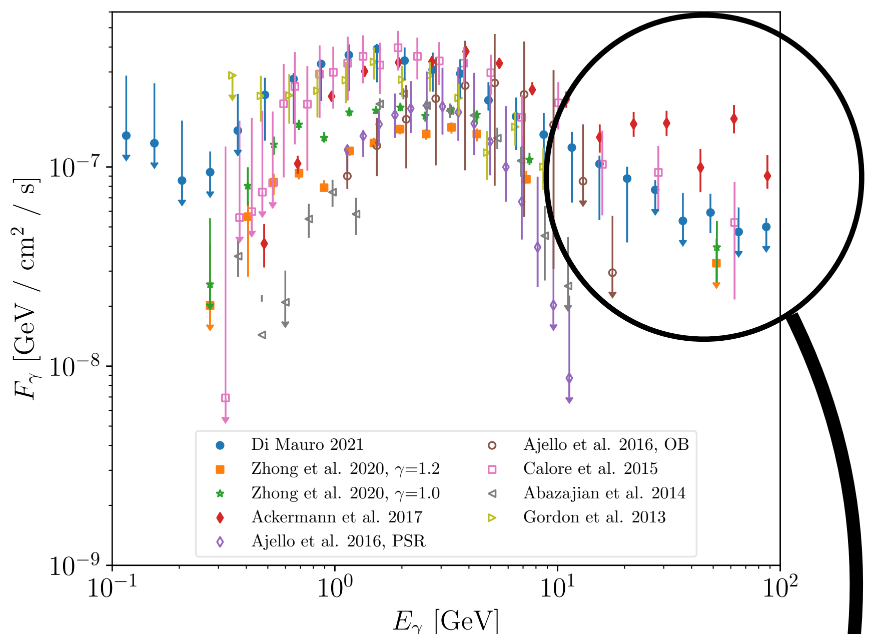
+ Gaussian processes to inject **large-scale variations** of diffuse model

Point-source contribution to GCE >= 40%.

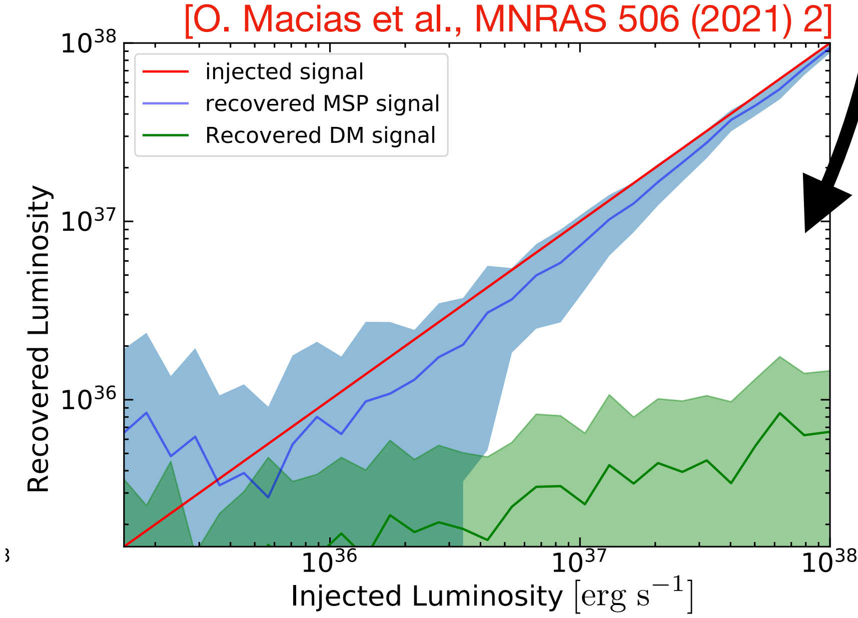
The GCE beyond GeV gamma rays

A TeV inverse-Compton tail:

[J.T. Dinsmore & T.R. Slatyer, JCAP 06 (2022) 06]



[O. Macias et al., MNRAS 506 (2021) 2]

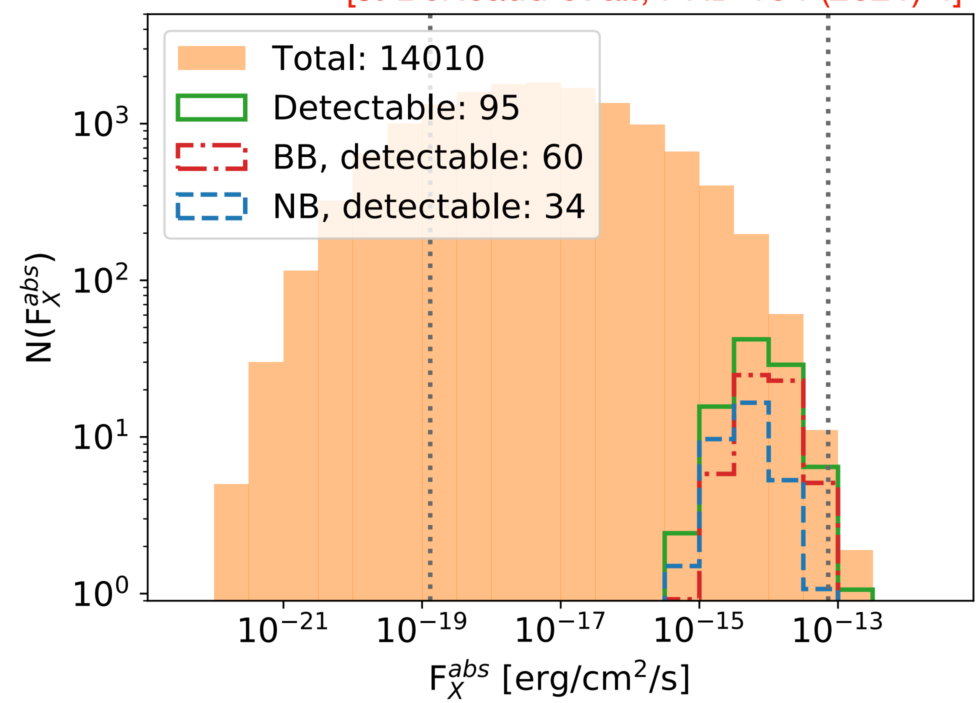


MSP population could produce detectable TeV emission via inverse Compton processes, CTA should see it (see also [C: Keith et al., arXiv:2212.08080]).

Finding bulge MSPs in radio data:

Despite being undetected in gamma rays, bulge MSPs may shine bright enough to be resolved in X rays/radio!

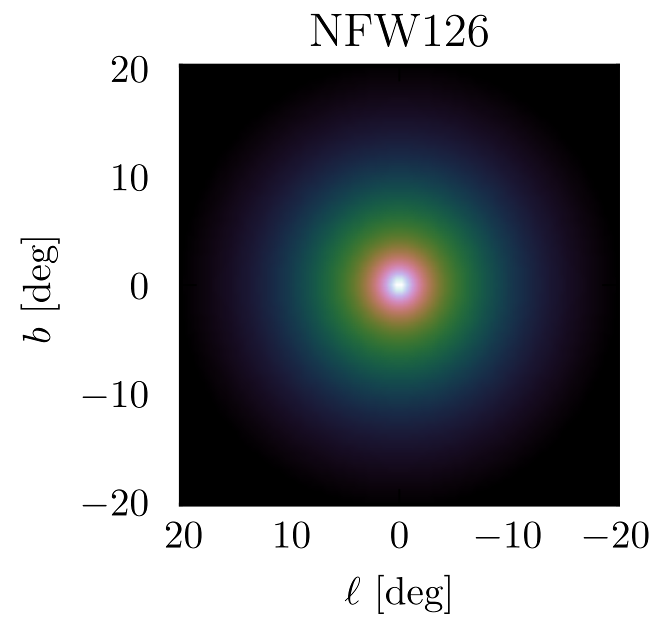
[J. Berteaud et al., PRD 104 (2021) 4]



- Expectations for X-ray emission of bulge MSPs consistent with Chandra.
- There are promising detected X-ray objects that may be MSPs in the bulge —> observe them with radio telescopes to measure pulsation period!
- MeerKAT and SKA are potent enough to observe the population with $\mathcal{O}(100h)$ of observation time.

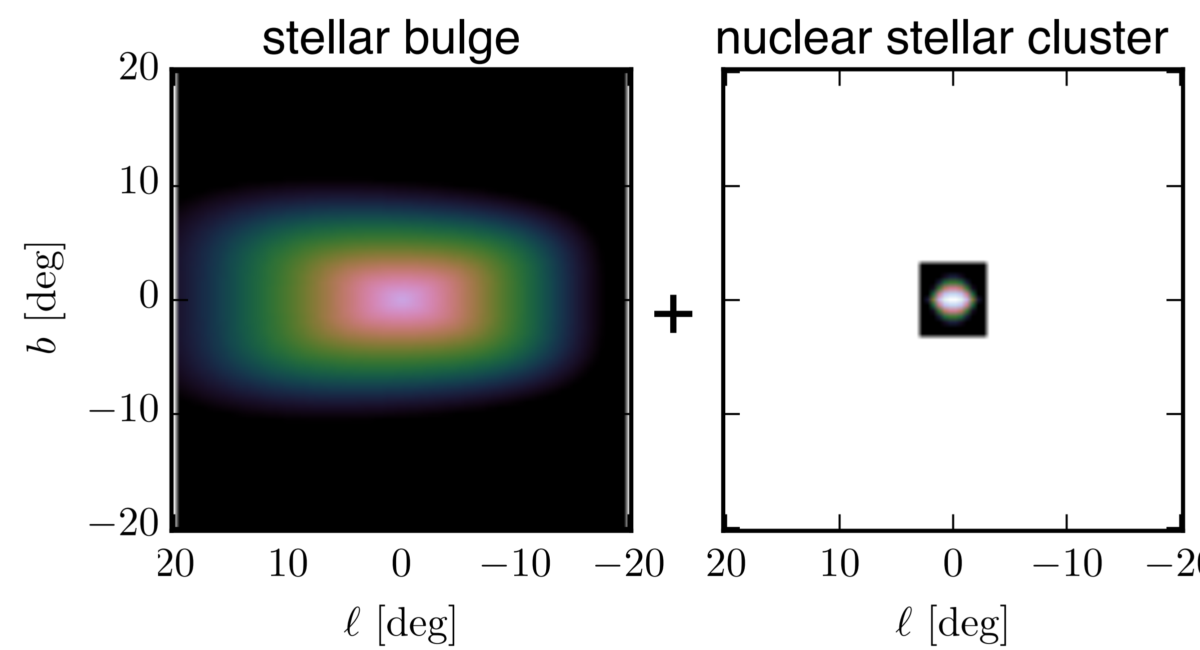
The recent GCE morphology debate

Doubts have recently been cast against the historically grown evidence for a preference in the data for a GCE tracing the distribution of old stellar populations.



supported by (incomplete):

[M. di Mauro, PRD 103 (2021) 6]
[I. Cholis et al., PRD 105 (2022) 10]
[S. D. McDermott et al., arXiv:2209.00006]



supported by (incomplete):

[R. Bartels et al., Nature Astron. 2 (2018) 10];
[O. Macias et al., JCAP 09 (2019) 042];
[F. Calore et al., PRL 127 (2021) 16]
[Pohl et al., ApJ 929 (2022) 2]

- Early works finding spherical symmetry usually applied non-refined diffuse templates and only tried one spatial profile.
- The more recent works employ a large model variety (cosmic-ray propagation parameters or diffuse components).

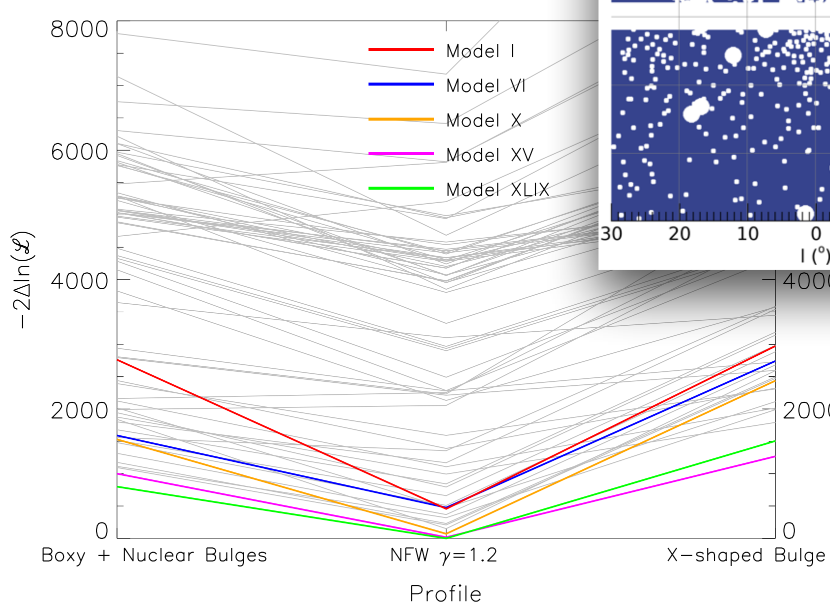
- Preference even stable when optimising the diffuse model with SkyFACT or advanced hydrodynamical simulations of gas in the MW.

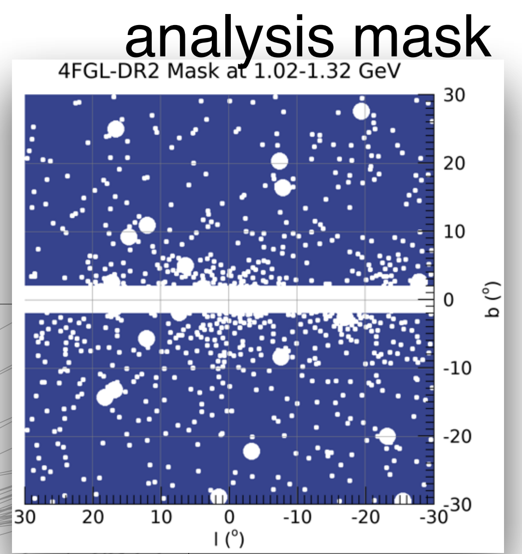
The recent GCE morphology debate

Where is the problem here?

On the NFW preference side:

[I. Cholis et al., PRD 105 (2022) 10]

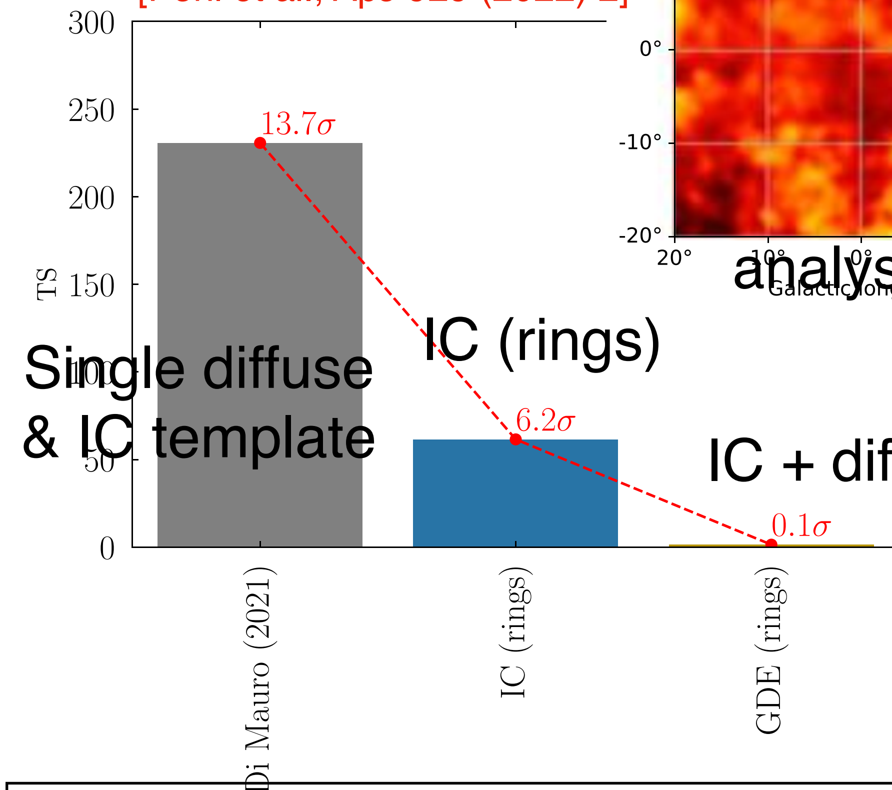


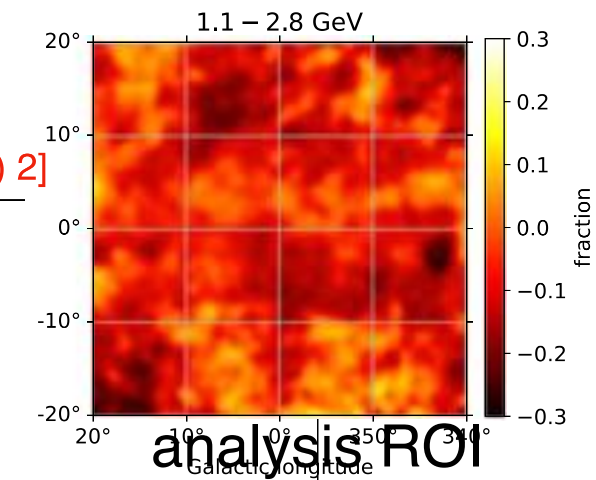


- Masked 4FGL sources and strip around Galactic plane opposed to [Pohl et al., ApJ 929 (2022) 2]
- Single diffuse template (although much freedom regarding cosmic-ray propagation parameters.

On the stellar density profile side:

[Pohl et al., ApJ 929 (2022) 2]





- Model of [M. di Mauro, PRD 103 (2021) 6] using stellar and nuclear bulge as baseline.
- Splitting diffuse templates in rings removes the preference of NFW profile.

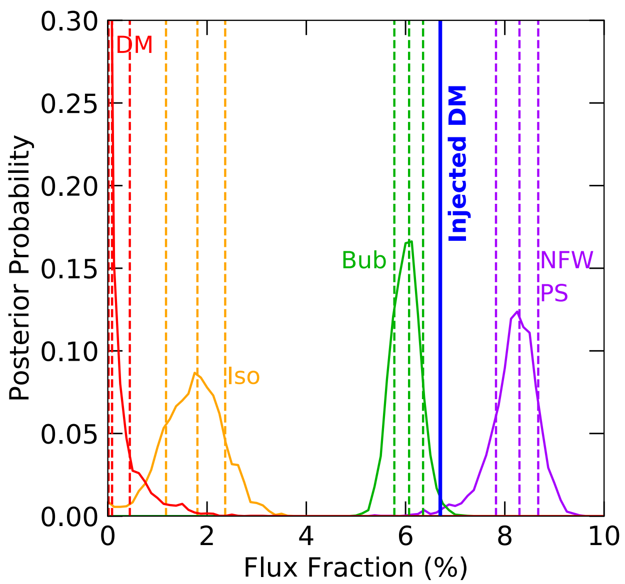
Bottom line: Template fitting without accounting for small-scale uncertainties is not viable anymore!

Mind the Gap: Why background uncertainties need to be overcome to progress on the GCE.

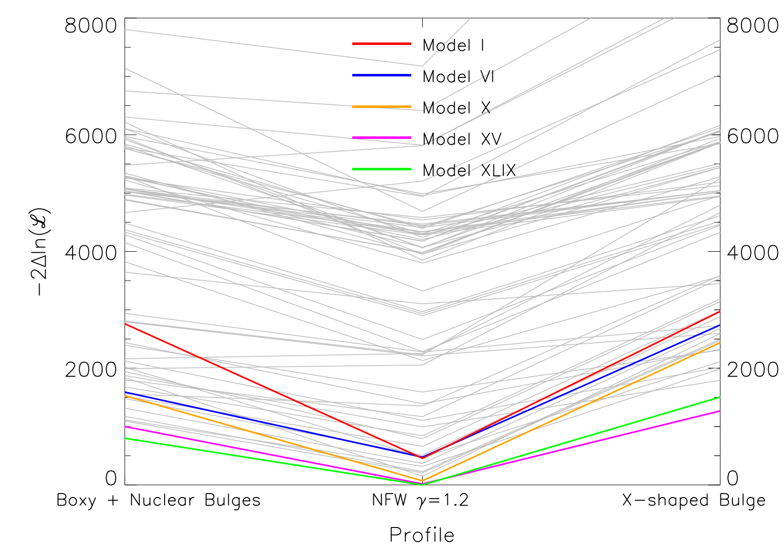
The road so far …

As I stated in the beginning:

We all agree: There is a significant excess of GeV gamma rays towards the Galactic centre measured by Fermi-LAT **above known astrophysical backgrounds**.



The history of non-Poissonian template fitting and the debate it stirred.



The debate around the GCE's morphology.

Our question: Can we quantify when we may trust results?

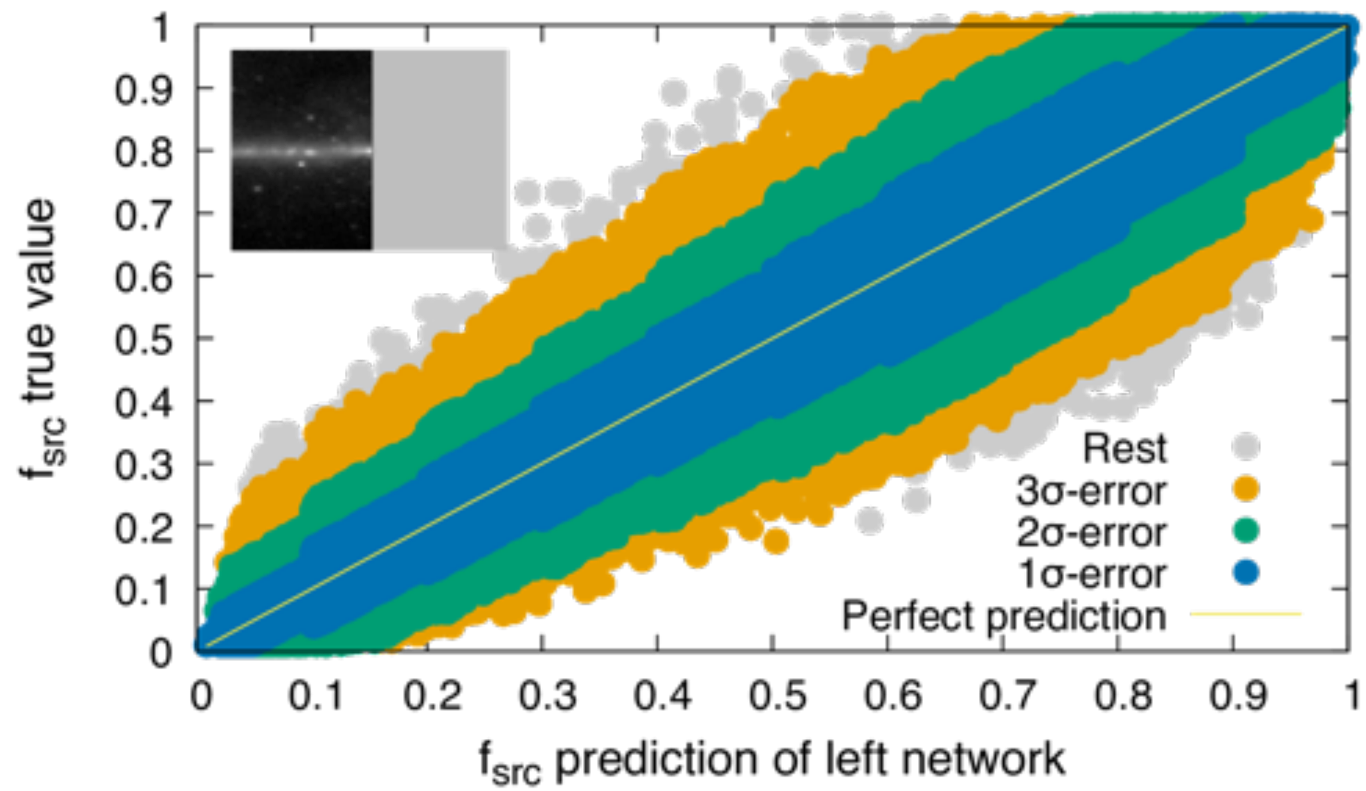
How machine learning can help

Machine learning techniques have already been applied to the GCE. **Convolutional neural networks are great at identifying specific small- and large-scale features in images.**

Pioneering: [S. Caron et al., JCAP 05 (2018) 058] —> The work presented here extends this pioneering study.

Recap of their approach:

1. Fit a given diffuse background model, FL8Y sources and DM template with fixed slope parameter $\gamma = 1.1$ to real Fermi-LAT data (a single energy bin from 1 - 6 GeV) to obtain total GCE luminosity.
2. Create MSP population whose total gamma-ray luminosity is a certain fraction (f_{src}) of the determined GCE luminosity.
3. Train a network on diffuse background template + FL8Y sources + f_{src} * MSP template + (1-f_{src}) DM template with output f_{src}.



f_{src} is 0.887 ± 0.105.

Quite simplistic and prone to be affected by background uncertainties given the previous discussion. Let's do it better!

How machine learning can help

Machine learning techniques have already been applied to the GCE. **Convolutional neural networks are great at identifying specific small- and large-scale features in images.**

Pioneering: [S. Caron et al., JCAP 05 (2018) 058] —> The work presented here extends this pioneering study.

Our approach:
Convolutional neural networks trained on images of the Galactic centre composed of background and signal templates to reconstruct the parameters of the GeV excess.

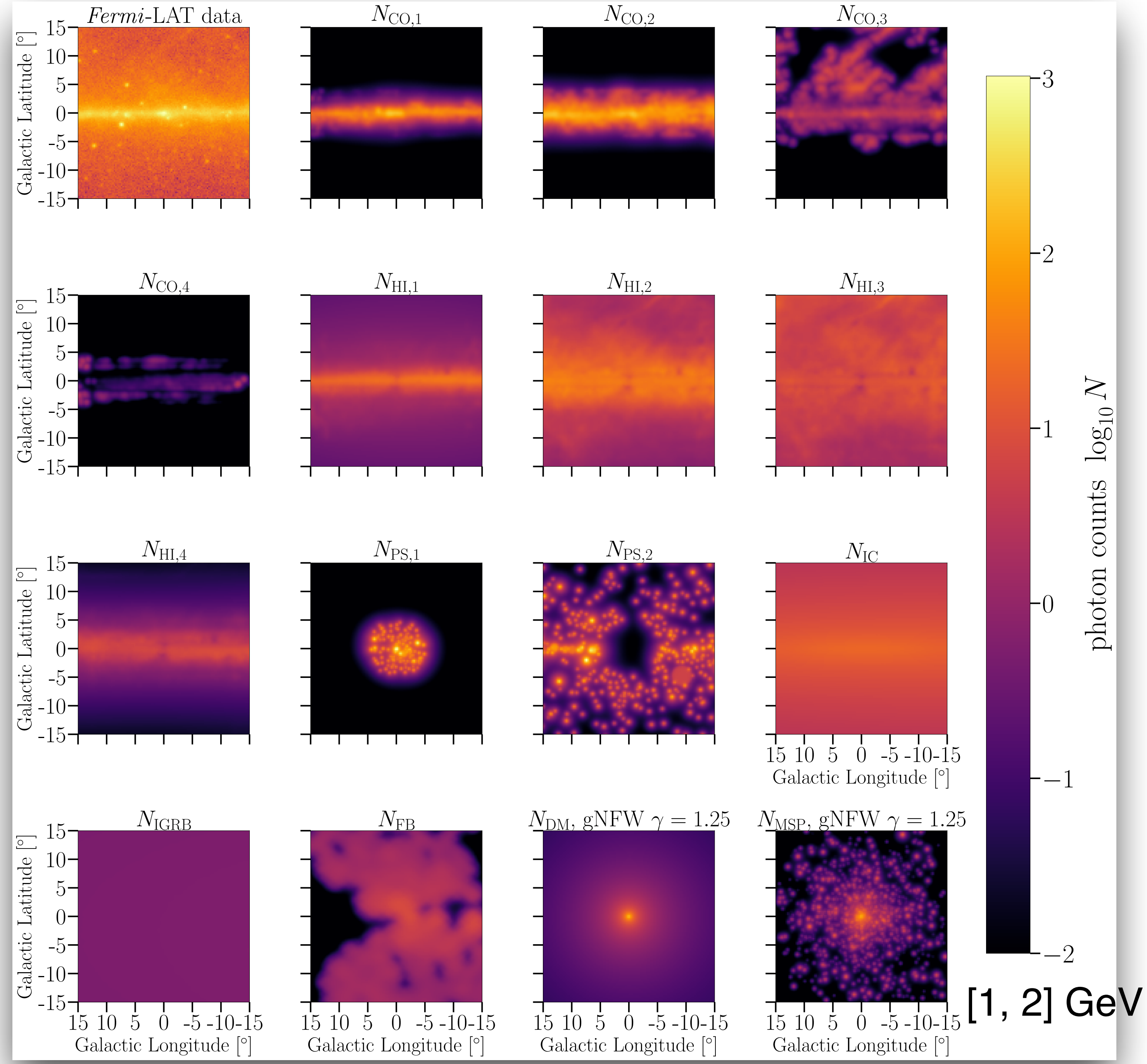
— 4 rings for diffuse background model (split into HI and CO contribution) + inverse-Compton (single) [Fermi collab., ApJ. Suppl. 224 (2016) 1];

— all 4FGL-DR2 point sources within 20°x20° region of interest (matches period selected for real data) [Fermi collab., ApJ.Suppl. 247 (2020) 1];

— Fermi Bubbles [Fermi collab. ApJ 840 (2017) 1];

— isotropic component

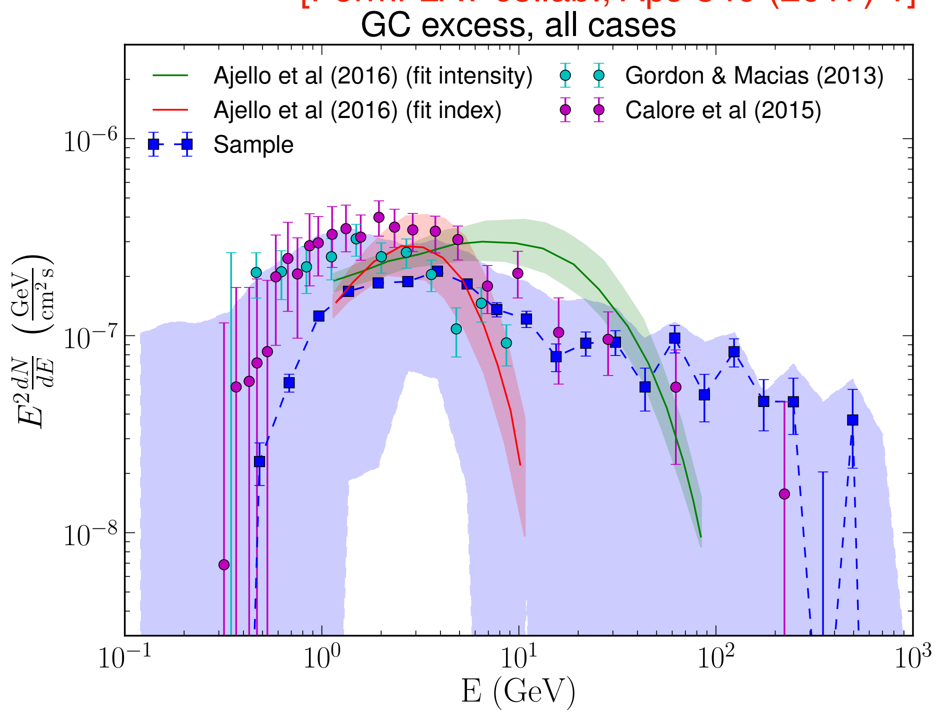
— GCE: smooth DM component + individually drawn MSPs both following a gNFW profile



A word on the millisecond pulsar population

MSP population generated from spatial profile (same as DM) and luminosity function + common spectral shape.

[Fermi-LAT collab., ApJ 840 (2017) 1]



minimal and maximal GCE luminosity (**blue-shaded band**) defines **lower and upper limit for the total luminosity** from a population of **(unresolved) MSP population**

1. Draw a total GCE luminosity (0.1 - 100 GeV) compatible with the uncertainty band.

$$\frac{\mathrm{d}N}{\mathrm{d}L} \propto \frac{1}{L} \exp\left[-\frac{\left(\log_{10} L - \log_{10} L_0\right)^2}{2\sigma_L^2}\right]$$

Draw individual pulsars:

2. Draw (ℓ, b) from gNFW profile,
3. Draw MSP luminosity from the luminosity function with current σ_L,
4. Derive spectrum normalisation according to the total luminosity,
5. Check if the resulting flux < Fermi-LAT detection threshold (3PC version) at this position
6. If so, add the pulsar to the MSP catalogue if the total GCE luminosity is not yet reached **(scenario A)**

   —> prepare **scenario B**: all drawn MSPs are included

Neural network architecture and scope

Network architecture:

— **Bayesian convolutional networks** to determine the simulation parameters in an inference task
—> input: (5, 120, 120) images of the GC: 5 energy bins ([0.5, 1], [1, 2], [2, 7],[7, 20], >20 GeV)
—> predict every template parameter with uncertainty

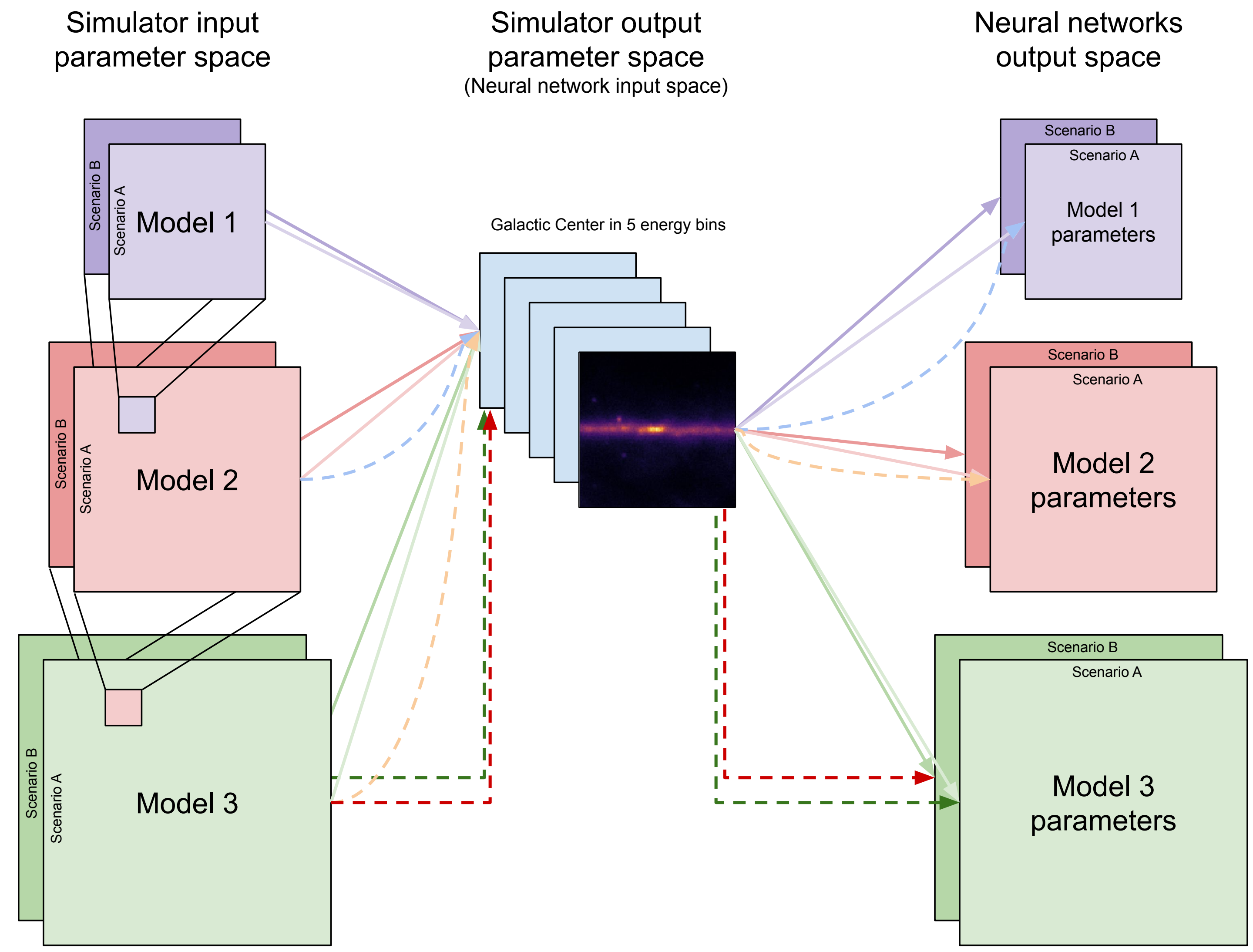
— **Deep ensemble networks** to also predict ***uncertainties due to the network itself***: What would have happened if the network was initialised differently (initial parameters, ordering of images, etc.)?
—> mean and scatter evaluated per network run and later combined

Additional uncertainties:

— DM and MSP templates follow the same spatial morphology. We are only interested in the fractional contribution of both components.

— The ‘**reality gap**’ — the discrepancy between modelled and real data — may dominate all reconstruction uncertainties!

— **Increase/decrease complexity of the model and check if results remain stable.**



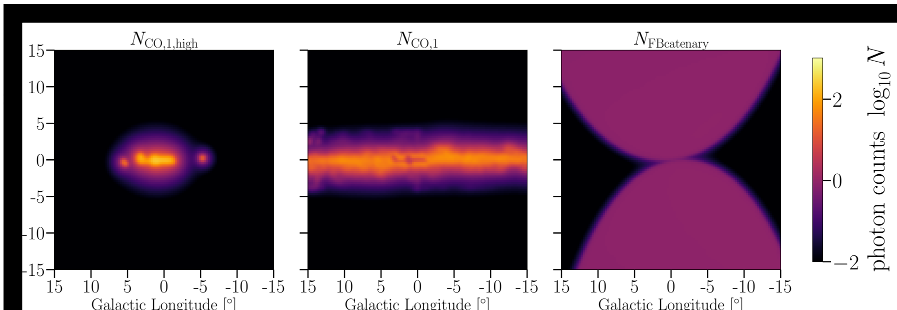
Neural network architecture and scope

Model setup to explore the impact of the **background model complexity** on the interpretation of the GCE with **Bayesian convolutional neural networks** used in a **DeepEnsembles** setup. We probe the '**reality gap**' — the discrepancy between modelled and real data.

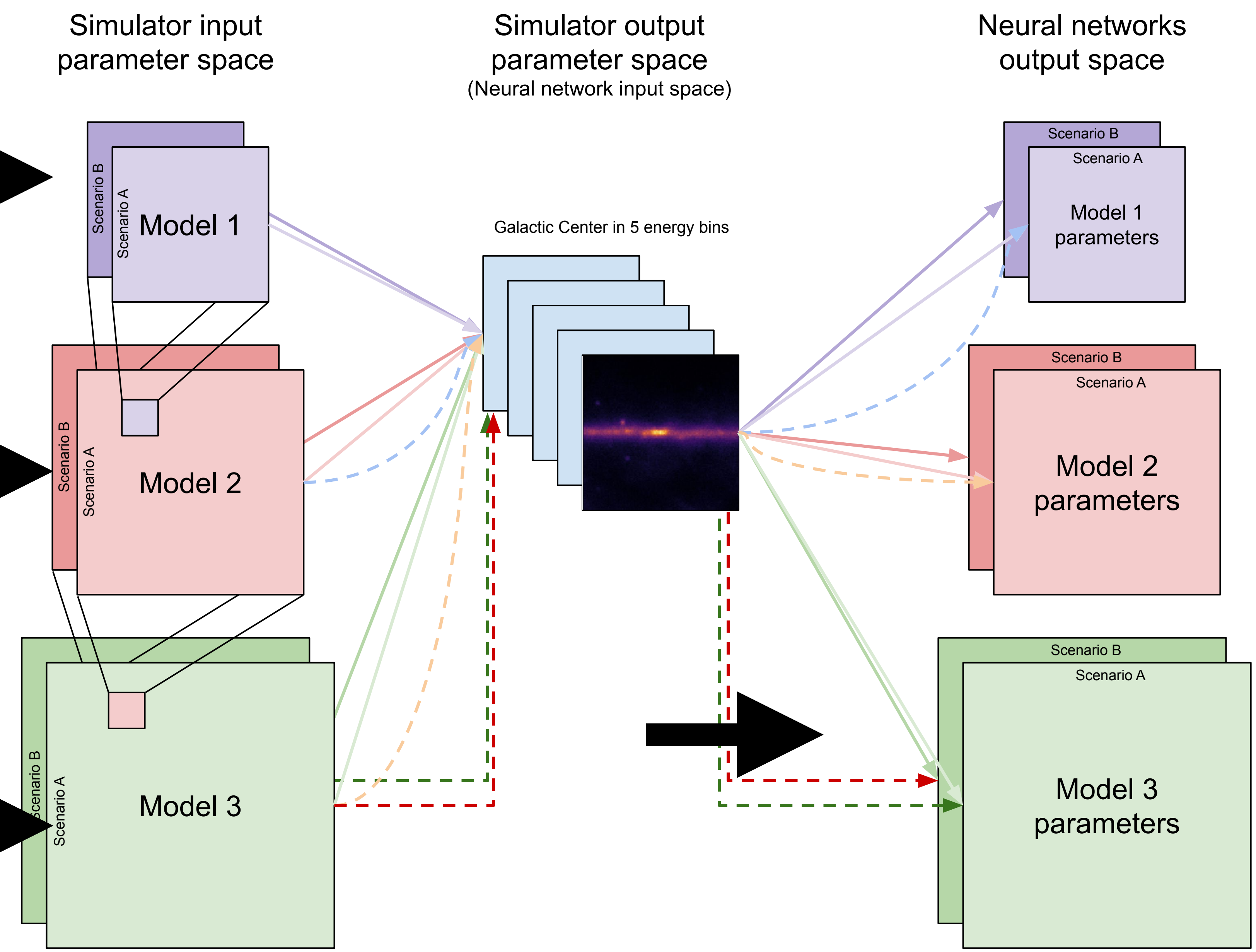
Reduce to 7 parameters, only one for astro diffuse emission (relative norm. of astro components fixed by likelihood fit)

compares to [S. Caron et al., JCAP 05 (2018) 058]

Initial 18 parameters



+ spectral freedom —> 25 parameters



scenario A: discard all MSPs above the Fermi LAT detection threshold

scenario B: keep all MSPs

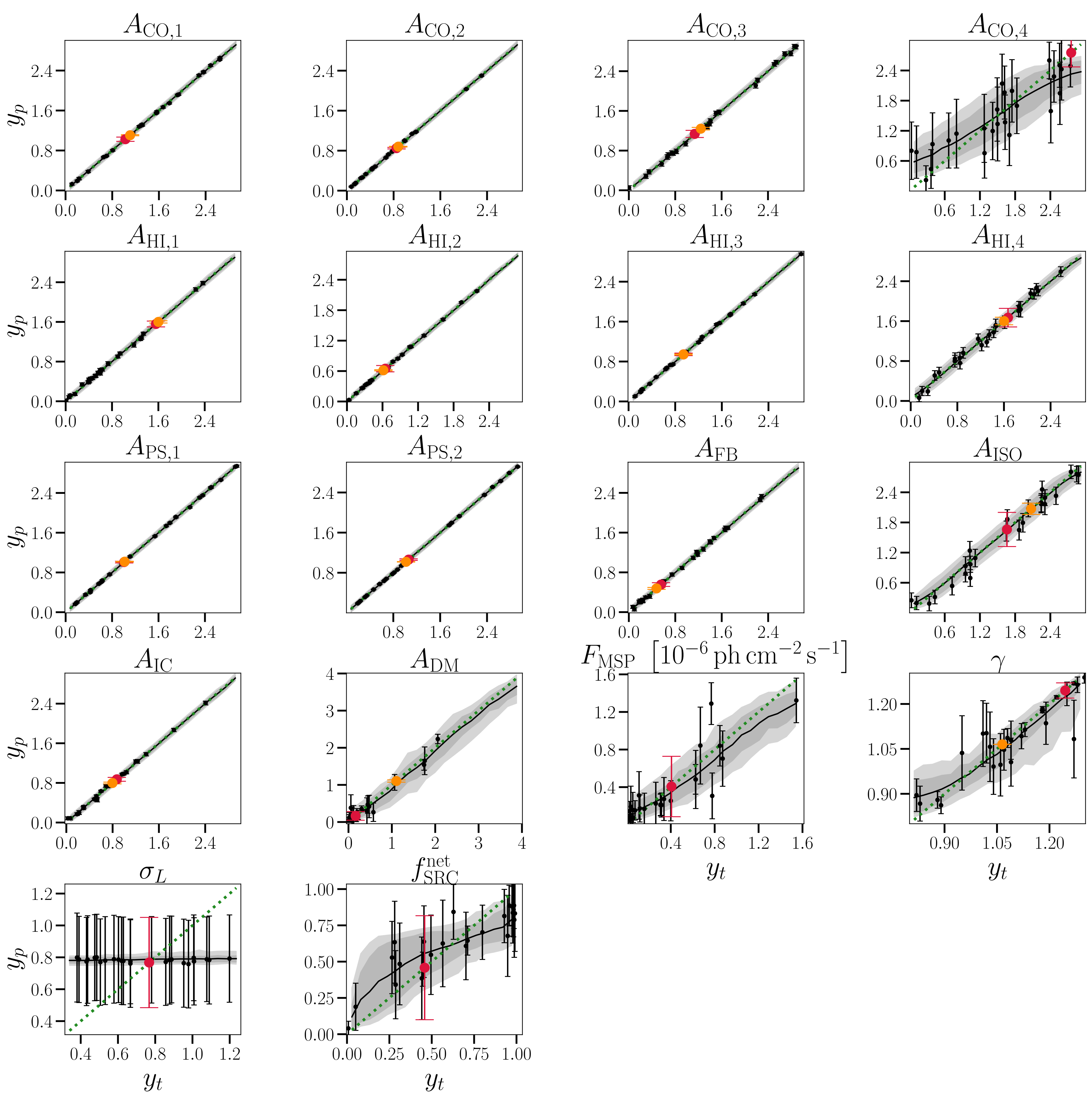
Results on fiducial model (2)

Gray band: 1(2) σ scatter of the mean prediction in the validation dataset

Error bars: network predicted uncertainty for selected points

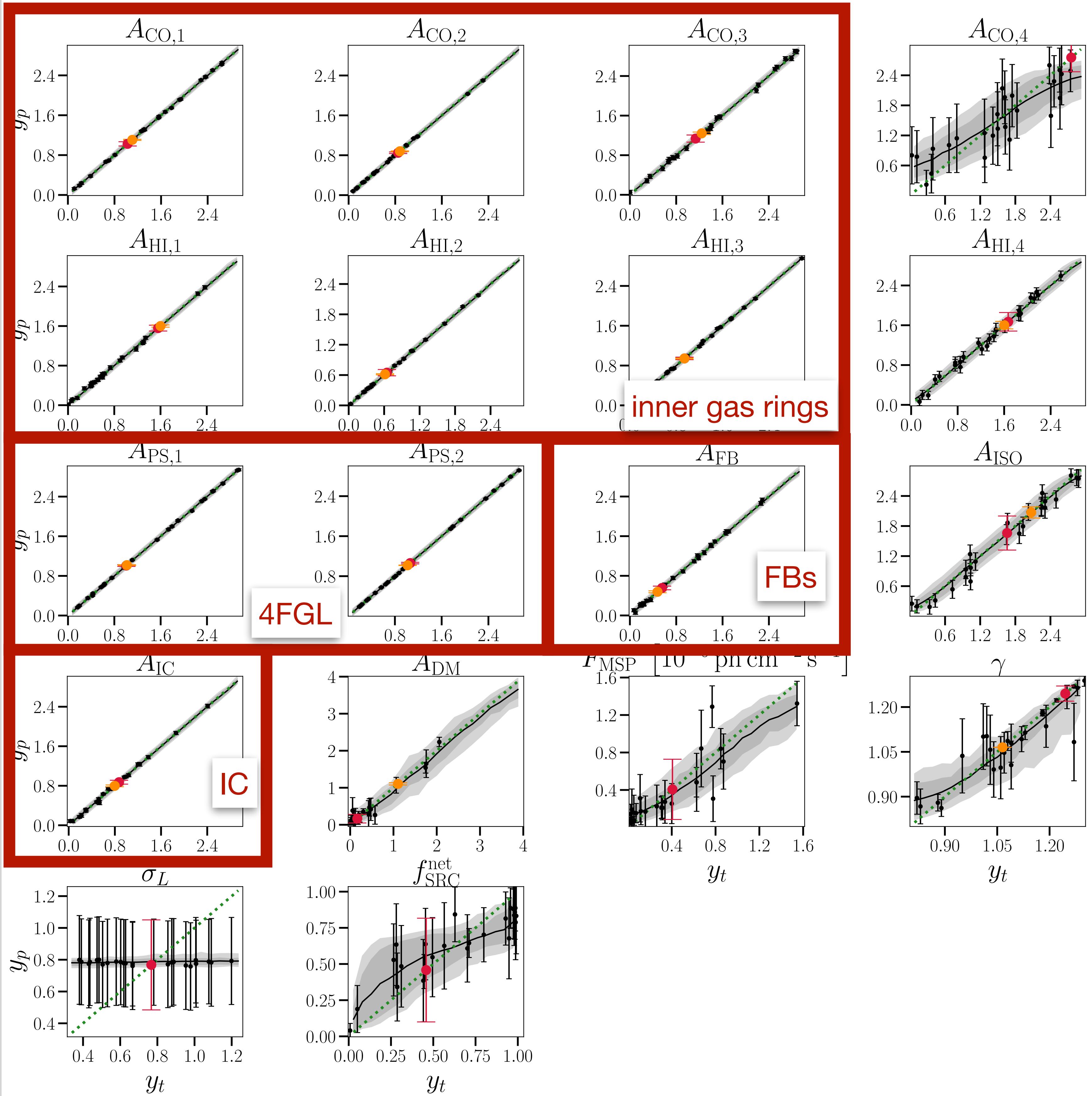
Red point: results on real data
Orange point: results of maximum likelihood fit

X/Y-axis: True/predicted value (means should fall on the diagonal



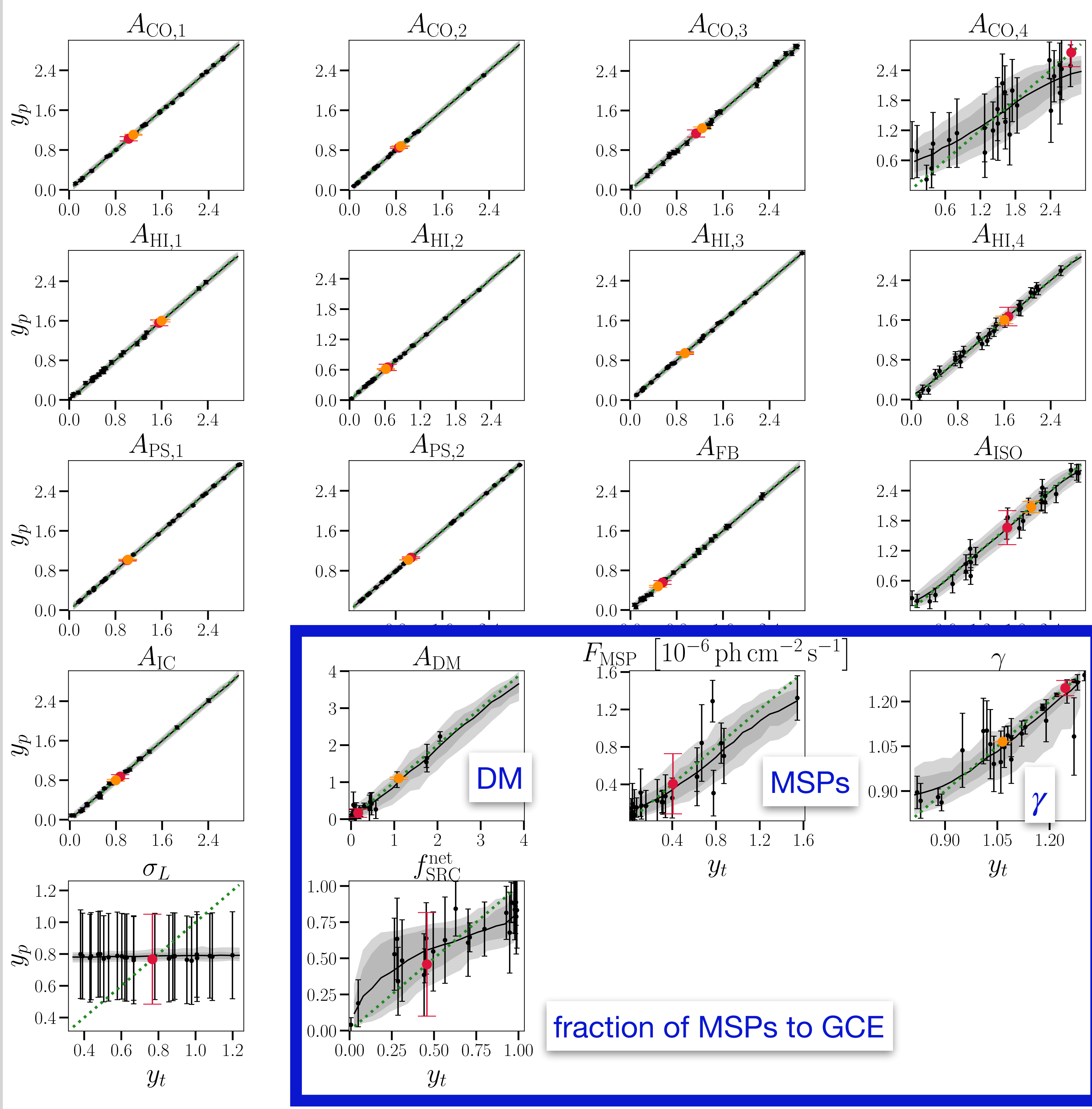
Results on fiducial model (2)

Bright components predicted with high precision and consistent with the traditional likelihood analysis.



Results on fiducial model (2)

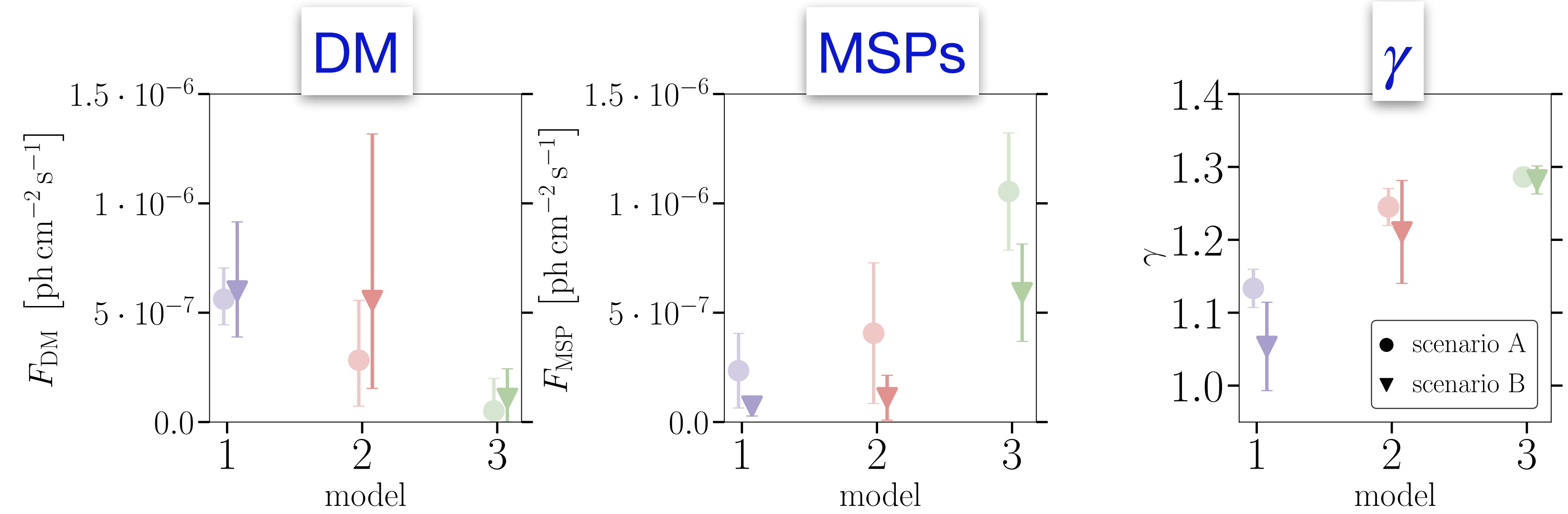
GCE related components predicted robustly, though fSRC has some bias and significant uncertainty

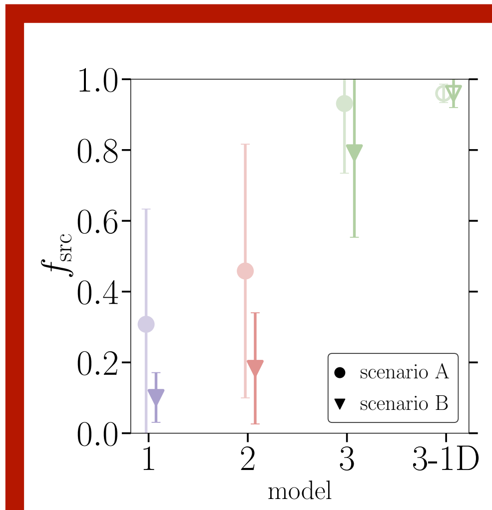


Evolution of results with model iteration

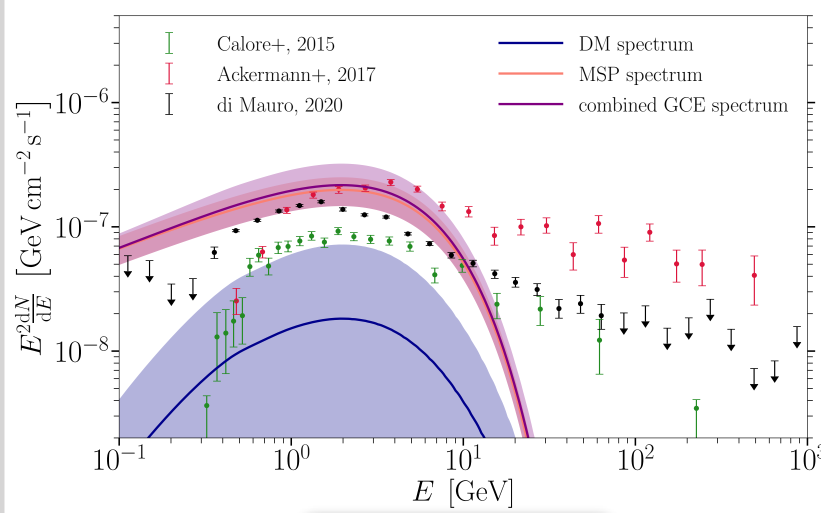
Predictions broadly consistent among the models (also conventional astro components),

BUT there seems to be a trend towards lower values of DM and high values of MSP normalisation when increasing the complexity of the models.

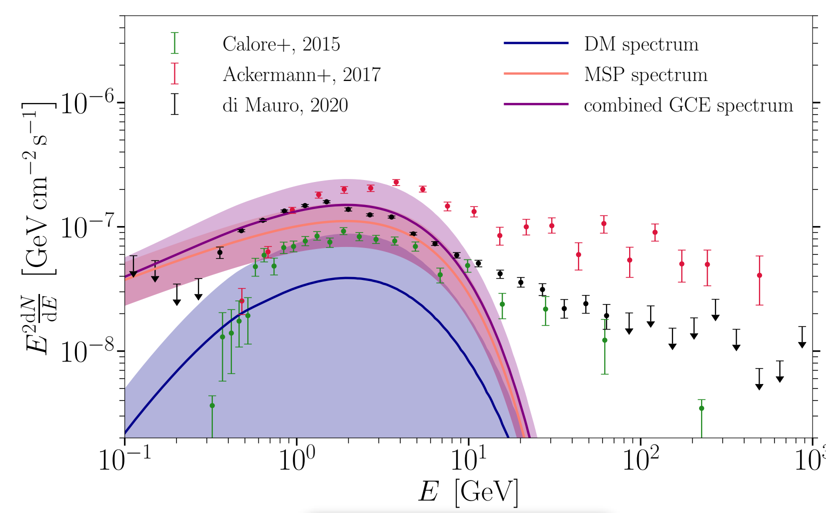




fraction of MSPs to GCE



scenario A



scenario B

Mind the Gap

There is only one gamma-ray sky…
Machine learning (as any analysis) yields a result, but how to check if this particular result is robust?

\+ Such checks especially relevant when using a new (black box?) tool

Two ways to probe the robustness:

training data

validation data

How biased is the network when it tries to parameterise a wolf in the universe of dogs?

Here: Validation dataset comprised of 4FGL diffuse background model:
— contains smoothed version of FBs and GeV excess
— other astro components as usual (Model 3A).

How far is a wolf from dogs in the world of the neural network?

Machine learning technique called: One-Class Deep Support Vector Data Description (Deep SVDD)

Classifies images as simulation-like or not

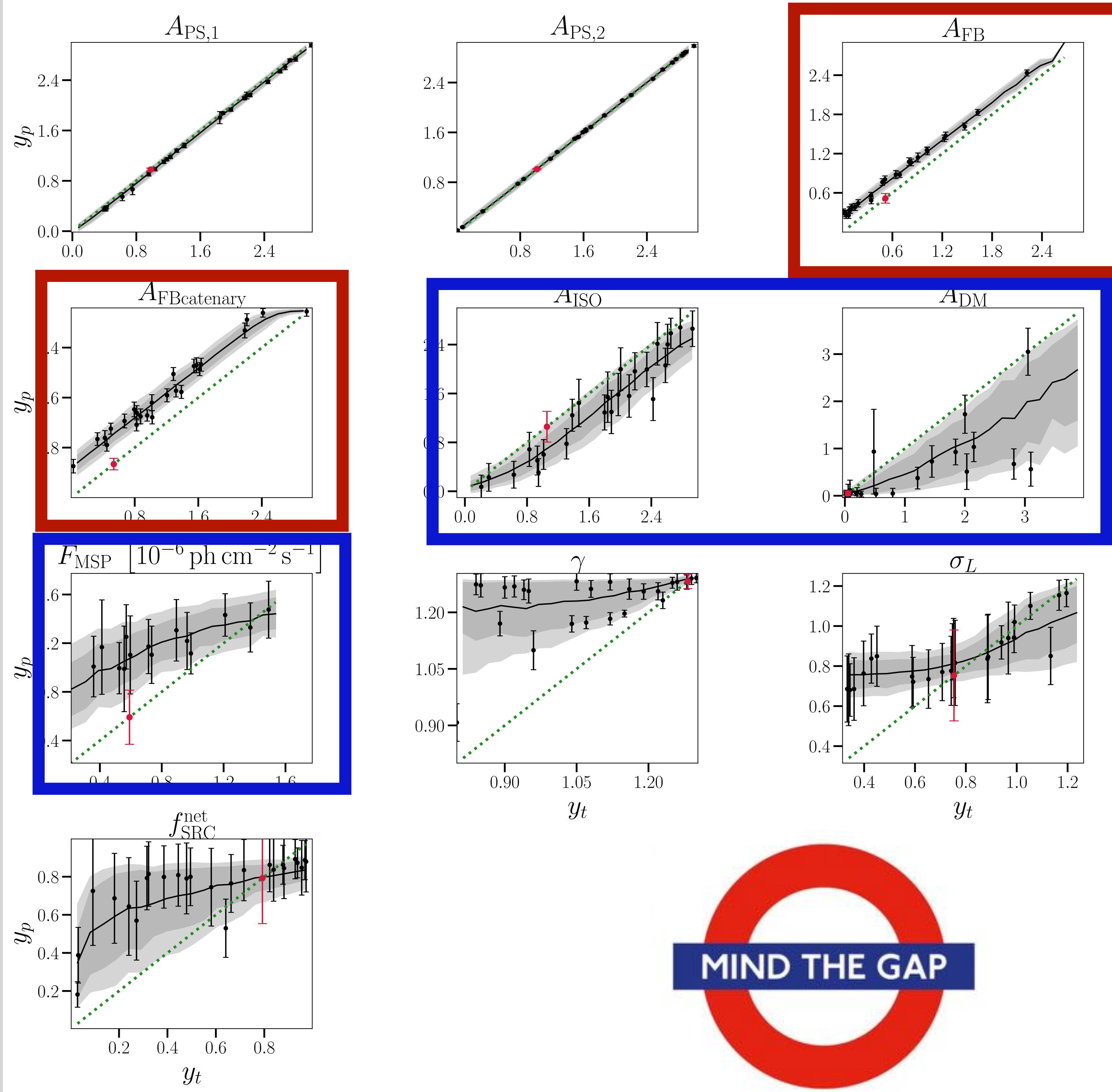
Mind the Gap — Fermi diffuse background model

Network trained on **Model 3A verified on the Fermi diffuse model**

FB templates are over predicted as expected

However, ISO and DM components are under predicted and (consequentially?) MSP template is over predicted

ILLUSTRATES THAT IN THE CASE THERE IS A ‘GAP’ BETWEEN THE TRAINING AND VERIFICATION DATA, RESULTS ARE UNRELIABLE!



MIND THE GAP

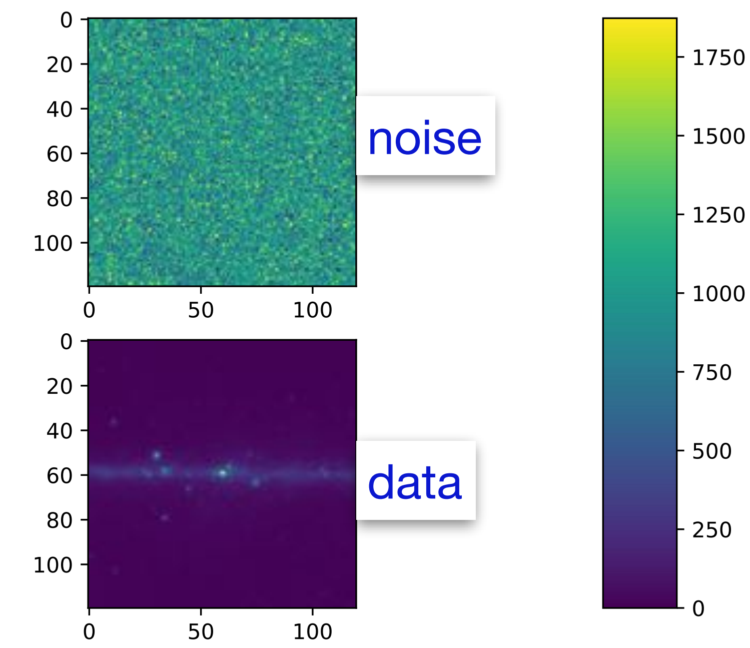
Mind the Gap — A means to quantify it

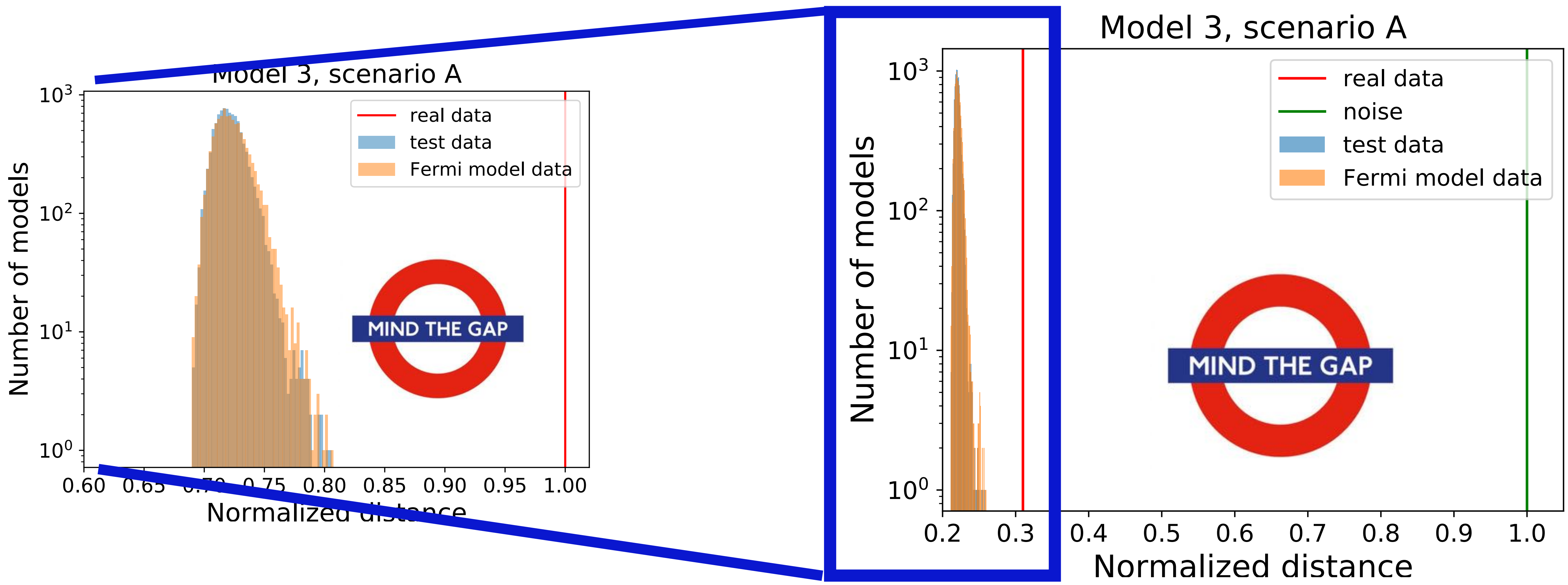
For Deep SVDD we stay in the framework of Model 3:

— A network trained on Model 3 is then shown validation data from: Model 3, the Fermi diffuse background model, real data and pure Gaussian noise.

— It quantifies the difference (*in latent space*) between what it 'knows' (its universe derived from the simulated training data) and what is shown to it.

— It answers the question: **Can this universe produce such an image?**





Conclusions

DeepEnsemble Networks are capable of recovering the background and the presence of the GCE. We found that:

- Bright components are detected robustly and consistently between our models. They are also detected consistently with the prediction from the traditional likelihood method.

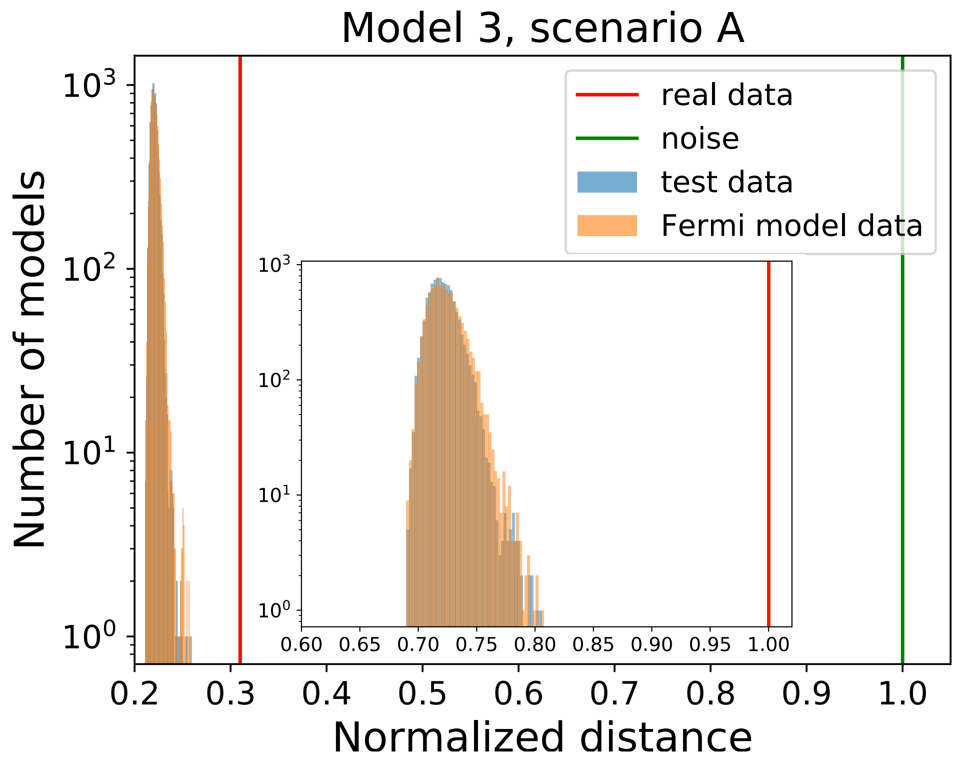
- The networks robustly detect the presence of the GCE in all our models, with the properties (flux and spatial distribution) consistent with other works.

However, the picture is not as clear as we (and everyone else!) wished:

- The nature of the GCE however, while well predicted within each model, **does not appear to be robust when networks are applied outside of their domain. We can predict anything from no DM to no MSPs by selecting a fitting background model.**

- ***Mind the gap: - the fact that reality is not part of the (background) model has been a limiting factor of many (all?) current works. What results can we trust at the moment?***

- **Deep SVDDs** offer a possibility to test severity of the reality gap. We are currently probing state-of-the-art models of the GC in this way. **Stay tuned!**



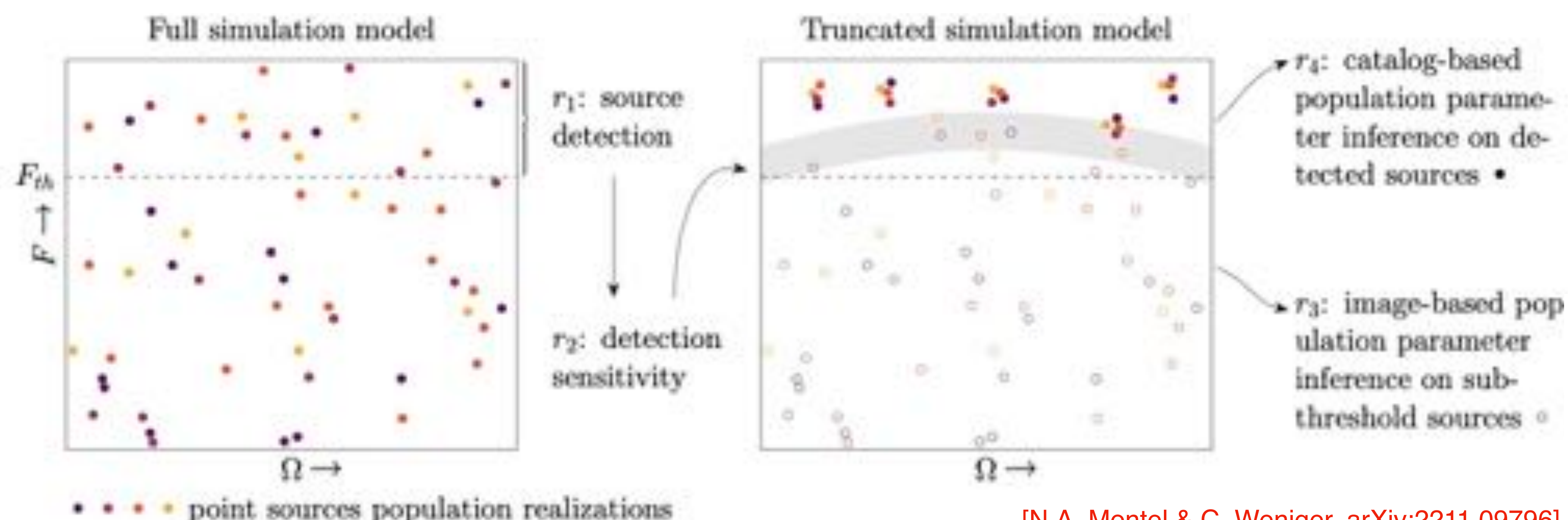
How to proceed? My humble opinion.

Truncated Marginal Neural Ratio Estimation

Avenues of future progress:

- Dedicated **searches for bulge millisecond pulsars in the X-ray and/or radio** band (MeerKAT, SKA).
- Gain understanding of the **small-scale uncertainties** of the Milky Way's **gas distribution** and update models in this regard.
- Comprehensive Bayesian study of the GCE using a **high-parametric model** taking into account the **full list of gamma-ray components and their associated uncertainties.**
  —> **Simulation-based inference** to beat the curse of dimensionality!
  (toy Fermi-LAT analysis with *swyft** seems promising)

**https://swyft.readthedocs.io/en/latest/*



[N.A. Montel & C. Weniger, arXiv:2211.09796]

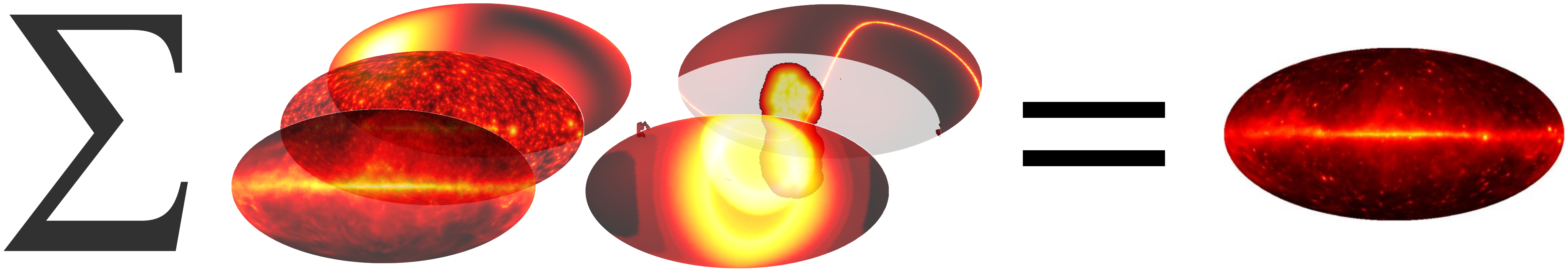
Backup slides

The essence of template-fitting

Many rigid background and signal templates spatially and spectrally binned.

Free parameters: Energy bin normalisation for each template.

$$N_{\text{params}} = N_{\text{ebins}} \times N_{\text{comp}}$$



Fit to Fermi-LAT data usually based on Poisson likelihood function:

$$\mathcal{L}(\boldsymbol{\mu} | \boldsymbol{n}) = \prod_i \frac{\mu_i^{n_i}}{(n_i)!} e^{-\mu_i}$$

Large variety of modifications to the basic approach:
— modify likelihood function (weighted Poisson likelihood to include instrumental uncertainties),
— inject non-Poissonian noise (as in NPTF),
— deform spatial morphology on the fly (Gaussian processes, etc.).

Combining templates and image reconstruction

SkyFACT accounts for systematic uncertainties by combining template fitting with methods from image reconstruction; it is a hybrid approach.

observed gamma-ray photons in a particular pixel

Spatial template **Spectral template**

$$\phi_{pb} = \sum_k T_p^{(k)} \tau_p^{(k)} \cdot S_b^{(k)} \sigma_b^{(k)} \cdot \nu^{(k)}$$

Nuisance parameters

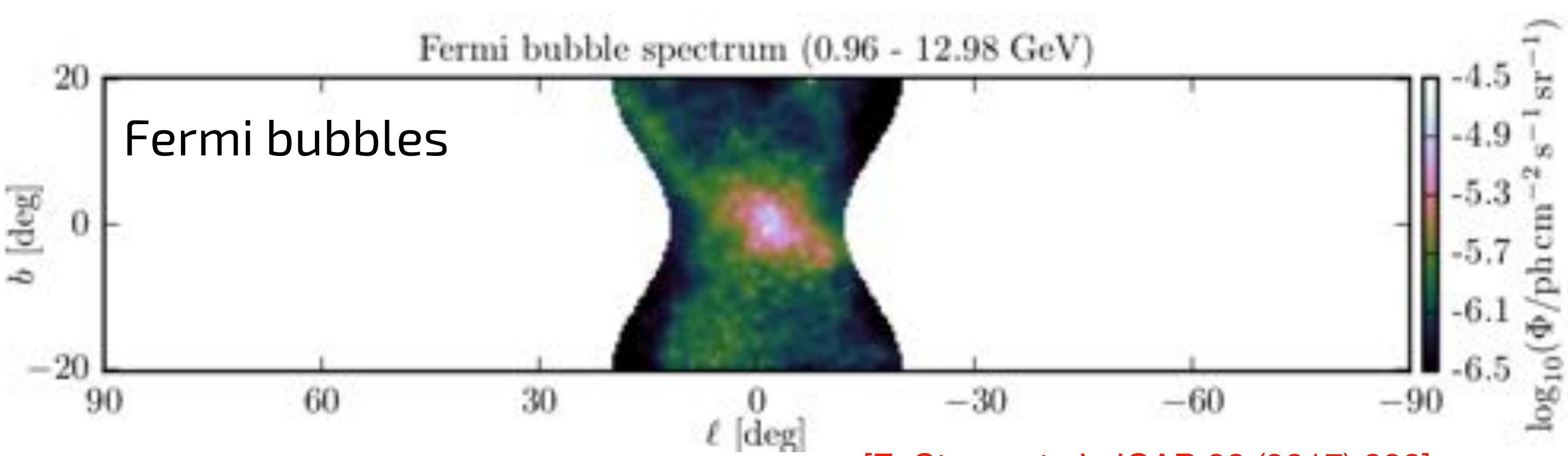
Idea:
- Each pixel can be individually varied according to the nuisance parameters.
- Range of the nuisance parameters is controlled by hyper parameters given by the user.
- The problem is constrained to stay within the uncertainty of the employed components.

Likelihood function:

$$\ln \mathcal{L} = \ln \mathcal{L}_{\mathrm{P}} + \ln \mathcal{L}_{\mathrm{R}}$$

Regularization of nuisance parameters

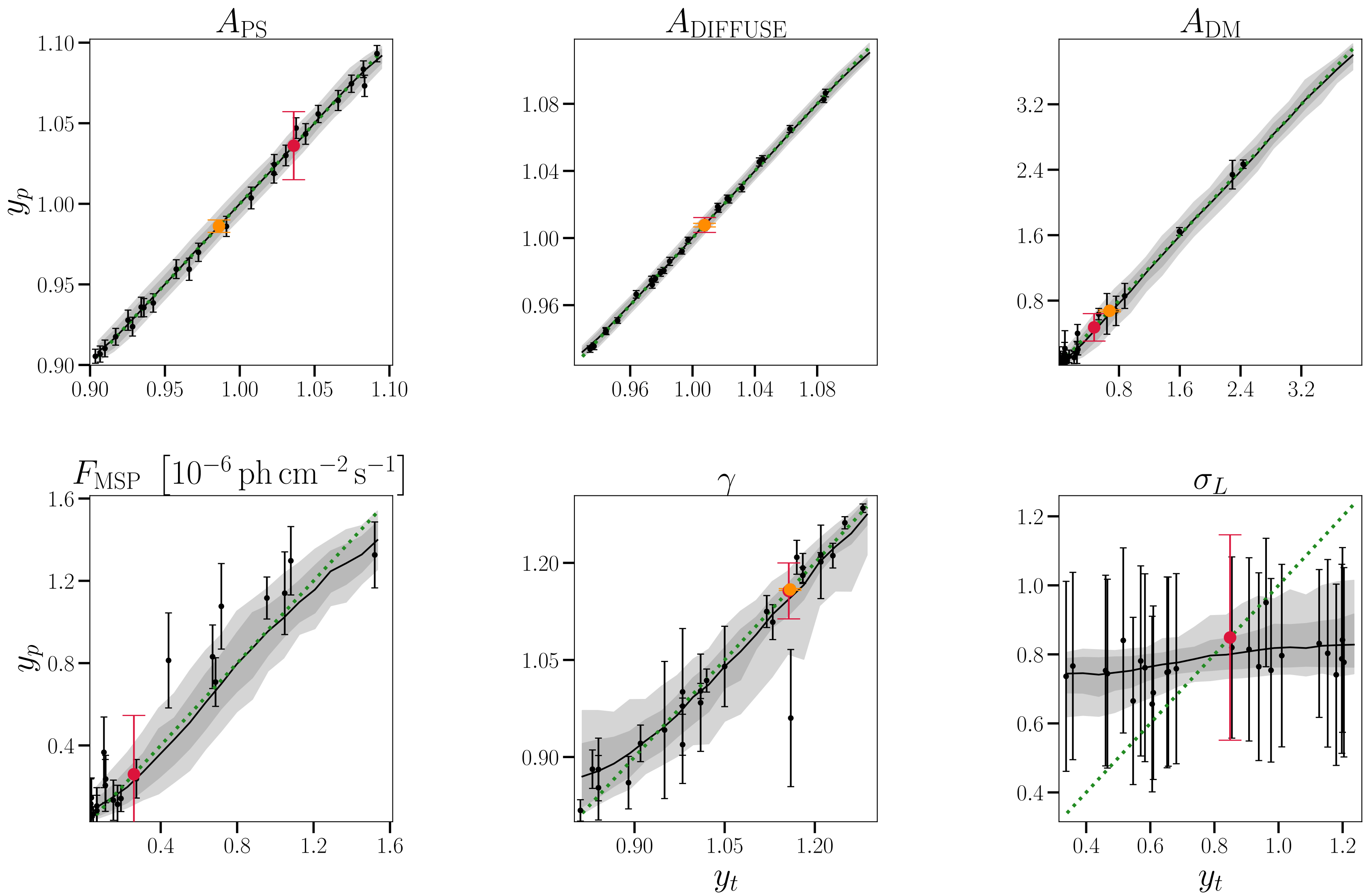
$$-2\ln \mathcal{L}_R = \sum_k \lambda_k \mathcal{R}_X(\boldsymbol{\tau}^{(k)}) + \lambda'_k \mathcal{R}_X(\boldsymbol{\sigma}^{(k)}) + \lambda''_k \mathcal{R}_X(\nu^{(k)}) + \eta_k \mathcal{S}_1(\boldsymbol{\tau}^{(k)}) + \eta'_k \mathcal{S}_2(\boldsymbol{\sigma}^{(k)})$$
$$+ \sum_s \lambda'_s \mathcal{R}_X(\boldsymbol{\sigma}^{(s)}) + \lambda''_s \mathcal{R}_X(\nu^{(s)}) + \eta'_s \mathcal{S}_2(\boldsymbol{\sigma}^{(s)}) \ ,$$

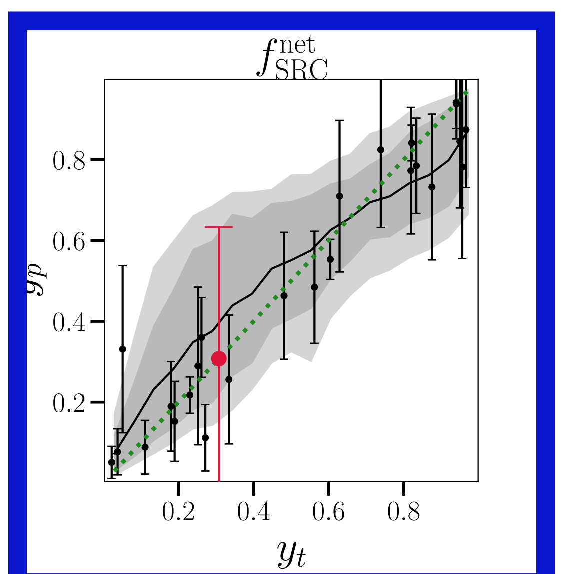


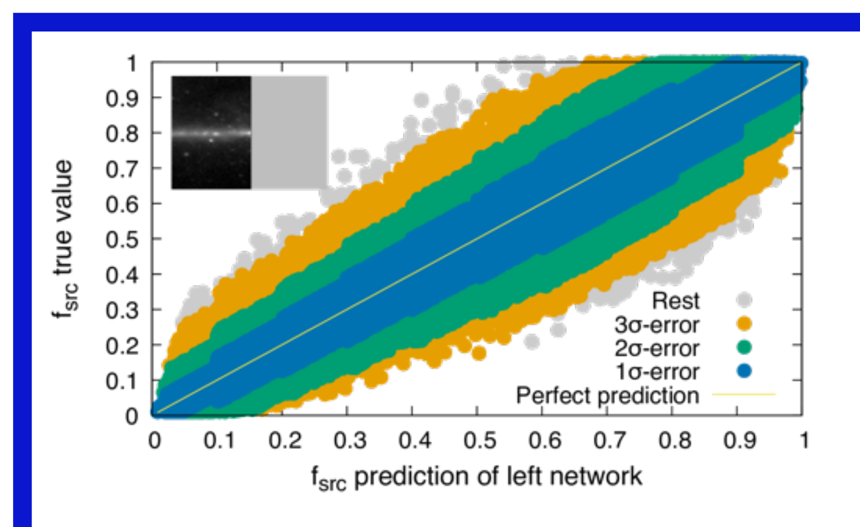
[E. Storm et al. JCAP 08 (2017) 022]

Notes:
- Typically > 10^5 parameters
- Problem is typically convex —> only one minimum
- Minimisation via L-BFGS-B algorithm

Comparable error estimates.

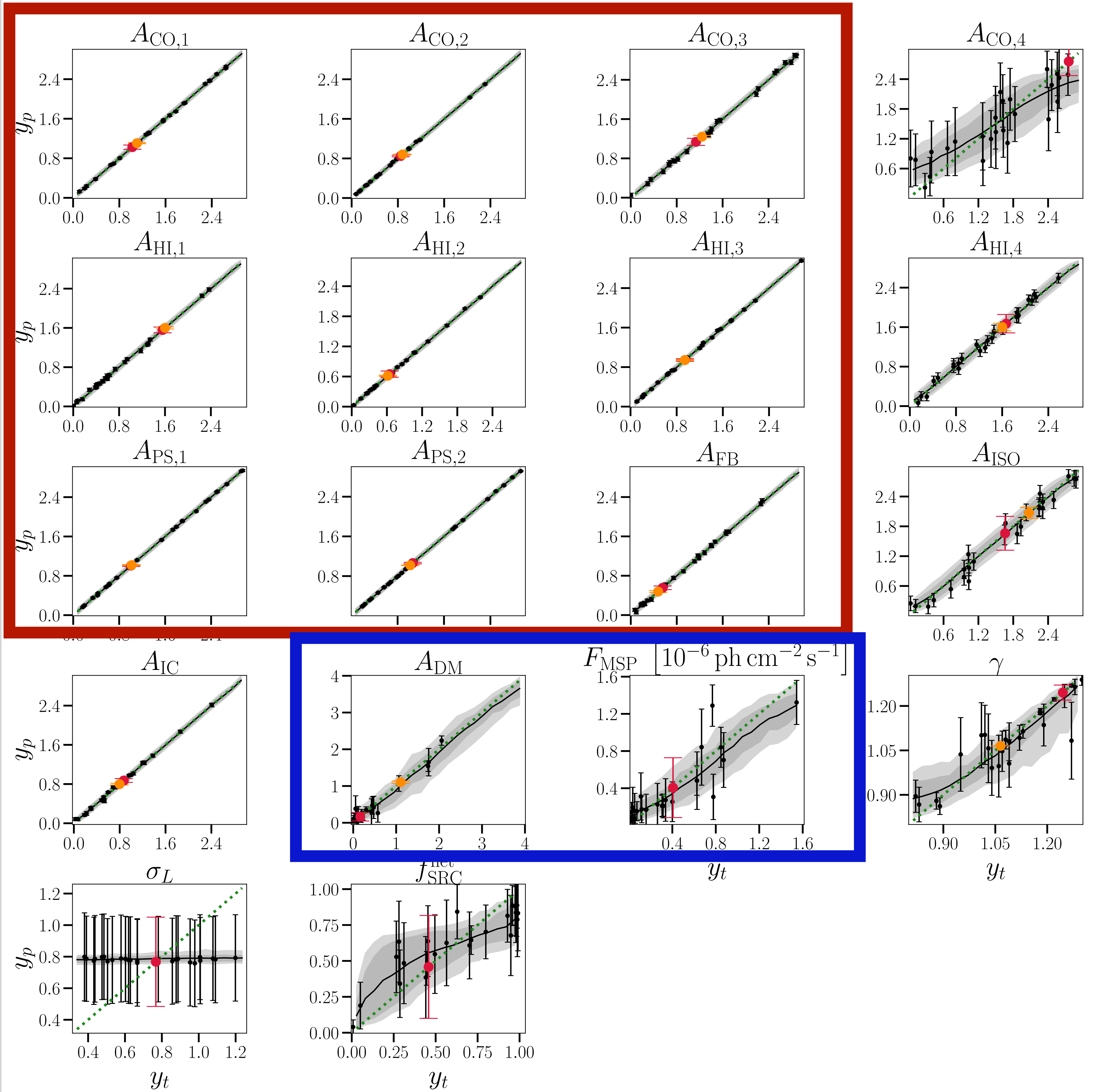






Results for a more complex model (3)

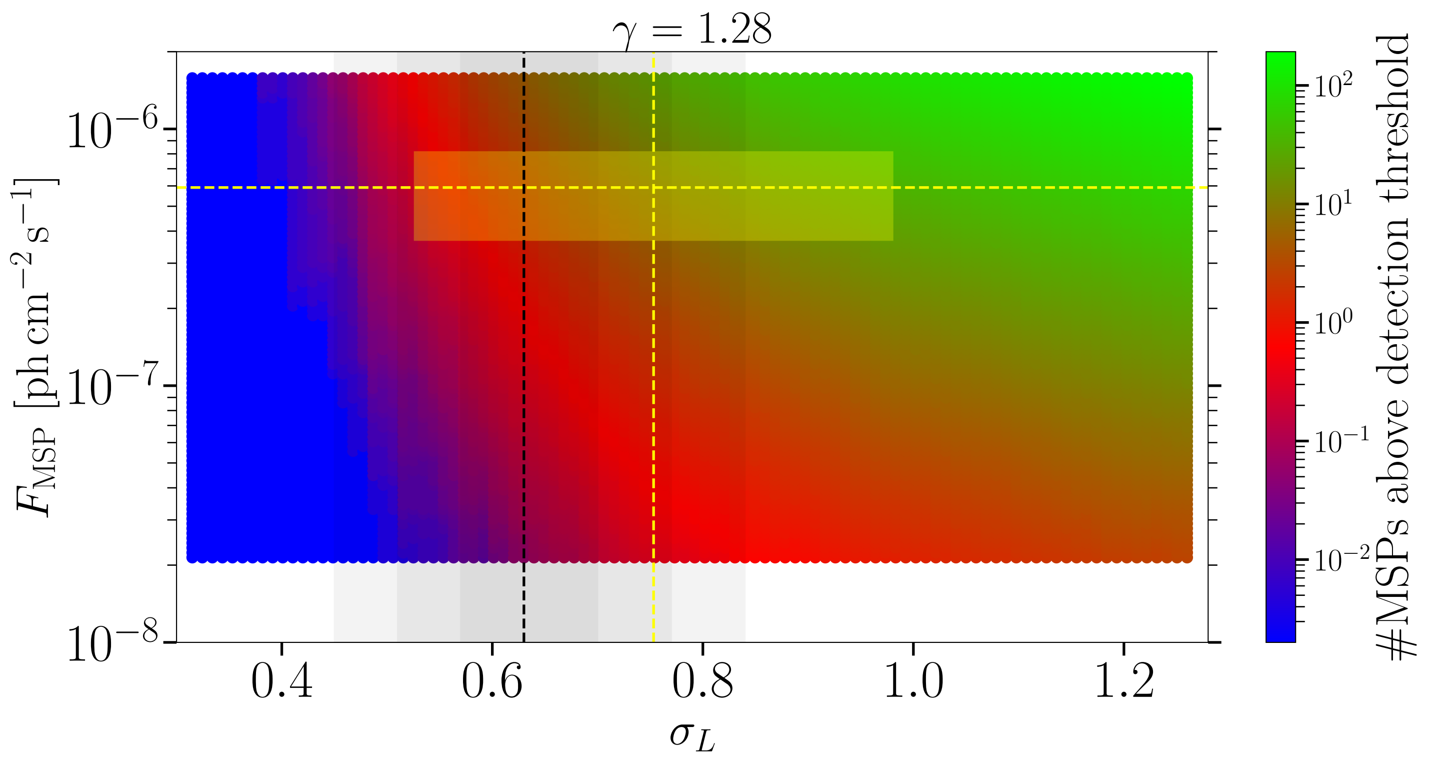
On simulated data the previous findings also apply to the most complex model we devised.

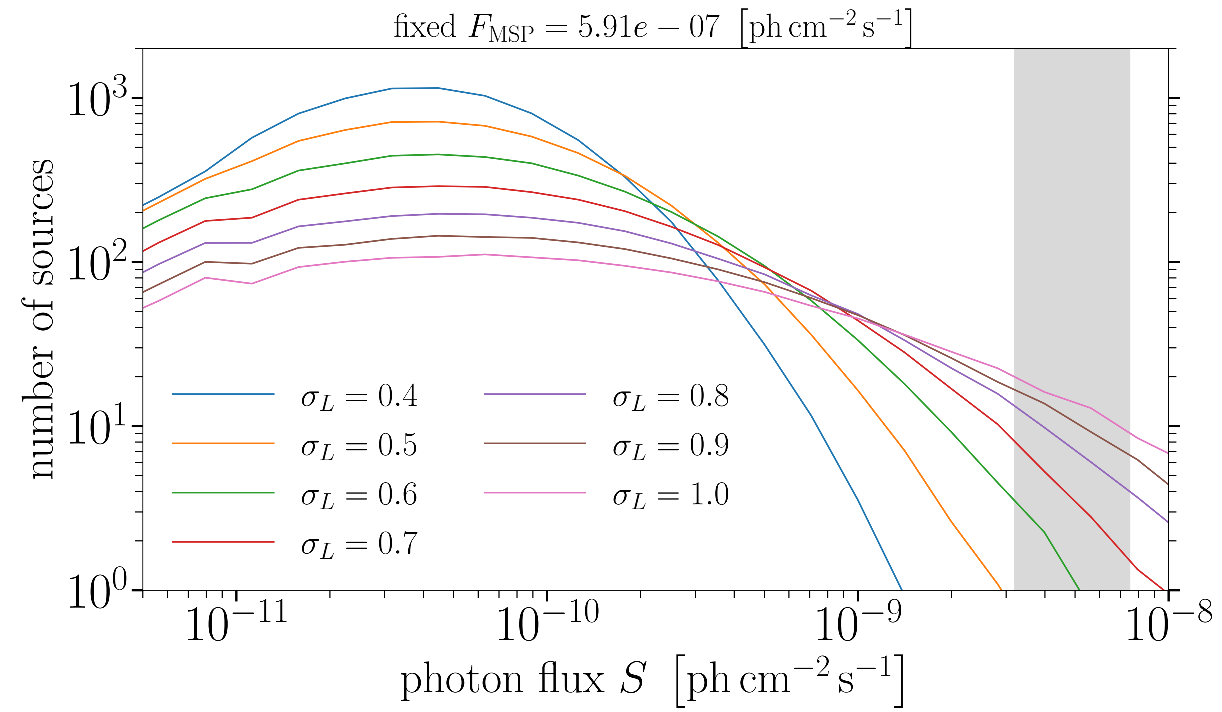


The detected & unresolved MSP population

The priors on our MSP luminosity Function covers two distinct regimes and anything in-between:

1. Entirely unresolved MSPs
2. Large number of MSPs above the Fermi-LAT detection threshold.



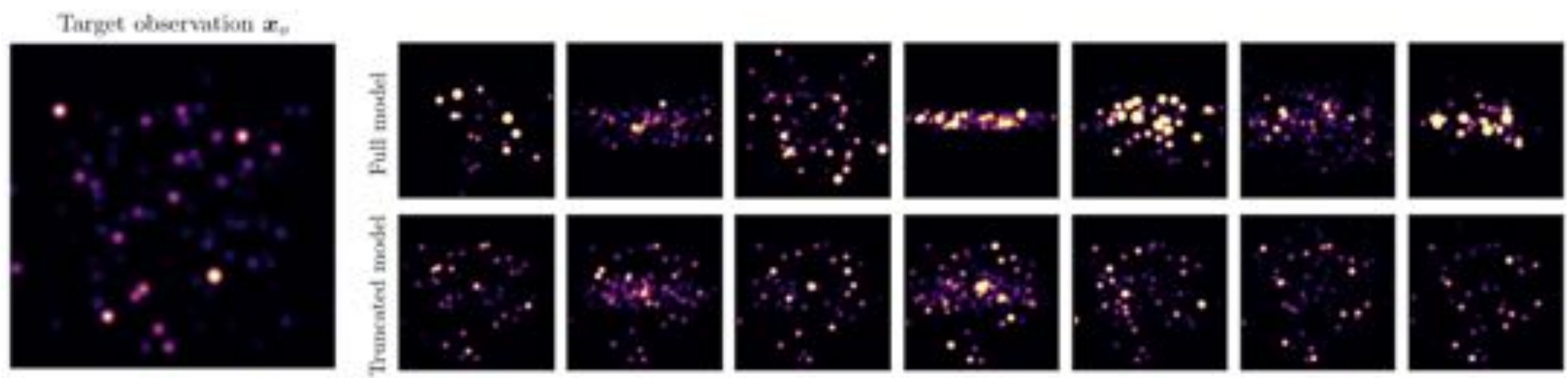


Detection is truncation: Inferring source population parameters

Constraining population parameters

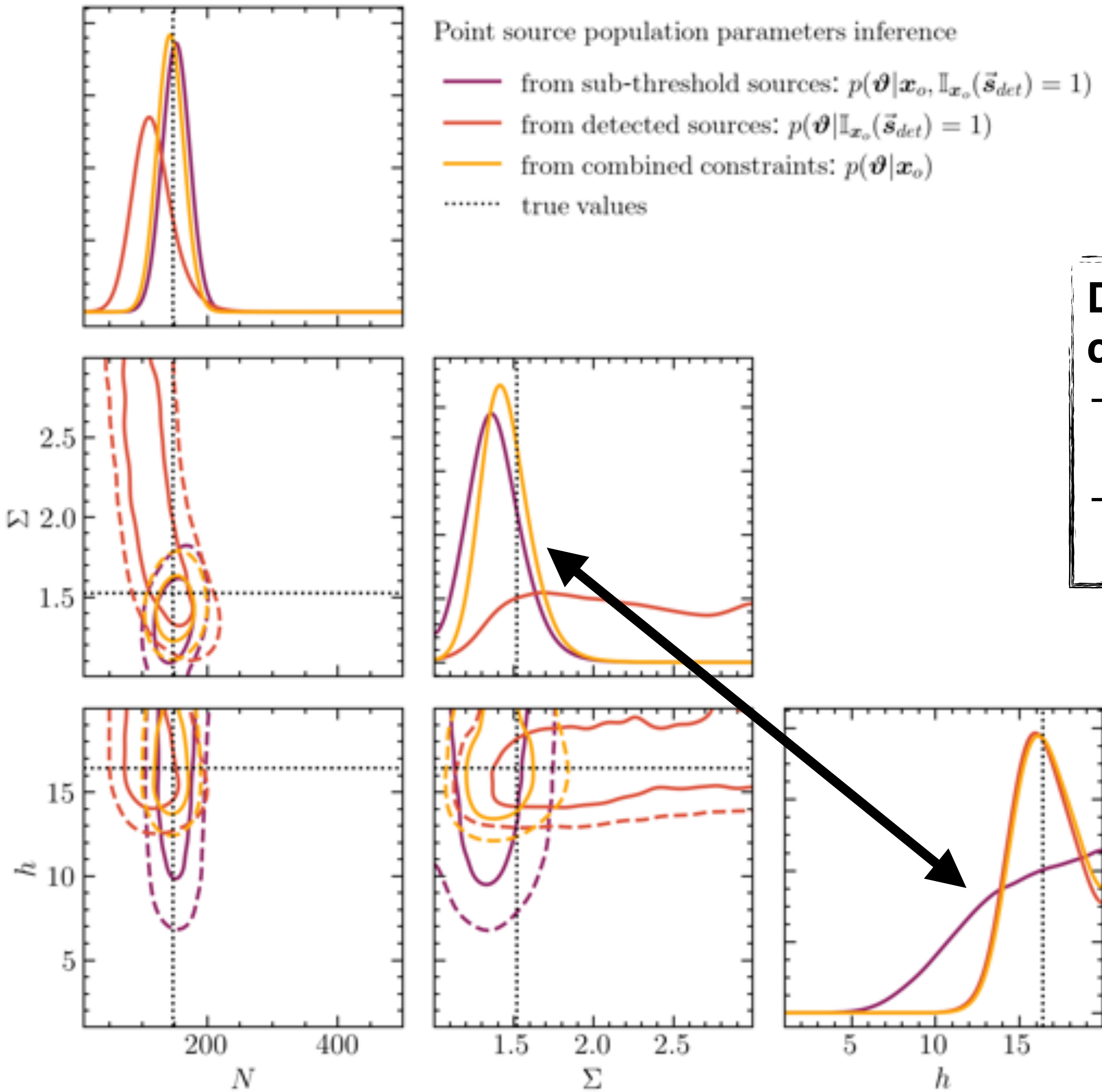
1. Run ratio estimator r_1 to obtain the detected sources for the observed sky x_0.
2. Now label these detected sources $(1 \ldots N_{\mathrm{det}})$ since we know they must be there.
3. Truncate their priors such that: $1 = \mathbb{1}_{x_0}(\vec{s}_{\mathrm{det}}) = \prod_{i=1}^{N_{\mathrm{det}}} \mathbb{1}_{x_0}\big(\Omega_i \in \mathcal{R}_i\big)\, \mathbb{1}_{x_0}\big(F_i \geq F_{\mathrm{th}}\big)$

   where $\mathcal{R}_i$ is a compact region around the source where $r_1 > 5$.
4. Define a truncated model simulator that varies sub-threshold sources as before but only varies the properties of detected sources according to the truncated priors.



Detection is truncation: Inferring source population parameters

Performance on toy simulations:



Detected and sub-threshold sources complement each other:
— luminosity fct. accessible via sub-threshold sources
— spatial distribution more constrained by detected sources