Leveraging neural simulation-based inference for astrophysical dark matter searches

Based on:

J. Brehmer*, SM*, J. Hermans, G. Louppe, K. Cranmer [ApJ, 1909.02005]
SM, K. Cranmer [PRD, 2110.06931]

Siddharth Mishra-Sharma

NSF Institute for Artificial Intelligence and Fundamental Interactions

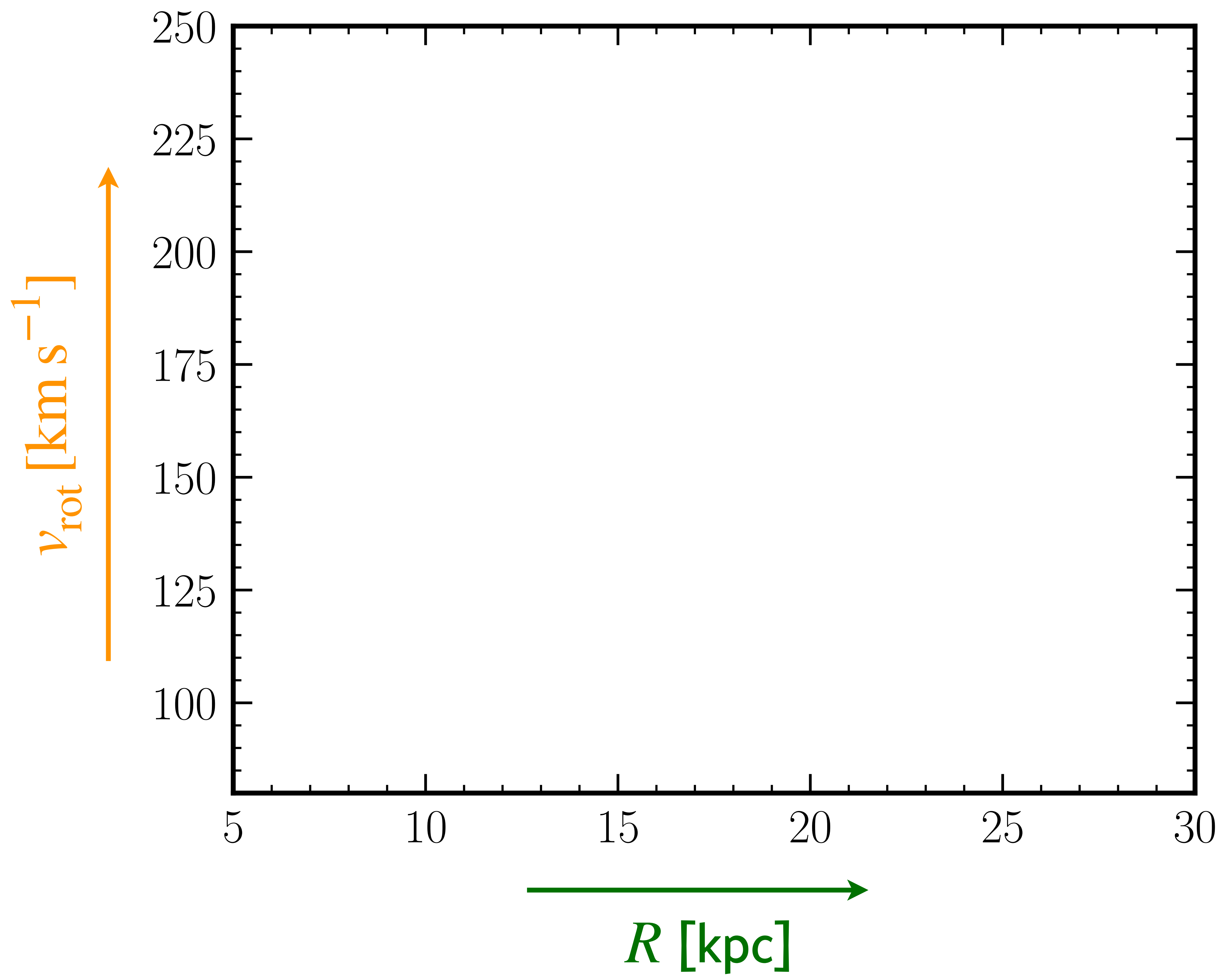
CHASC Astrostatistics Seminar

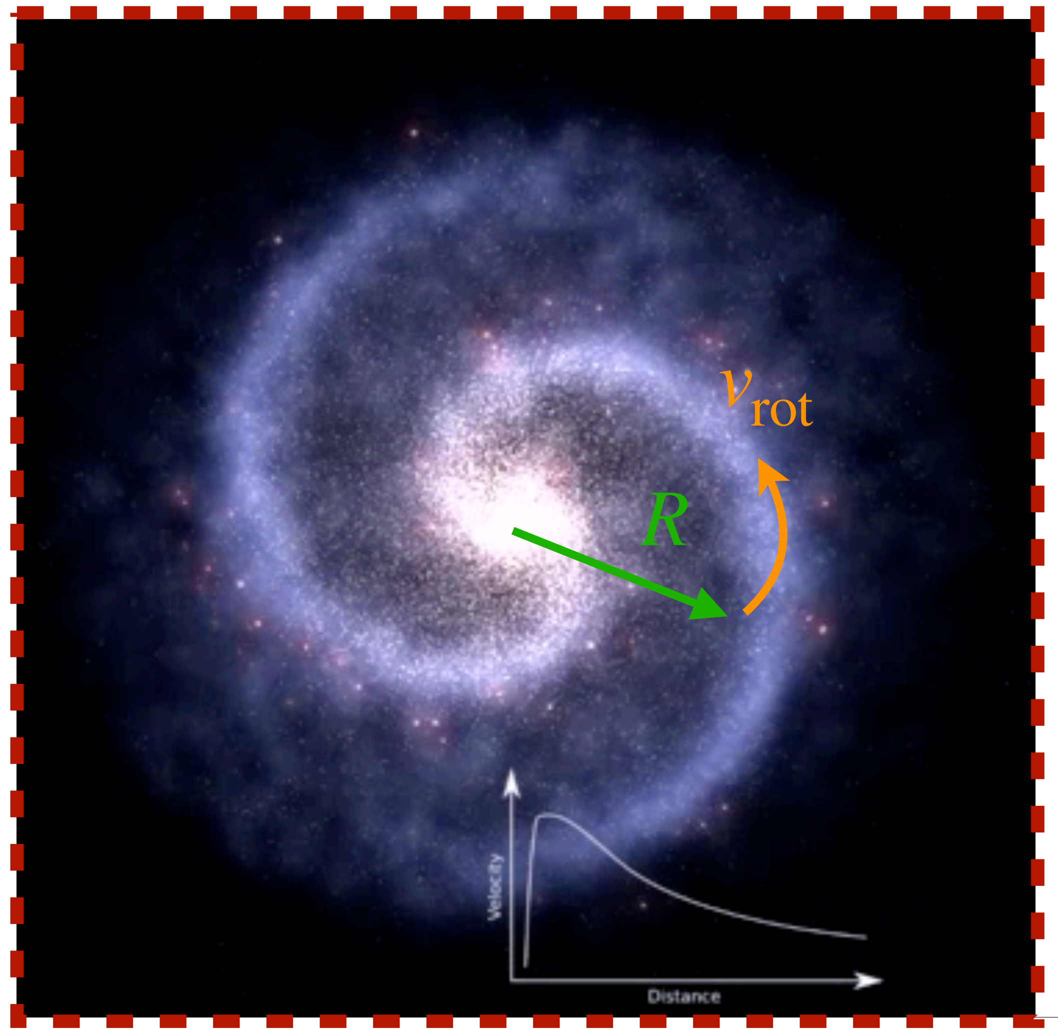
April 5, 2022

Evidence for dark matter

$$v_{\rm rot}(r) = \sqrt{\frac{GM_{\rm enc}}{r}}$$

Galactic rotation curves



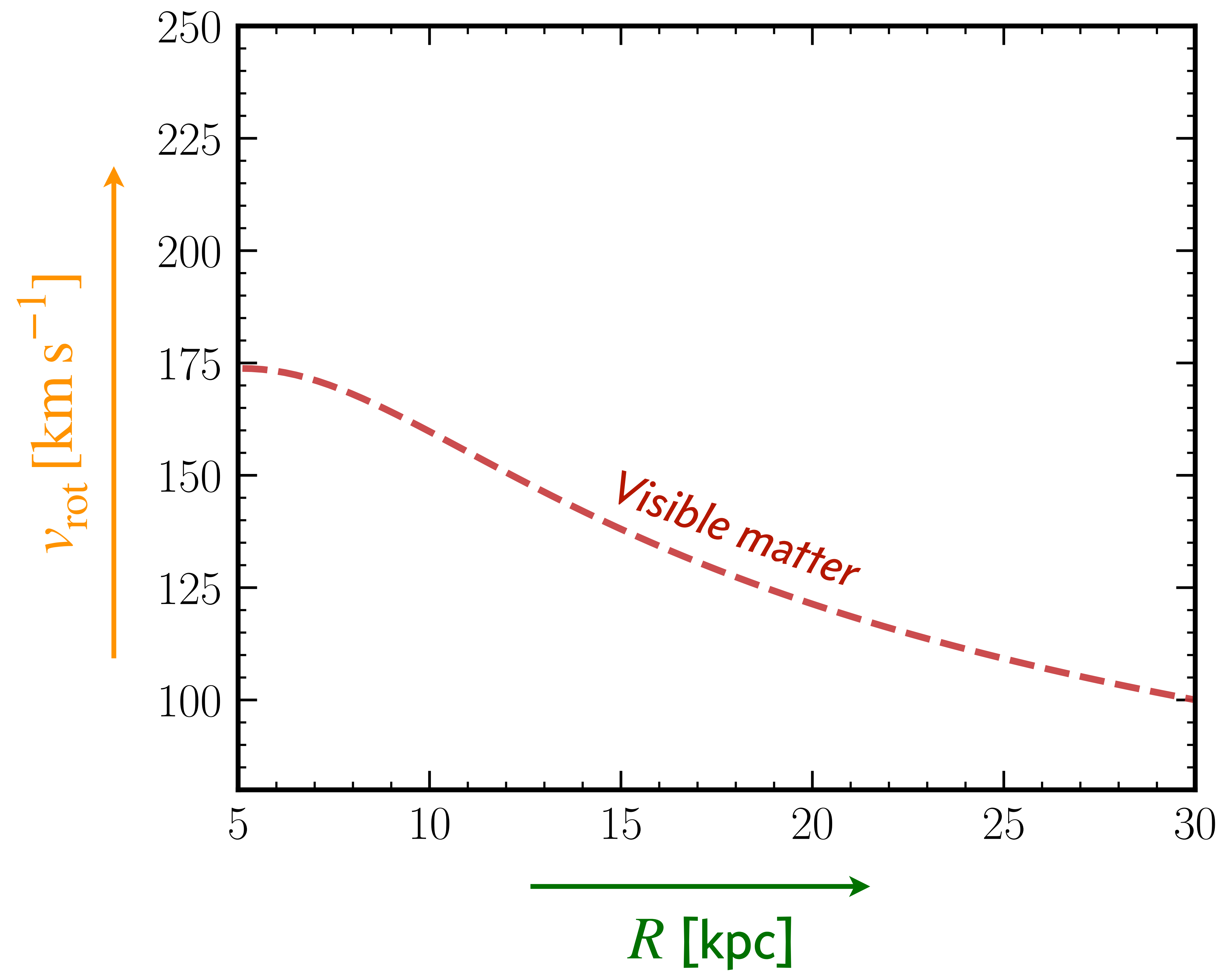


https://beltoforion.de/en/spiral_galaxy_renderer/

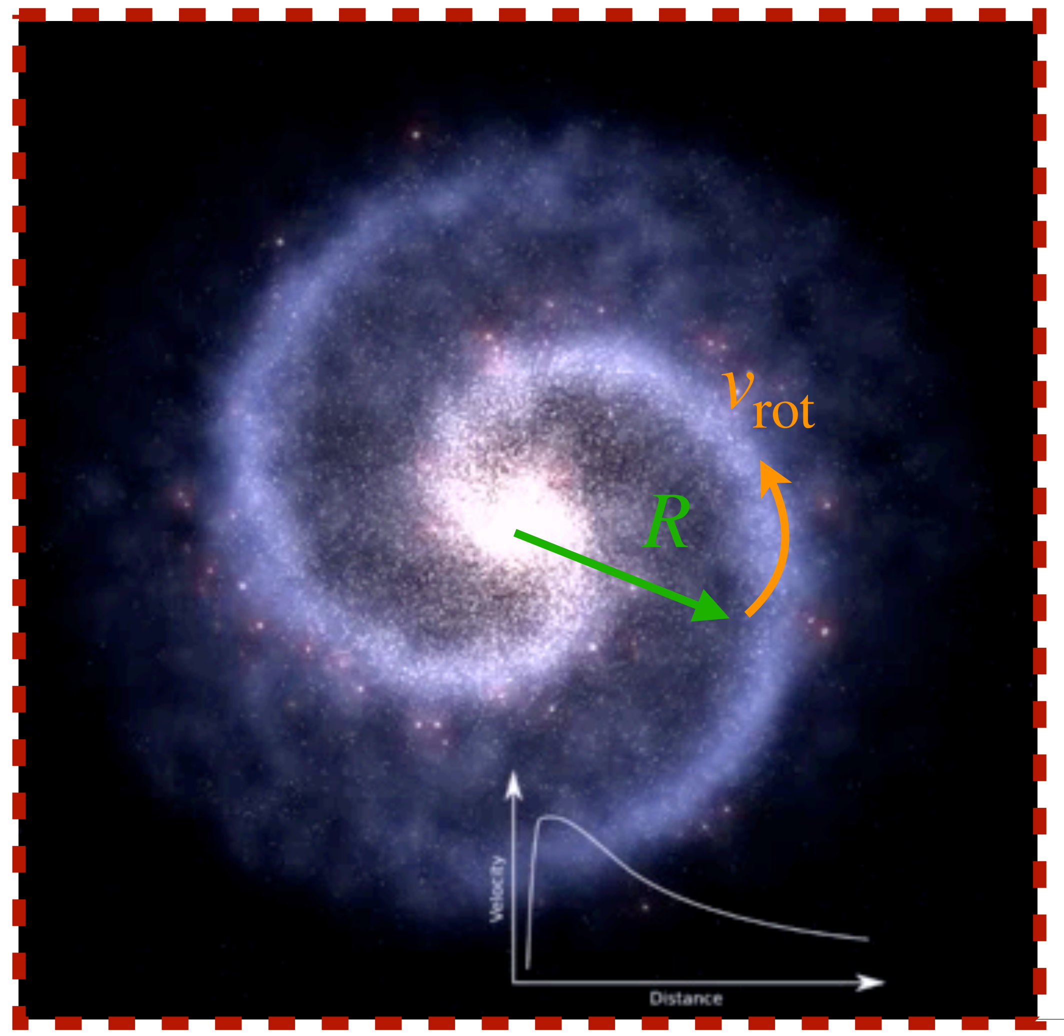
Evidence for dark matter

$$v_{\rm rot}(r) = \sqrt{\frac{GM_{\rm enc}}{r}}$$

Galactic rotation curves



Only visible matter

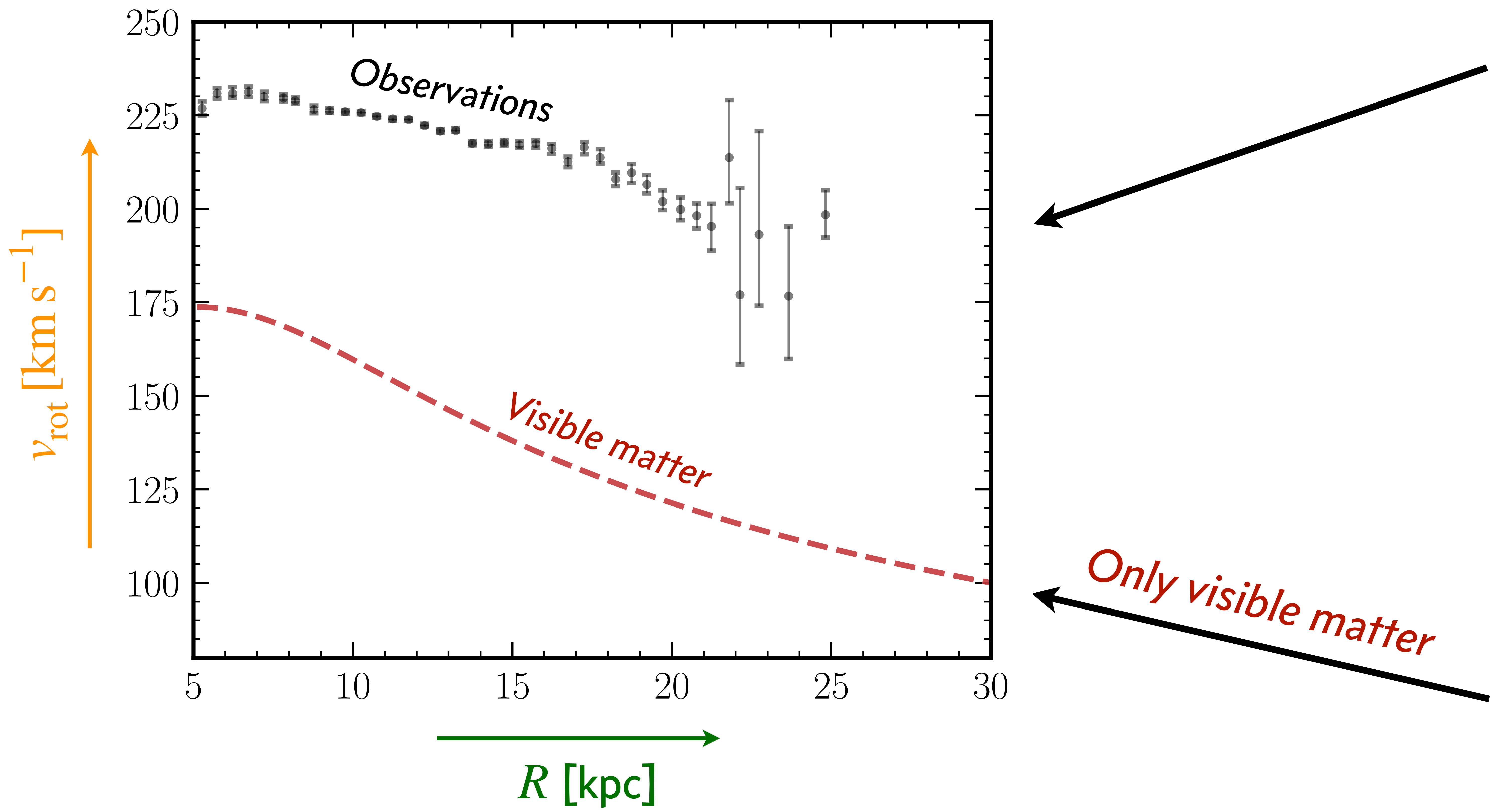


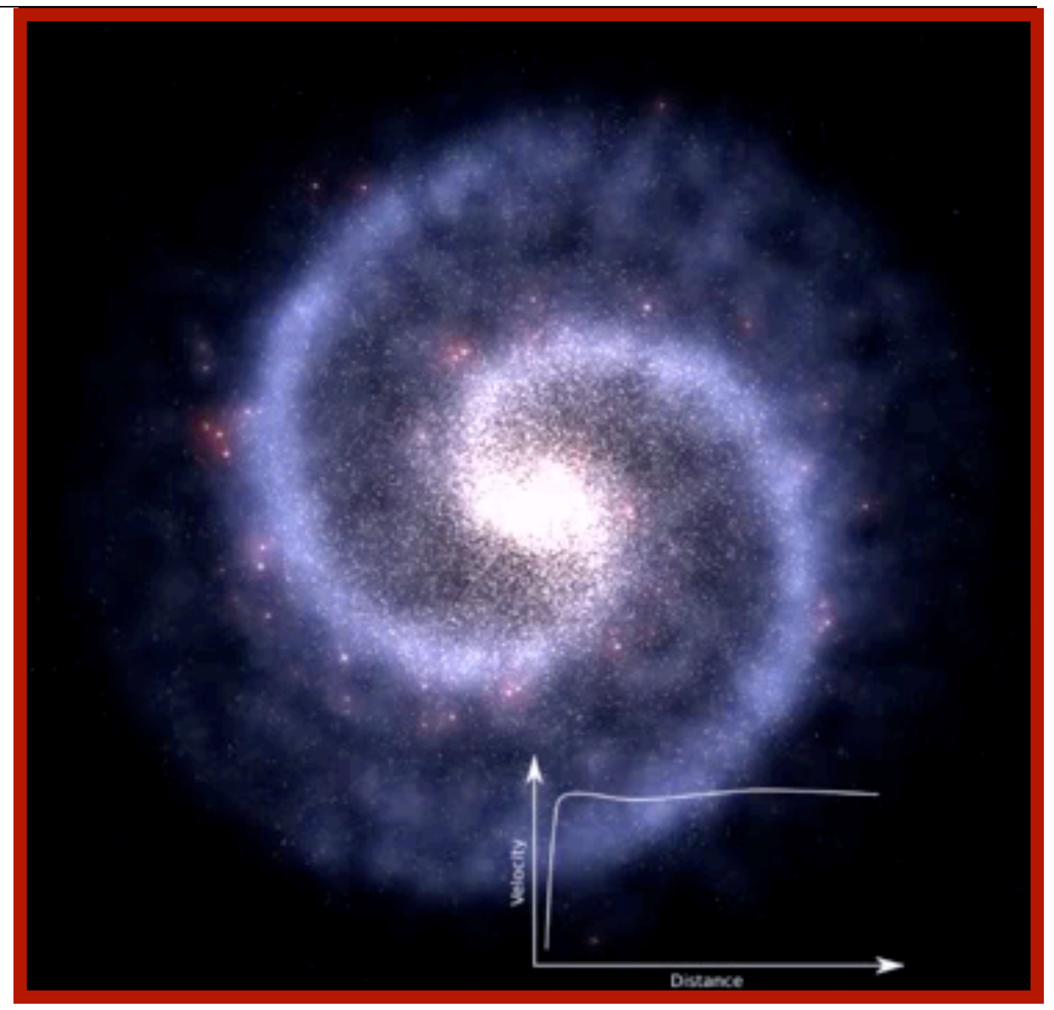
https://beltoforion.de/en/spiral_galaxy_renderer/

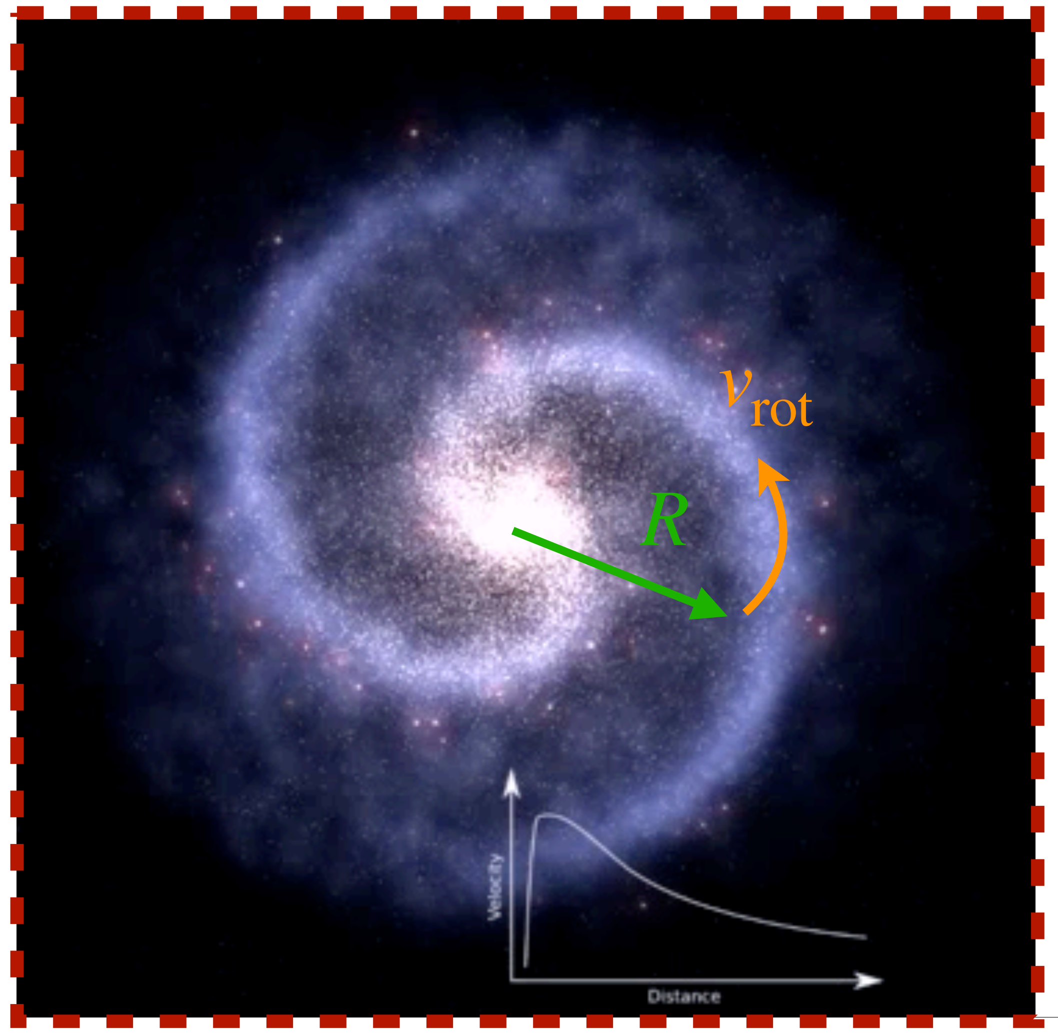
Evidence for dark matter

Galactic rotation curves

$$v_{\mathrm{rot}}(r) = \sqrt{\frac{GM_{\mathrm{enc}}}{r}}$$





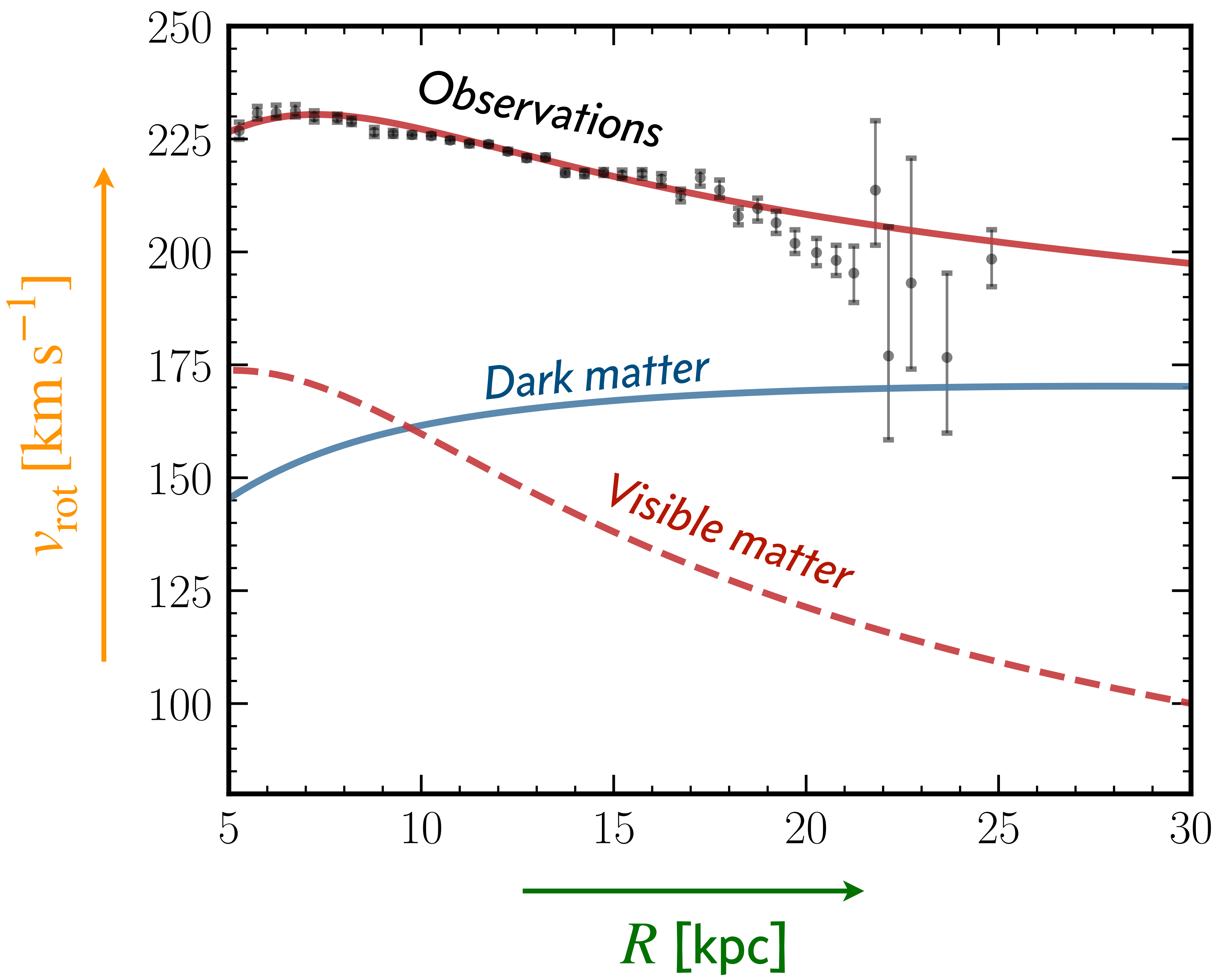


https://beltoforion.de/en/spiral_galaxy_renderer/

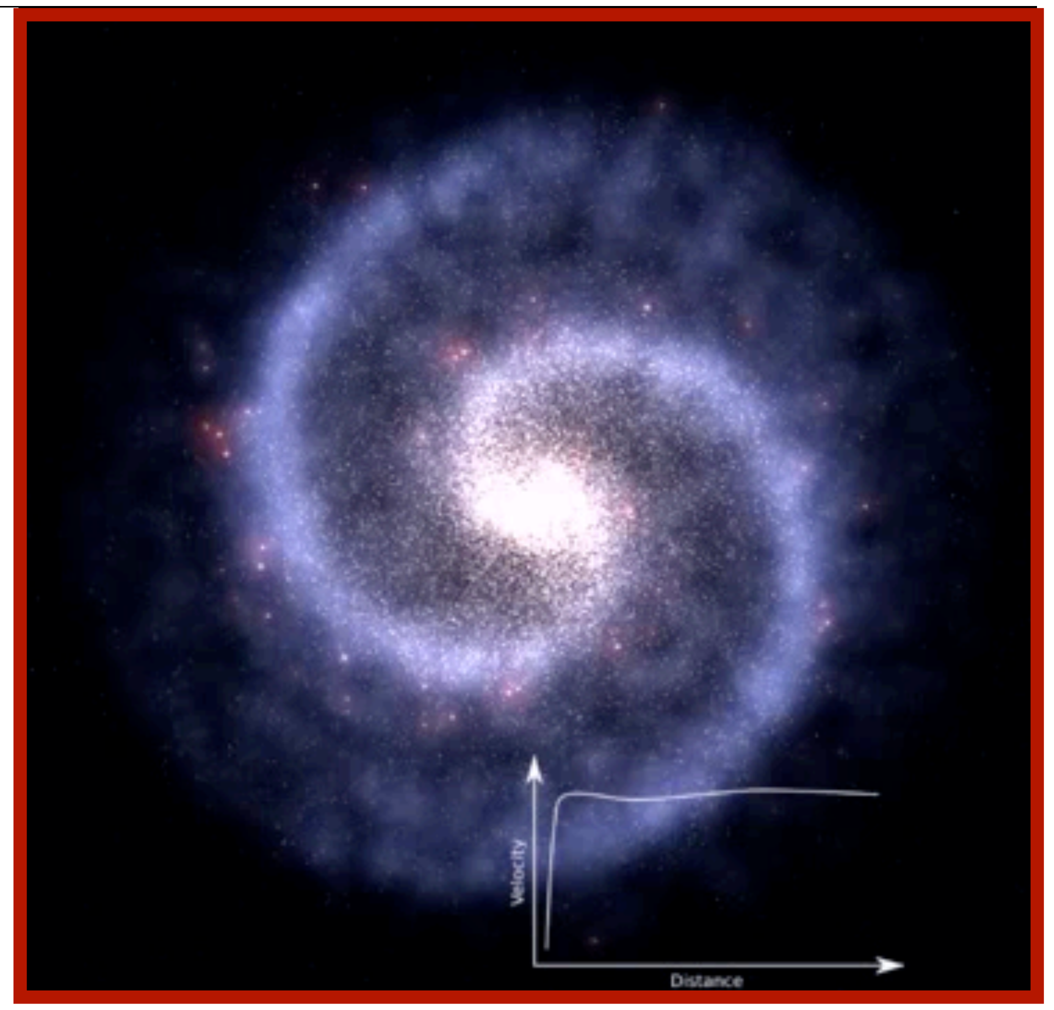
Evidence for dark matter

$$v_{\rm rot}(r) = \sqrt{\frac{GM_{\rm enc}}{r}}$$

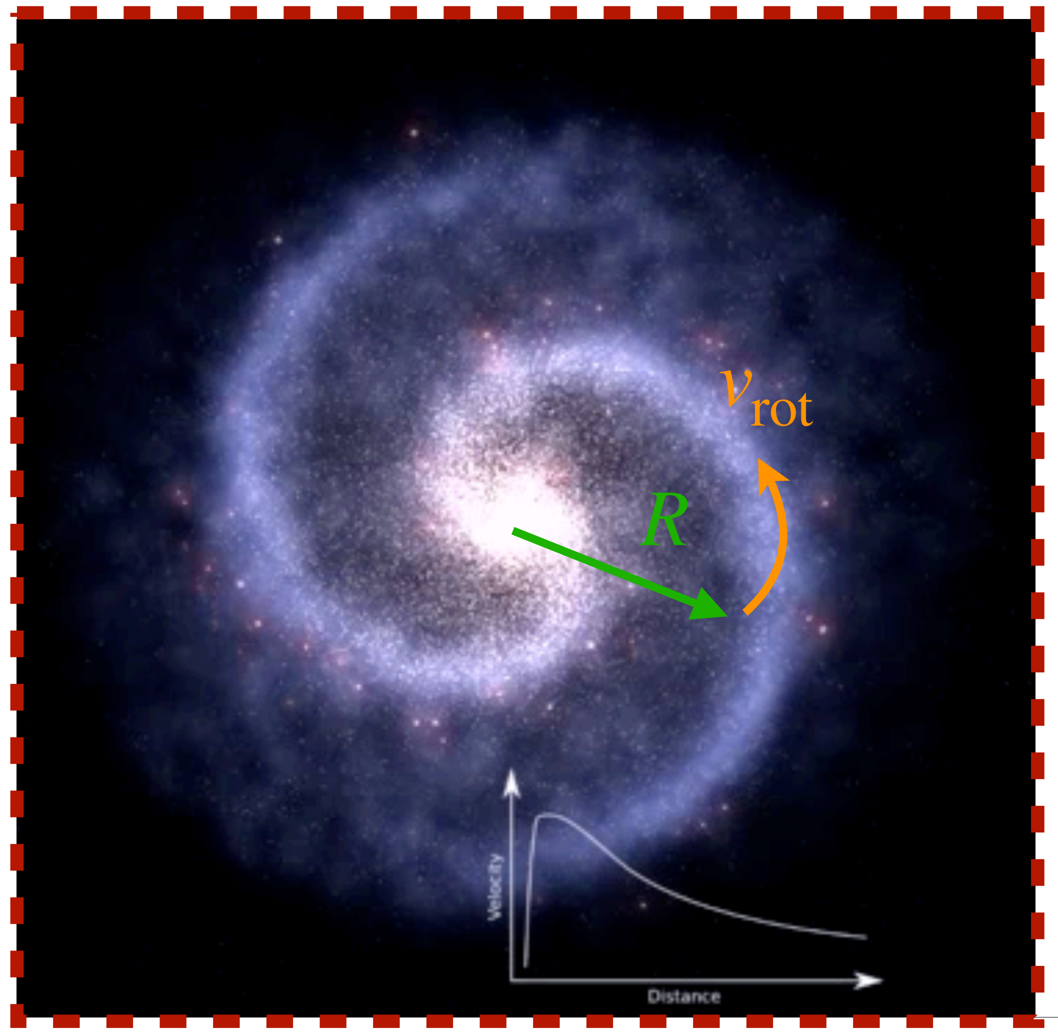
Galactic rotation curves



With dark matter



Only visible matter

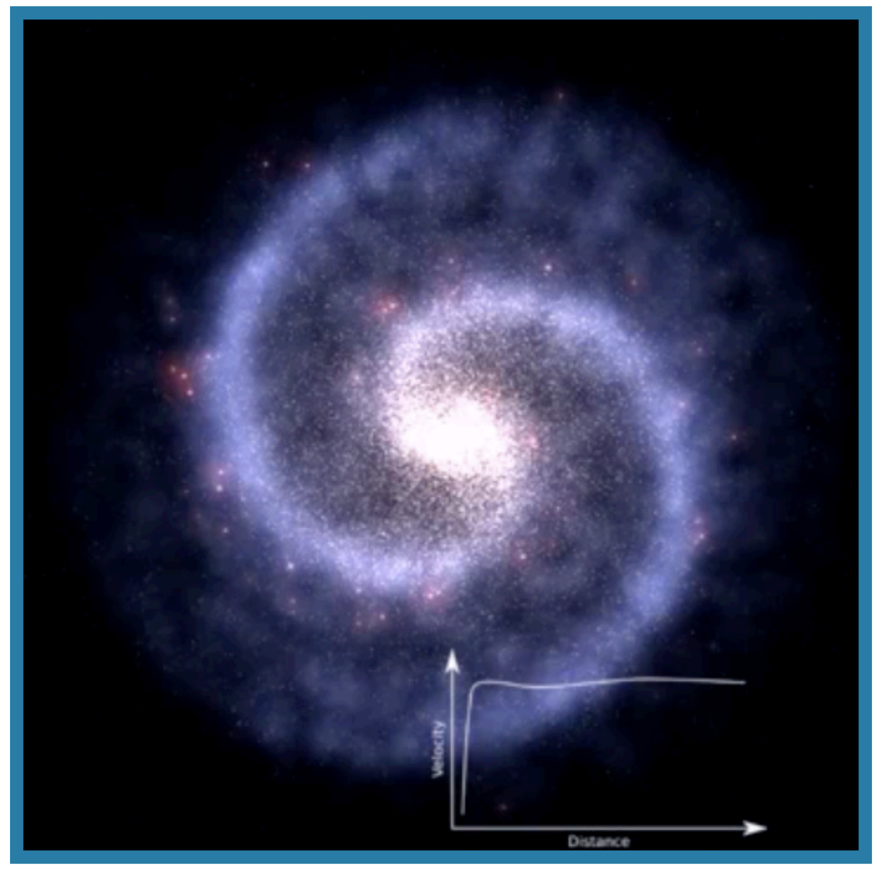


https://beltoforion.de/en/spiral_galaxy_renderer/

Evidence for dark matter…

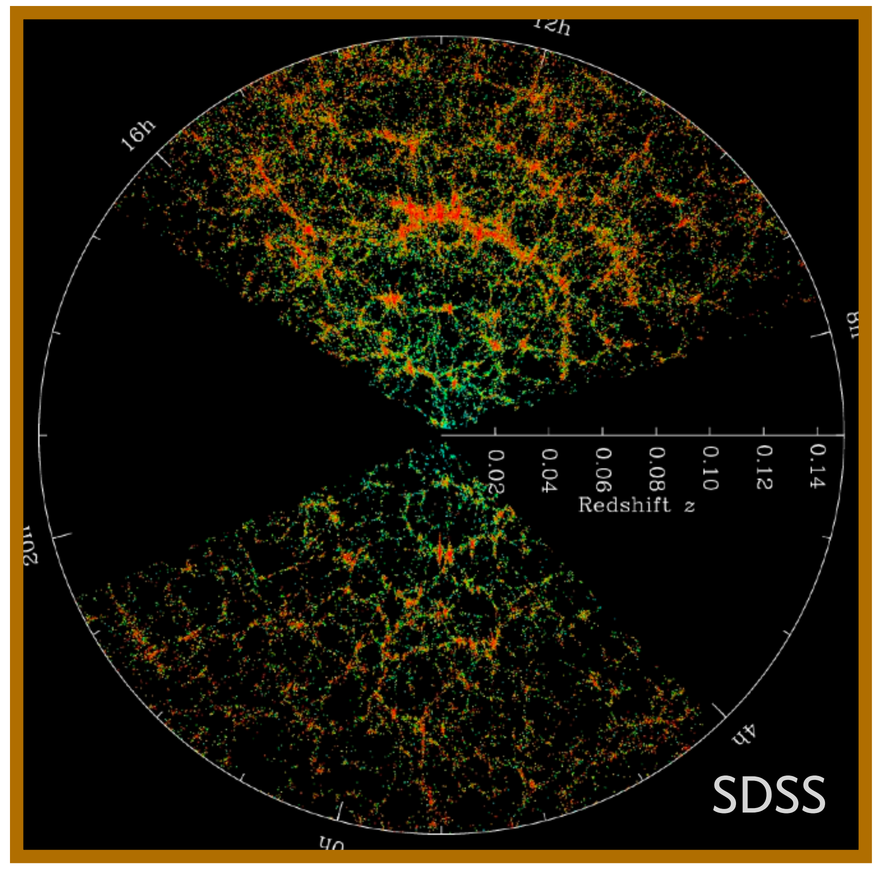
…exists over a diversity of scales and physical systems

Galaxy rotation curves

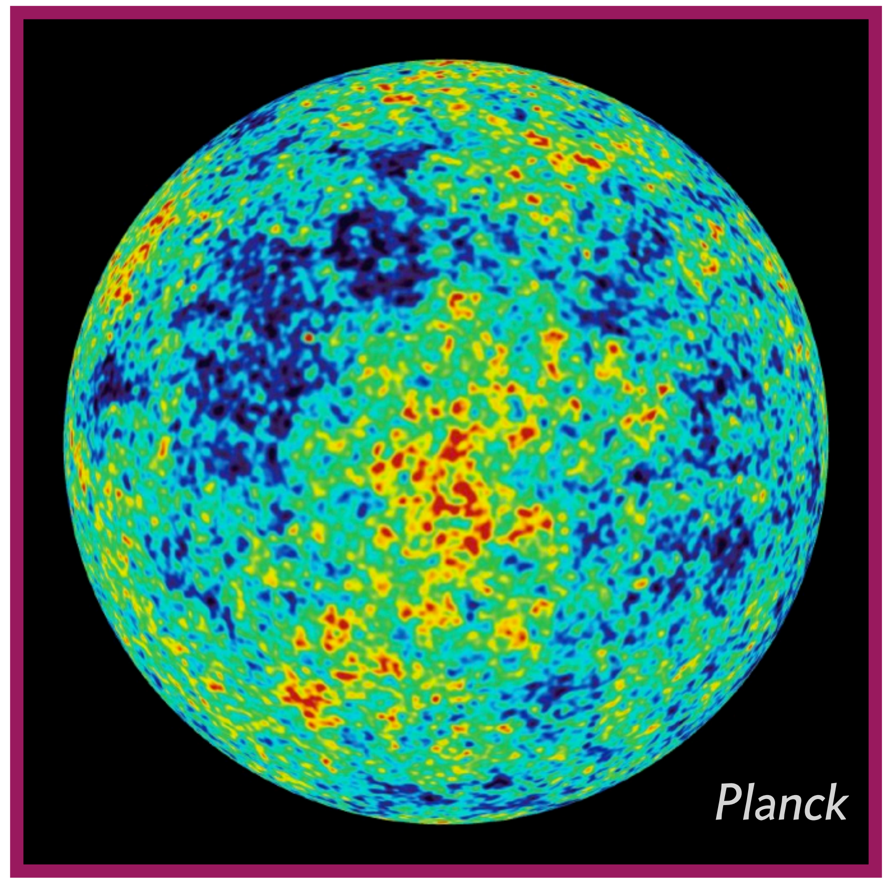


Galaxy clusters

Large-scale structure

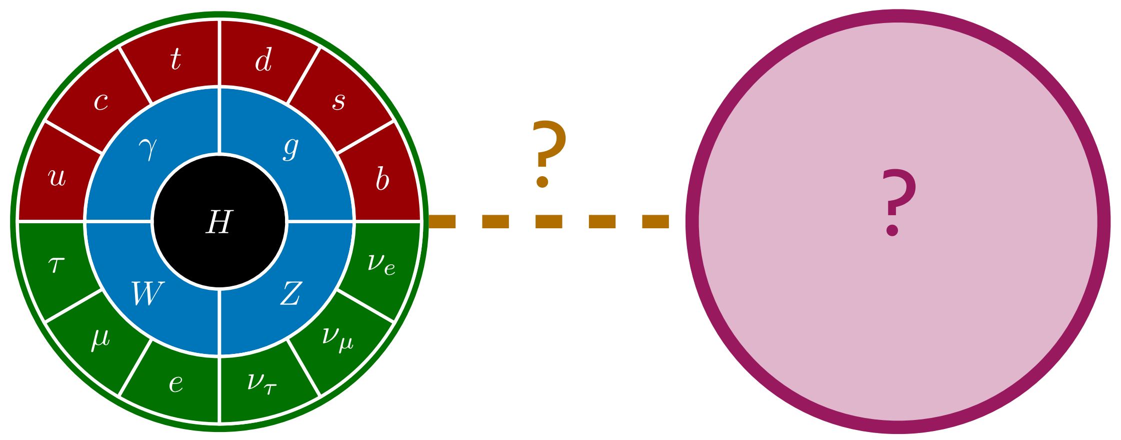


Cosmic Microwave Background

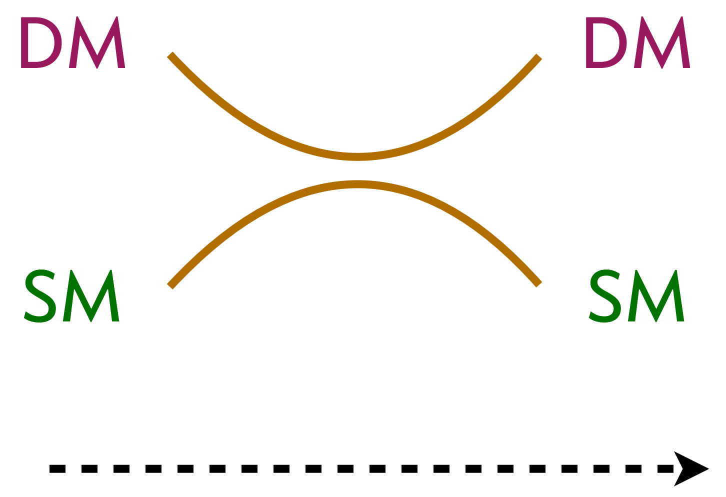


Increasing scale

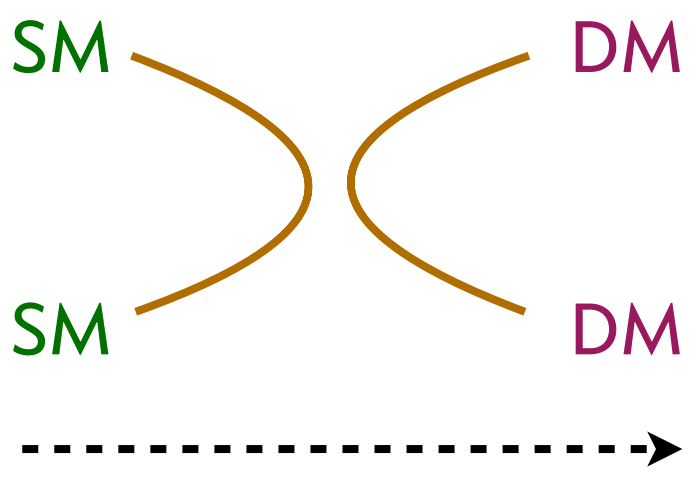
Searches for dark matter



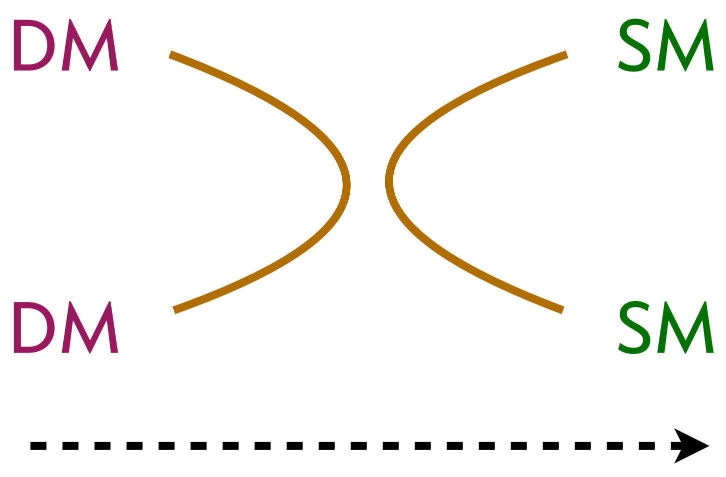
Scattering of DM against SM



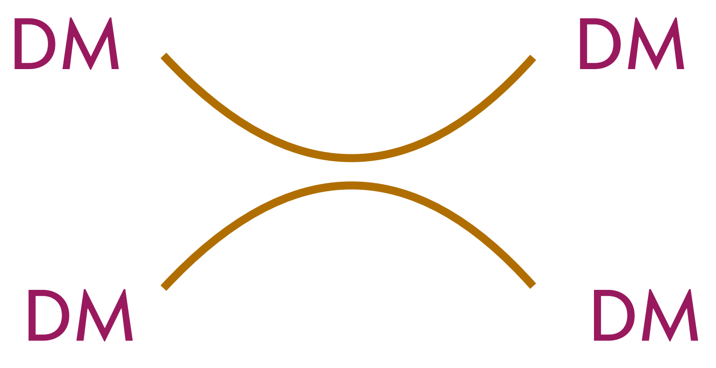
Production of DM from SM



Production of SM from DM

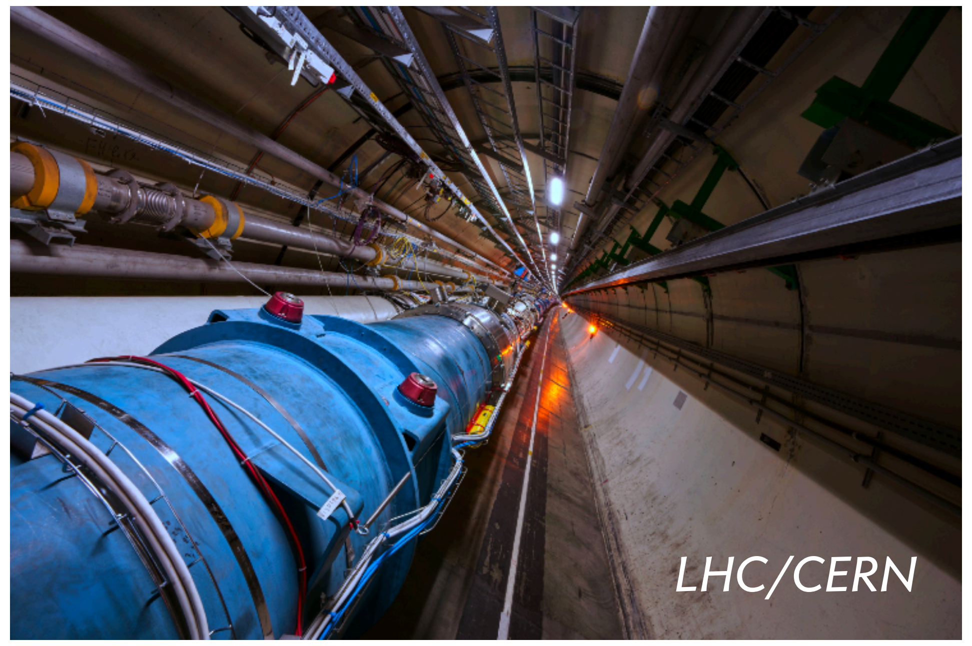


Gravitational effects of DM interaction

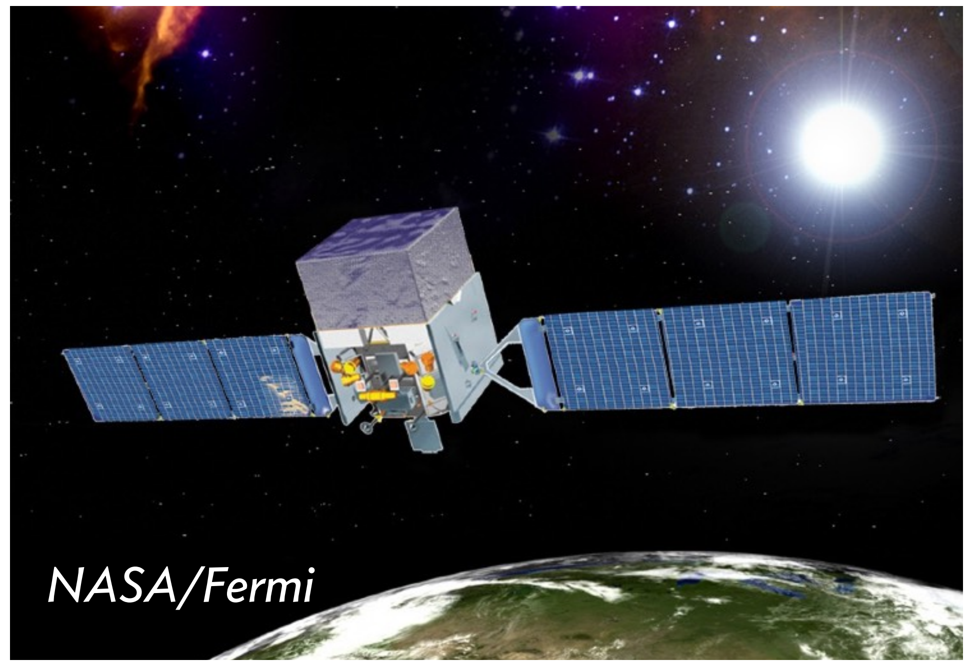


Direct detection

Particle colliders

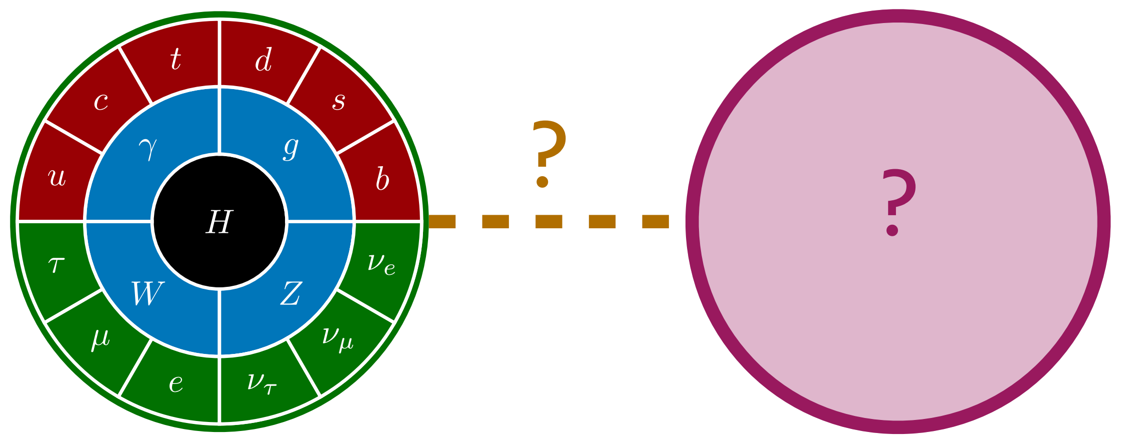


Indirect detection

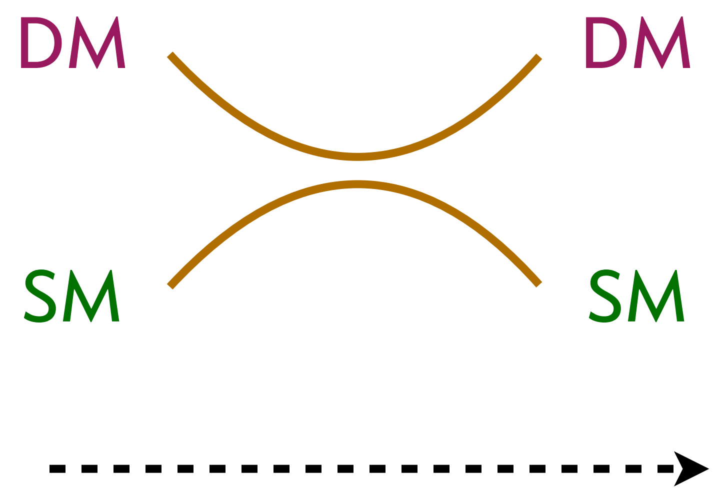


Astrophysical probes

Searches for dark matter

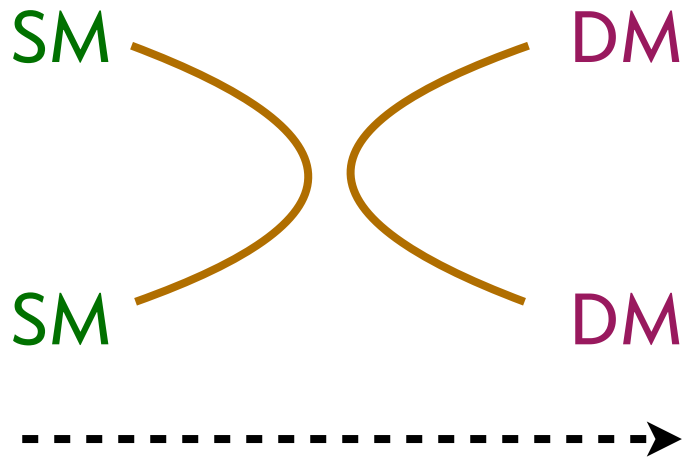


Scattering of DM against SM

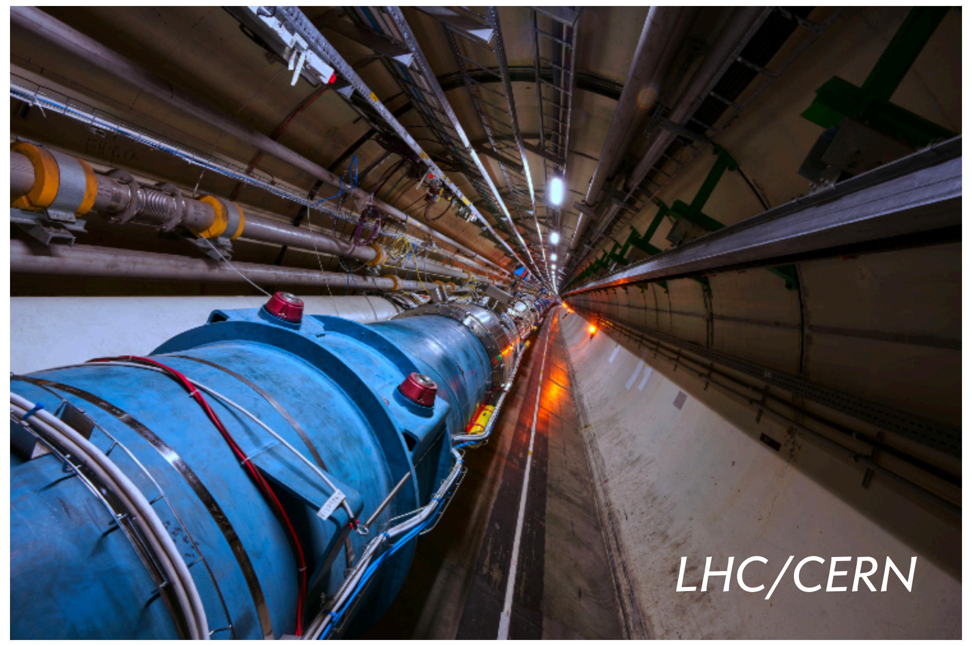


Direct detection

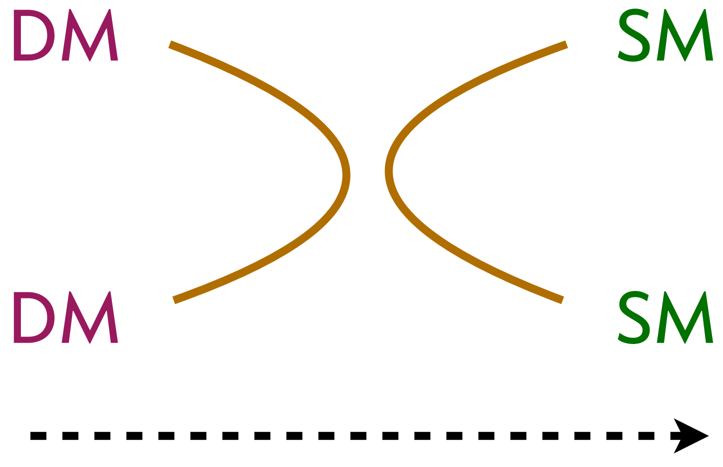
Production of DM from SM



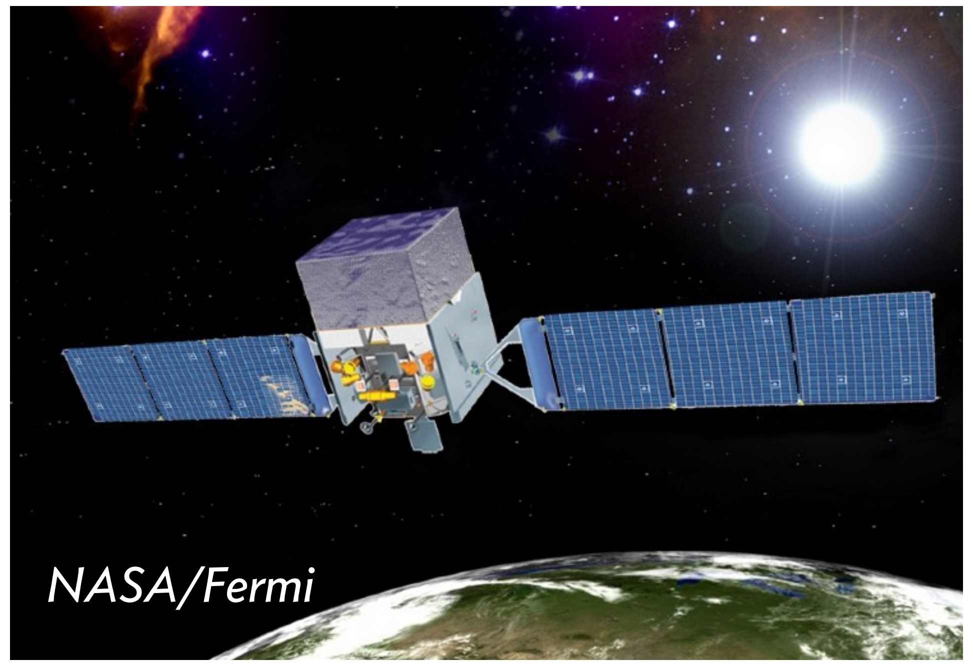
Particle colliders



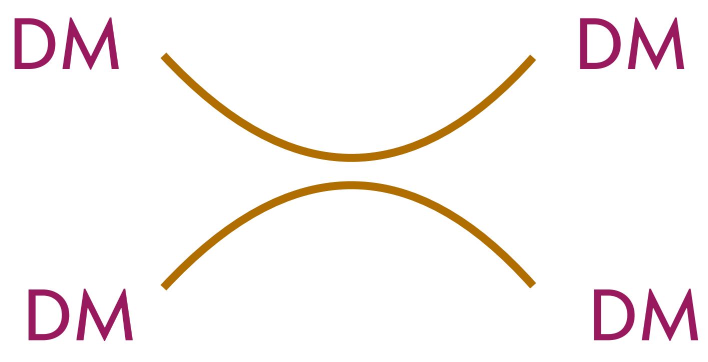
Production of SM from DM



Indirect detection

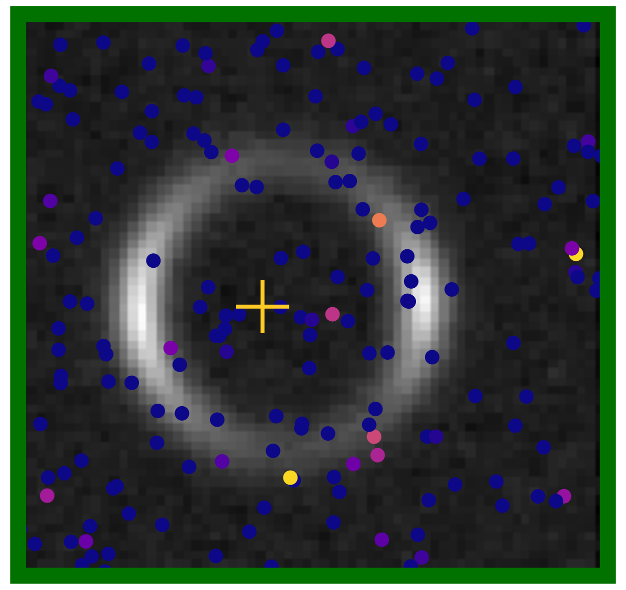


Gravitational effects of DM interaction



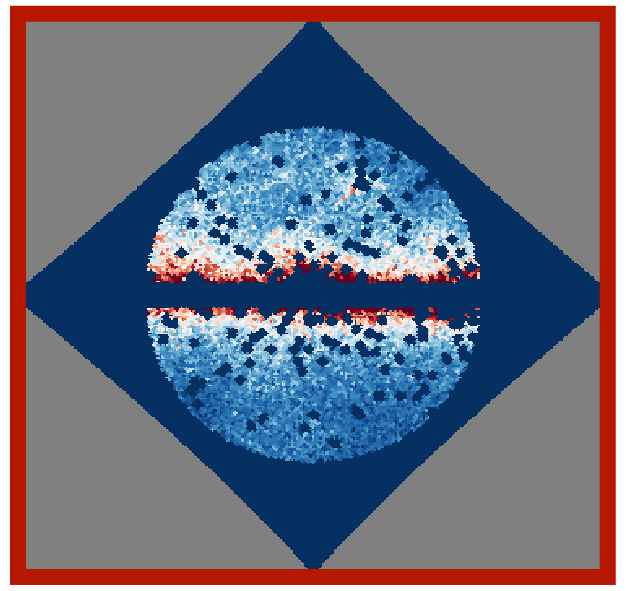
Astrophysical probes

Outline



Detecting extragalactic dark matter in strong lenses

Combining information from thousands of systems

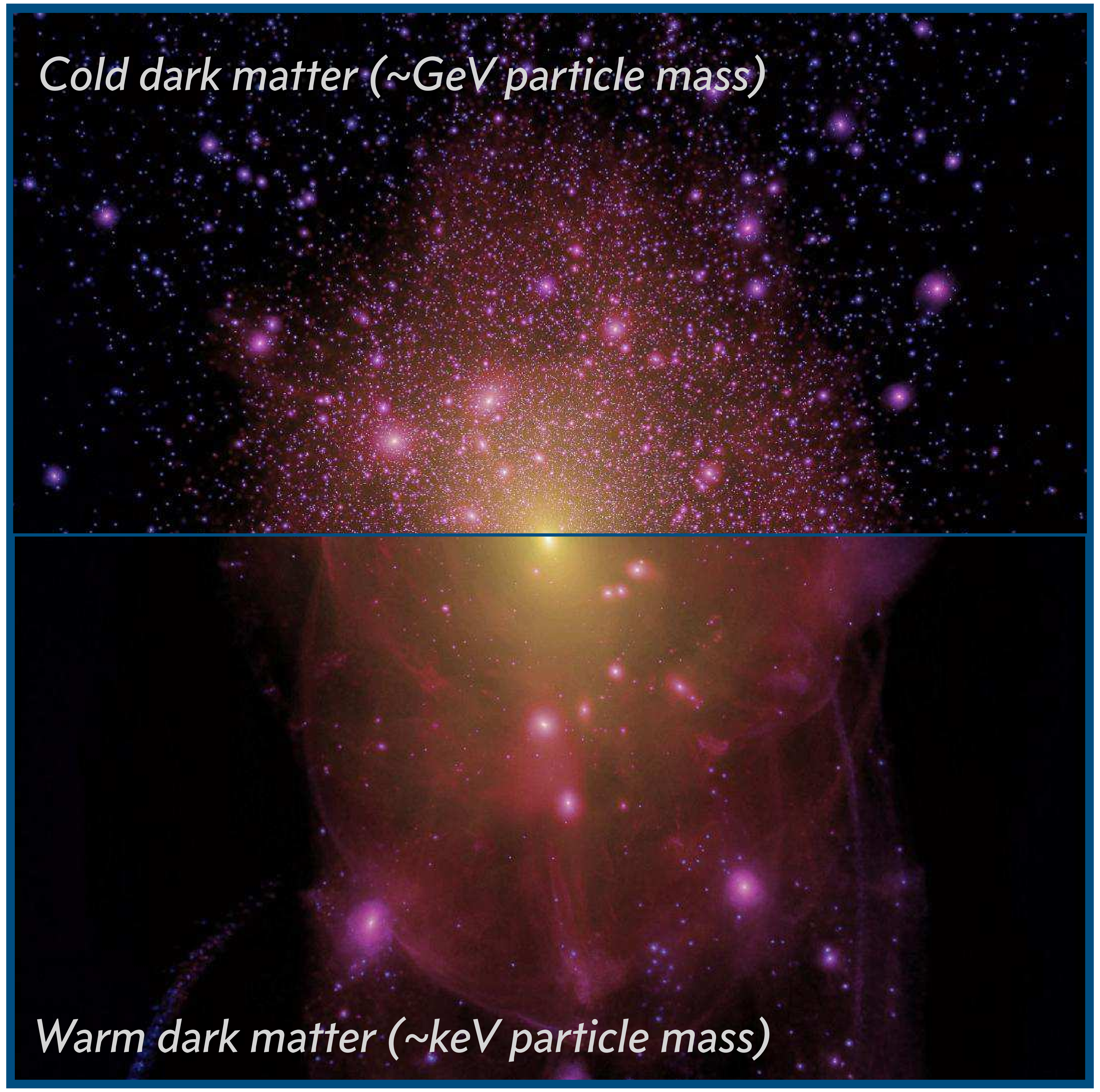


Characterizing γ-ray point sources in the Galactic Center

Exploiting more information to reduce model misspecification

Microphysics from macrophysics

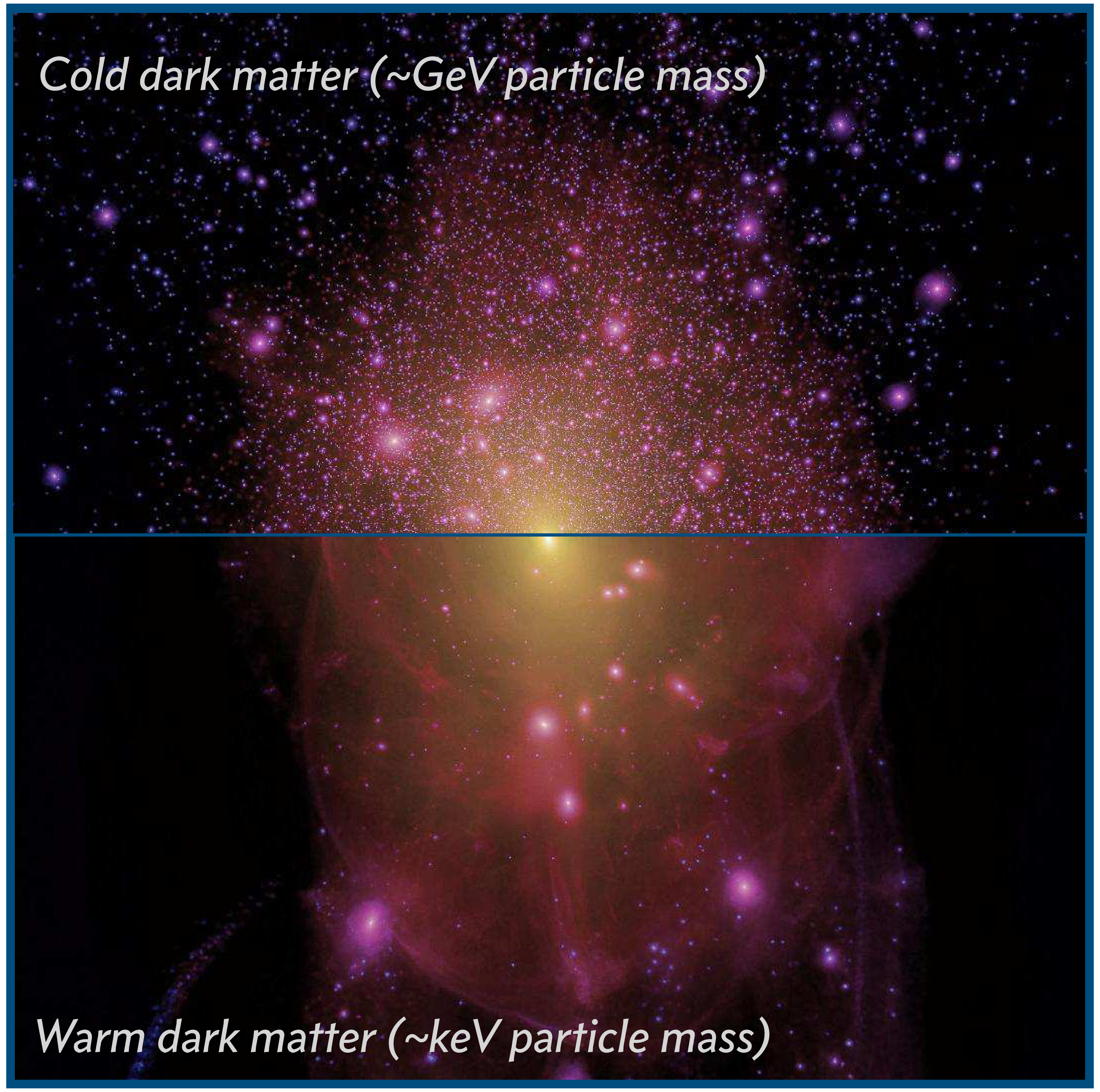
Signs of new physics can show up in the macroscopic distribution of matter



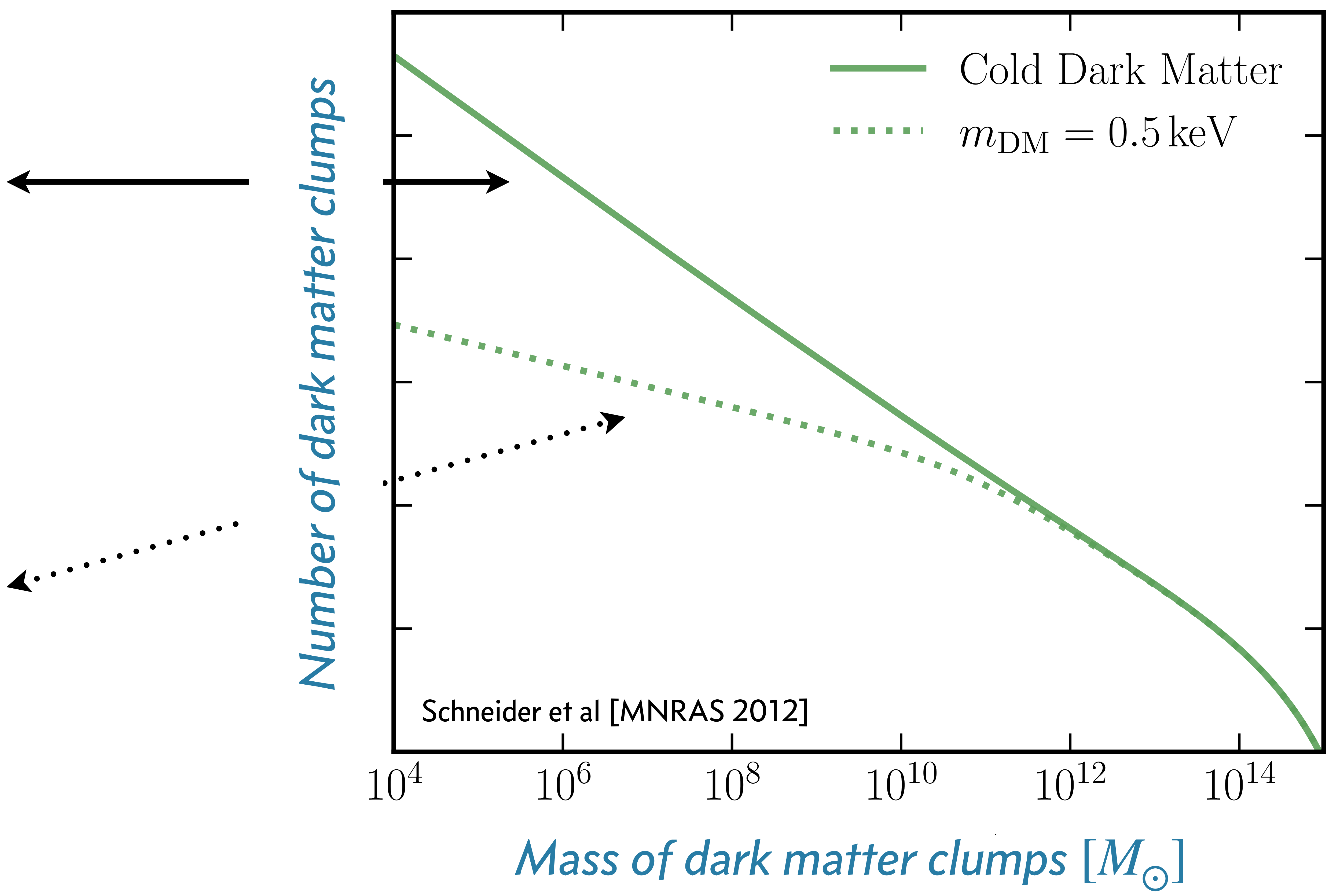
Aquarius simulation

Microphysics from macrophysics

Signs of new physics can show up in the macroscopic distribution of matter

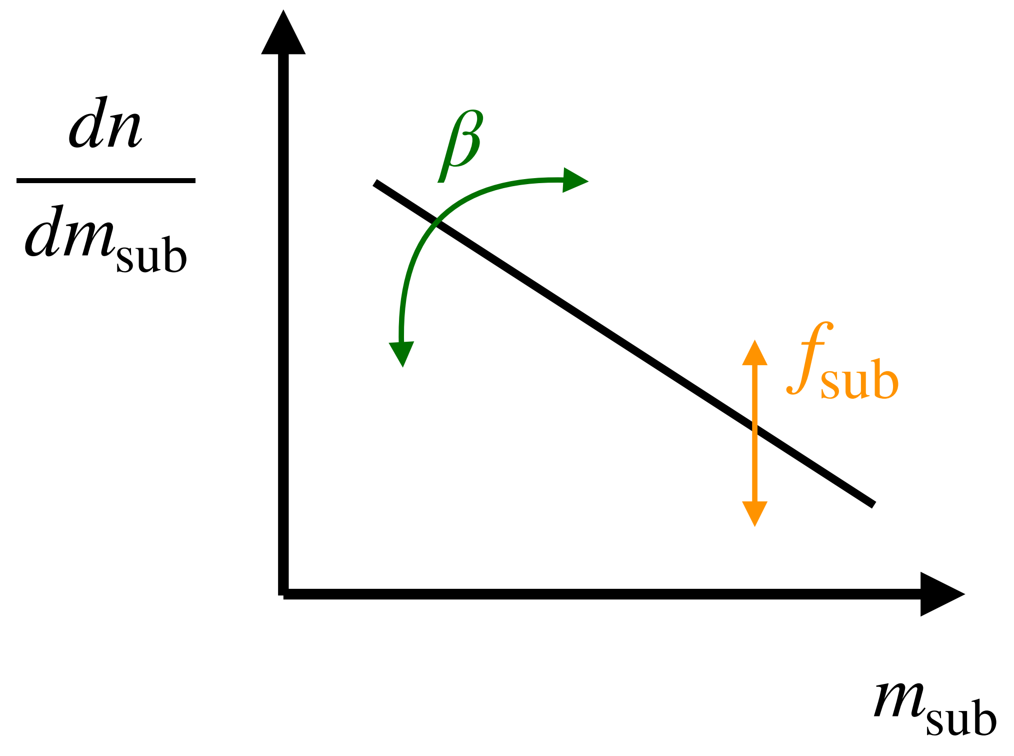


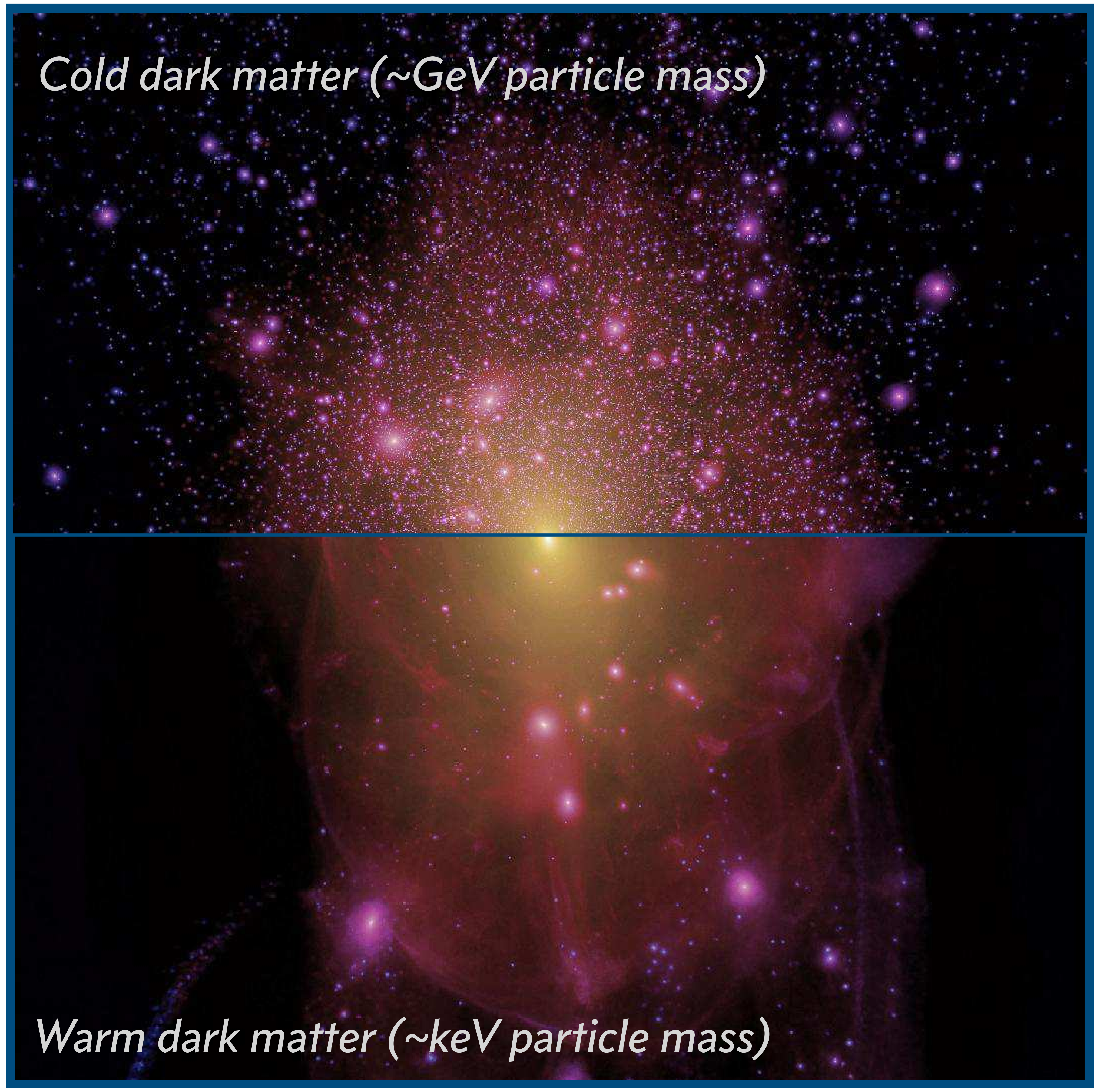
Aquarius simulation



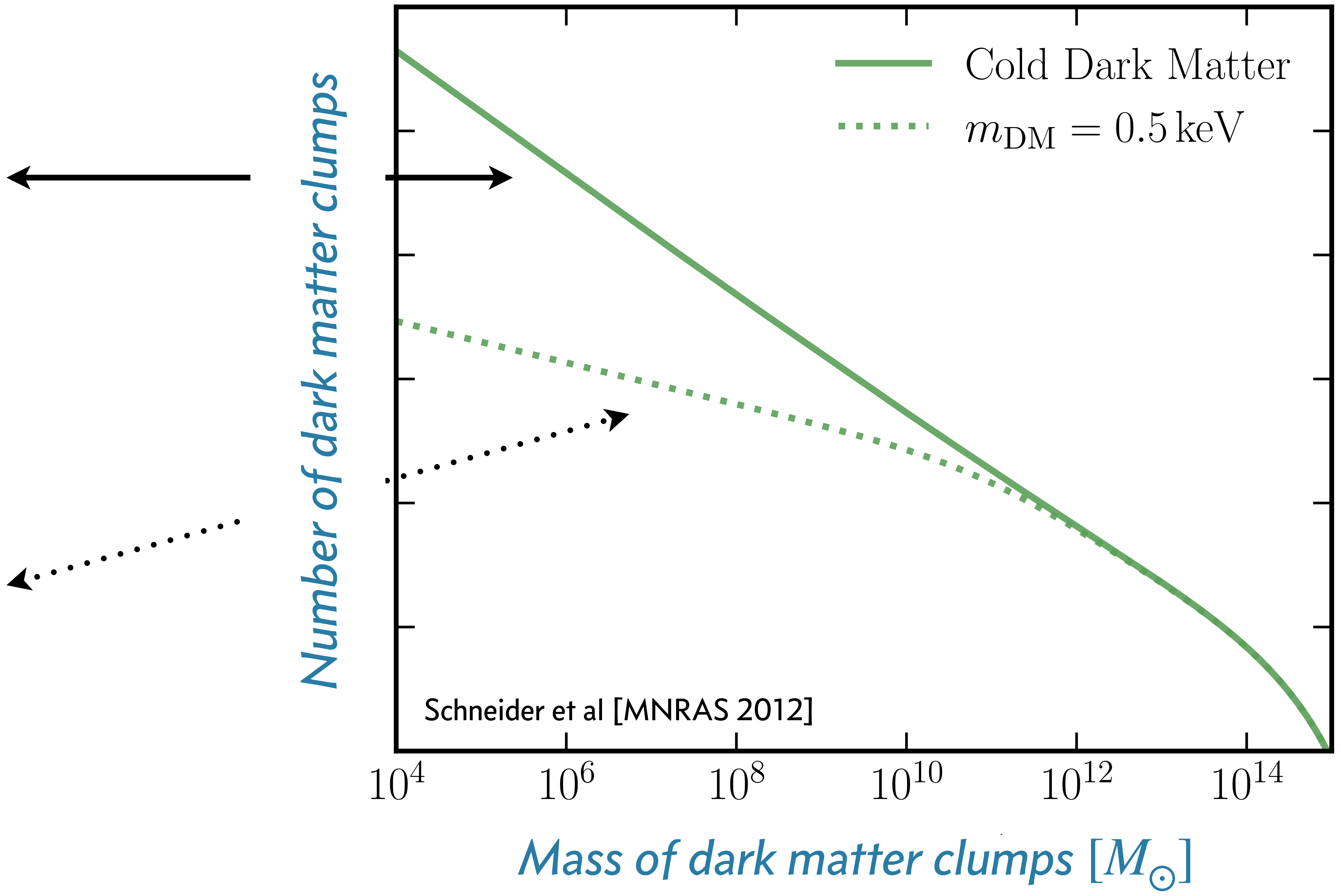
Microphysics from macrophysics

Signs of new physics can show up in the macroscopic distribution of matter

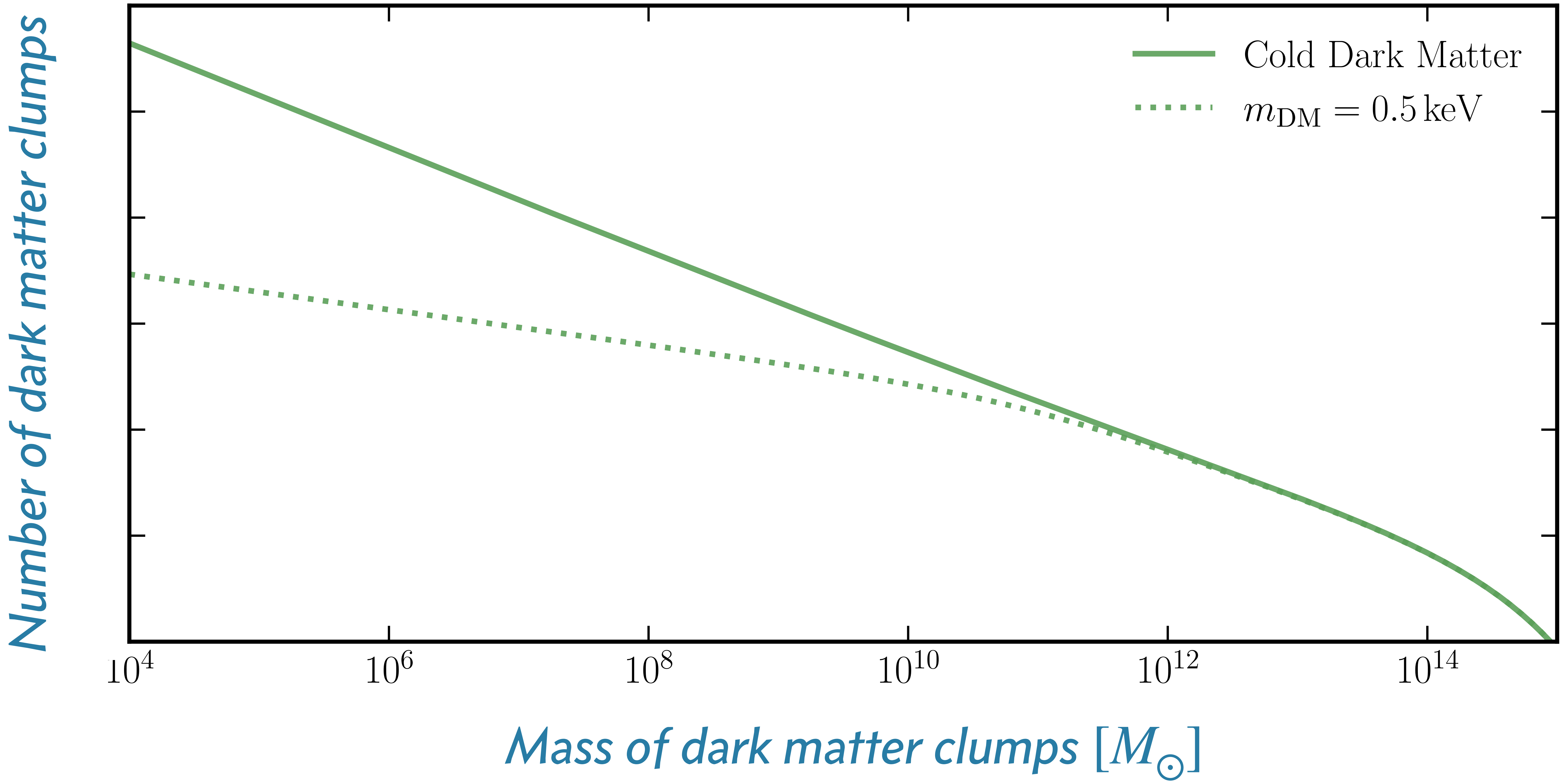




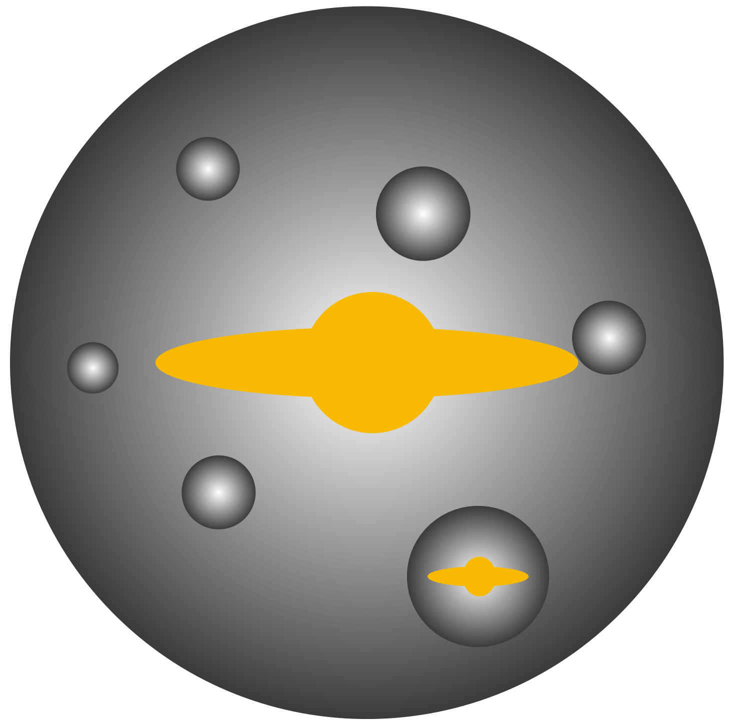
Aquarius simulation



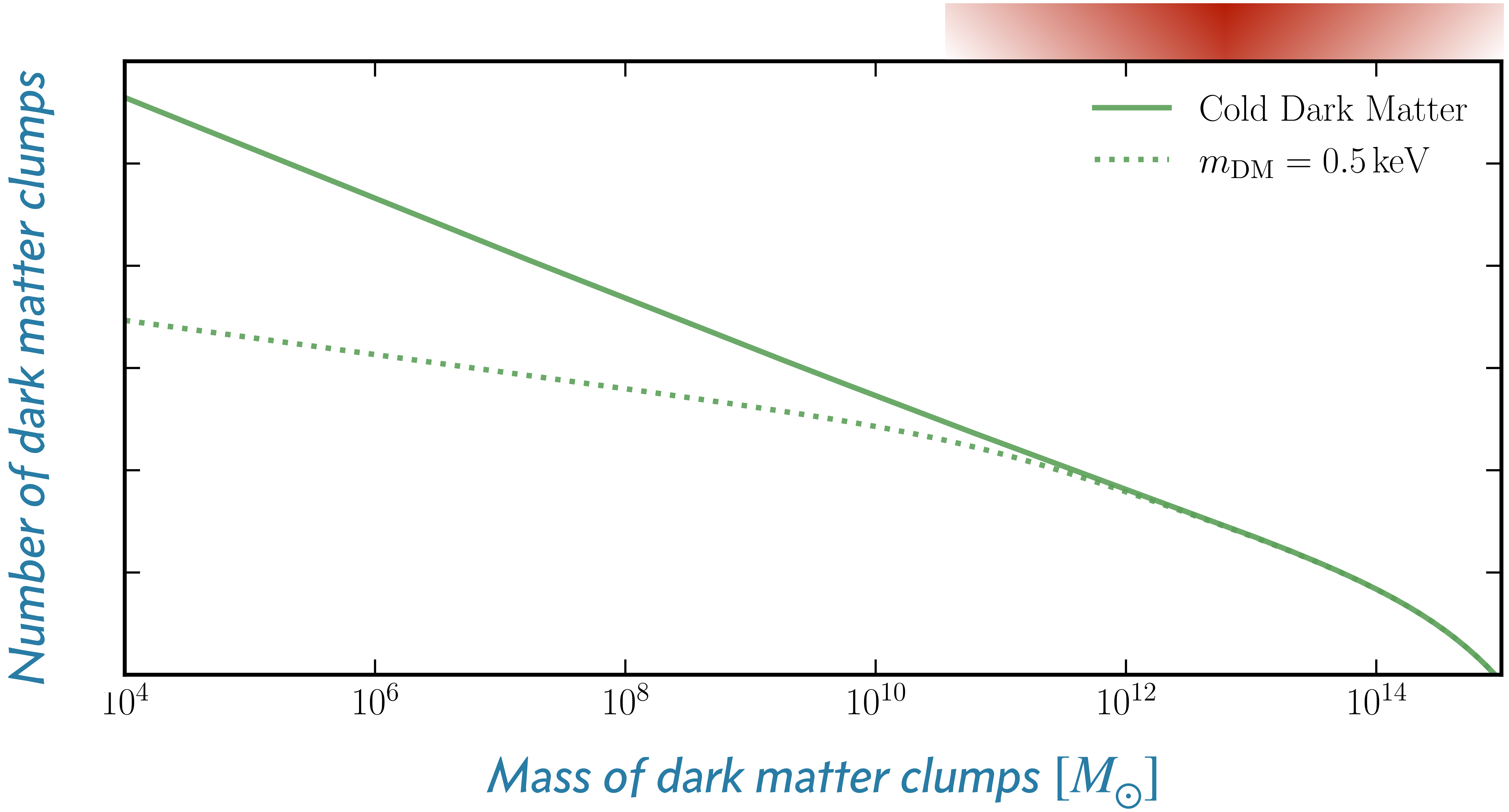
Finding dark matter subhalos



Finding dark matter subhalos



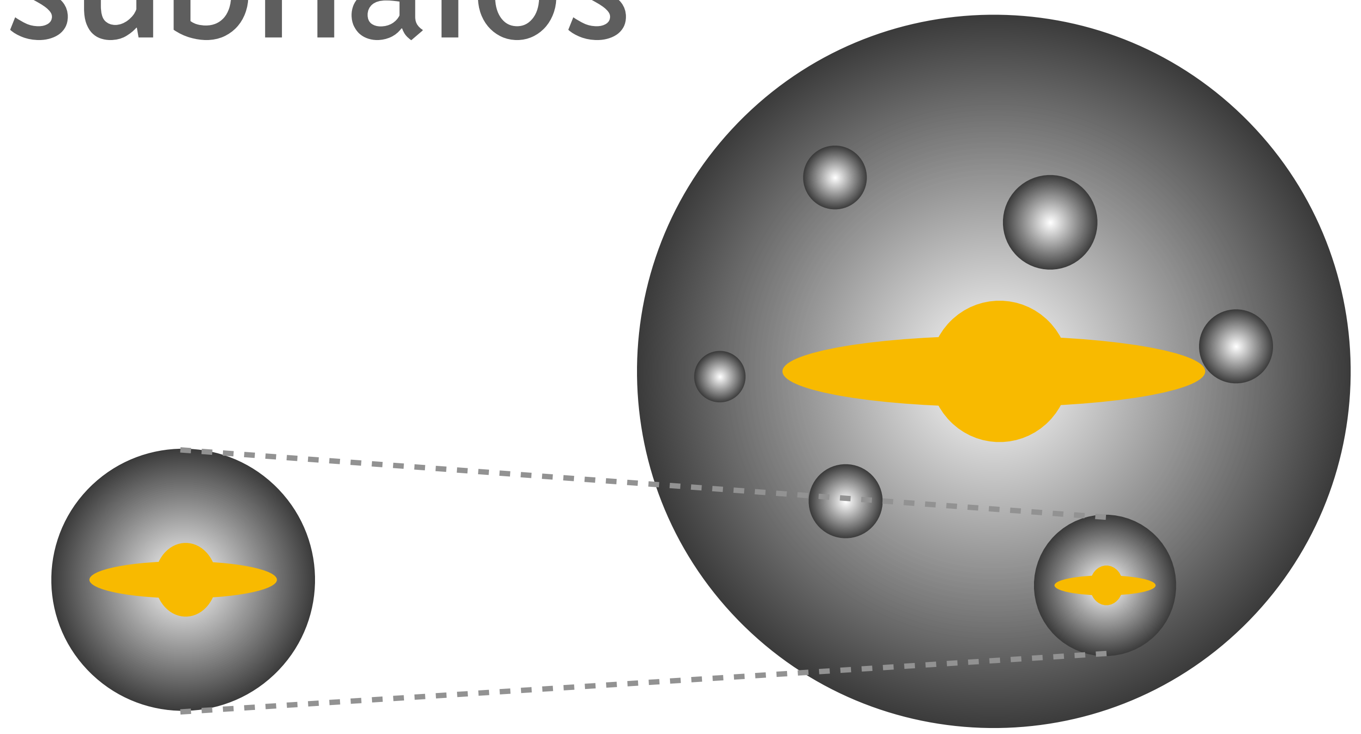
Galaxies and clusters

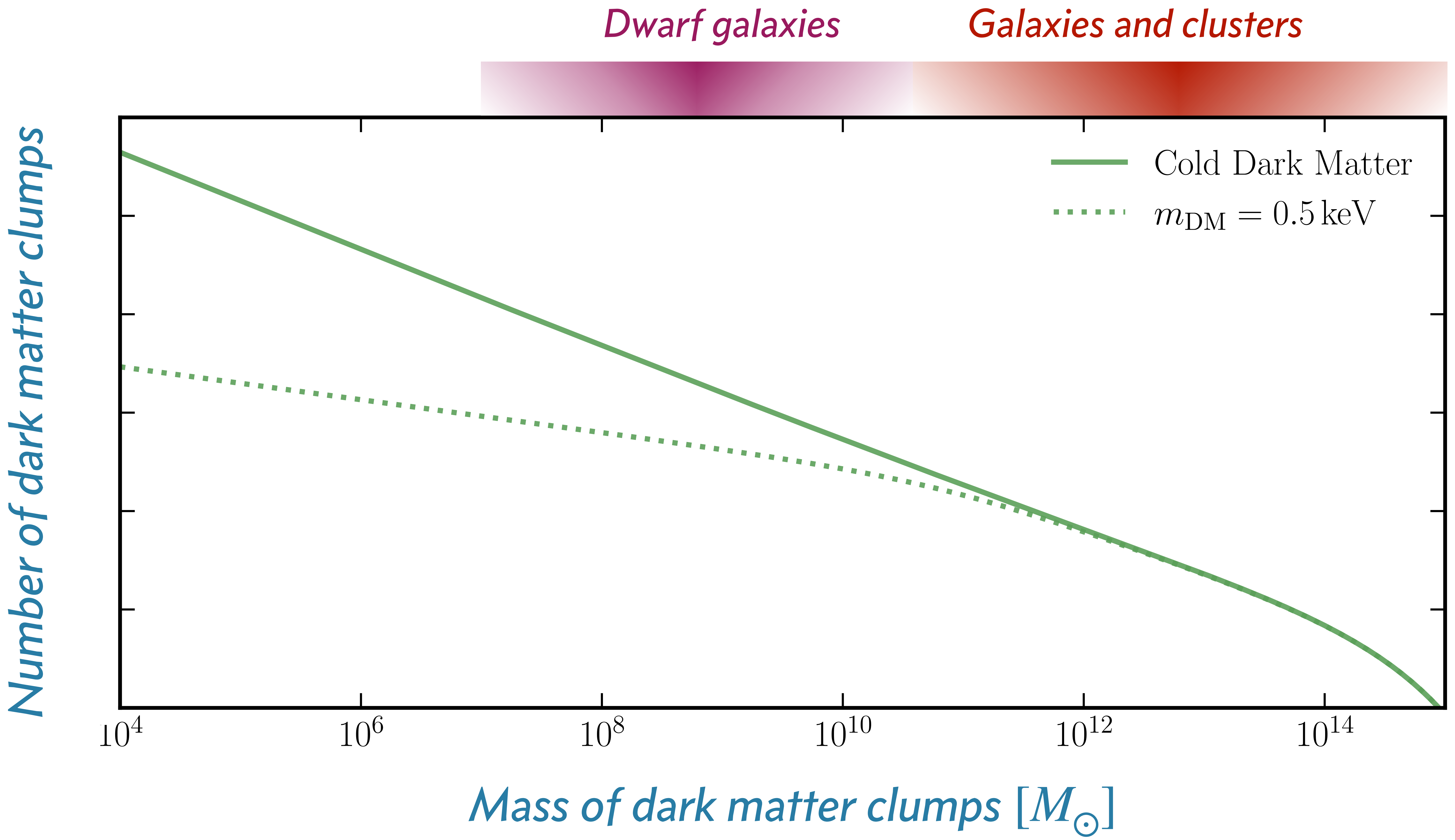


Clusters

Galaxies

Finding dark matter subhalos



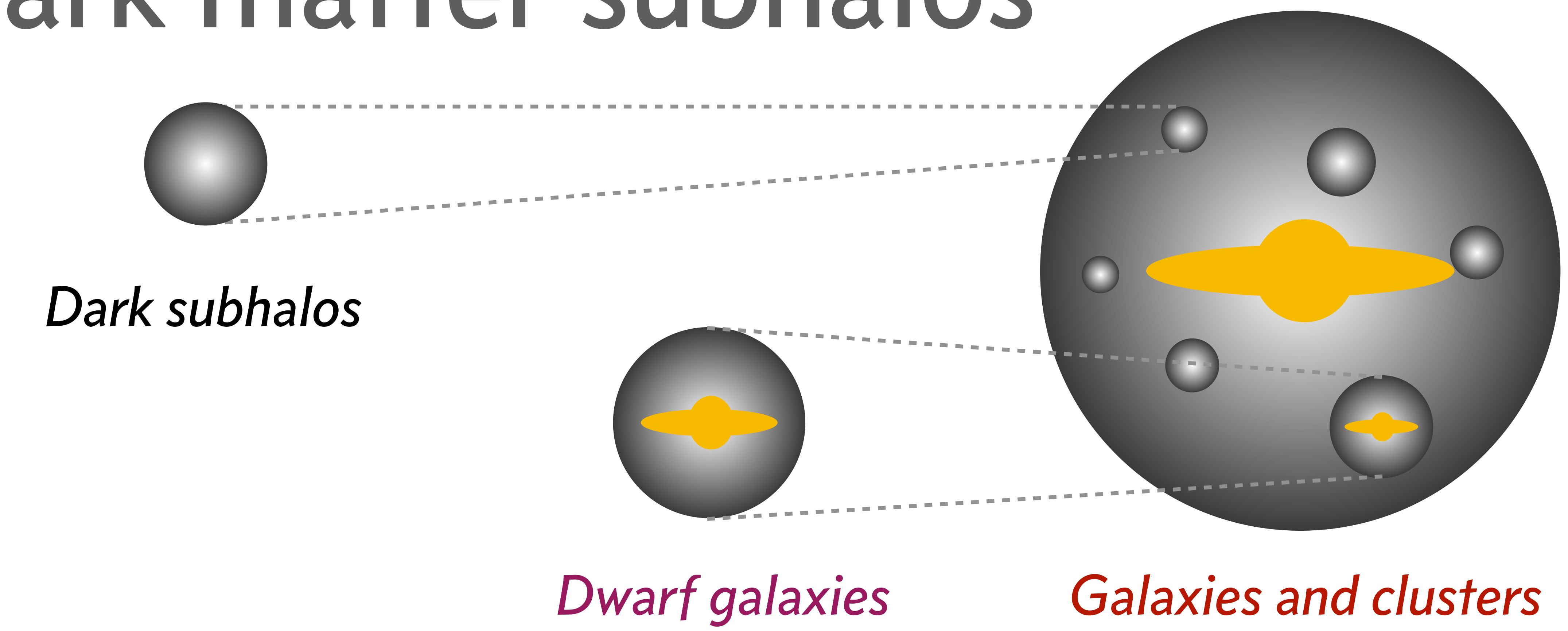


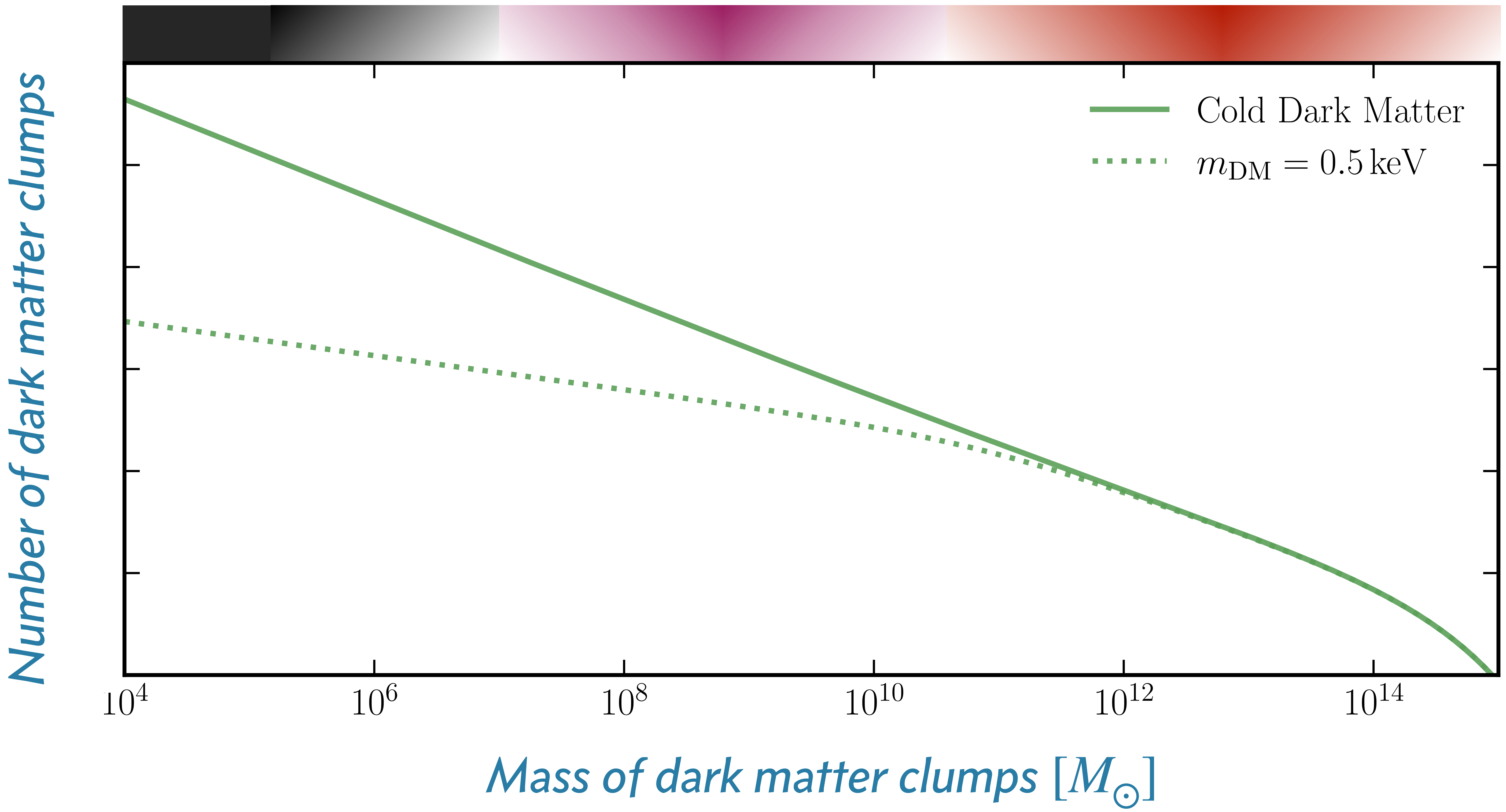
Clusters

Galaxies

Dwarf galaxies

Finding dark matter subhalos



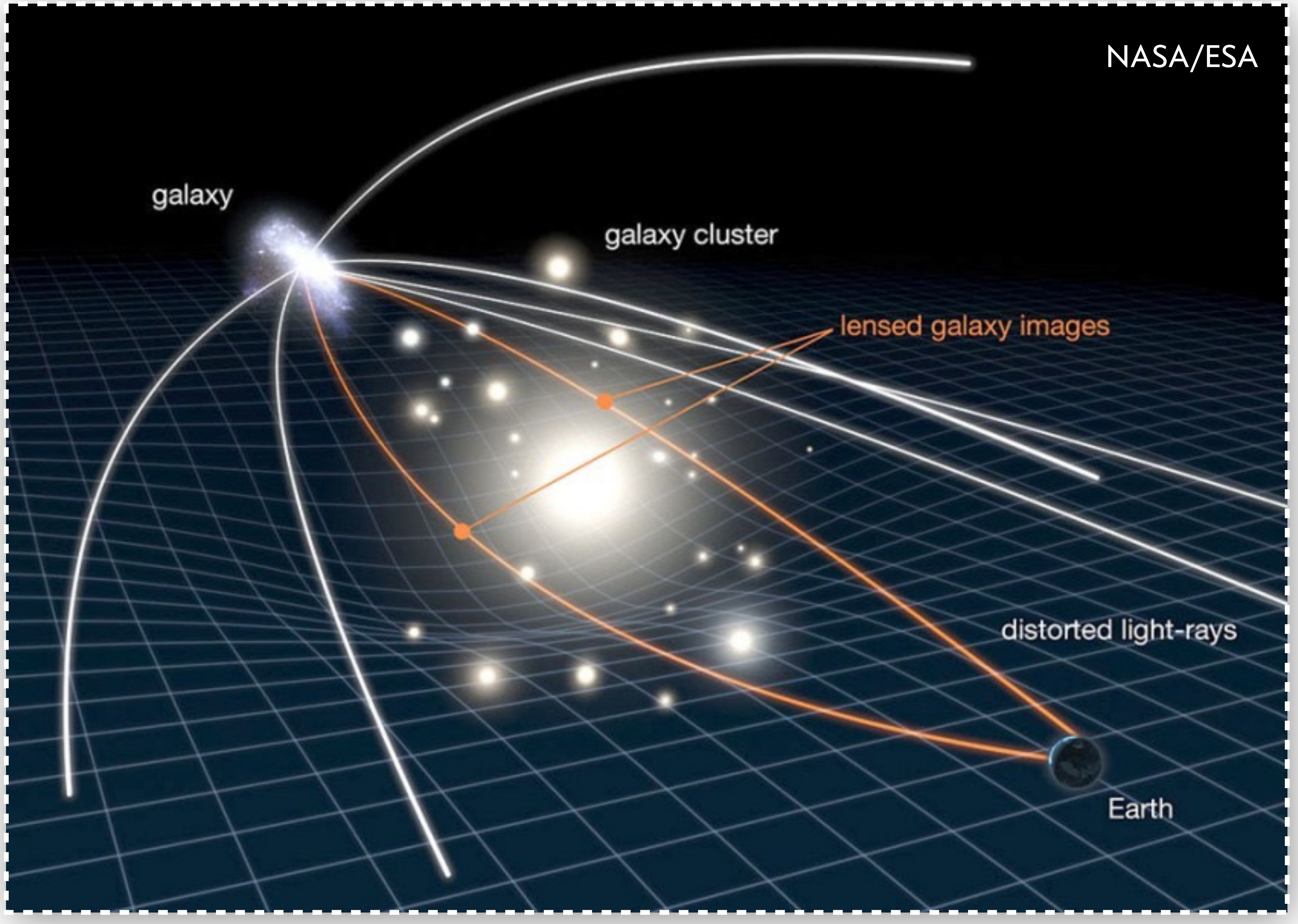


Clusters

Galaxies

Dwarf galaxies

Gravitational lensing

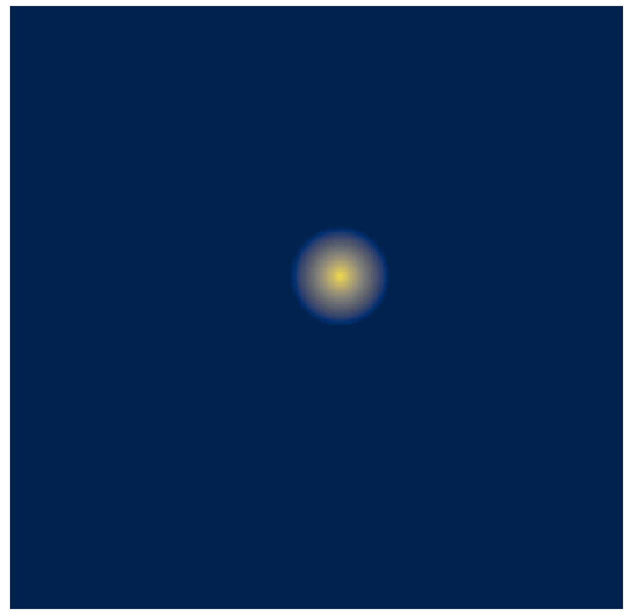


Intervening mass causes a shift in the *apparent* position of background light

$$\overrightarrow{\Delta\theta} \sim \vec{\nabla}_\theta \Psi_\mathrm{G}(\vec{r})$$

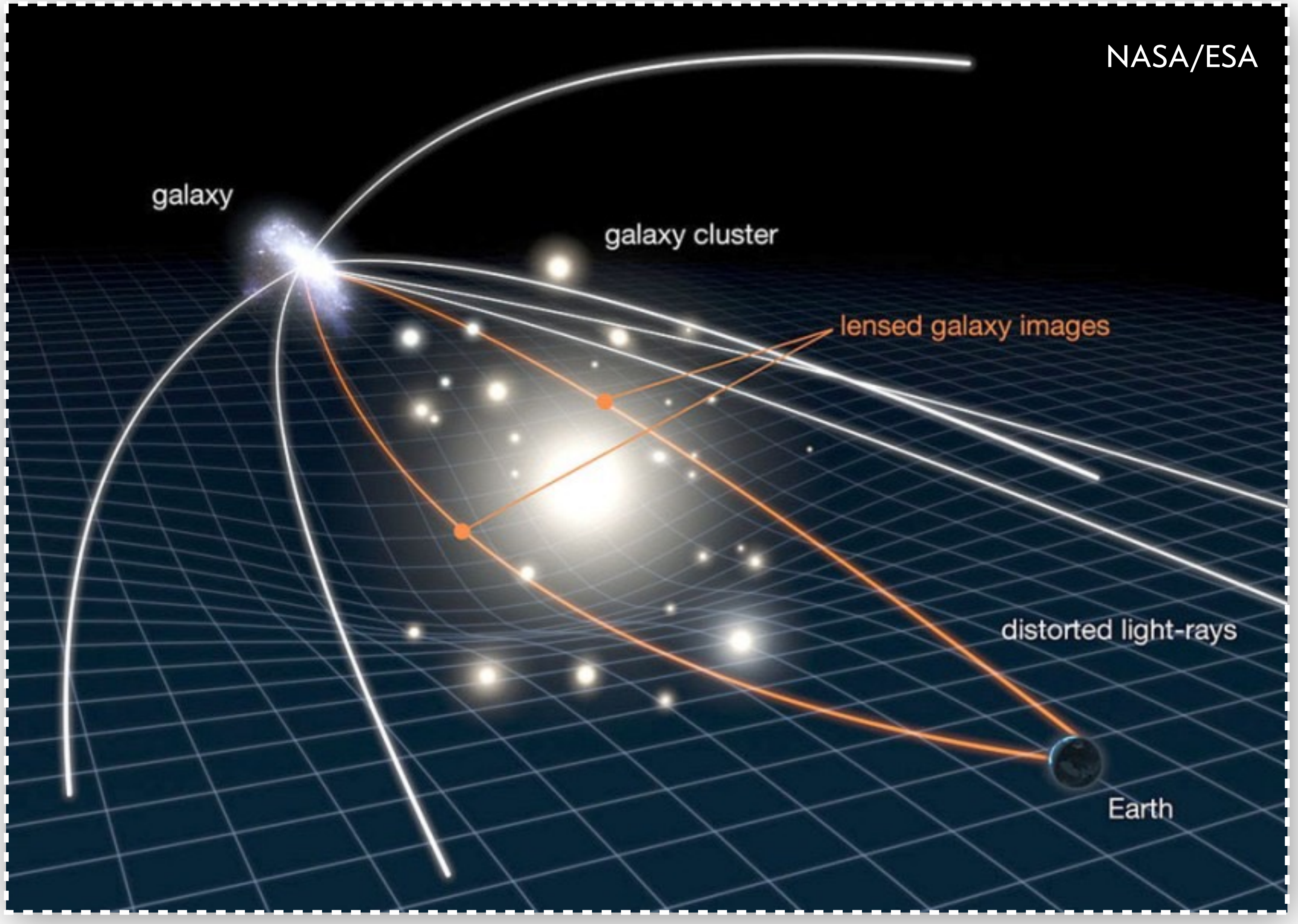
Strong lensing: extended arcs, multiple images

Original source



Lensed source

Gravitational lensing

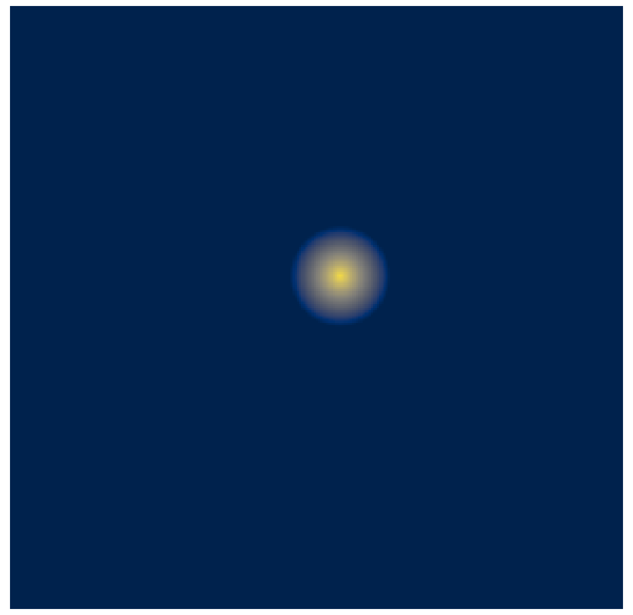


Intervening mass causes a shift in the *apparent* position of background light

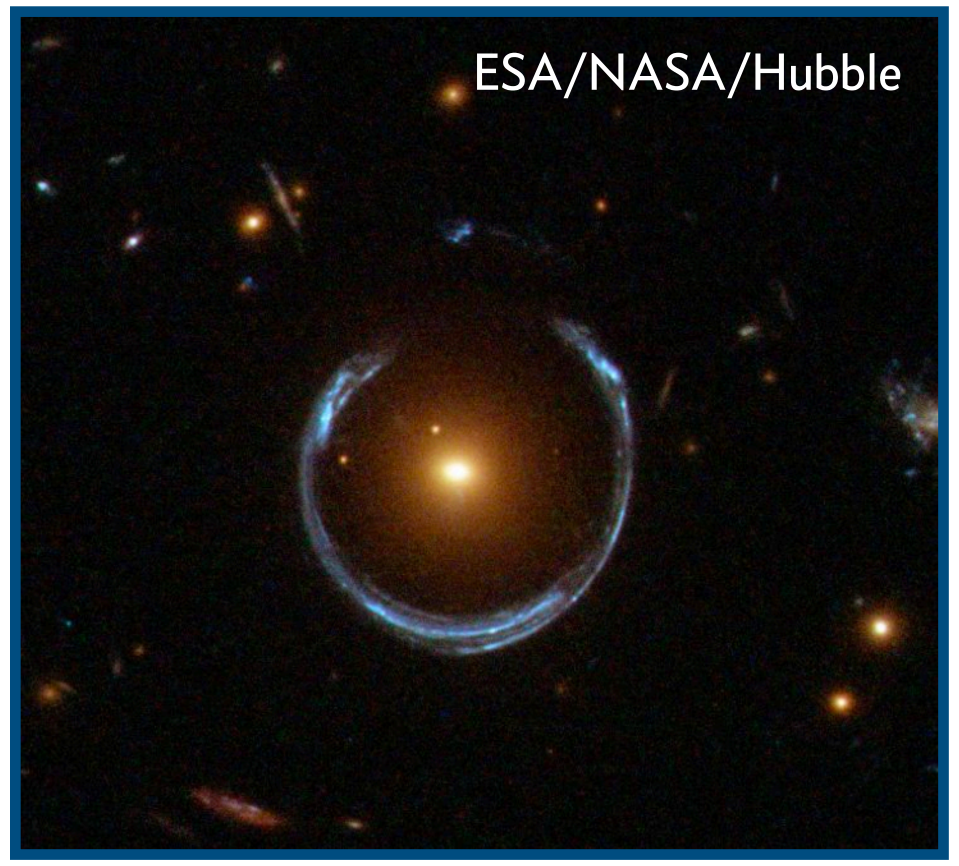
$$\overrightarrow{\Delta\theta} \sim \vec{\nabla}_\theta \Psi_{\mathrm{G}}(\vec{r})$$

Strong lensing: extended arcs, multiple images

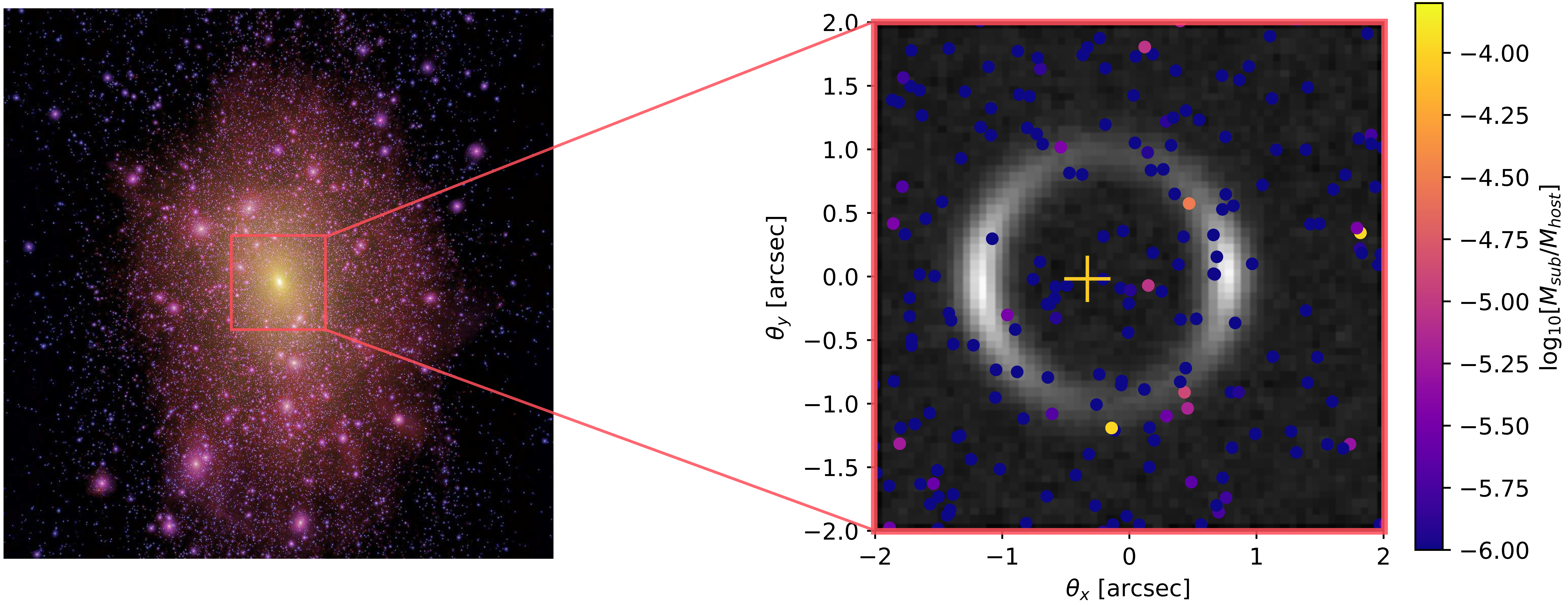
Original source



Lensed source



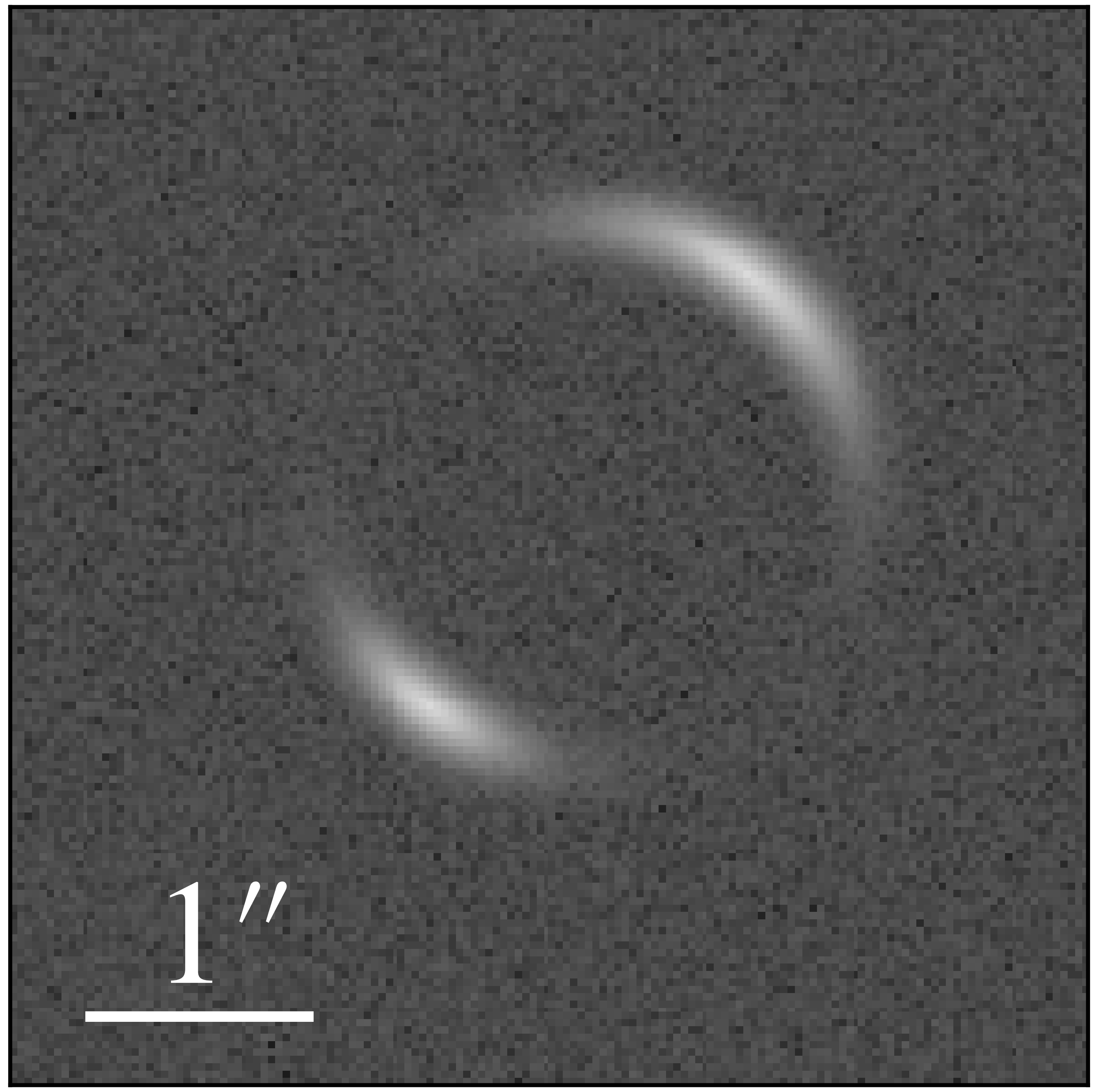
Strong lensing: effect of substructure



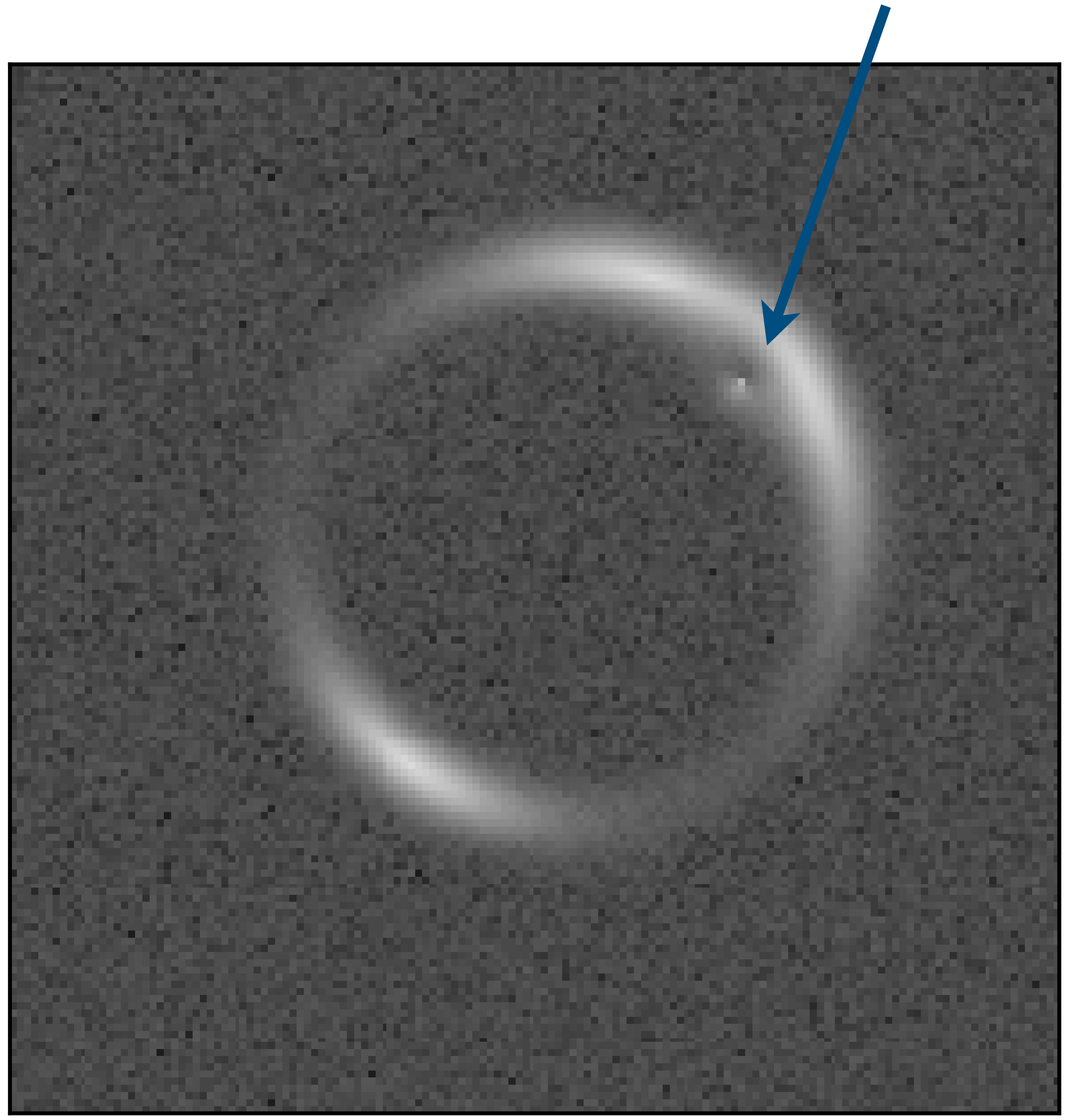
Substructure causes percent-level shifts in strongly lensed image

Strong lensing: effect of substructure

Smooth halo only



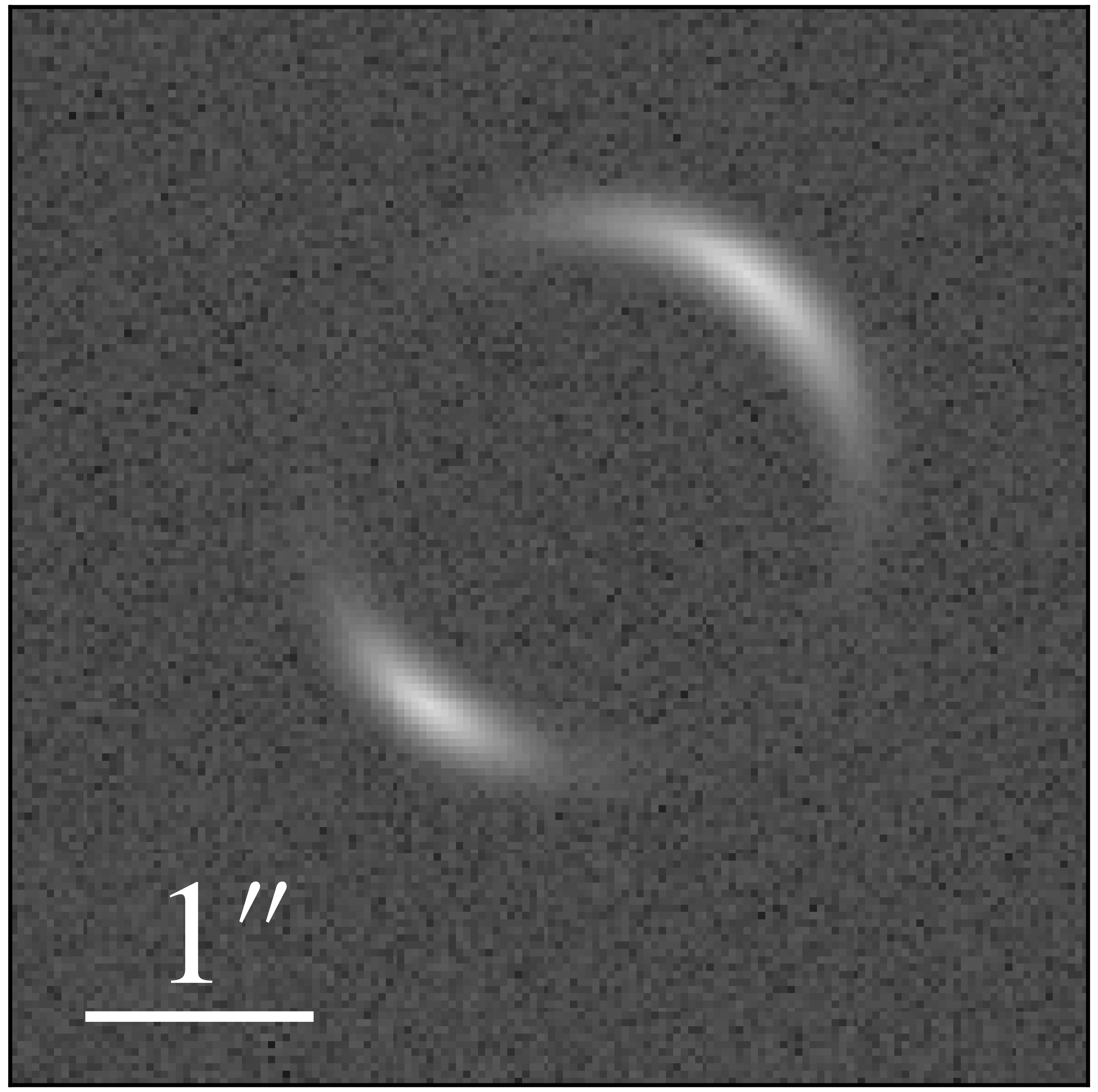
Smooth halo + *subhalo*



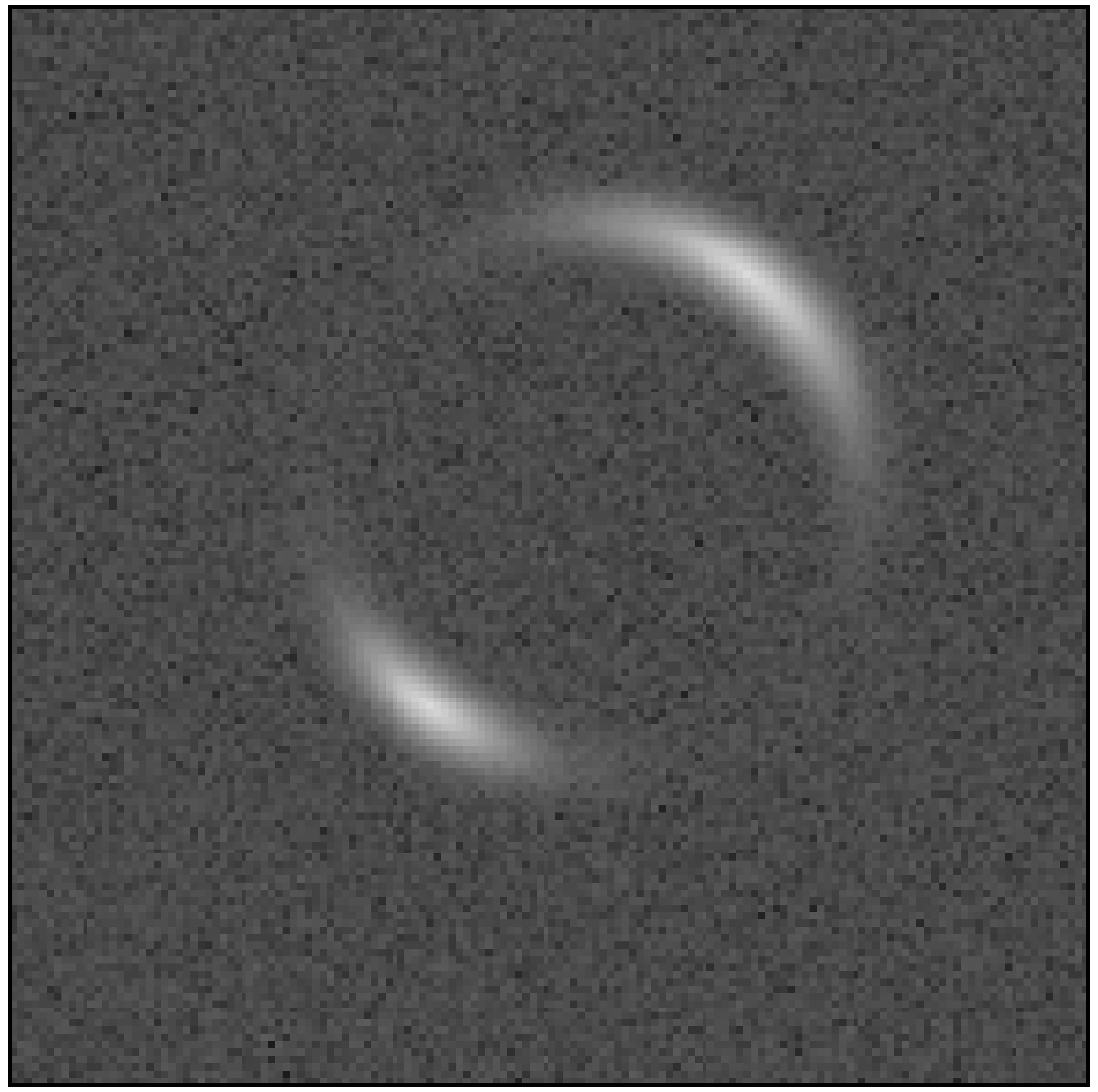
Substructure perturbs lensing rings compared to only smooth halo

Strong lensing: effect of substructure *in reality*

Smooth halo only



Smooth halo + *subhalos*



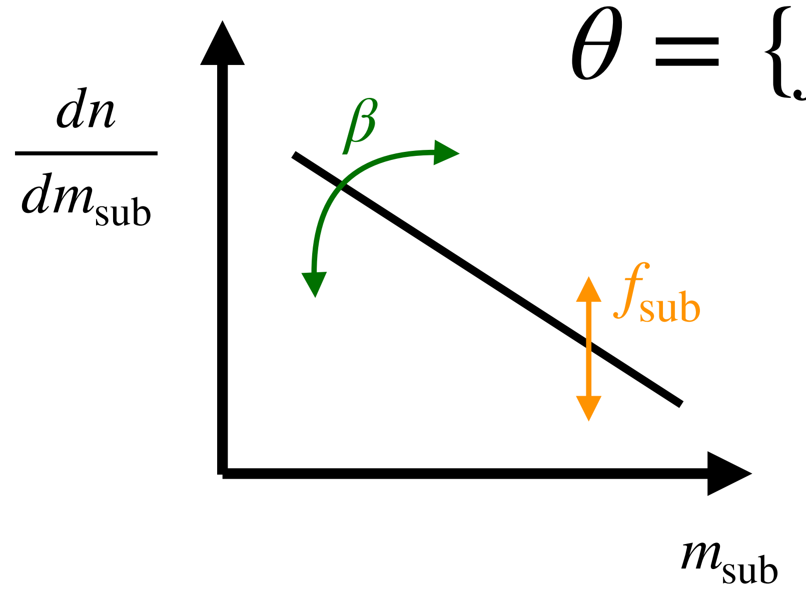
Effect is very subtle for realistic dark matter substructure

Modeling substructure in strong lenses

Parameters of interest

Subhalo population parameters

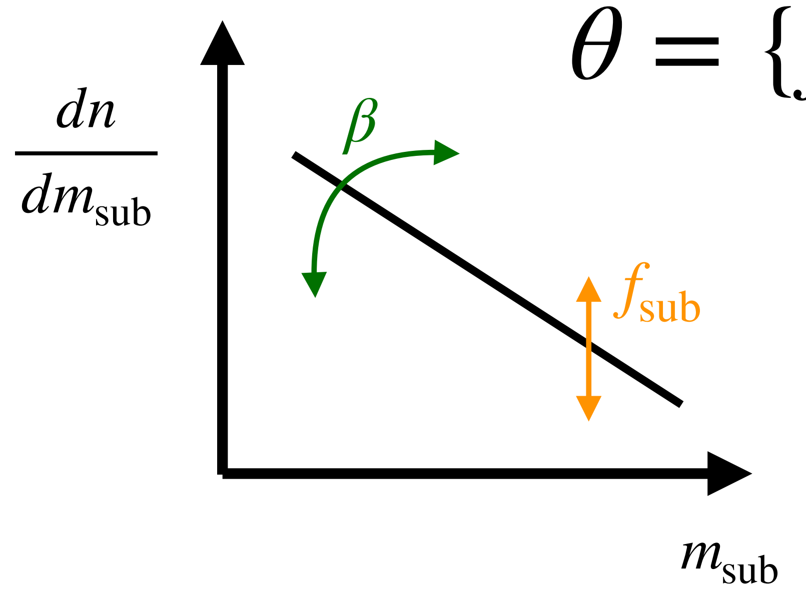
$$\theta = \{f_{\rm sub}, \beta\}$$



Modeling substructure in strong lenses

Parameters of interest

Subhalo population parameters



$$\theta = \{f_{\mathrm{sub}}, \beta\}$$

Latent variables

Source/host properties

$z_{\mathrm{src}}, z_{\mathrm{lens}}$

Subhalo properties

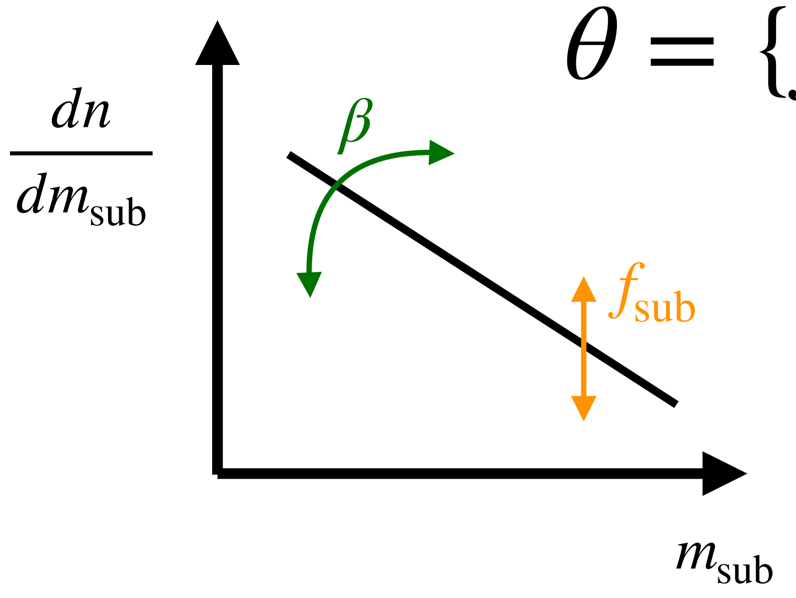
$n_{\mathrm{sub}}, \{z_{\mathrm{sub},\,i}\}$

$$p\left(z_{\mathrm{src}}\right)\ p\left(z_{\mathrm{lens}}\right)\ p\left(n_{\mathrm{sub}} \mid \theta\right)\prod_{i}^{n_{\mathrm{sub}}} p\left(z_{\mathrm{sub},\,i} \mid \theta\right)$$

Modeling substructure in strong lenses

Parameters of interest

Subhalo population parameters



$$\theta = \{f_{\rm sub}, \beta\}$$

Latent variables

Source/host properties

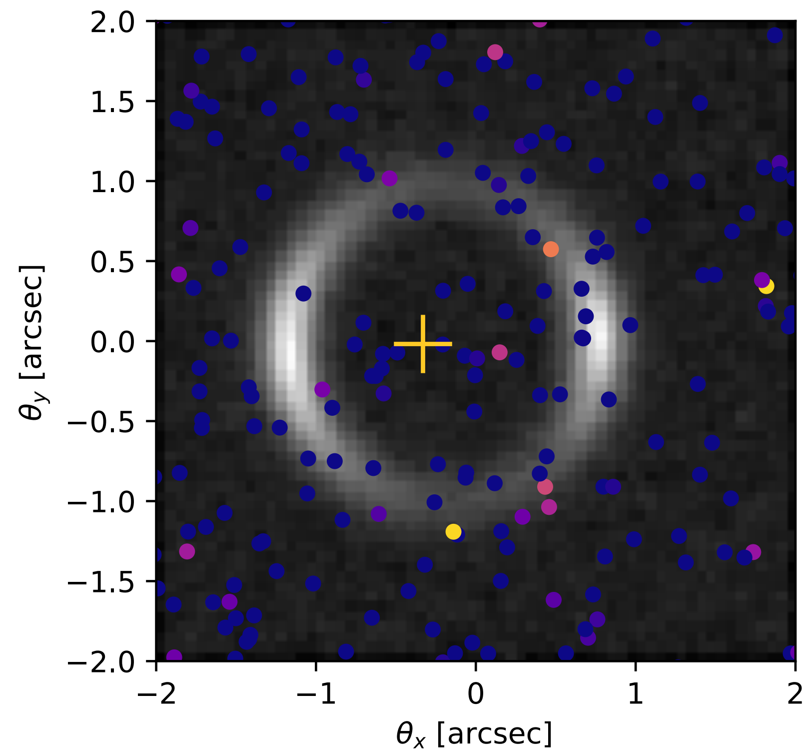
$$z_{\rm src}, z_{\rm lens}$$

Subhalo properties

$$n_{\rm sub}, \{z_{{\rm sub},\,i}\}$$

Observables

Lensing image x

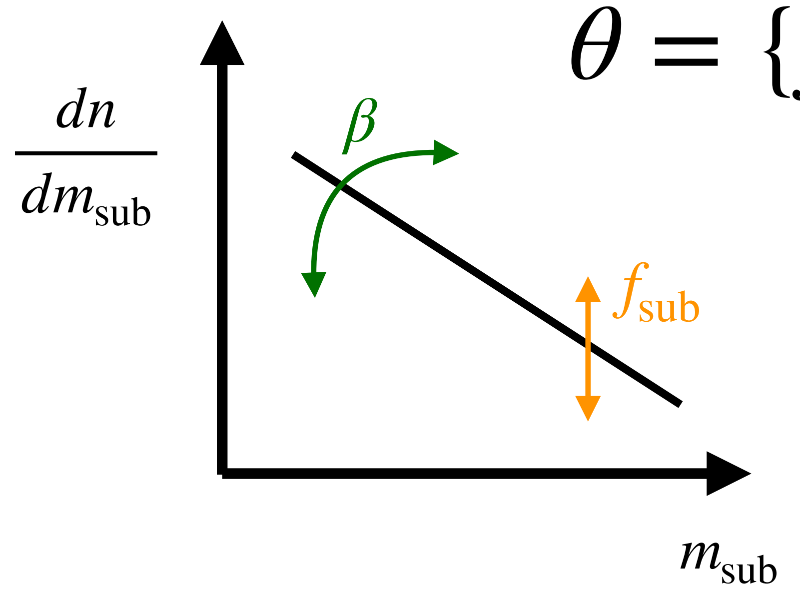


$$p\left(z_{\rm src}\right)\ p\left(z_{\rm lens}\right)\ p\left(n_{\rm sub} \mid \theta\right) \prod_i^{n_{\rm sub}} p\left(z_{{\rm sub},\,i} \mid \theta\right)$$

$$p\left(x \mid z_{\rm src}, z_{\rm lens}, \left\{z_{{\rm sub},\,i}\right\}\right)$$

Modeling substructure in strong lenses

Parameters of interest
Subhalo population parameters



$\theta = \{f_{\text{sub}}, \beta\}$

Latent variables

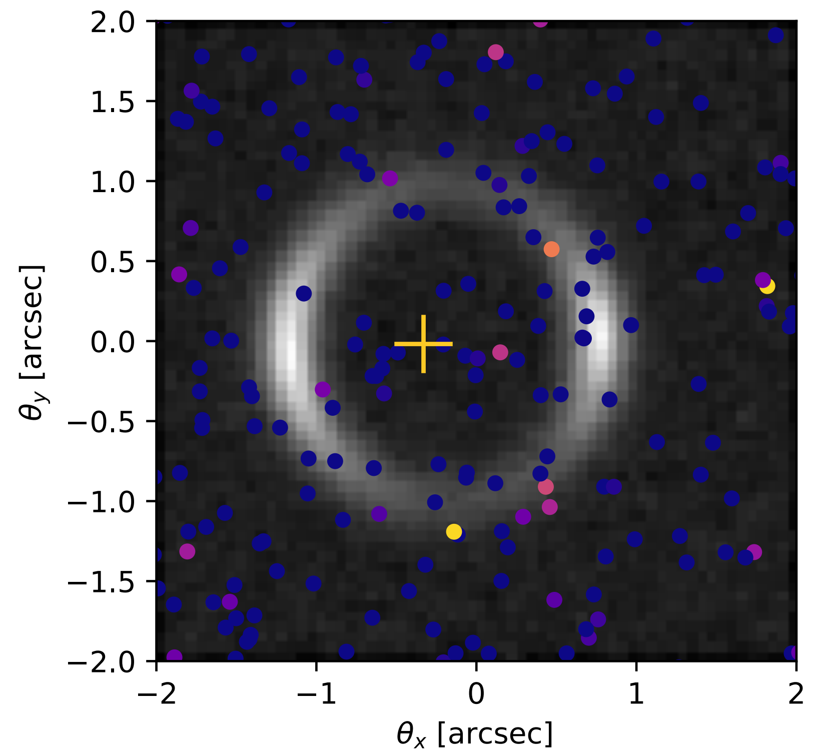
Source/host properties

$z_{\text{src}}, z_{\text{lens}}$

Subhalo properties

$n_{\text{sub}}, \{z_{\text{sub},\, i}\}$

Observables
Lensing image x



We can easily write a simulator to sample from

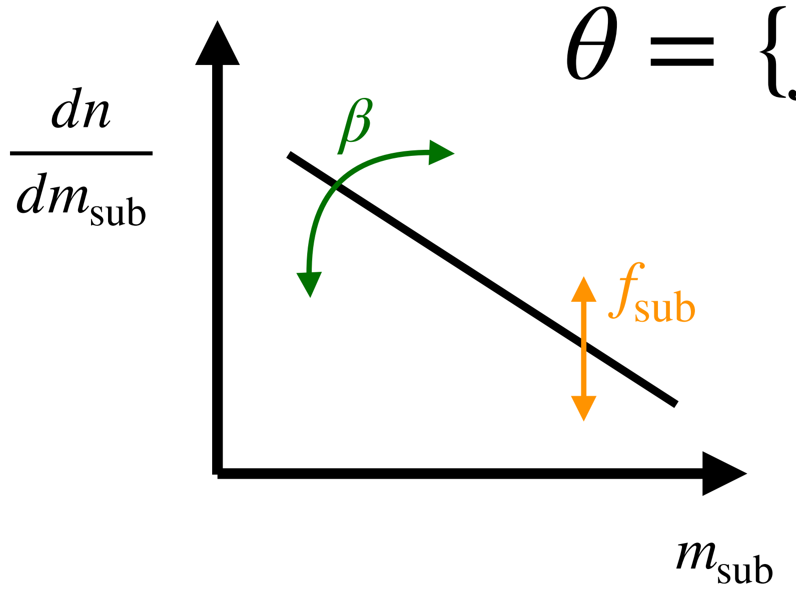
$$p(x, z \mid \theta) = \; p\left(z_{\text{src}}\right)\; p\left(z_{\text{lens}}\right)\; p\left(n_{\text{sub}} \mid \theta\right) \prod_{i}^{n_{\text{sub}}} p\left(z_{\text{sub},\, i} \mid \theta\right) \qquad p\left(x \mid z_{\text{src}}, z_{\text{lens}}, \left\{z_{\text{sub},\, i}\right\}\right)$$

Prediction (Simulation) $\longrightarrow$

Modeling substructure in strong lenses

Parameters of interest

Subhalo population parameters



$$\theta = \{f_{\mathrm{sub}}, \beta\}$$

Latent variables

Source/host properties

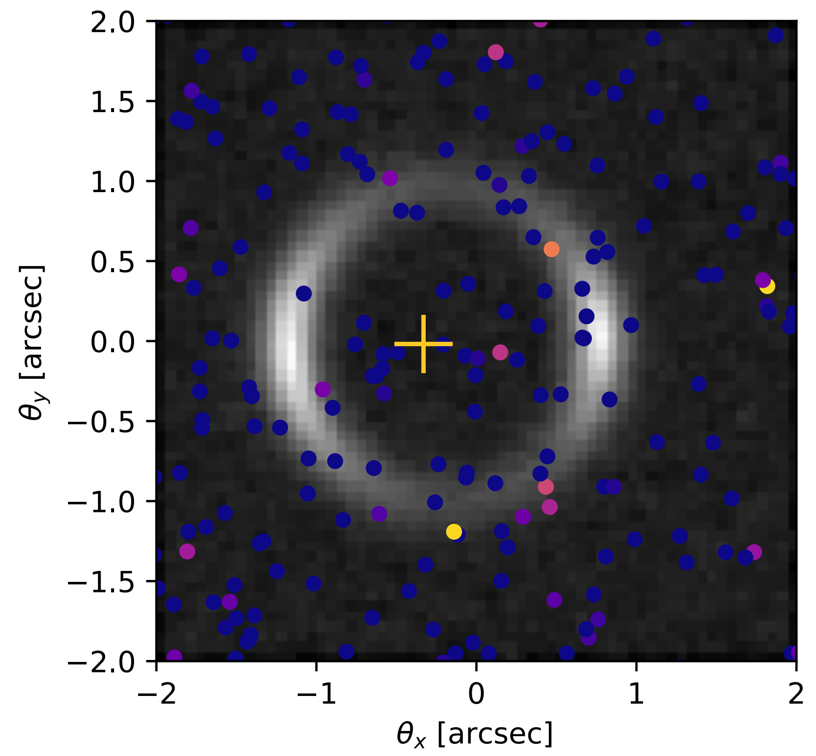
$$z_{\mathrm{src}}, z_{\mathrm{lens}}$$

Subhalo properties

$$n_{\mathrm{sub}}, \{z_{\mathrm{sub},\,i}\}$$

Observables

Lensing image x



The key quantity for inference is the marginal likelihood

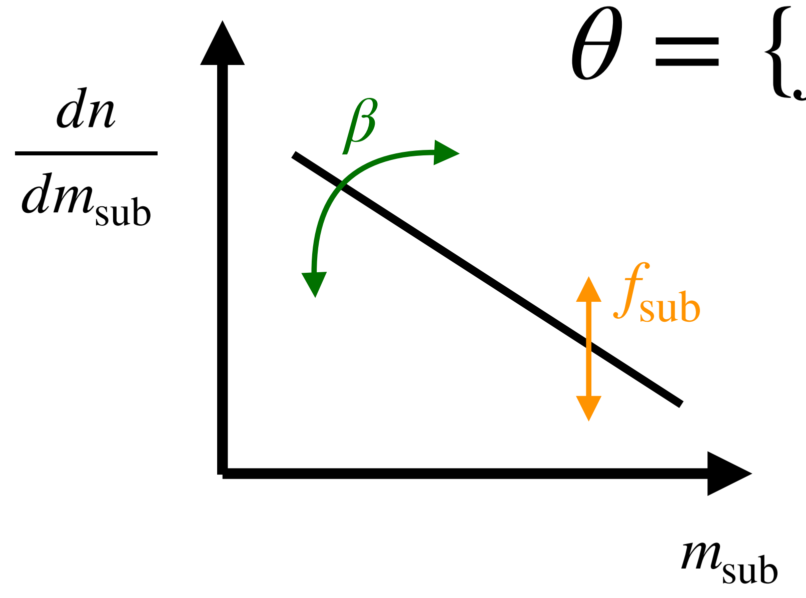
$$p(x \mid \theta) = \int \mathrm{d}z_{\mathrm{source}} \int \mathrm{d}z_{\mathrm{lens}} \sum_{n_{\mathrm{sub}}} \int \mathrm{d}^{n_{\mathrm{sub}}} z_{\mathrm{sub}}\; p\left(z_{\mathrm{src}}\right)\; p\left(z_{\mathrm{lens}}\right)\; p\left(n_{\mathrm{sub}} \mid \theta\right) \prod_{i}^{n_{\mathrm{sub}}} p\left(z_{\mathrm{sub},\,i} \mid \theta\right) \qquad p\left(x \mid z_{\mathrm{src}}, z_{\mathrm{lens}}, \left\{z_{\mathrm{sub},\,i}\right\}\right)$$

Inference

Modeling substructure in strong lenses

Parameters of interest

Subhalo population parameters



$\theta = \{f_{\mathrm{sub}}, \beta\}$

Latent variables

Source/host properties

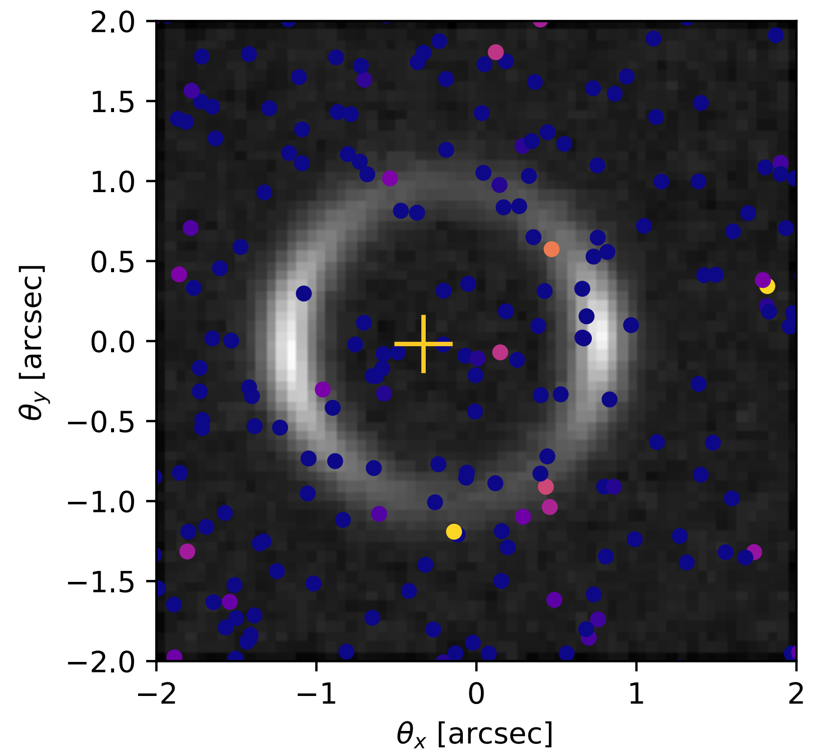
$z_{\mathrm{src}}, z_{\mathrm{lens}}$

Subhalo properties

$n_{\mathrm{sub}}, \{z_{\mathrm{sub},\,i}\}$

Observables

Lensing image x



The key quantity for inference is the marginal likelihood

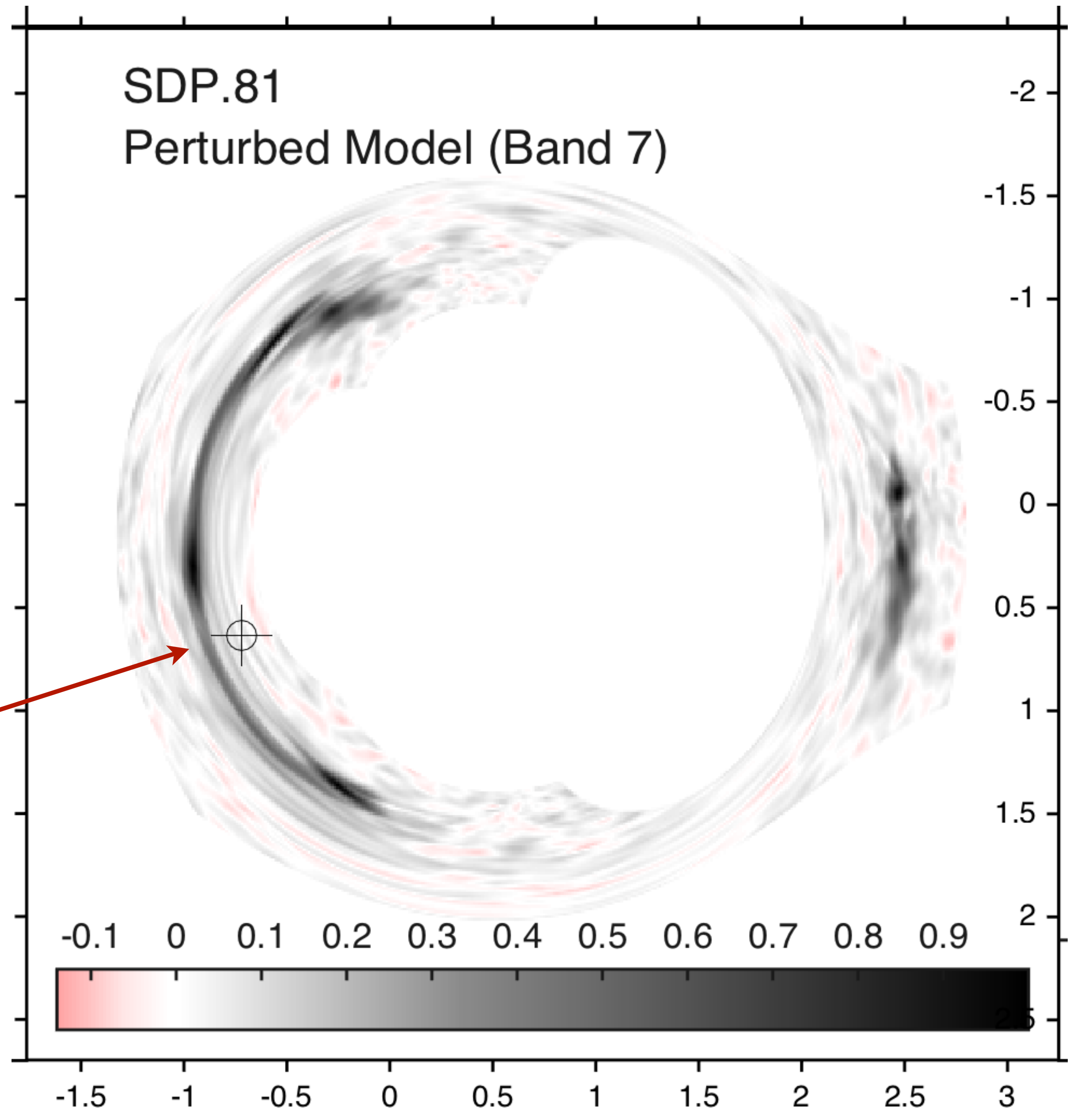
$$p(x \mid \theta) = \boxed{\int \mathrm{d}z_{\mathrm{source}} \int \mathrm{d}z_{\mathrm{lens}} \sum_{n_{\mathrm{sub}}} \int \mathrm{d}^{n_{\mathrm{sub}}} z_{\mathrm{sub}}}\; p\left(z_{\mathrm{src}}\right)\; p\left(z_{\mathrm{lens}}\right)\; p\left(n_{\mathrm{sub}} \mid \theta\right) \prod_{i}^{n_{\mathrm{sub}}} p\left(z_{\mathrm{sub},\,i} \mid \theta\right) \qquad p\left(x \mid z_{\mathrm{src}}, z_{\mathrm{lens}}, \left\{z_{\mathrm{sub},\,i}\right\}\right)$$

Integral over the huge latent space is intractable—*challenge for inference*

Inference

Searches for individual subhalos

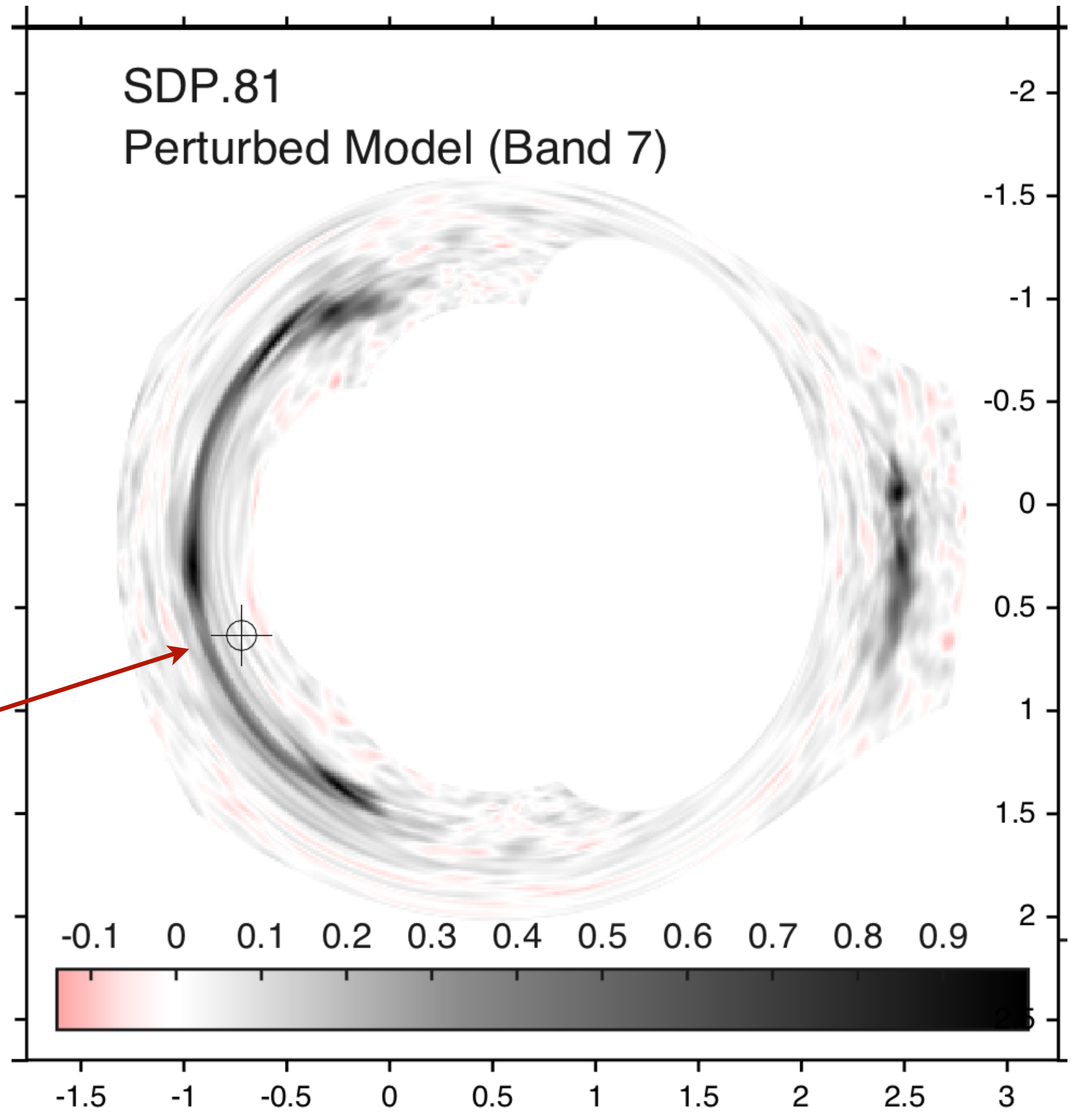
Constraints on subhalo mass function from detections of individual subhalos



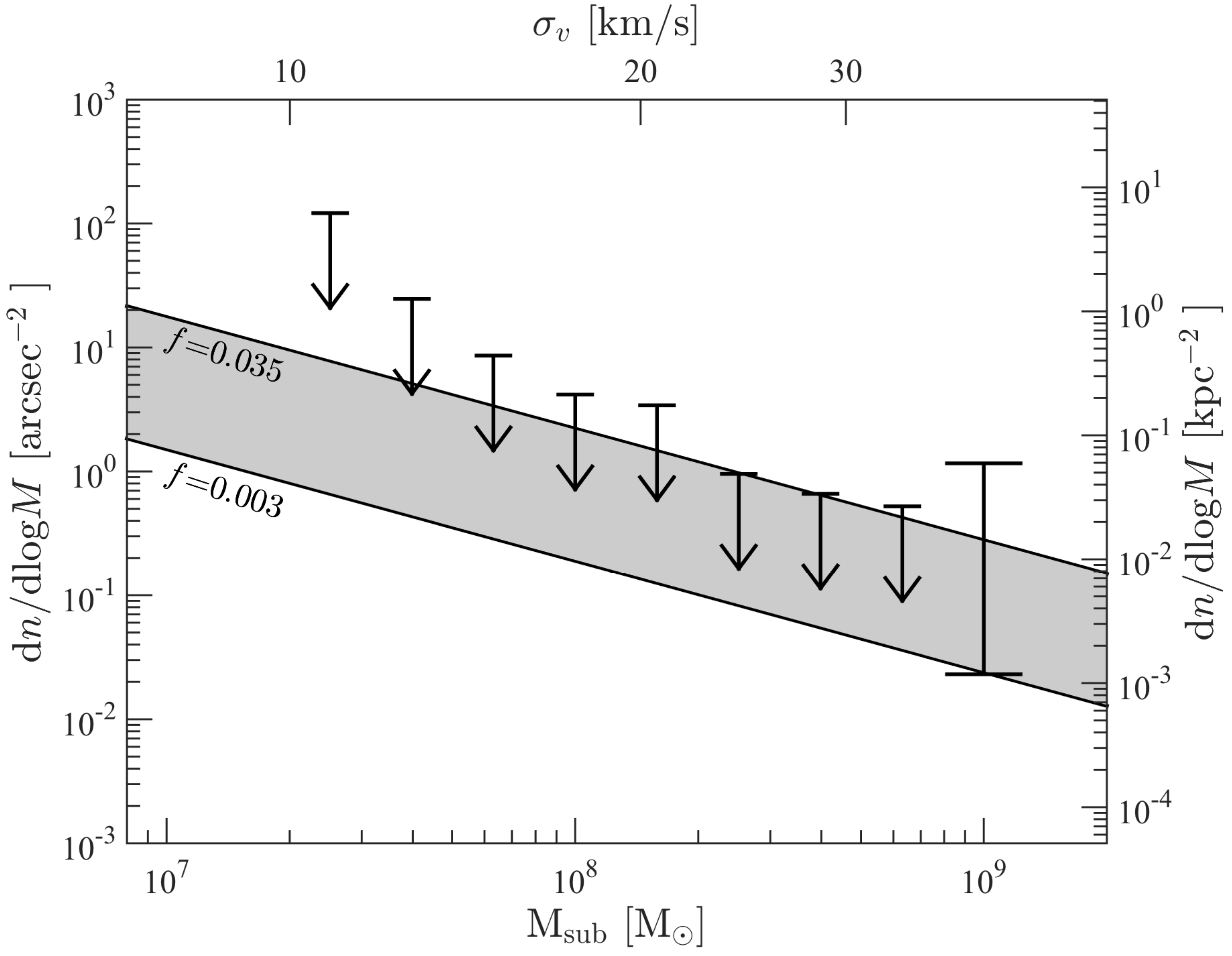
Single detected subhalo

Searches for individual subhalos

Constraints on subhalo mass function from detections of individual subhalos



Single detected subhalo



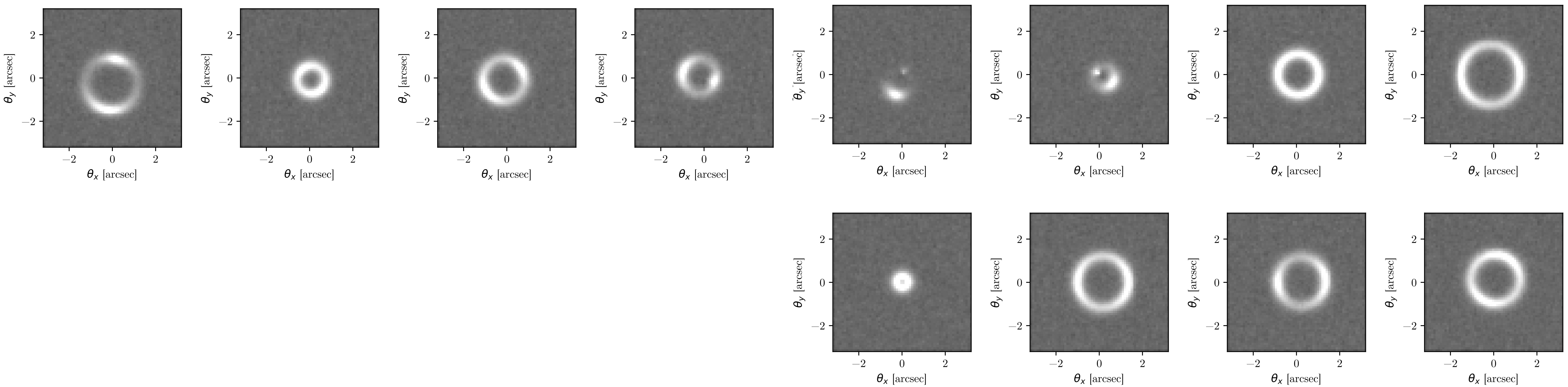
Sensitive to individual, more massive subhalos

Goal: scalable inference of substructure population

$\mathcal{O}(10{,}000)$

Future observatories like the *Euclid* are expected to deliver large samples of galaxy-galaxy lenses

Collett [ApJ 2015]

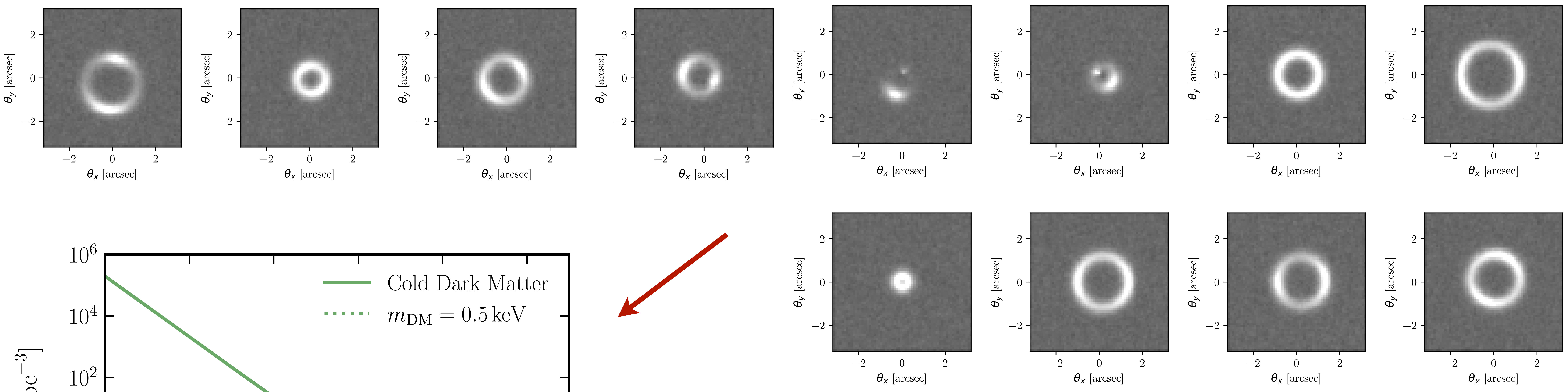


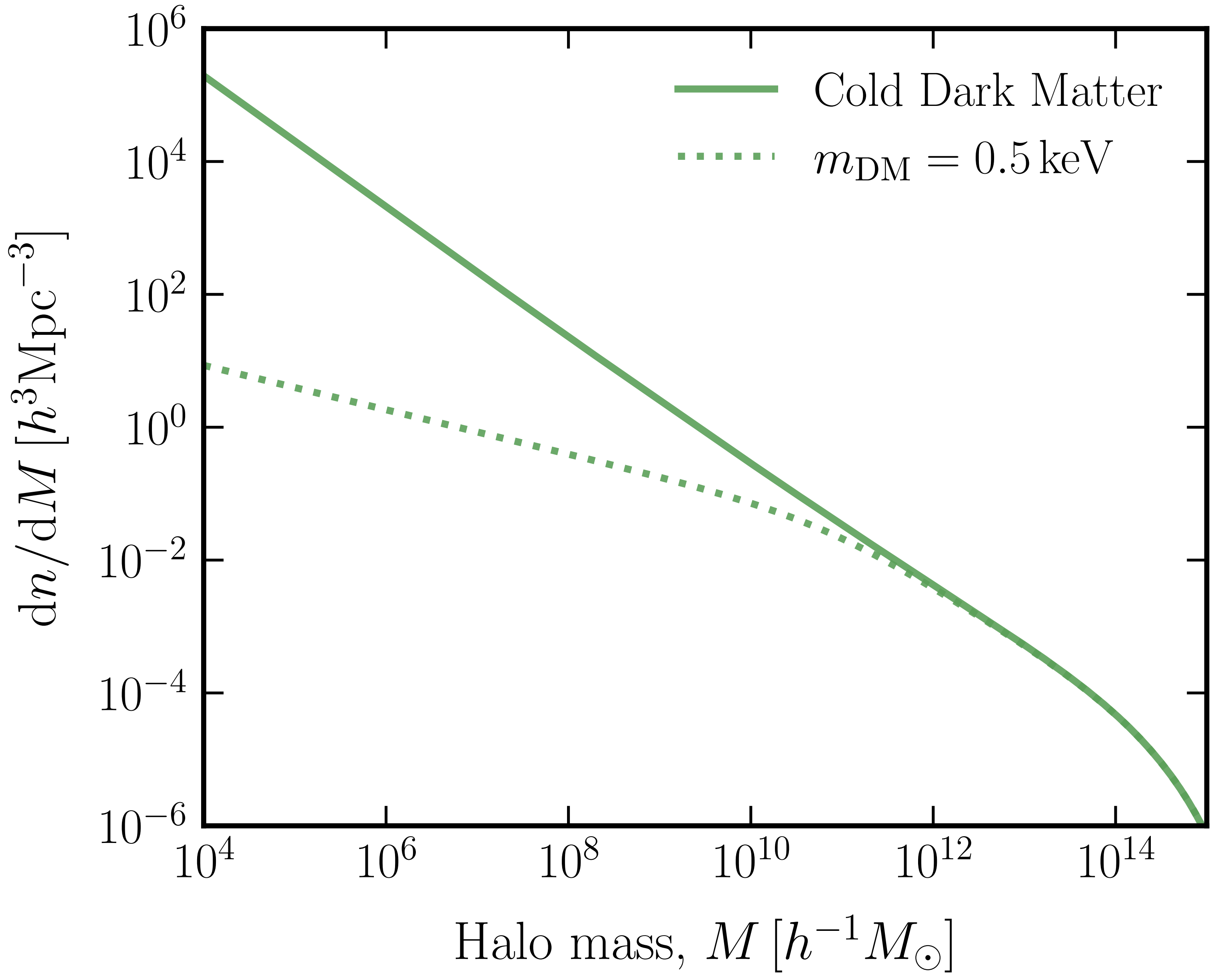
Goal: scalable inference of substructure population

$\mathcal{O}$(10,000)

Future observatories like the *Euclid* are expected to deliver large samples of galaxy-galaxy lenses

Collett [ApJ 2015]



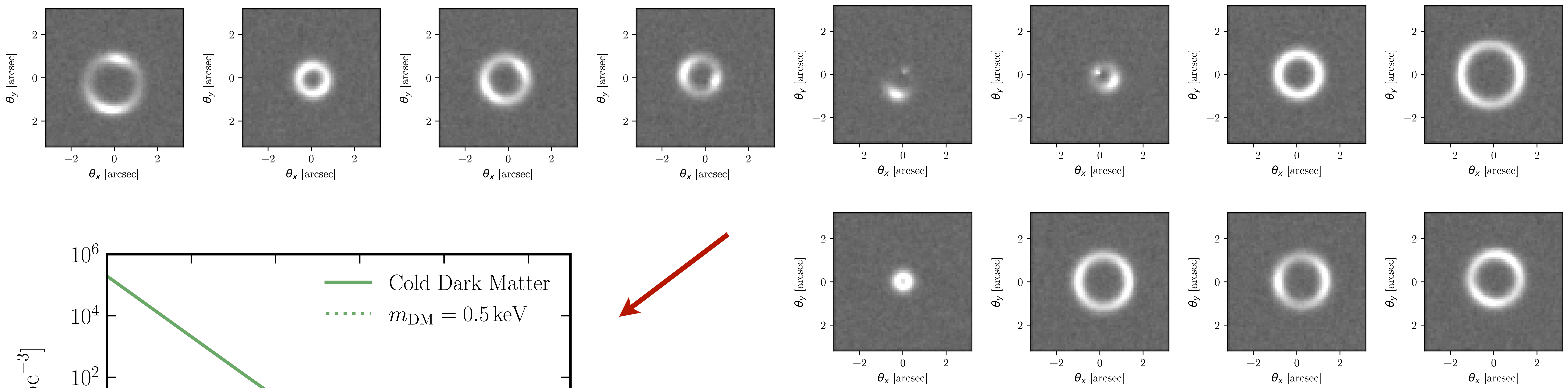


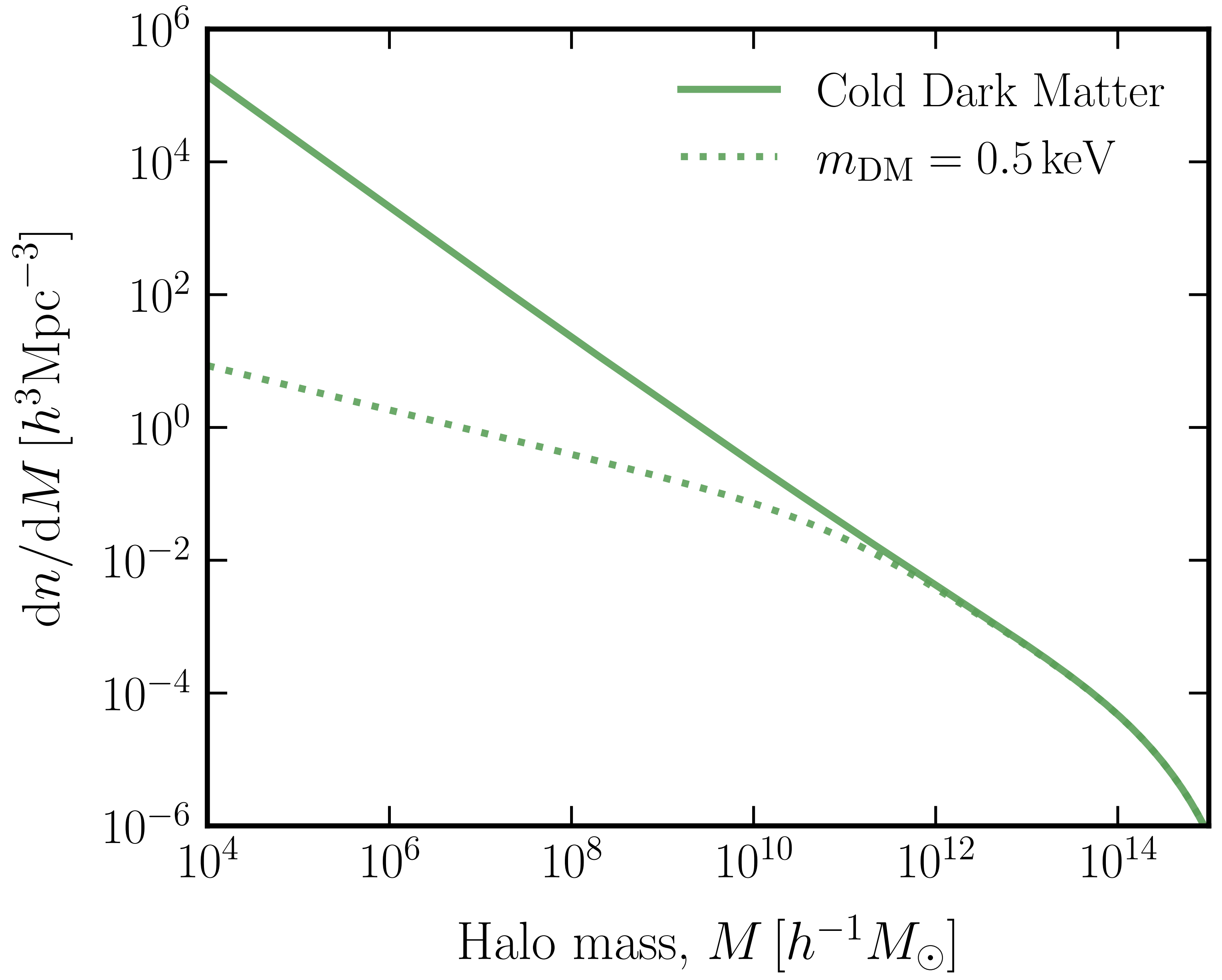
Goal: scalable inference of substructure population

$\mathcal{O}$(10,000)

Future observatories like the *Euclid* are expected to deliver large samples of galaxy-galaxy lenses

Collett [ApJ 2015]

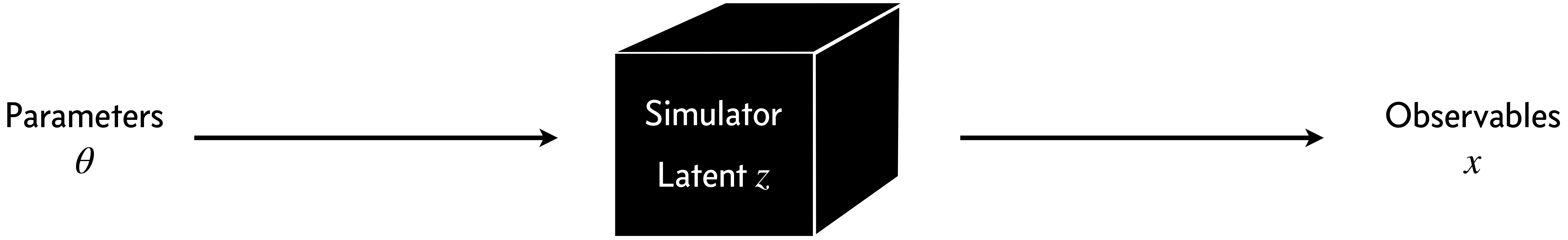




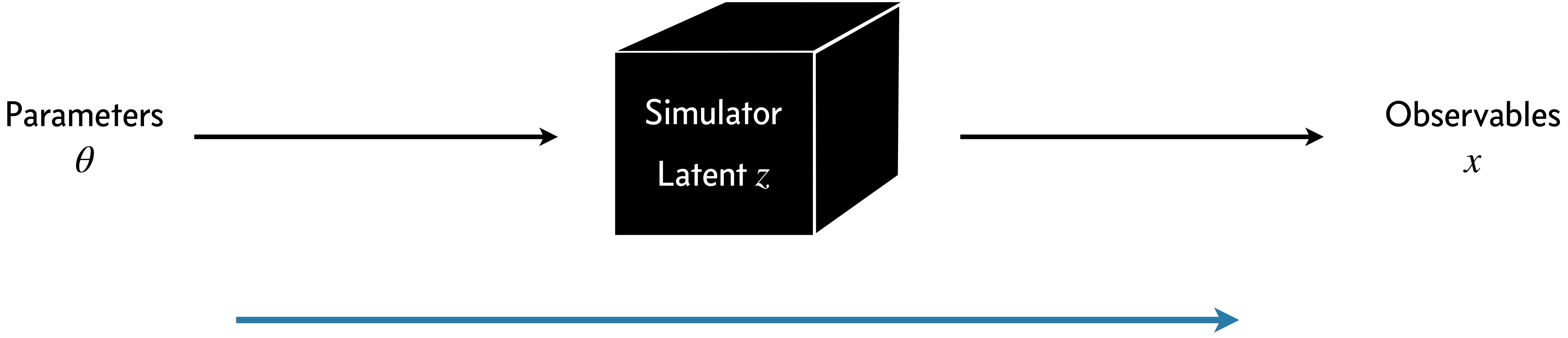
- *Capture maximum information from high-dim data*
- *Scalable to a large sample of lenses*
- *Can deal with a large number of nuisance/latent parameters*

Simulation-based inference (SBI)

Slides inspiration: Johann Brehmer



Simulation-based inference (SBI)

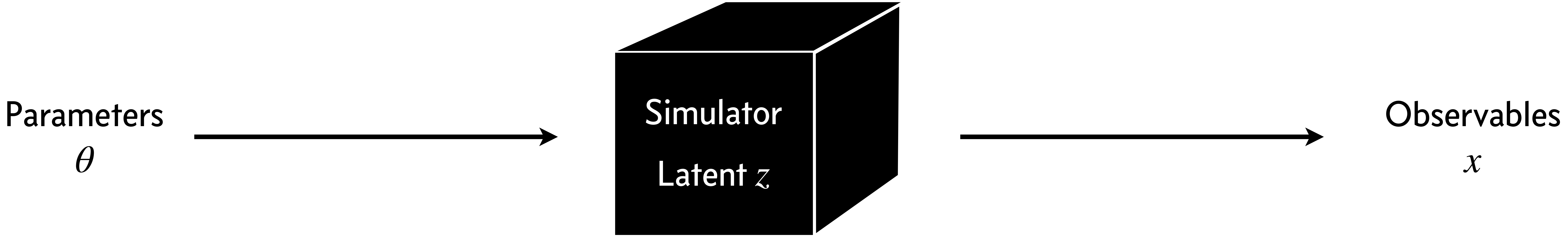


Prediction:

- Well-motivated mechanistic, causal model
- Simulator can generate samples $x \sim p(x \,|\, \theta)$

Simulation-based inference (SBI)

Slides inspiration: Johann Brehmer



Prediction:

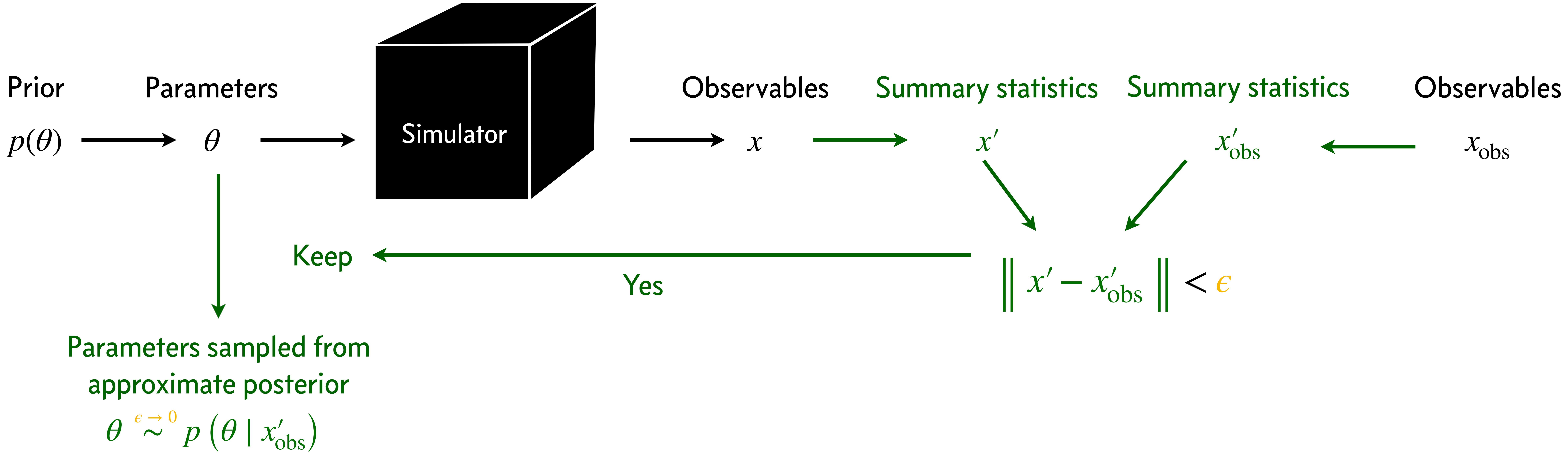
- Well-motivated mechanistic, causal model
- Simulator can generate samples $x \sim p(x \,|\, \theta)$

Inference:

- Likelihood $p(x \,|\, \theta) = \int \mathrm{d}z\, p(x, z \,|\, \theta)$ is intractable
- *Inference is challenging*

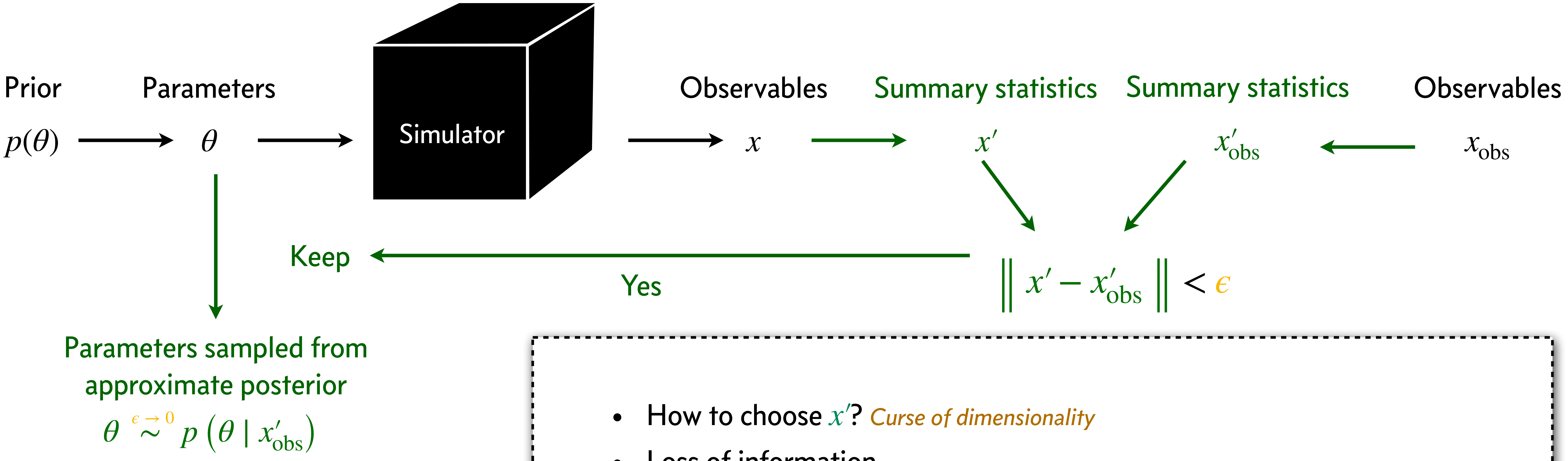
“Traditional” SBI: *Approximate Bayesian Computation*

[Rubin 1984]



"Traditional" SBI: *Approximate Bayesian Computation*

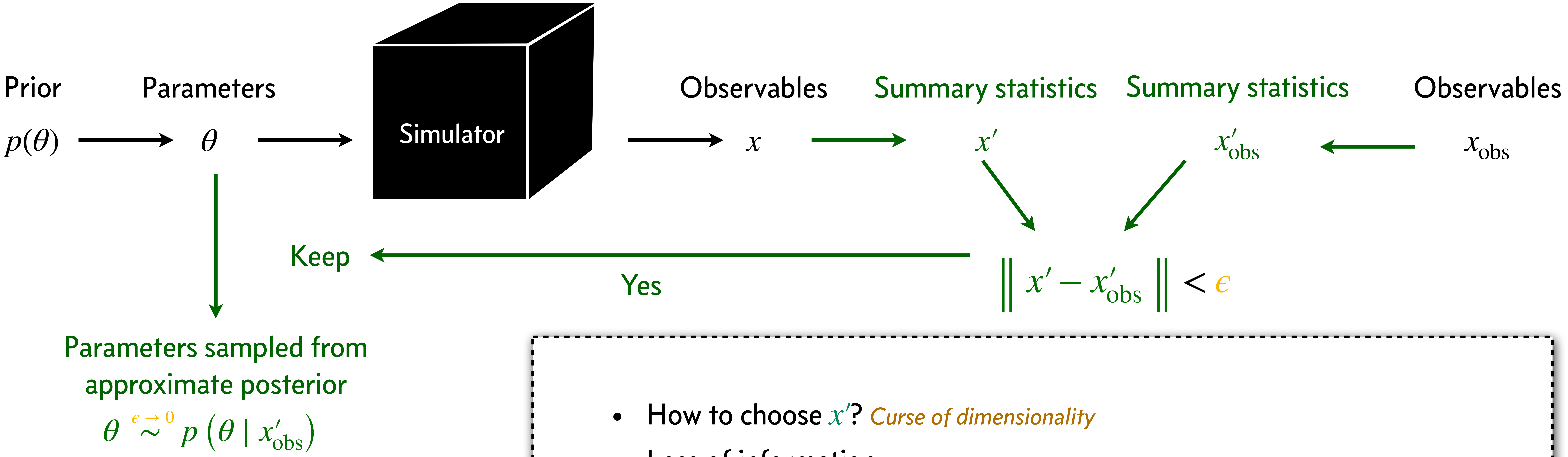
[Rubin 1984]



- How to choose x'? *Curse of dimensionality*
- Loss of information
- How to compare with data? *Likelihood may not be available*
- Need to re-run pipeline for new data or new prior

“Traditional” SBI: *Approximate Bayesian Computation*

[Rubin 1984]



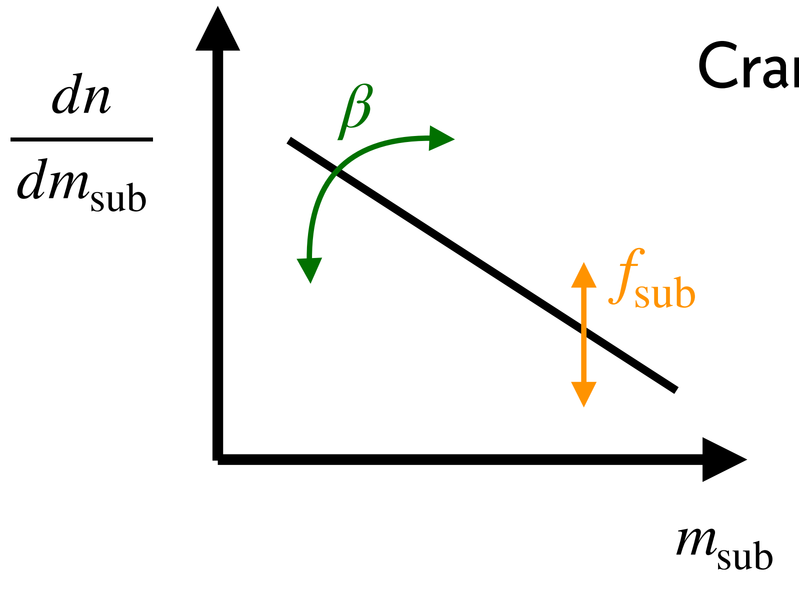
- How to choose x'? *Curse of dimensionality*
- Loss of information
- How to compare with data? *Likelihood may not be available*
- Need to re-run pipeline for new data or new prior

Lots of recent progress using ML

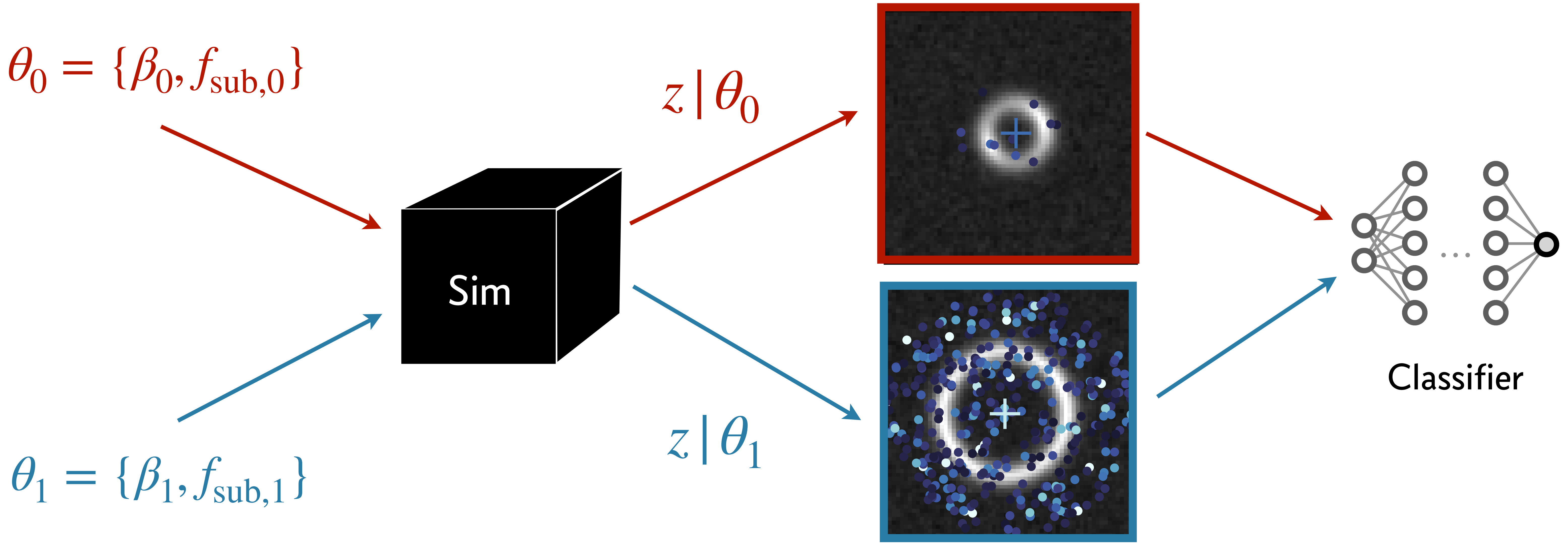
see Cranmer, Brehmer, Louppe [PNAS 2020] for a review

The likelihood-ratio trick

We can train a classifier between two sets of simulated samples

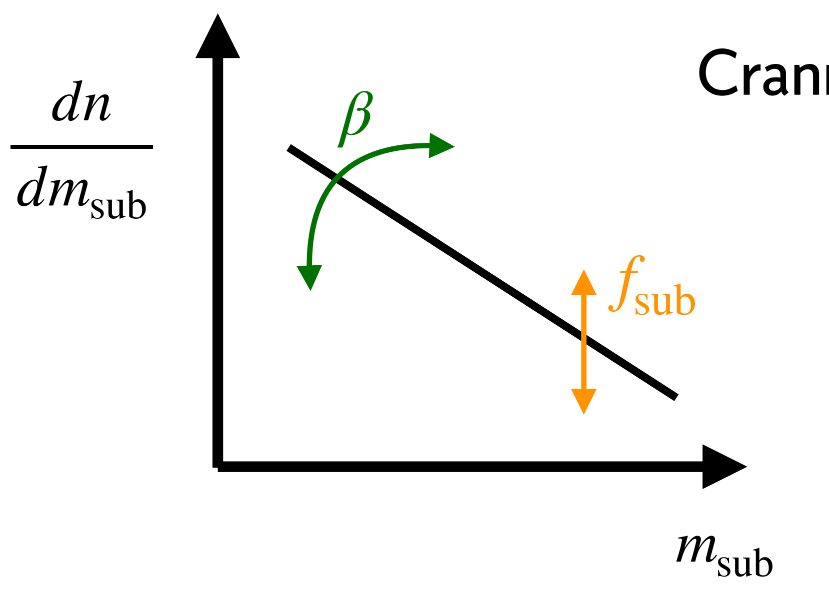


Cranmer, Pavez, Louppe [arXiv 2015]
Brehmer et al [PRL, PRD 2018]
Stoye et al [arXiv 2018]
Hermans et al [ICML 2020]
+ others

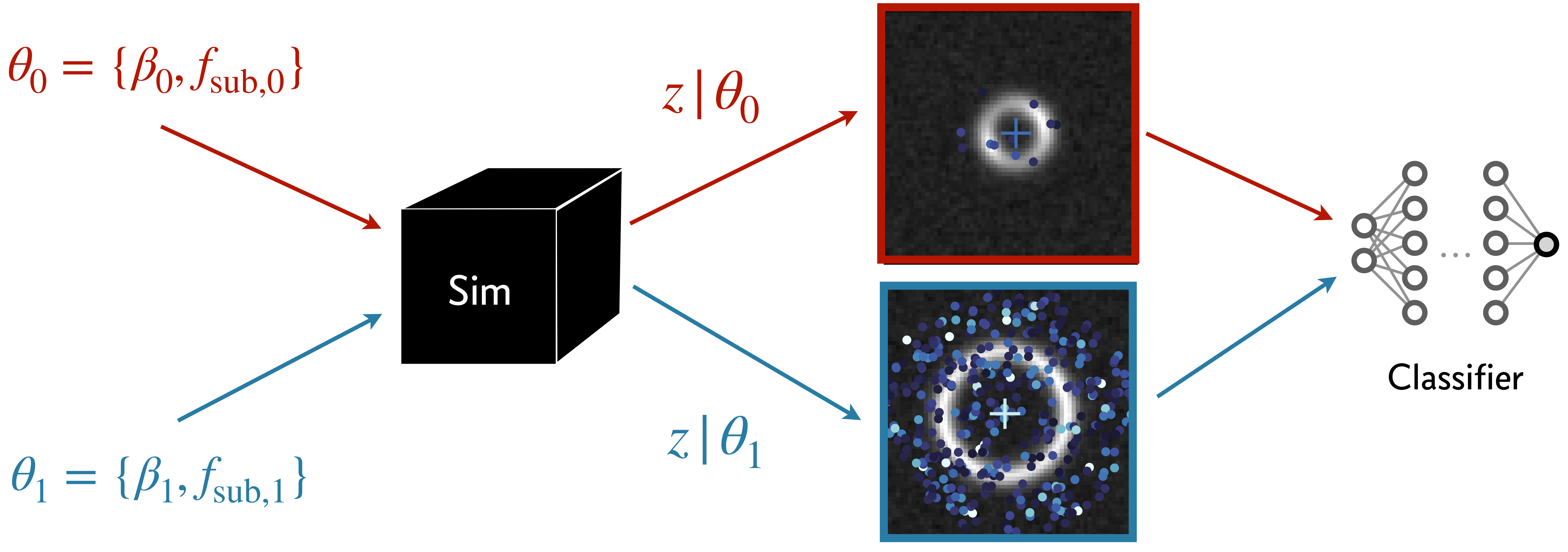


The likelihood-ratio trick

We can train a classifier between two sets of simulated samples

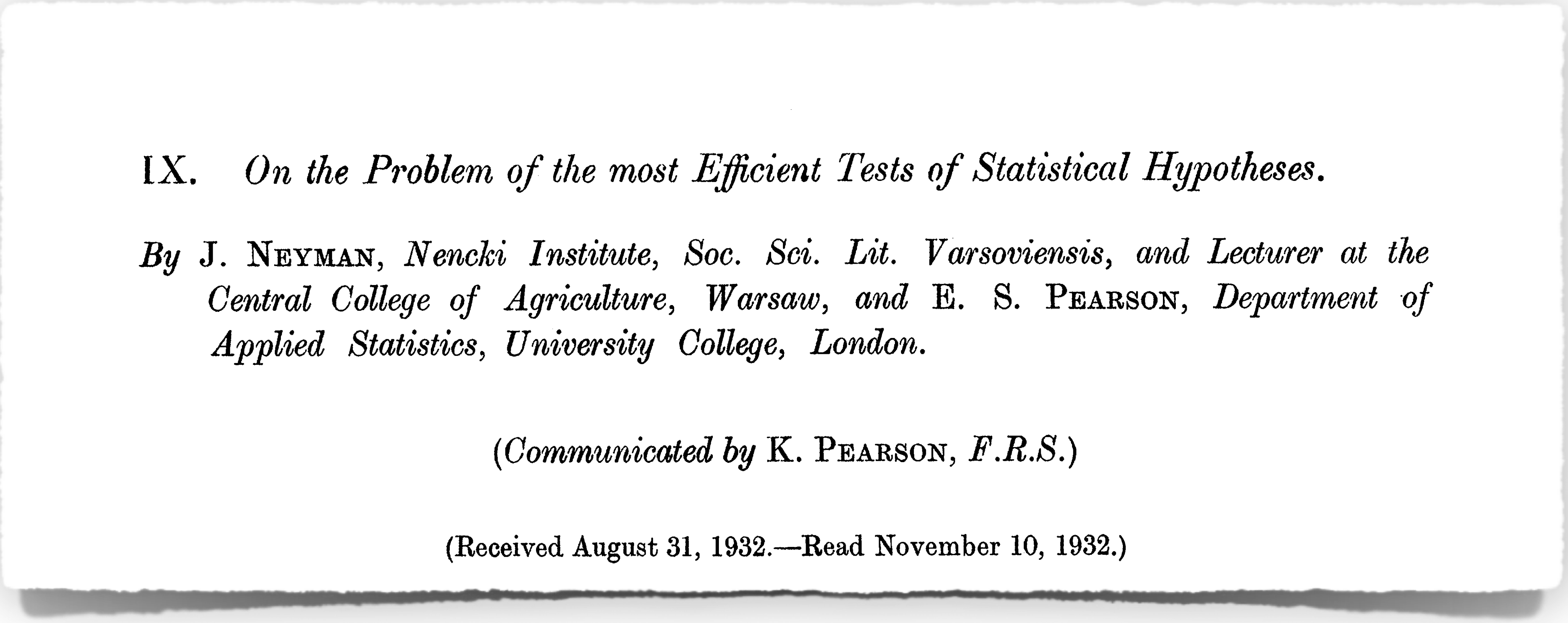


Cranmer, Pavez, Louppe [arXiv 2015]
Brehmer et al [PRL, PRD 2018]
Stoye et al [arXiv 2018]
Hermans et al [ICML 2020]
+ others



Decision function

$$d(x) \rightarrow \frac{p\left(x \mid \theta_0\right)}{p\left(x \mid \theta_0\right) + p\left(x \mid \theta_1\right)}$$

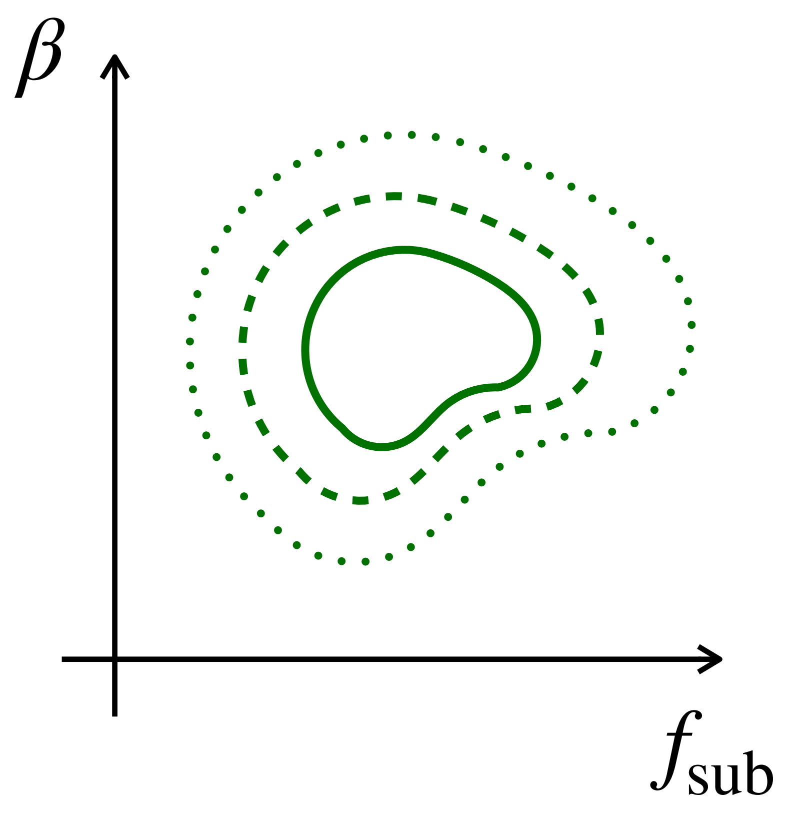


IX. *On the Problem of the most Efficient Tests of Statistical Hypotheses.*

By J. NEYMAN, *Nencki Institute, Soc. Sci. Lit. Varsoviensis, and Lecturer at the Central College of Agriculture, Warsaw, and* E. S. PEARSON, *Department of Applied Statistics, University College, London.*

(*Communicated by* K. PEARSON, *F.R.S.*)

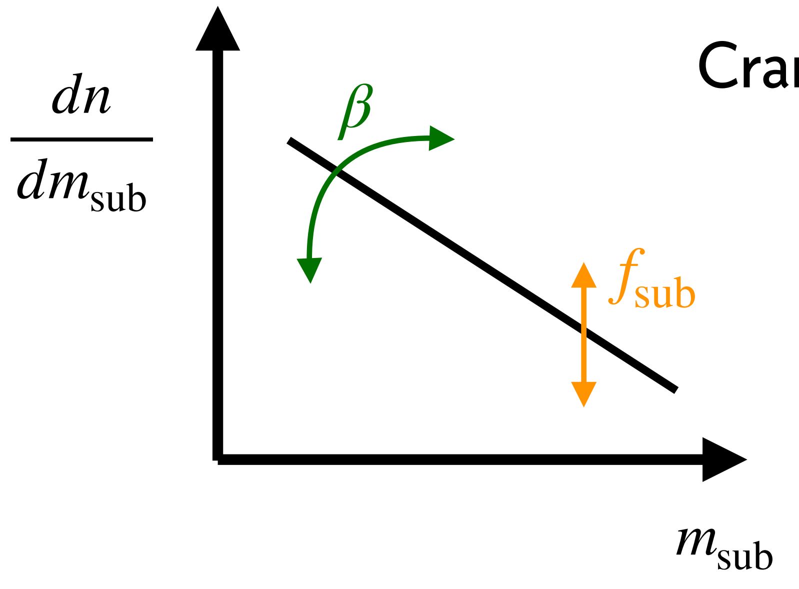
(Received August 31, 1932.—Read November 10, 1932.)



Estimator for likelihood ratio

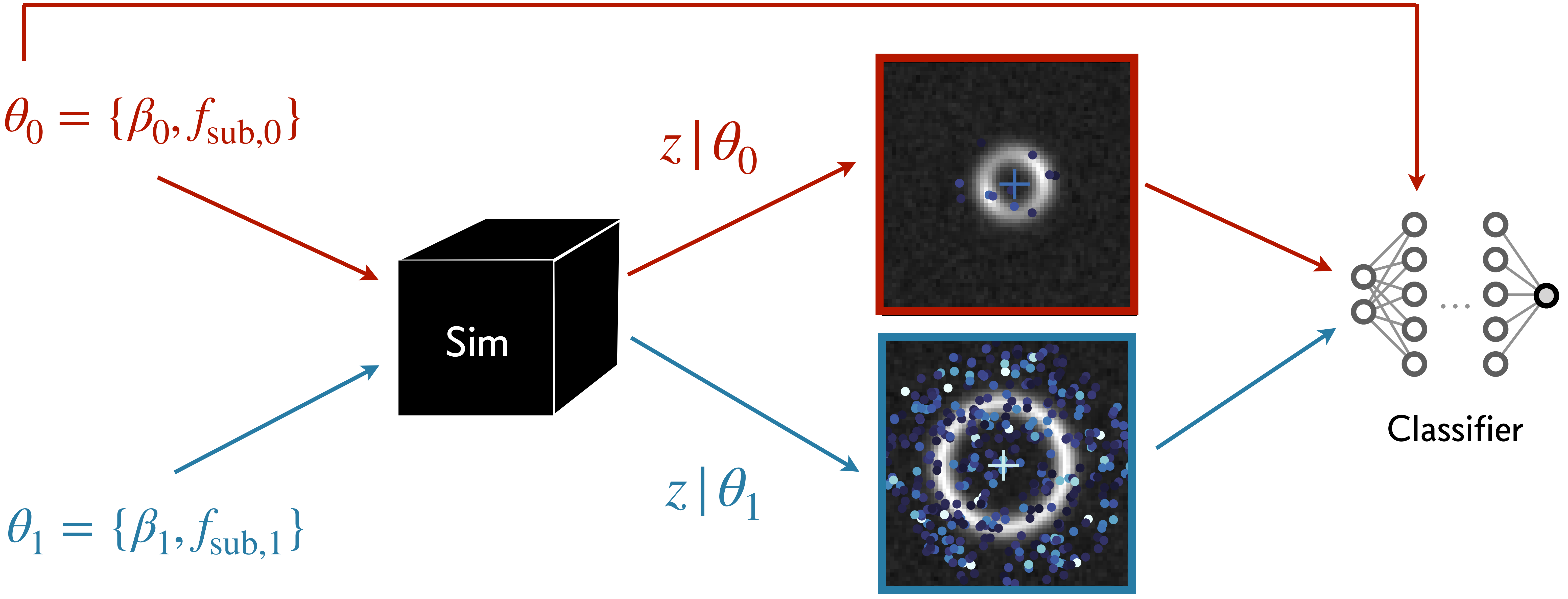
$$\hat{r}(x) = \frac{d(x)}{1 - d(x)} \rightarrow \frac{p\left(x \mid \theta_0\right)}{p\left(x \mid \theta_1\right)}$$

The likelihood-ratio trick



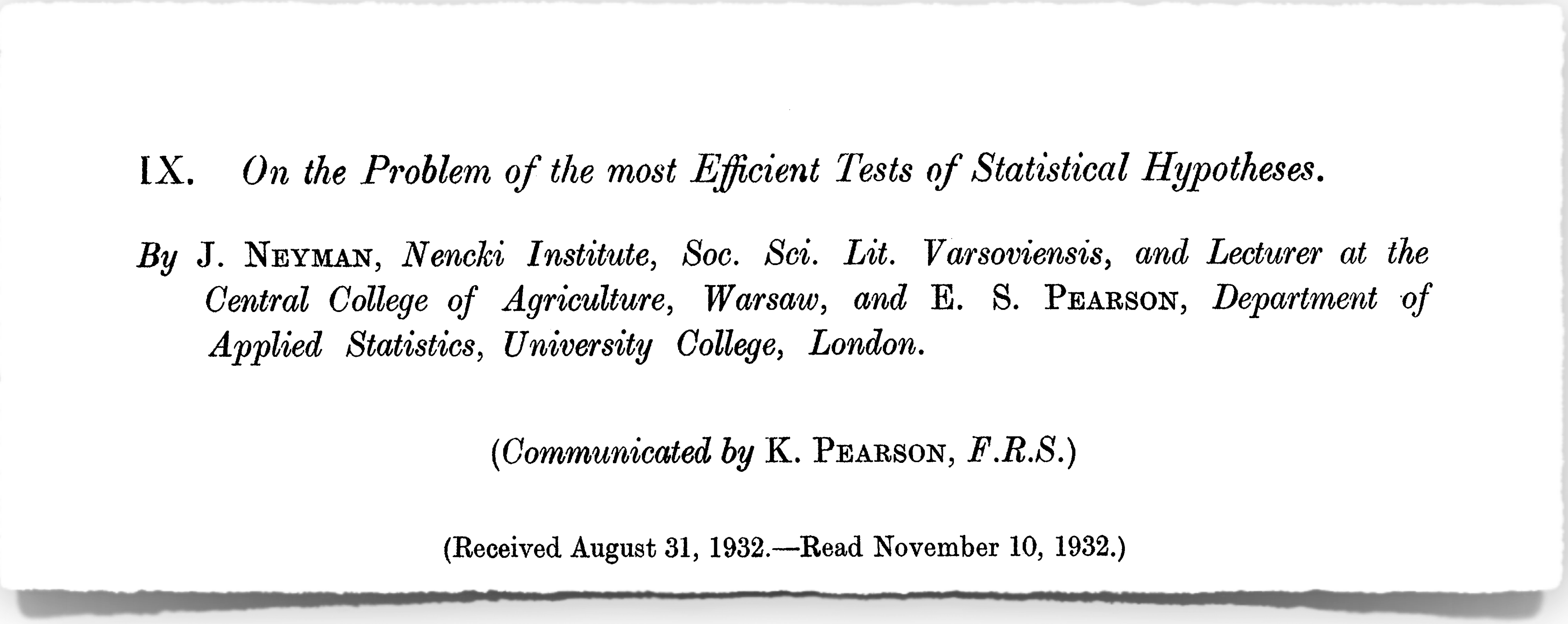
Cranmer, Pavez, Louppe [arXiv 2015]
Brehmer et al [PRL, PRD 2018]
Stoye et al [arXiv 2018]
Hermans et al [ICML 2020]
+ others

We can train a classifier between two sets of simulated samples



Decision function

$$d(x,\theta) \to \frac{p\left(x \mid \theta\right)}{p\left(x \mid \theta\right) + p\left(x \mid \theta_{\rm ref}\right)}$$

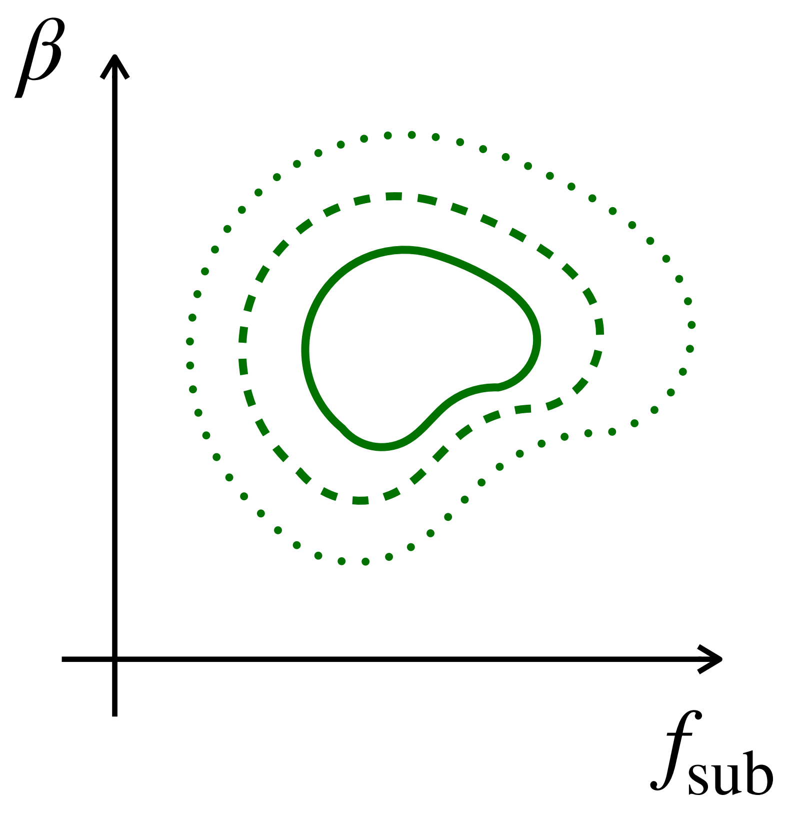


IX. *On the Problem of the most Efficient Tests of Statistical Hypotheses.*

By J. Neyman, *Nencki Institute, Soc. Sci. Lit. Varsoviensis, and Lecturer at the Central College of Agriculture, Warsaw, and* E. S. Pearson, *Department of Applied Statistics, University College, London.*

(*Communicated by* K. Pearson, *F.R.S.*)

(Received August 31, 1932.—Read November 10, 1932.)



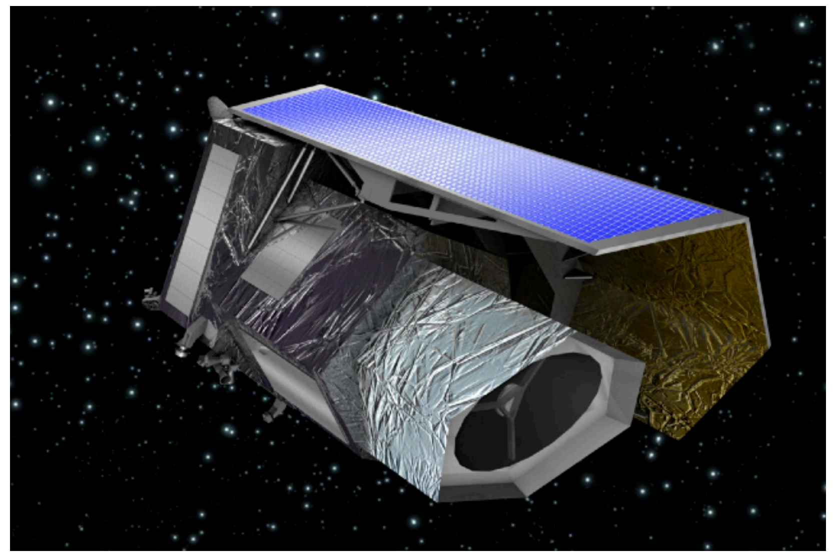
Estimator for likelihood ratio

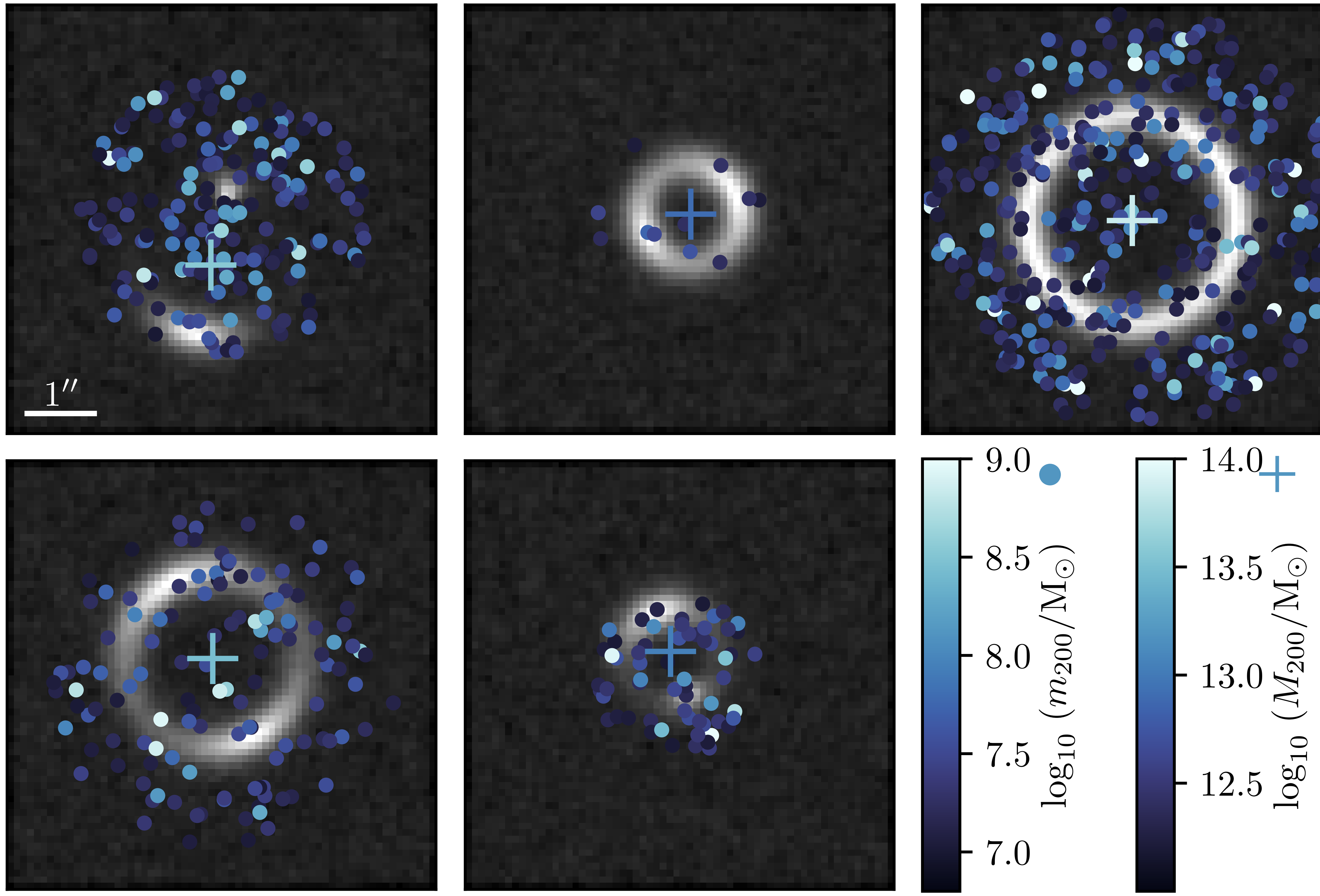
$$\hat{r}(x \mid \theta) = \frac{d(x,\theta)}{1 - d(x,\theta)} \to \frac{p\left(x \mid \theta\right)}{p\left(x \mid \theta_{\rm ref}\right)}$$

Proof of principle

Brehmer, SM, Hermans, Louppe, Cranmer [ApJ 2019]

Simulated ensemble of galaxy-galaxy lenses observable by *Euclid*



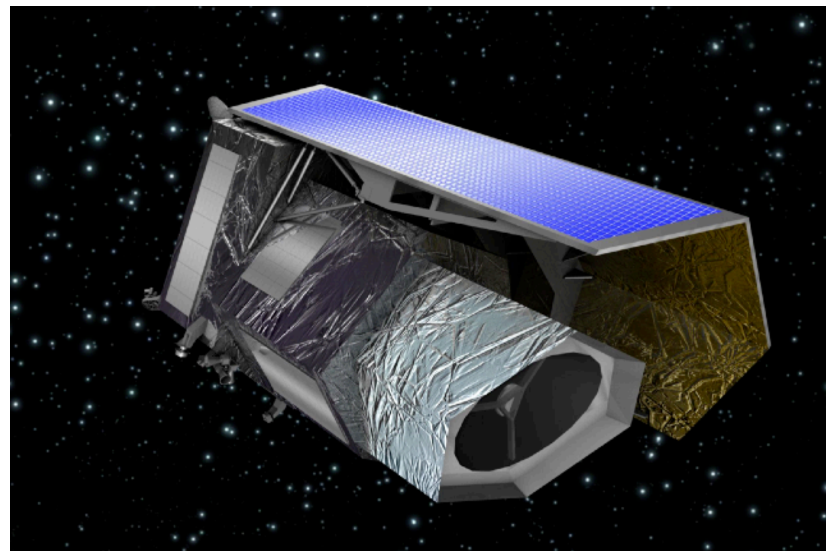


$f_{\mathrm{sub}} = 5\,\%\,,\ \beta = -0.9$

Proof of principle

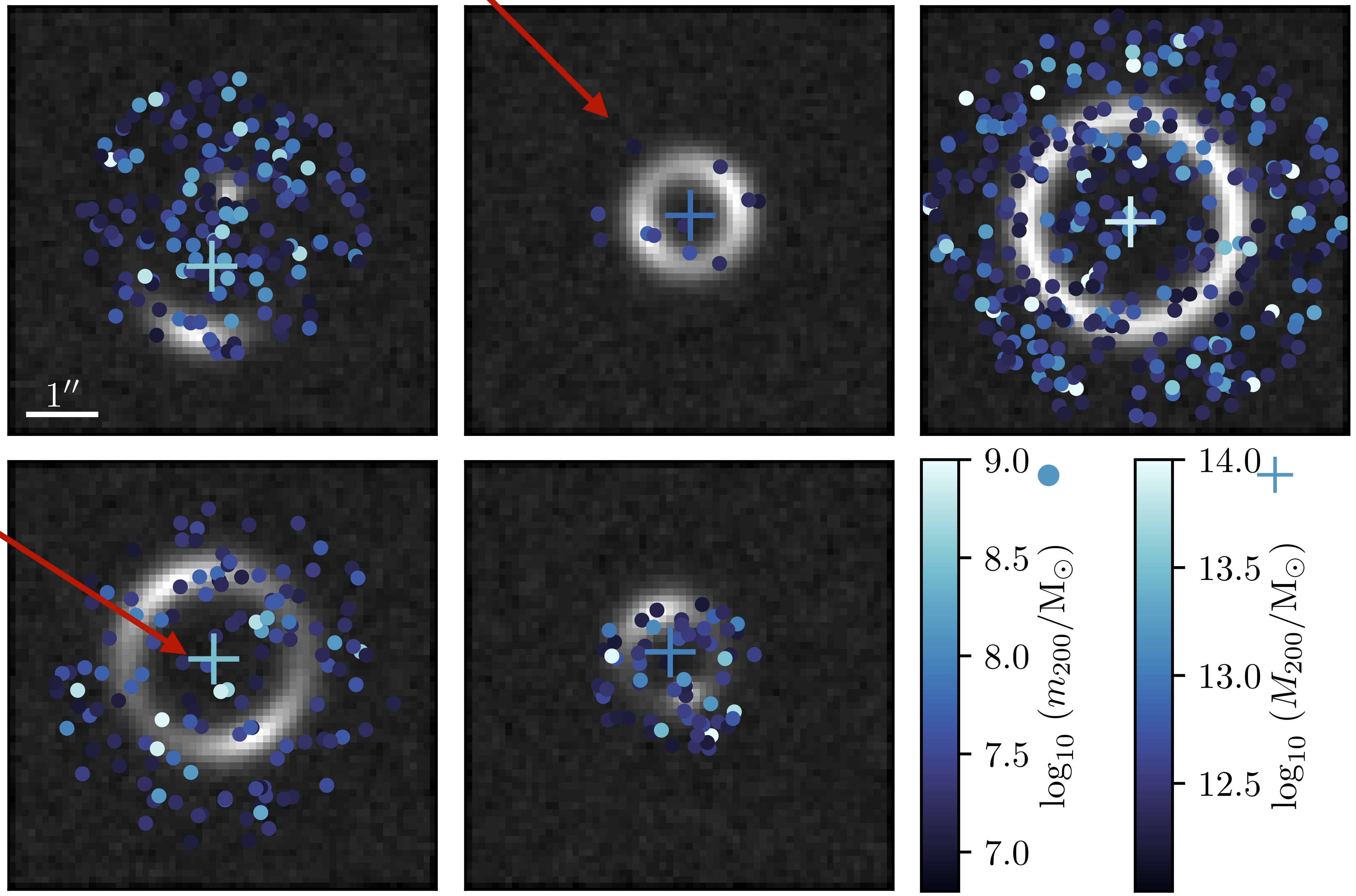
Brehmer, SM, Hermans, Louppe, Cranmer [ApJ 2019]

Simulated ensemble of galaxy-galaxy lenses observable by *Euclid*



Lensing host galaxies at $z \sim 0.5 - 1$

Galaxy sources at $z \sim 1.5$

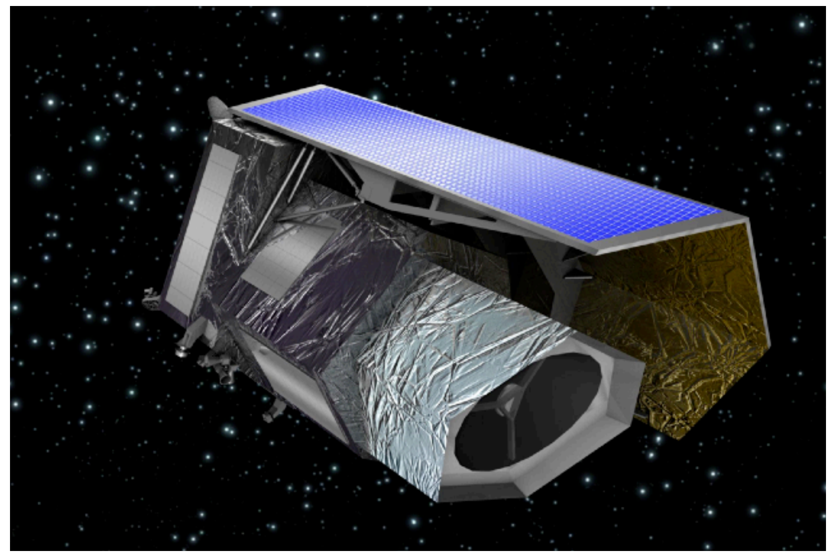


$f_{\mathrm{sub}} = 5\,\%\,,\ \beta = -0.9$

Proof of principle

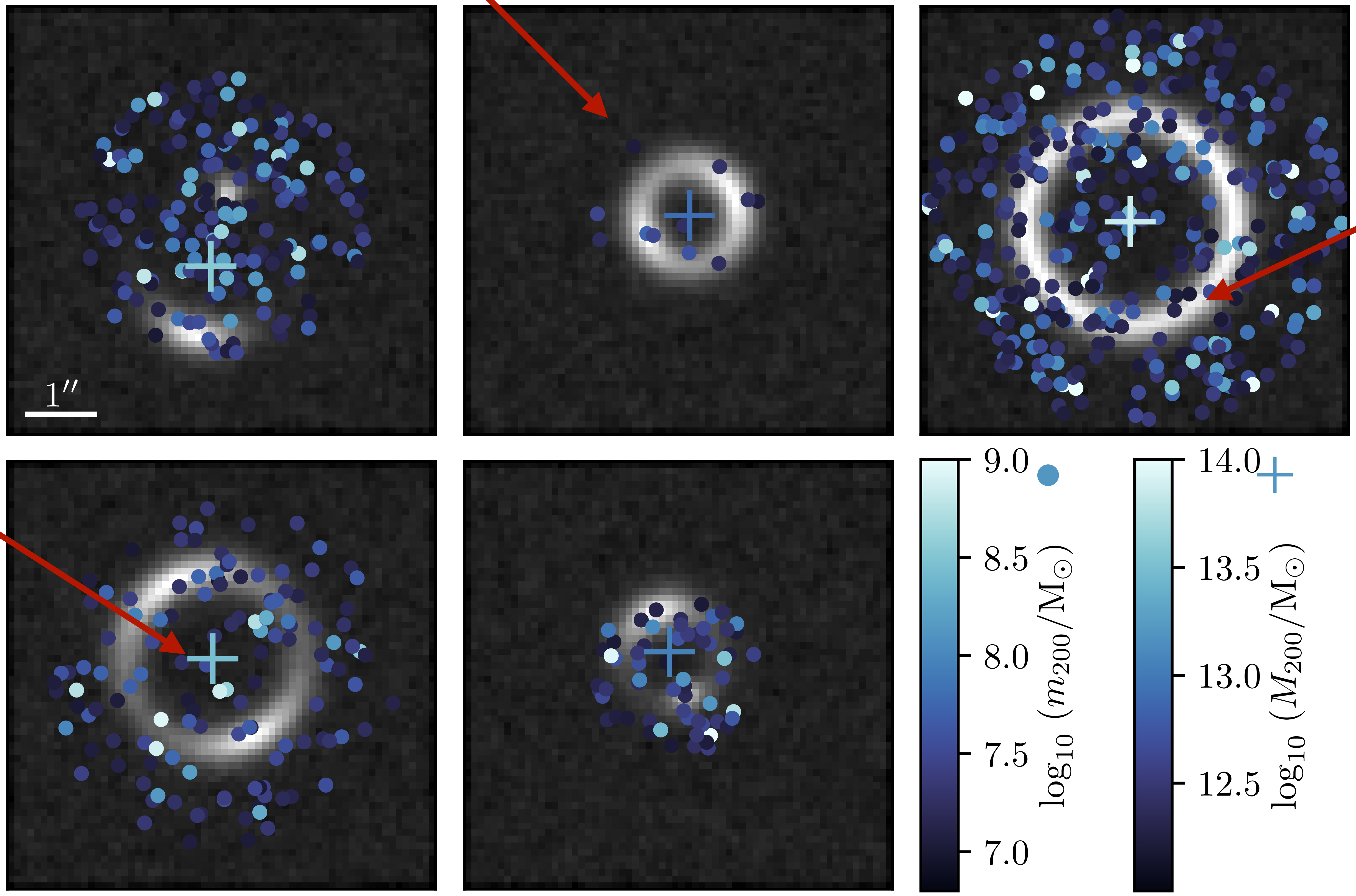
Brehmer, SM, Hermans, Louppe, Cranmer [ApJ 2019]

Simulated ensemble of galaxy-galaxy lenses observable by *Euclid*

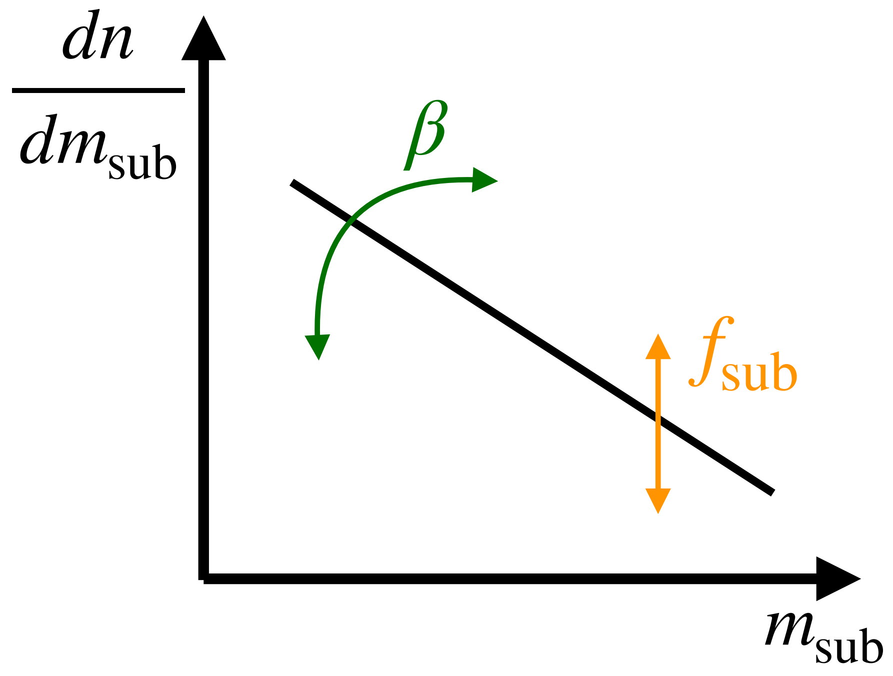


Lensing host galaxies at $z \sim 0.5 - 1$

Galaxy sources at $z \sim 1.5$



Subhalo mass function with two parameters

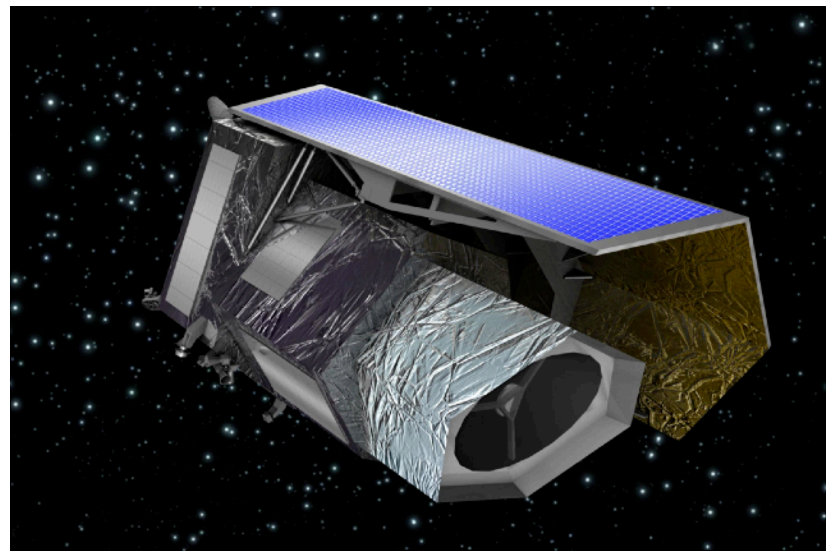


$f_{\mathrm{sub}} = 5\,\%\,,\ \beta = -0.9$

Proof of principle

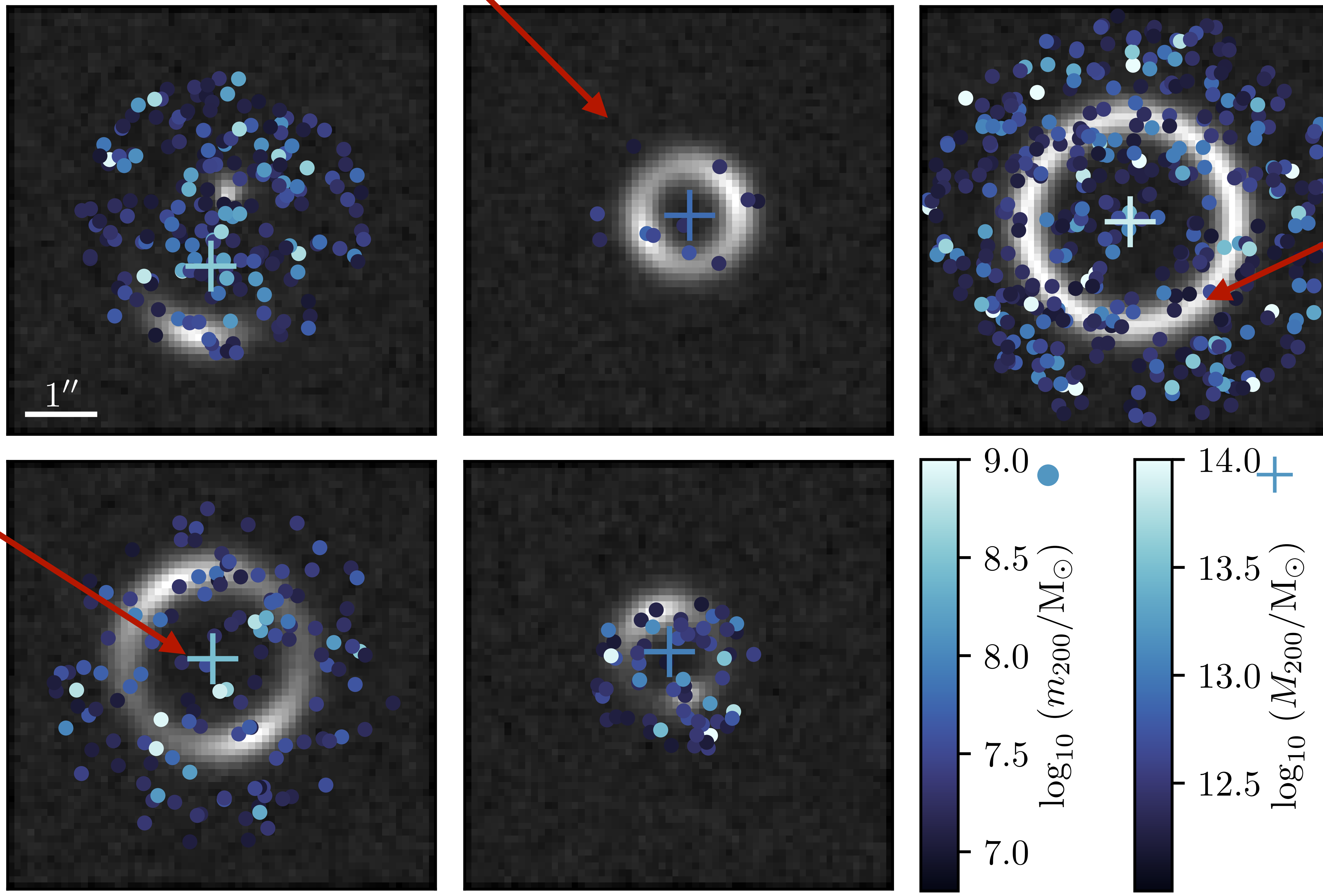
Brehmer, SM, Hermans, Louppe, Cranmer [ApJ 2019]

Simulated ensemble of galaxy-galaxy lenses observable by *Euclid*

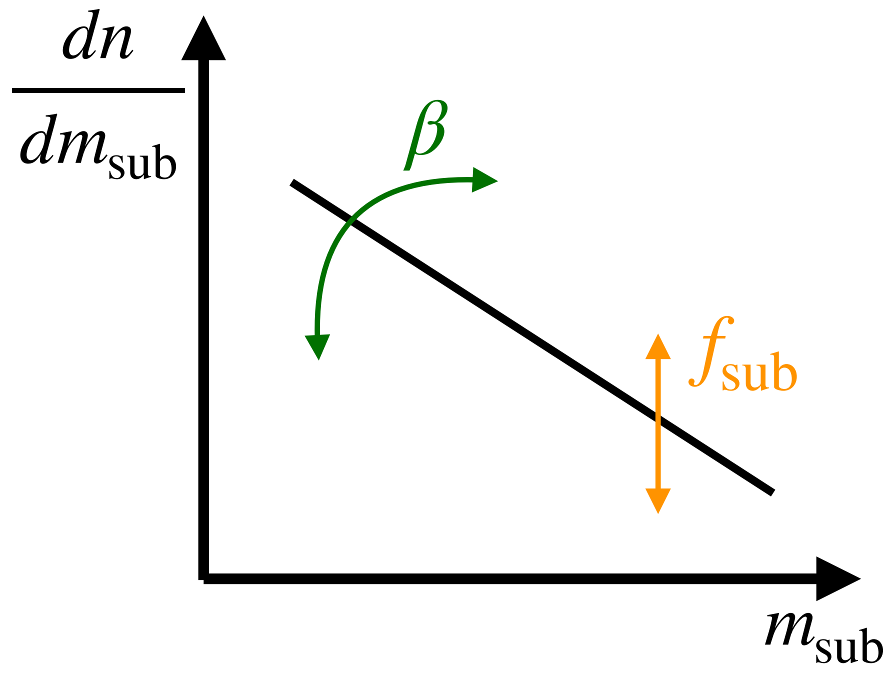


Lensing host galaxies at $z \sim 0.5 - 1$

Galaxy sources at $z \sim 1.5$

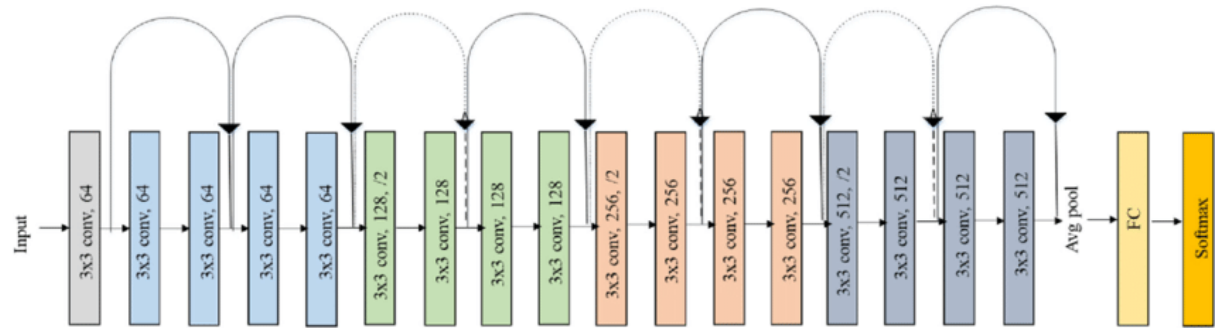


Subhalo mass function with two parameters



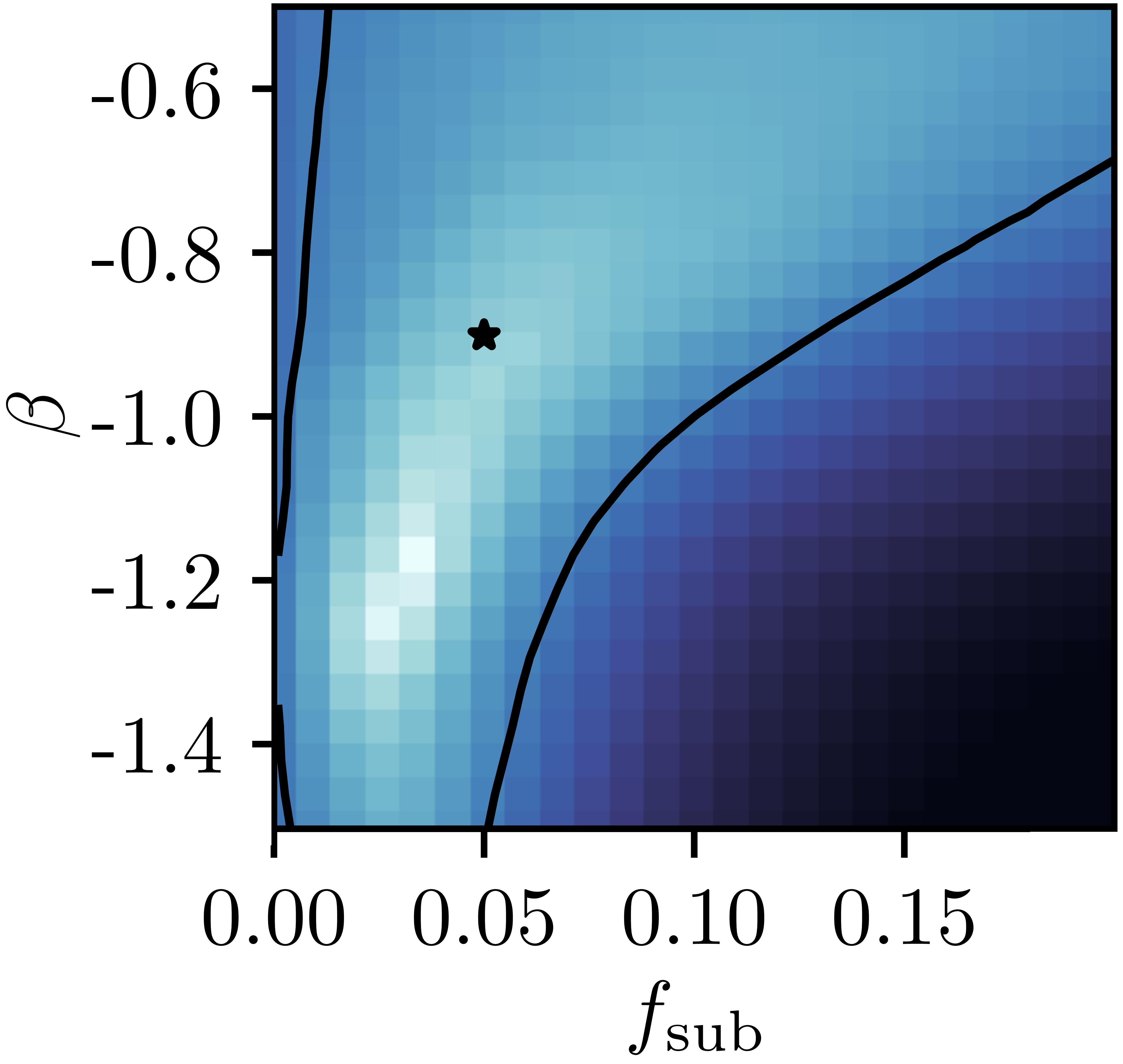
$f_{\mathrm{sub}} = 5\,\%\,,\ \beta = -0.9$

ResNet-18 architecture



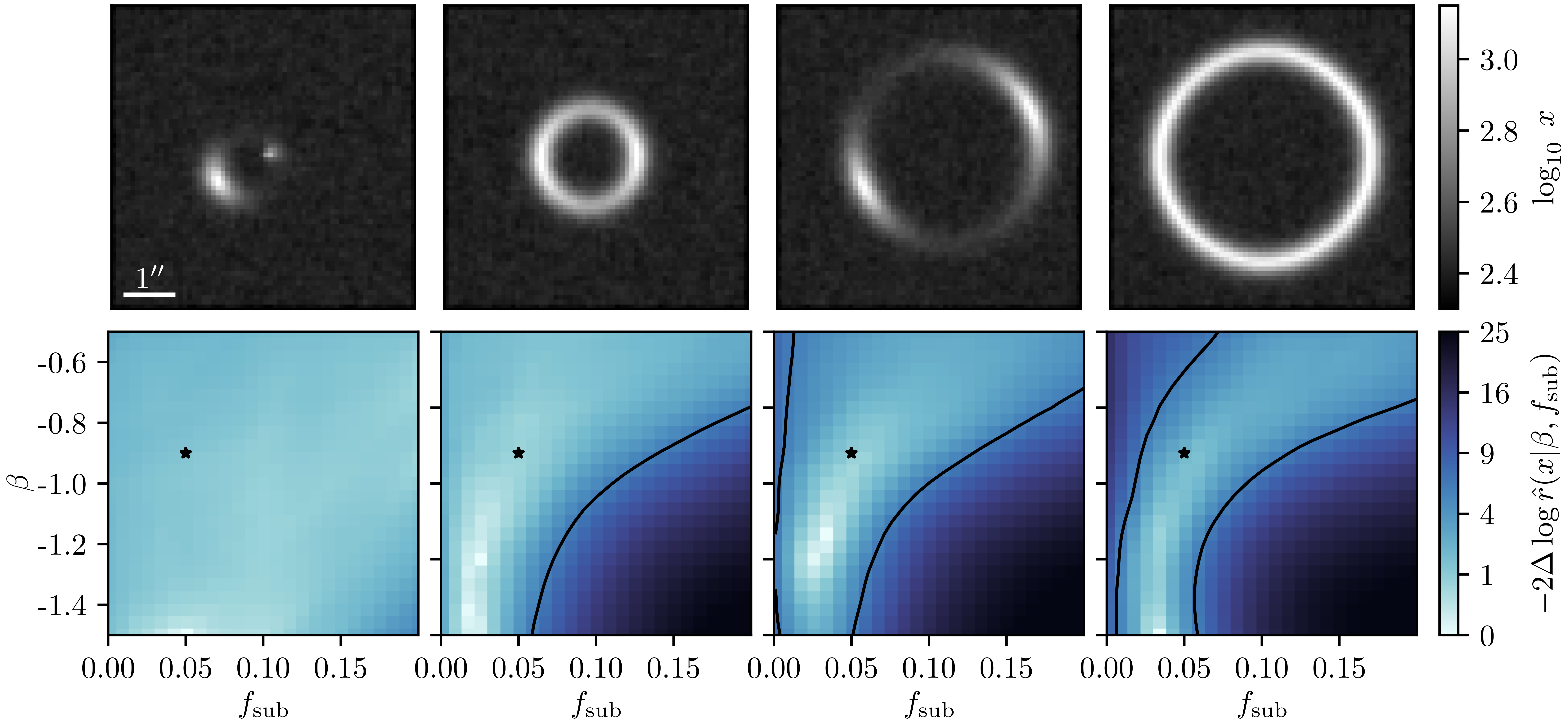
Inferred likelihood ratios

$f_{\rm sub} = 5\,\%\,,\, \beta = -\,0.9$



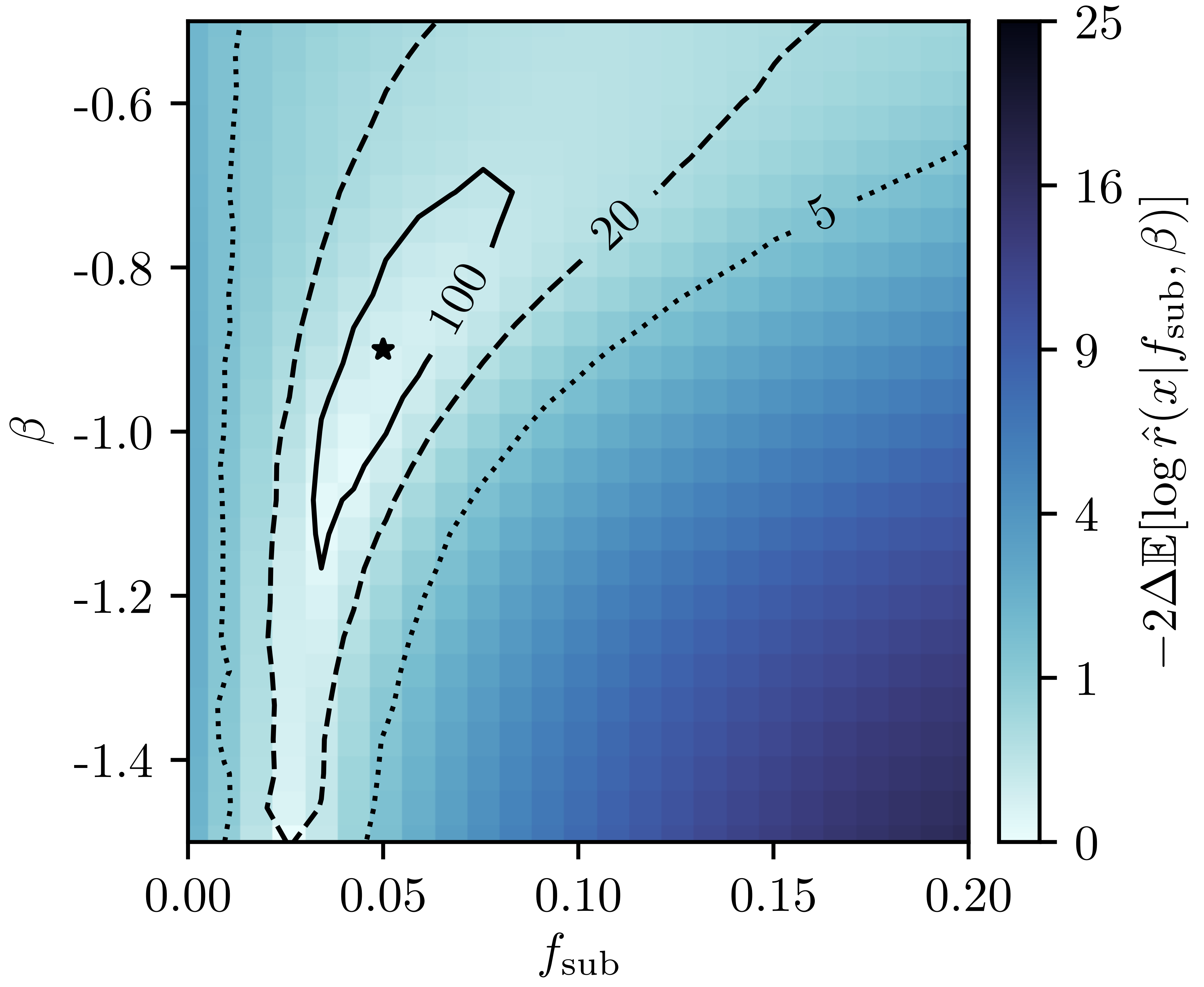
Inferred likelihood ratios: *individual images*

$f_{\rm sub} = 5\,\%\,,\ \beta = -\,0.9$



Inferred likelihood ratios: *stacking images*

$f_{\rm sub} = 5\,\%\,,\ \beta = -\,0.9$



Combination of lenses can place powerful constraints on subhalo mass function

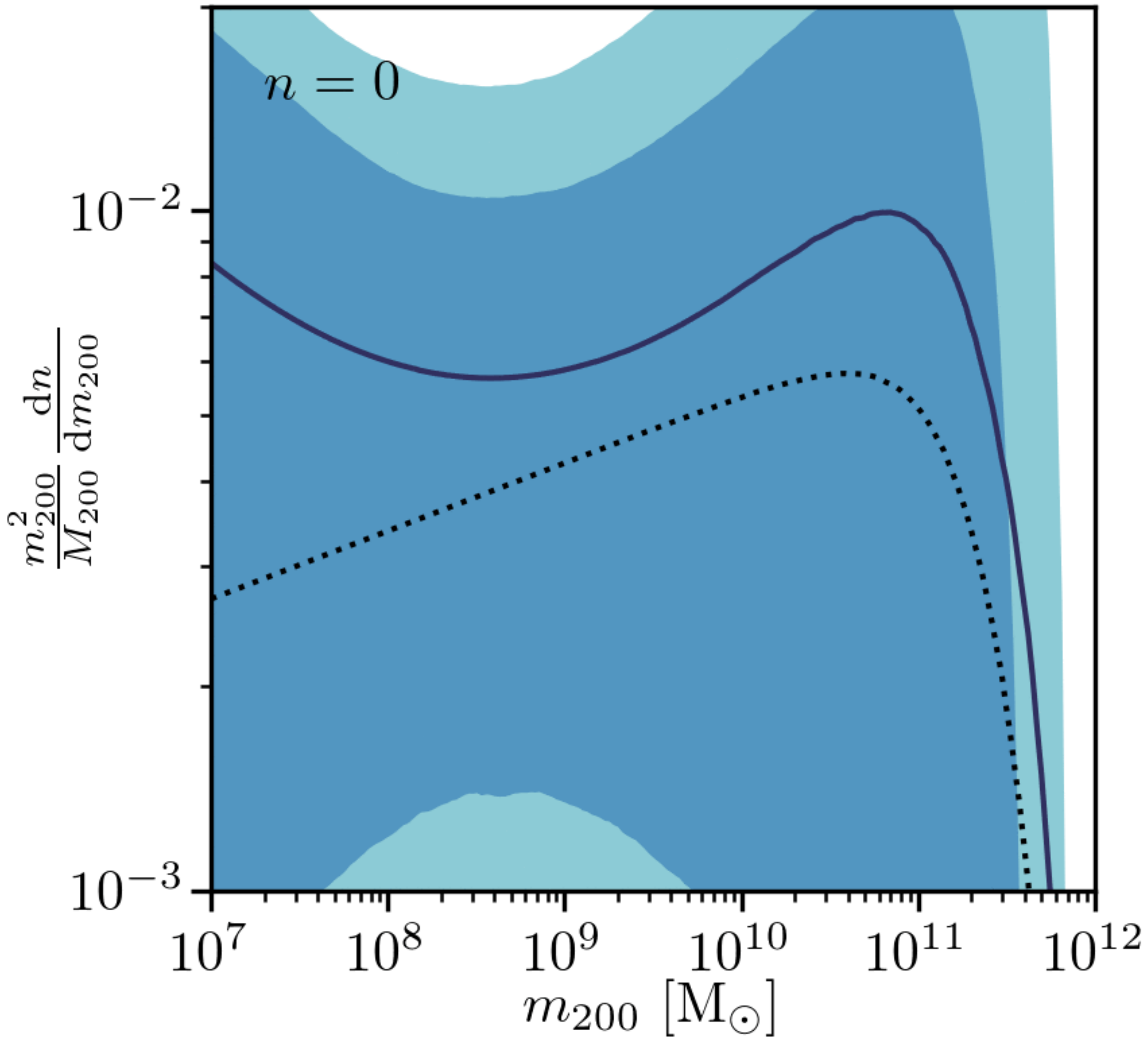
Bayesian inference

$f_{\rm sub} = 5\,\%\,,\, \beta = -0.9$

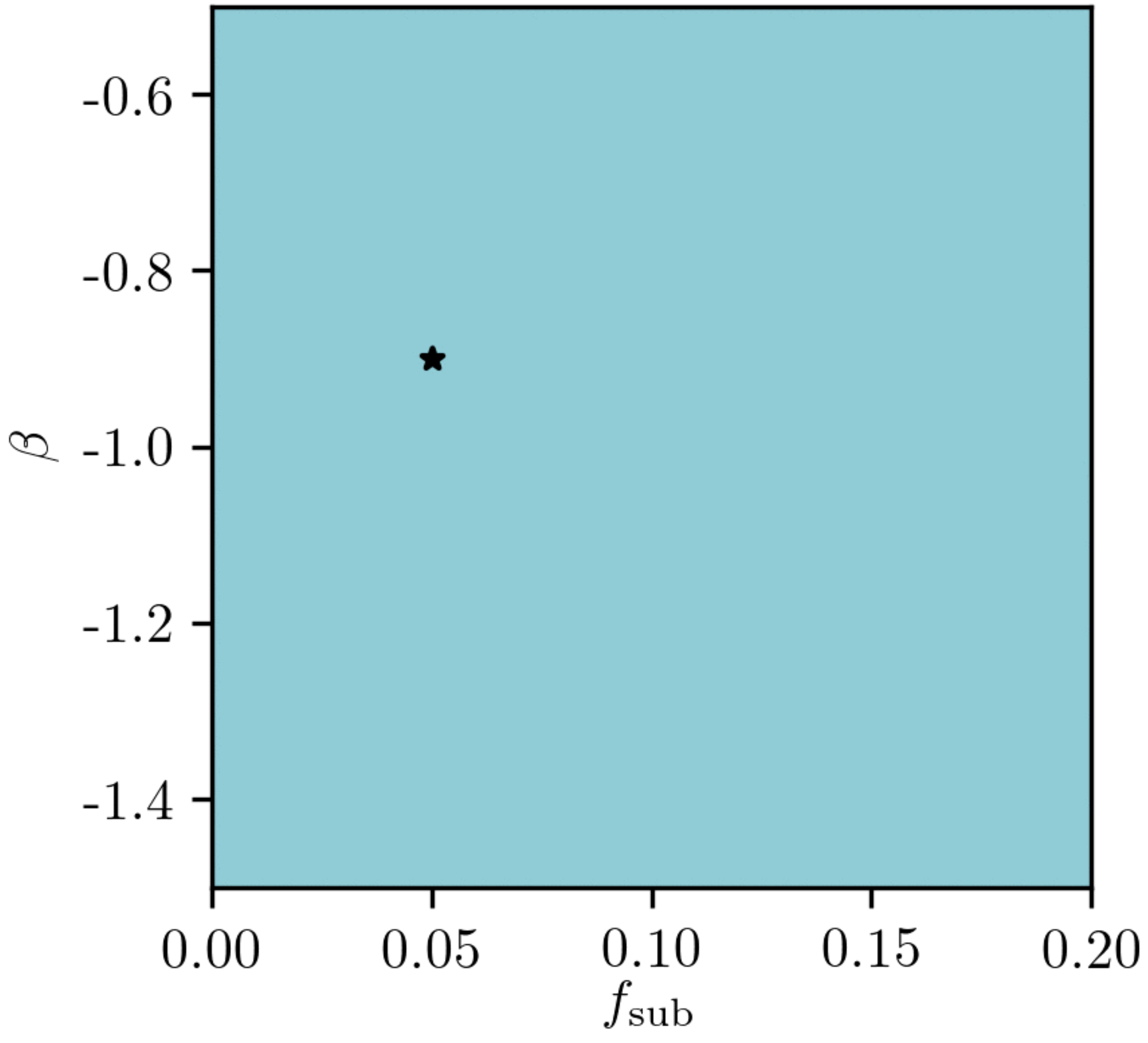
Uniform priors for $f_{\rm sub} \sim [0,\ 0.2]$, $\beta \sim [-1.5, -0.5]$

Lens sample

Mass function posterior



Parameter posteriors



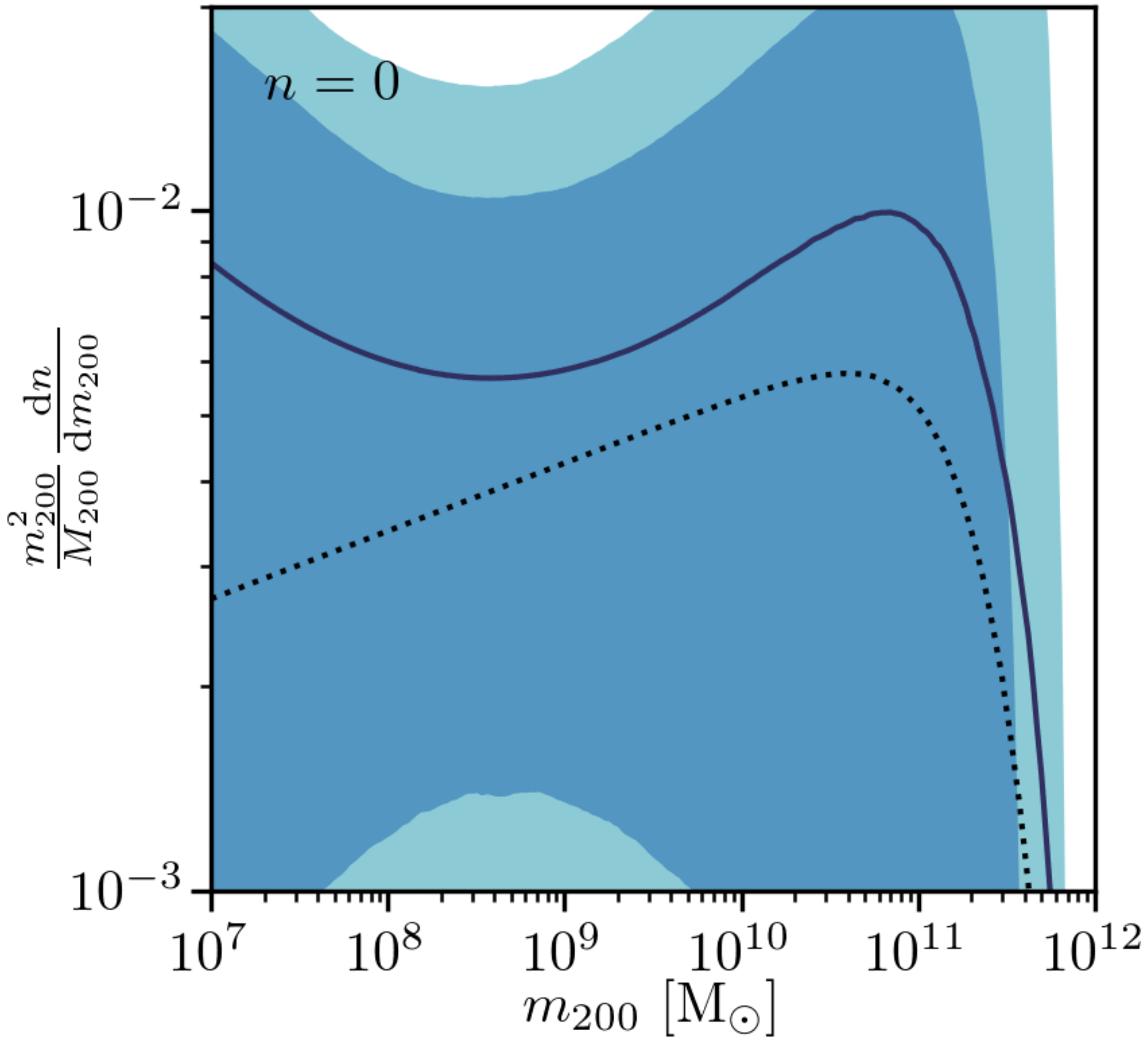
Bayesian inference

$f_{\rm sub} = 5\,\%\,,\, \beta = -0.9$

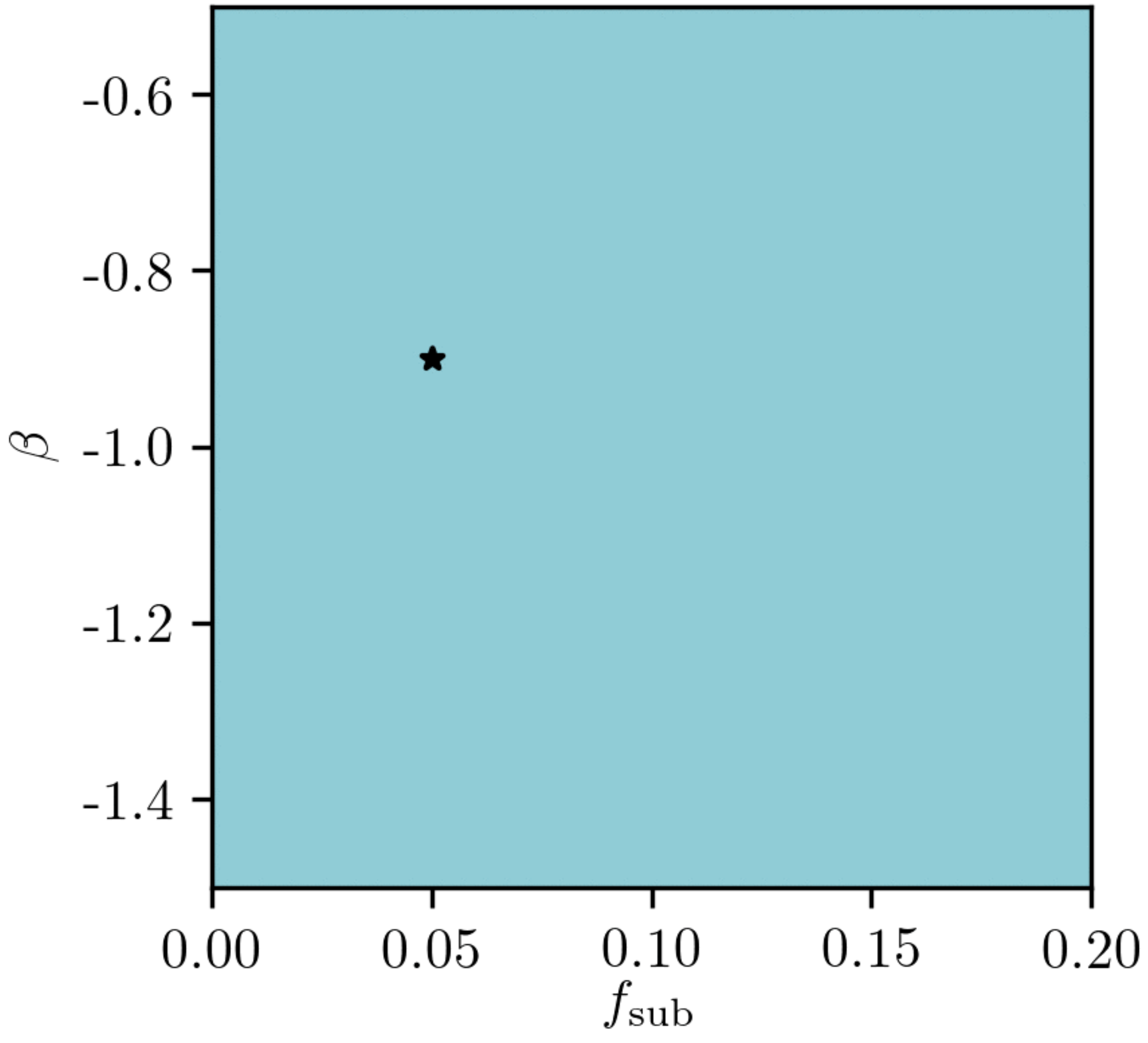
Uniform priors for $f_{\rm sub} \sim [0,\ 0.2],\ \beta \sim [-1.5, -0.5]$

Lens sample

Mass function posterior



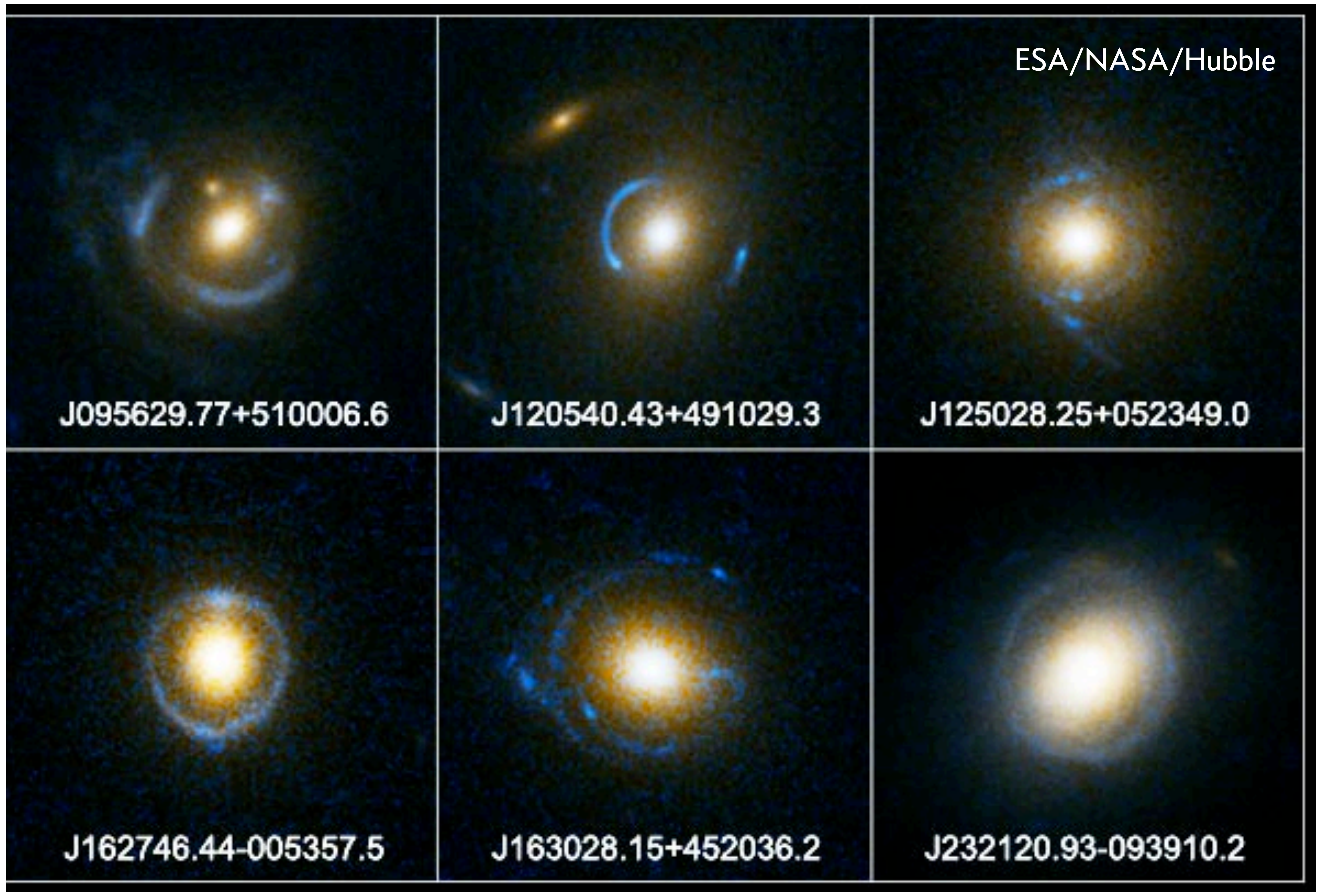
Parameter posteriors



Bayesian inference

$f_{\text{sub}} = 5\,\%\,,\ \beta = -0.9$

Uniform priors for $f_{\text{sub}} \sim [0,\ 0.2],\ \beta \sim [-1.5, -0.5]$



Analysis of real lensing data will require several extensions, e.g. accounting for complex morphology of lensed galaxies

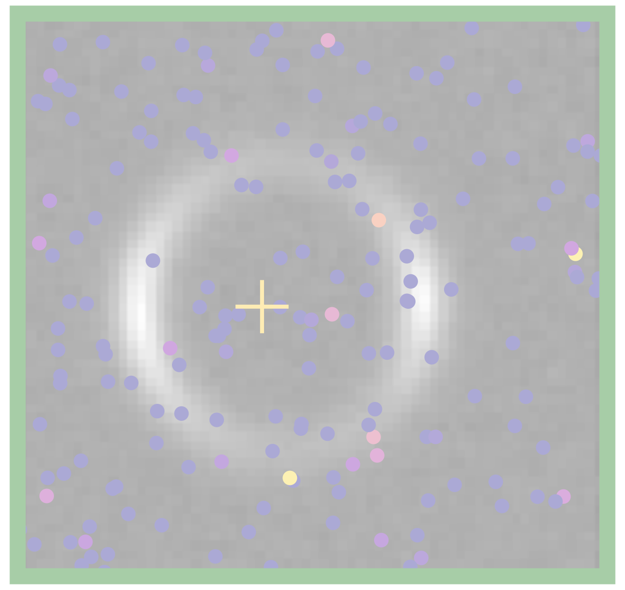
Lots of recent progress, e.g.

- Auto encoders Chianese et al [MNRAS 2020]
- Gaussian processes Coogan et al [NeurIPS ML4PS 2020], Karchev et al [MNRAS 2021]
- Real galaxies Wagner-Carena et al [2203.00690]

iors

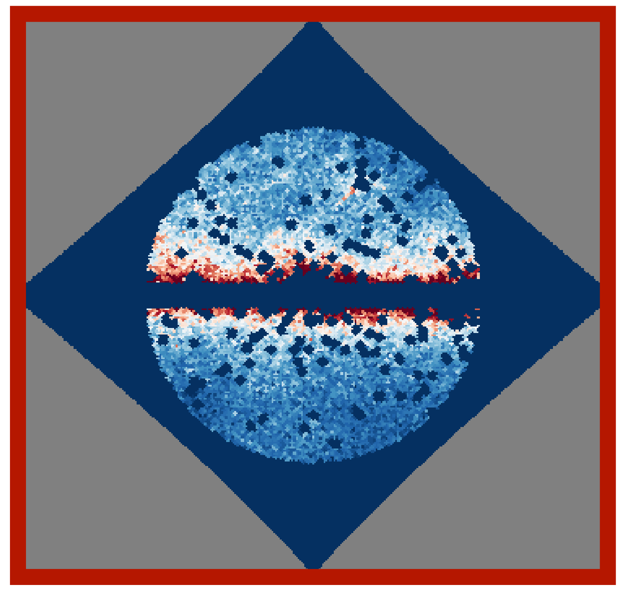
15 0.20

Outline



Detecting extragalactic dark matter in strong lenses

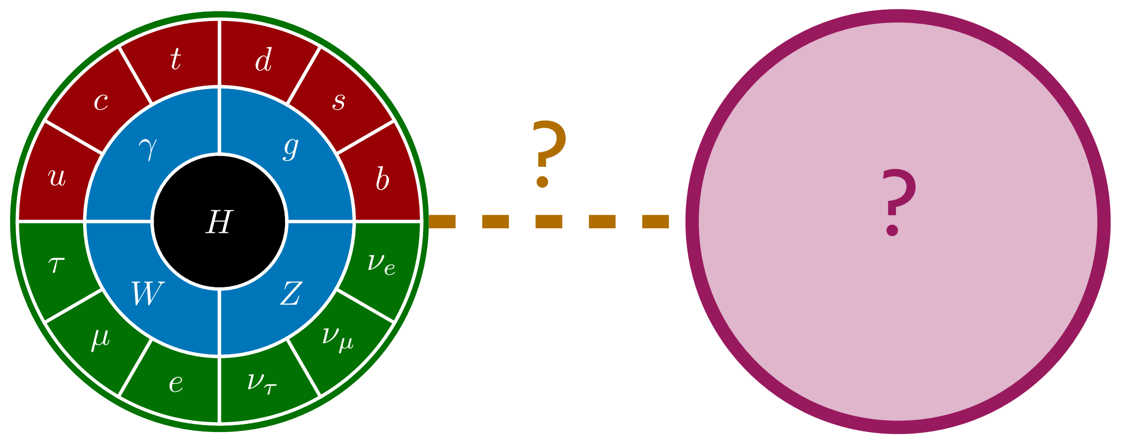
Combining information from thousands of systems



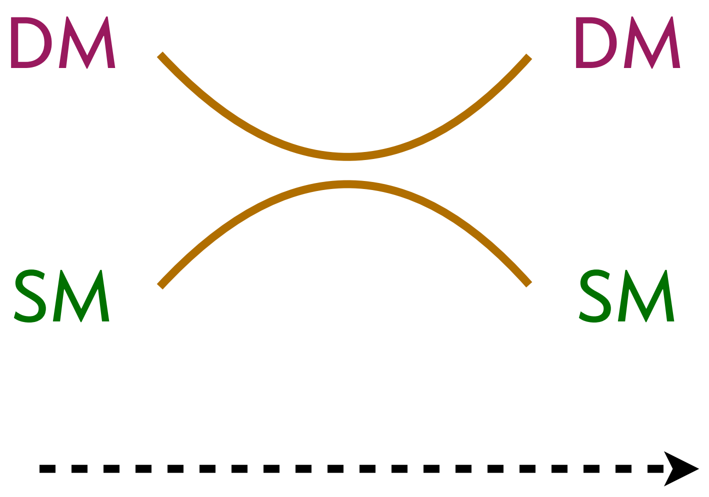
Characterizing γ-ray point sources in the Galactic Center

Exploiting more information to reduce model misspecification

Searches for dark matter

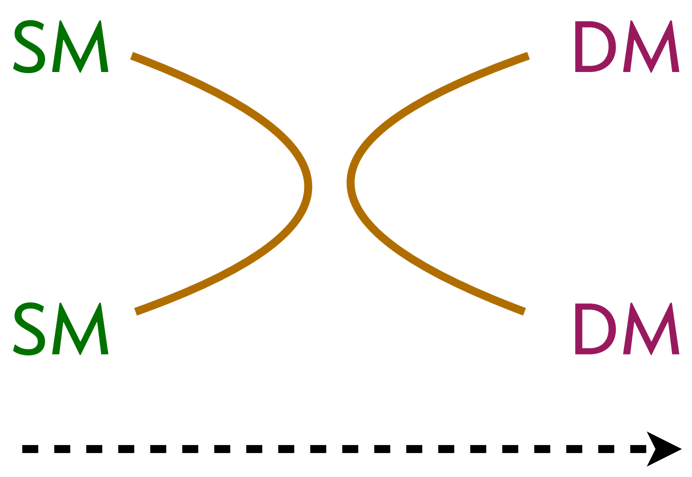


Scattering of DM against SM

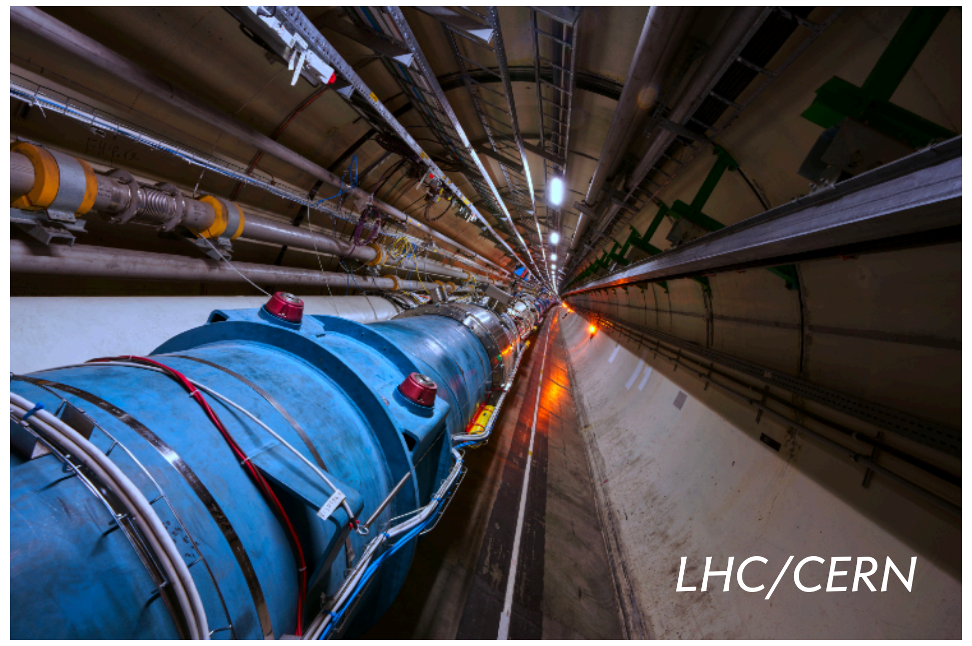


Direct detection

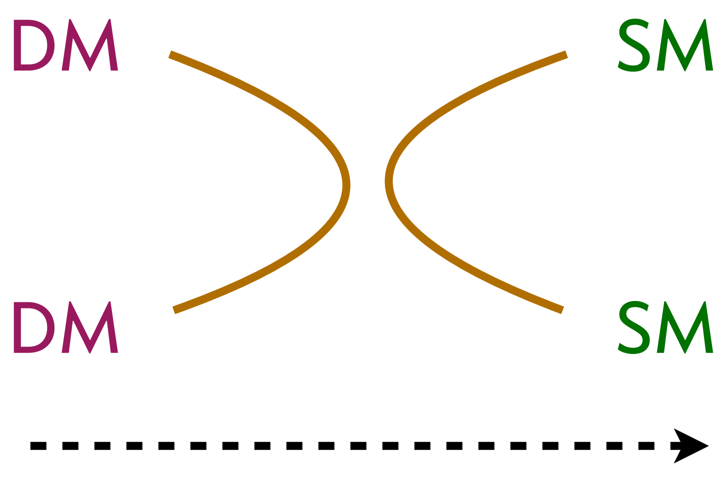
Production of DM from SM



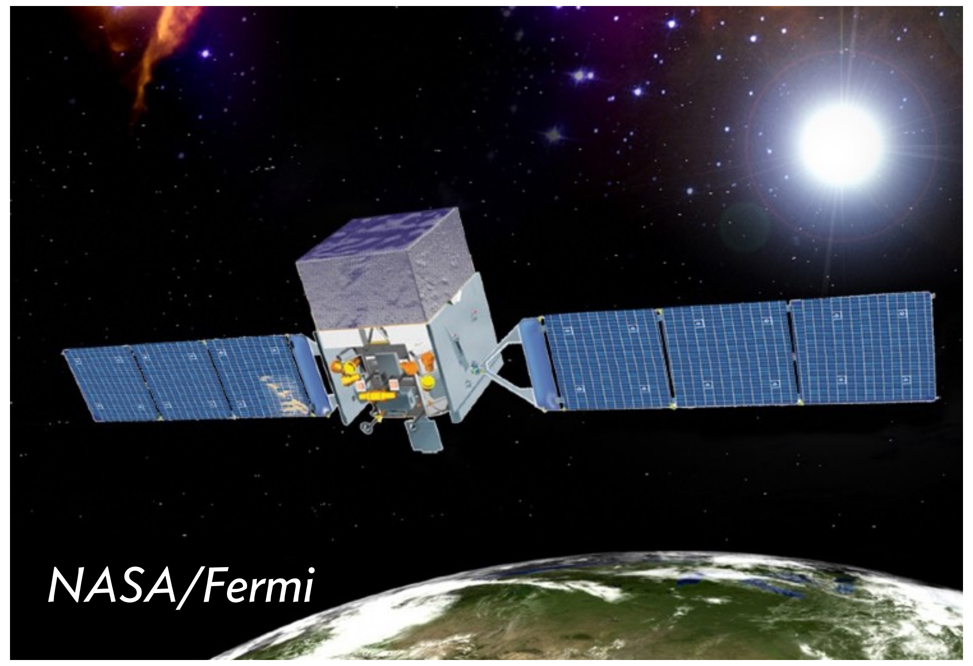
Particle colliders



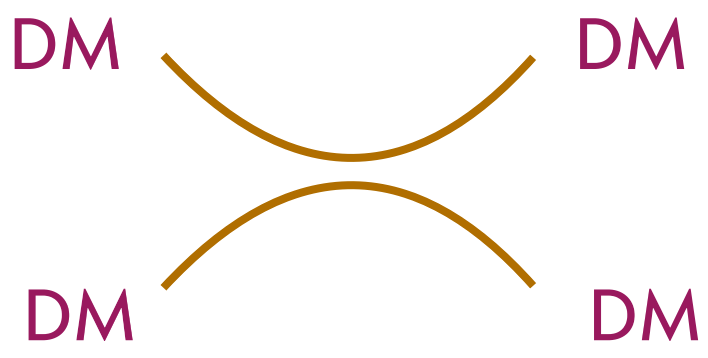
Production of SM from DM



Indirect detection

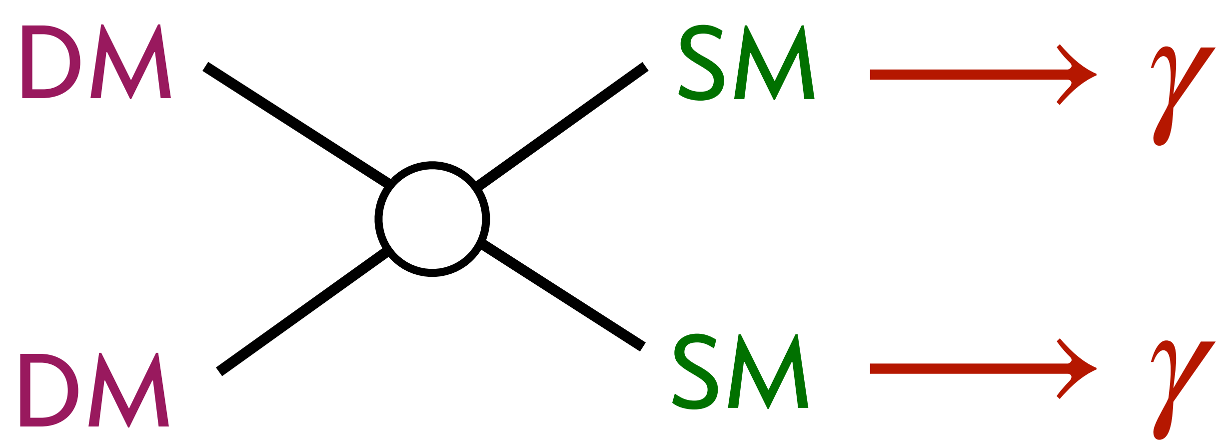


Gravitational effects of DM interaction



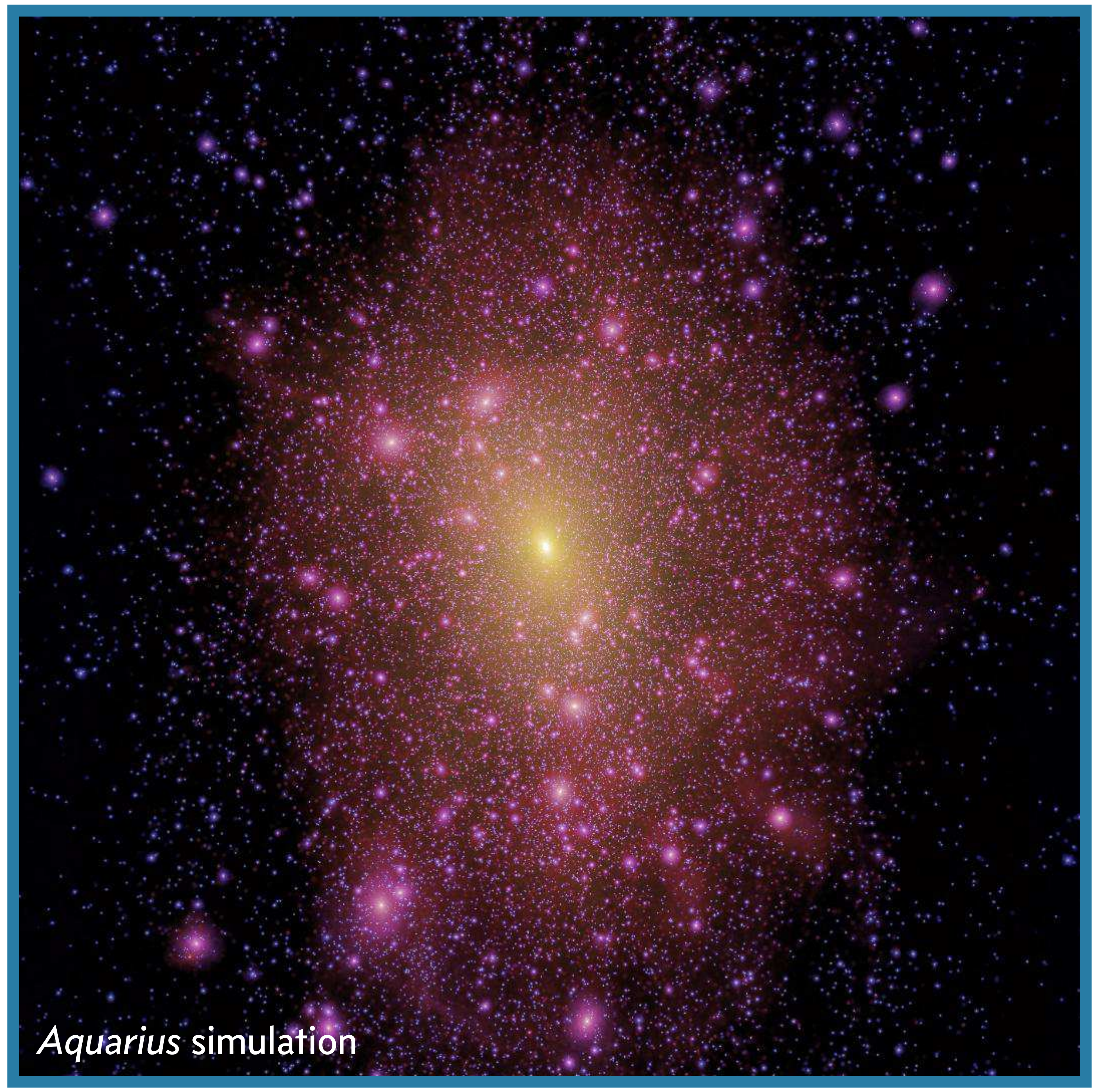
Astrophysical probes

Dark matter annihilation in the Galactic Center

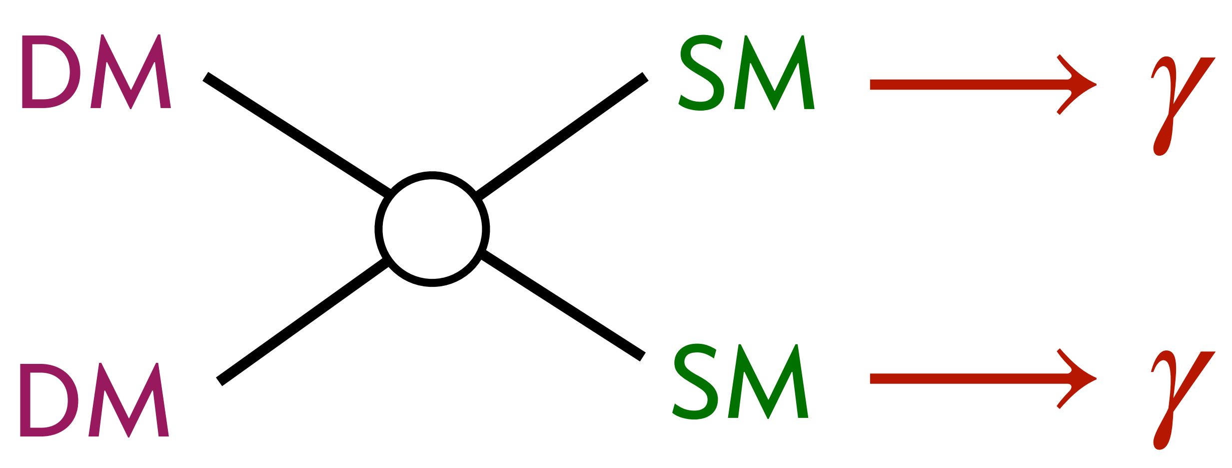


Annihilating WIMP DM would produce excess γ-rays from the Galactic Center

DM distribution

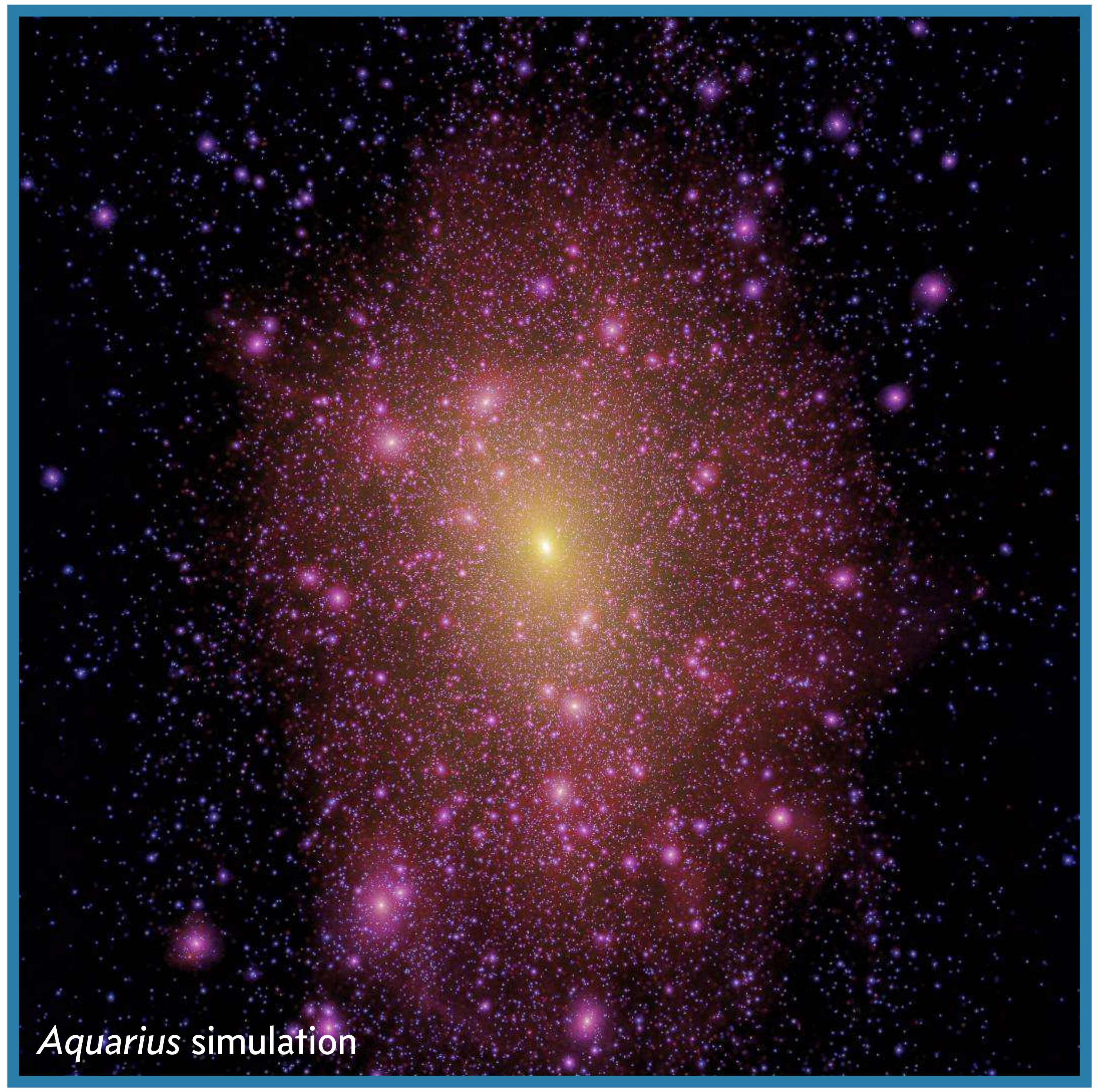


Dark matter annihilation in the Galactic Center

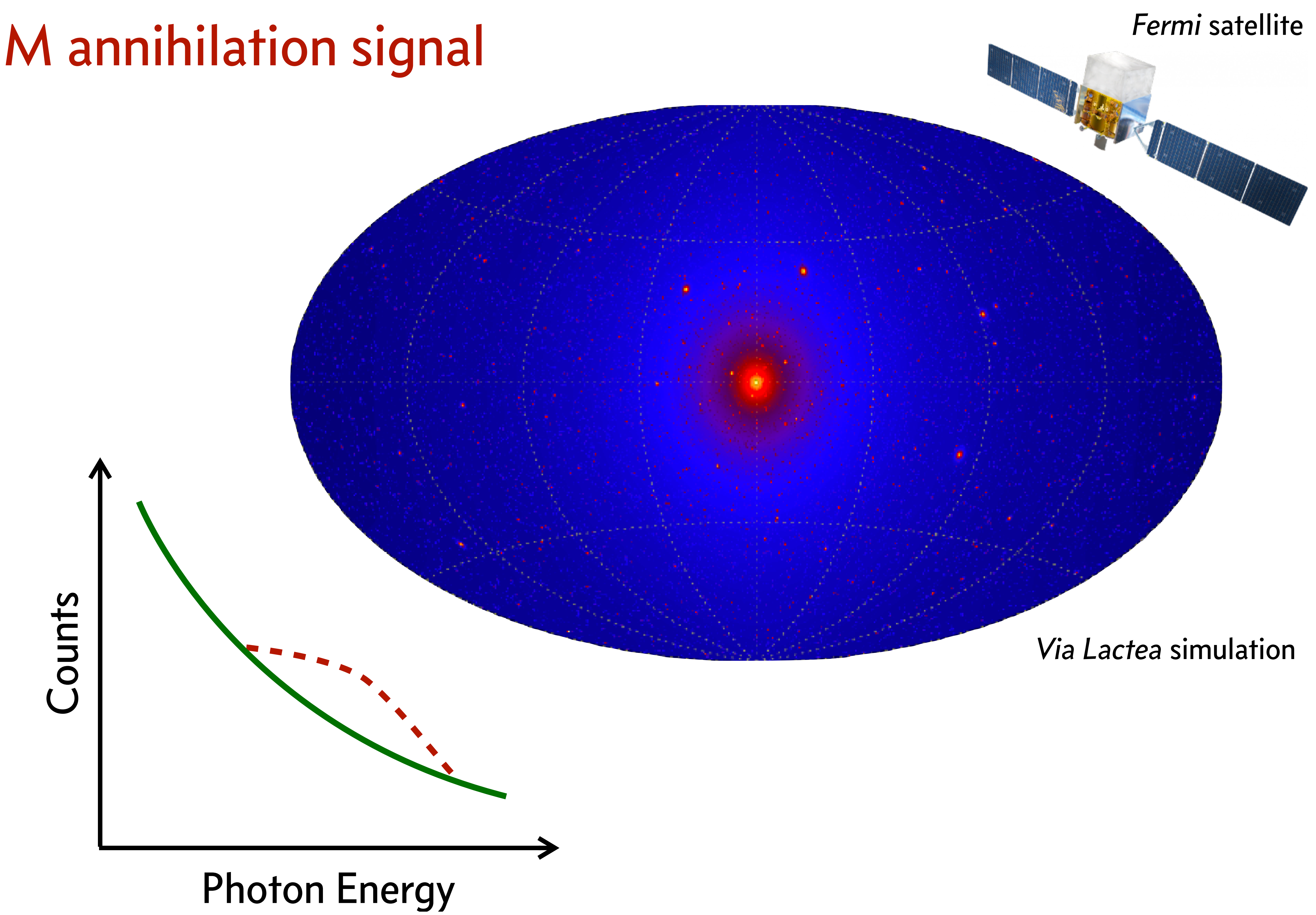


Annihilating WIMP DM would produce excess γ-rays from the Galactic Center

DM distribution



DM annihilation signal



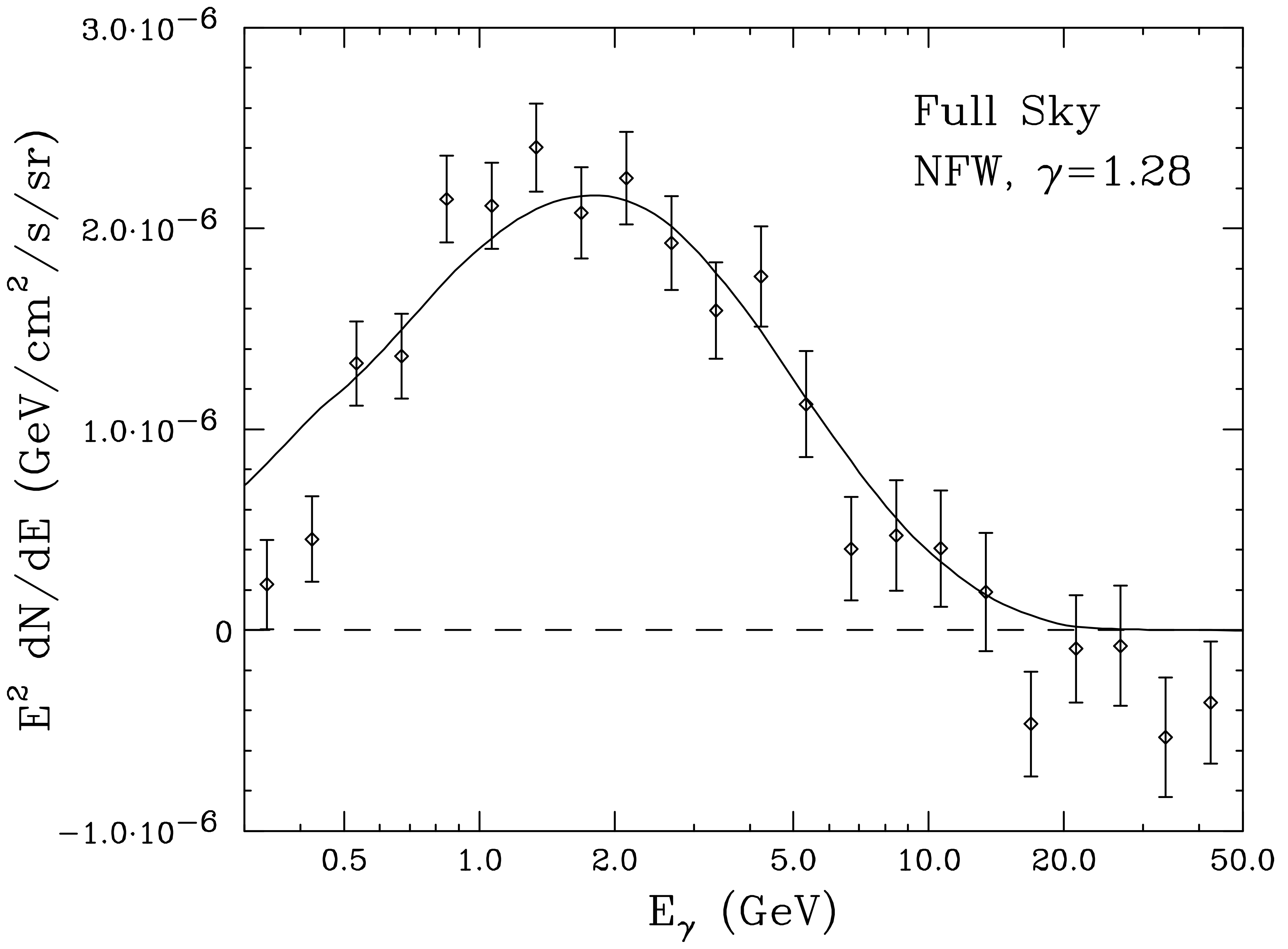
The *Fermi* Galactic Center Excess

Daylan et al [PDU 2016]

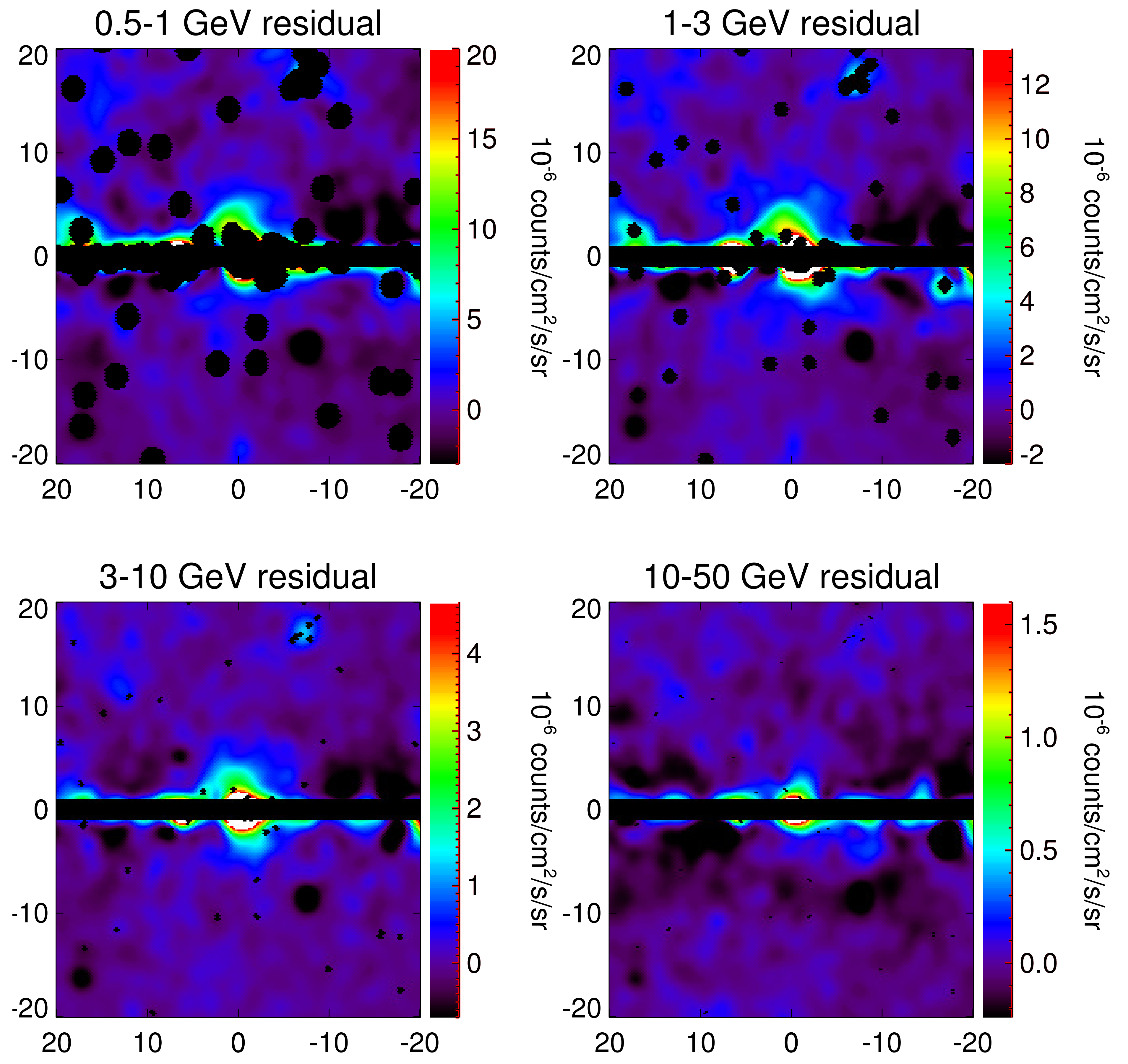
Some facts:

- ~ Spherically symmetric γ-ray excess in the Inner Galaxy
- Extends out ~10° from the center of Galaxy
- Constitutes ~10% total flux

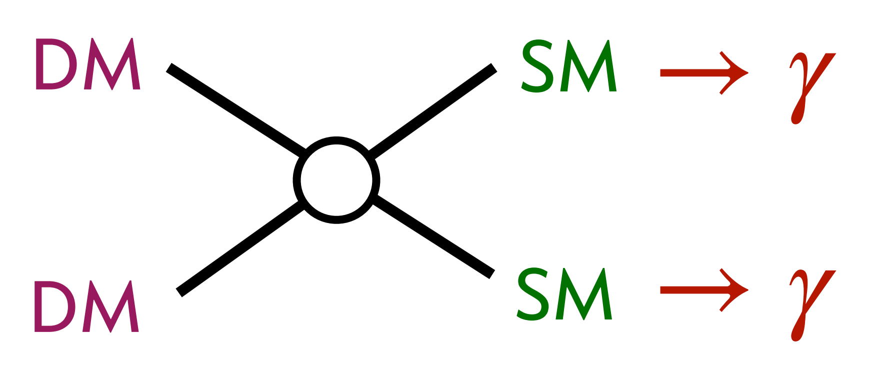
Energy spectrum



Spatial morphology

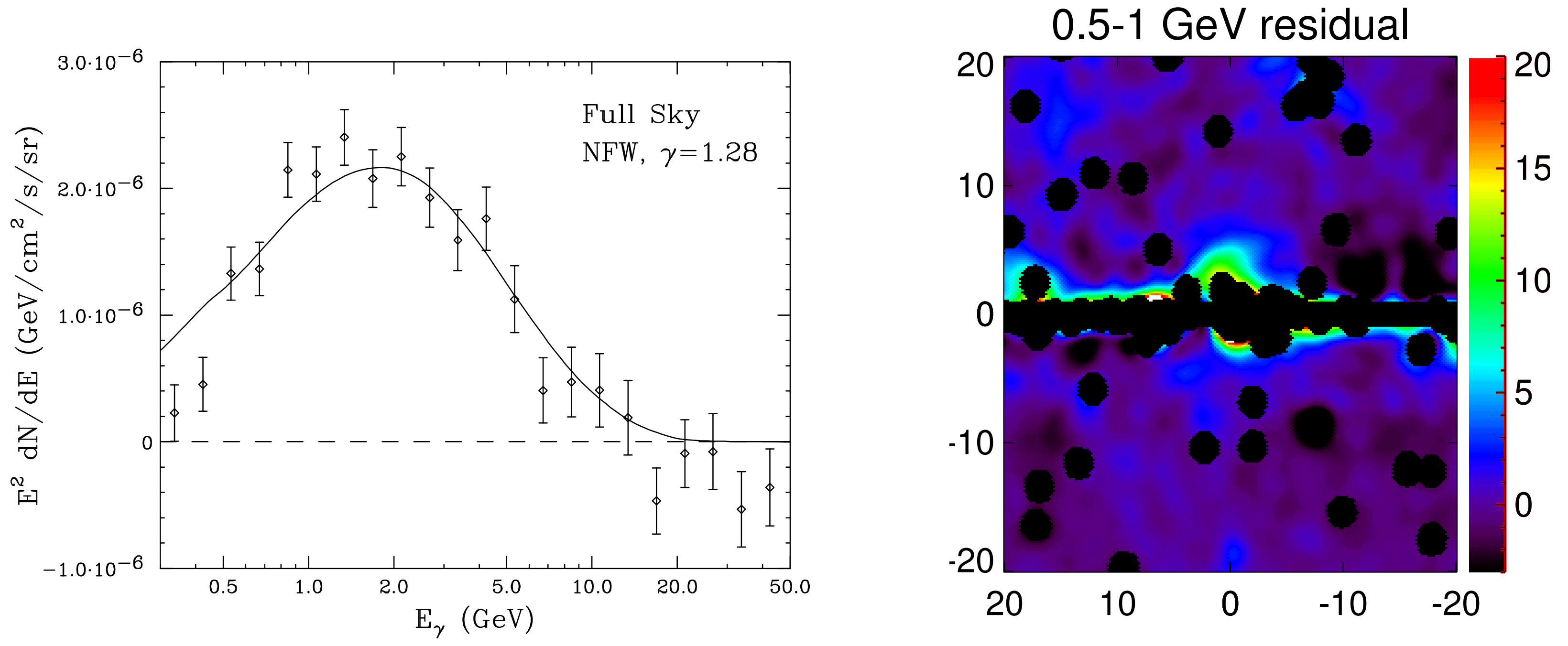


Possible explanations



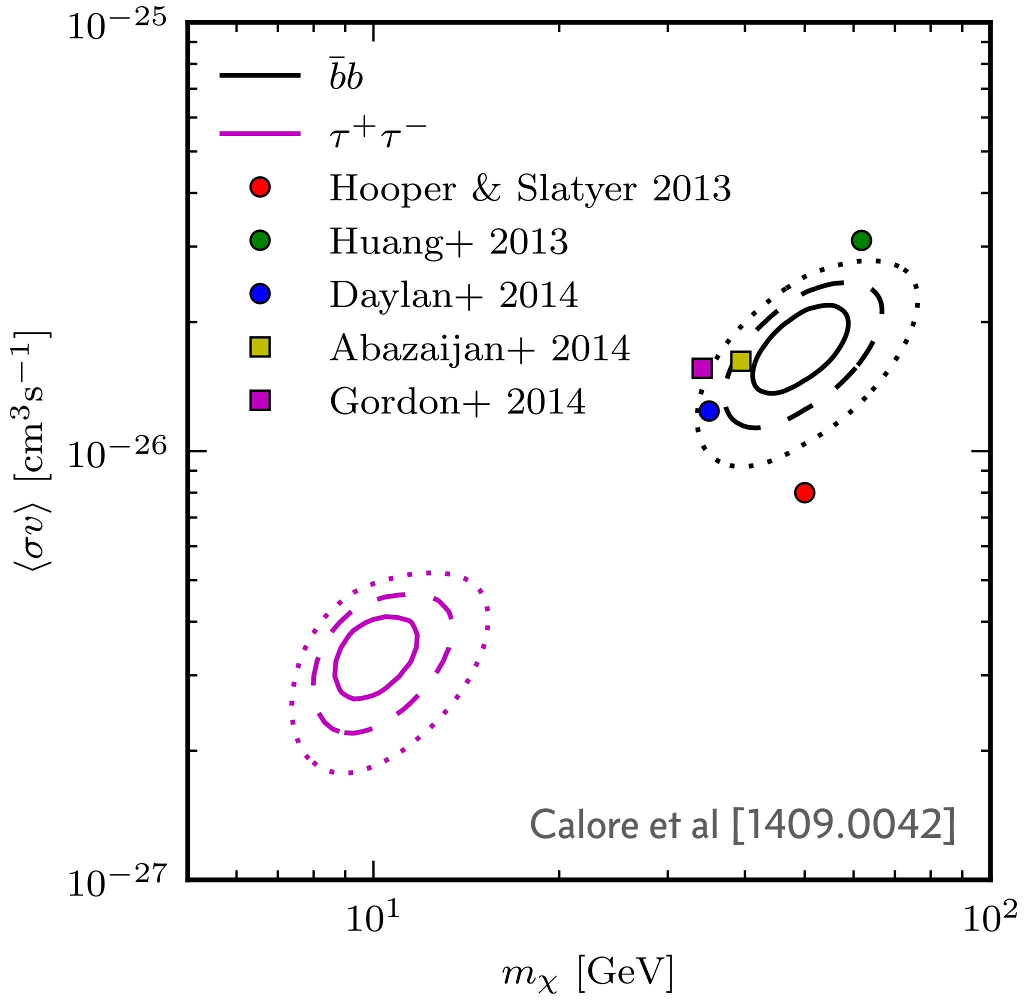
Dark Matter

Spectrum and spatial morphology consistent with DM expectation



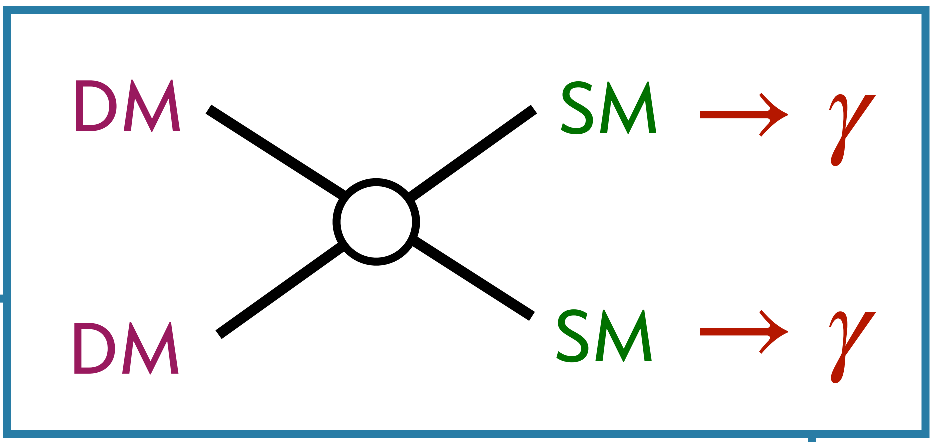
Daylan et al [1402.6703]

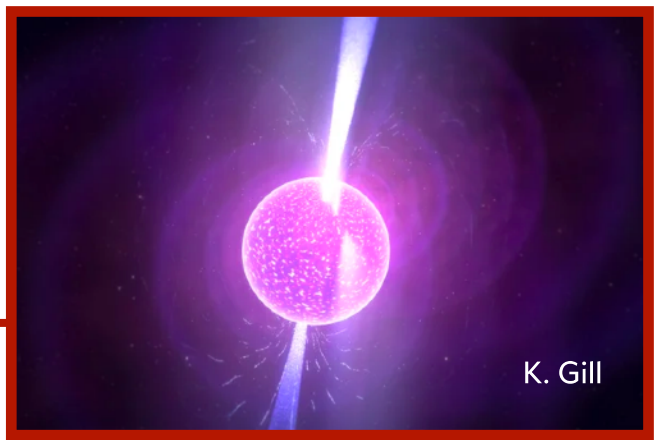
Consistent with thermal cross section expectation



Calore et al [1409.0042]

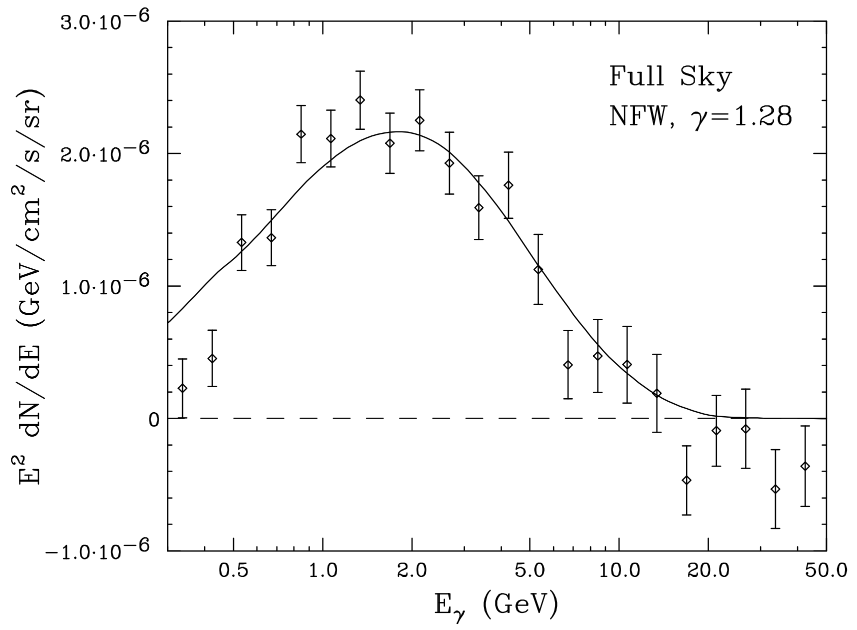
Possible explanations

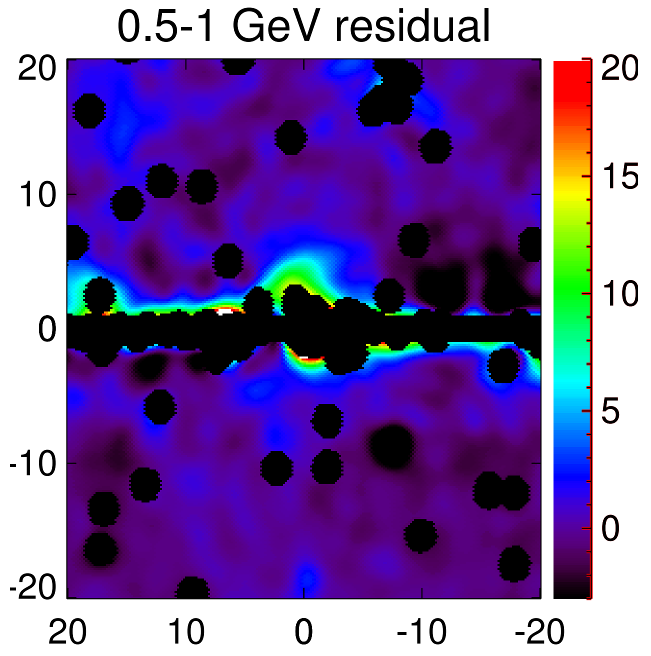




Dark Matter

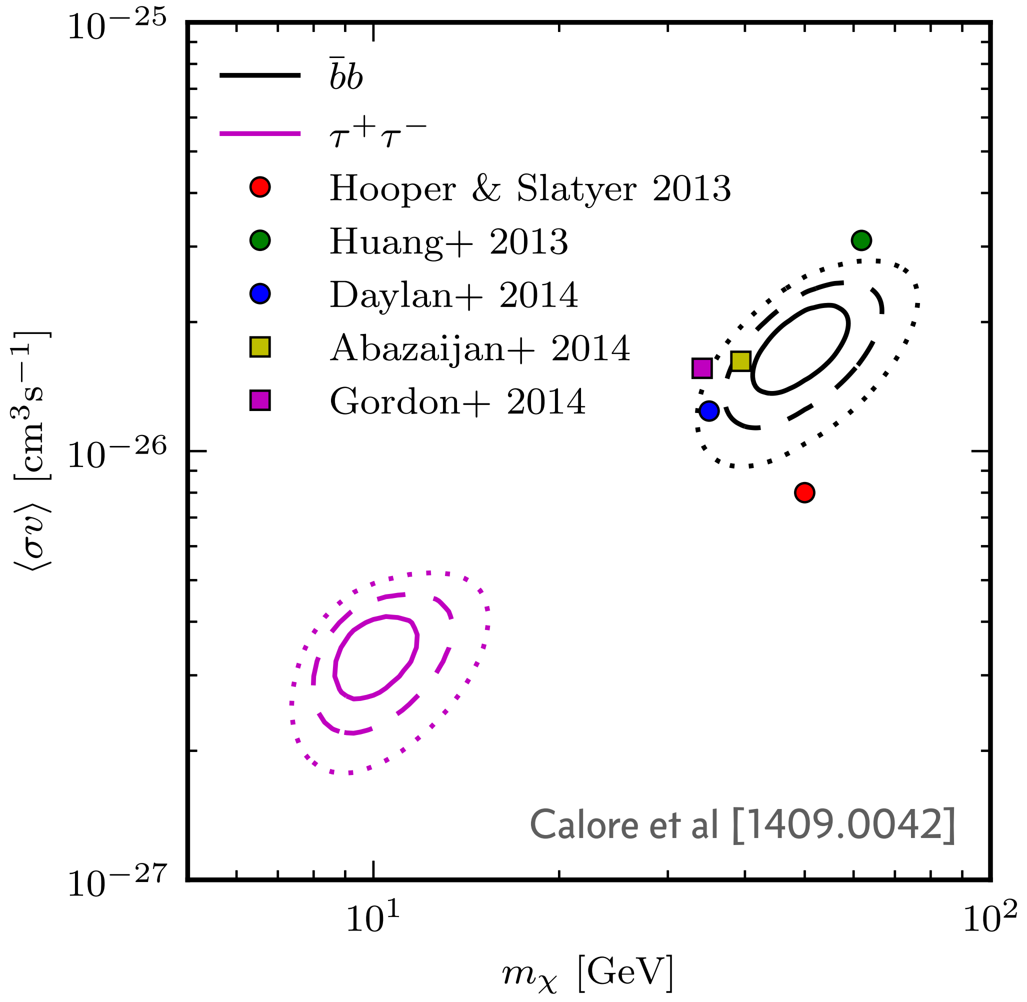
Spectrum and spatial morphology consistent with DM expectation





Daylan et al [1402.6703]

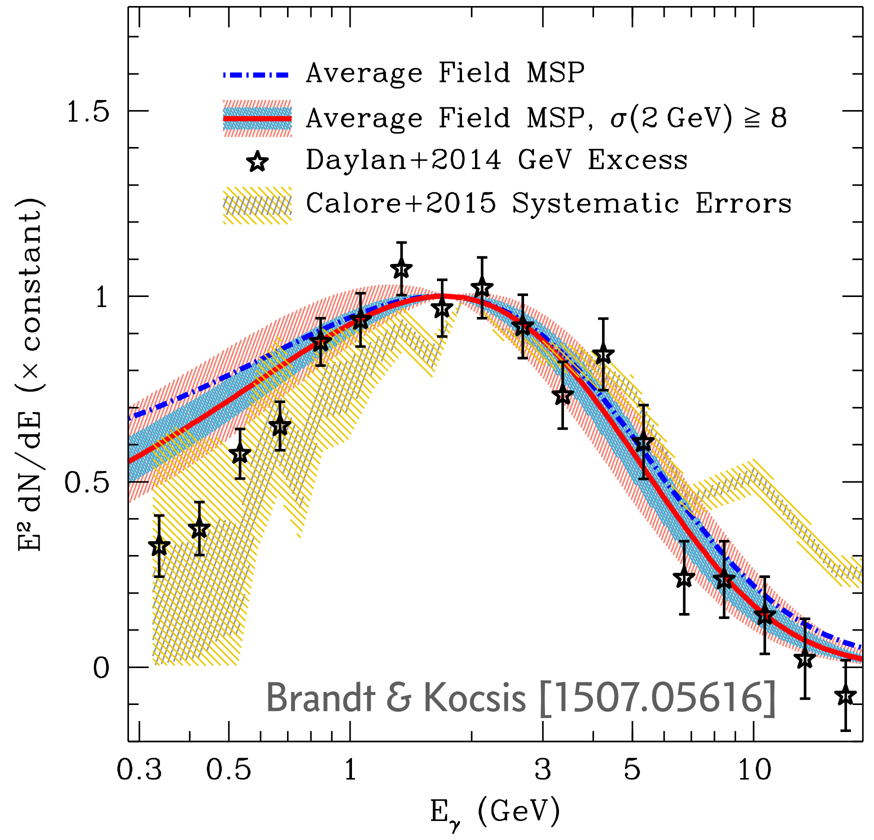
Consistent with thermal cross section expectation



Calore et al [1409.0042]

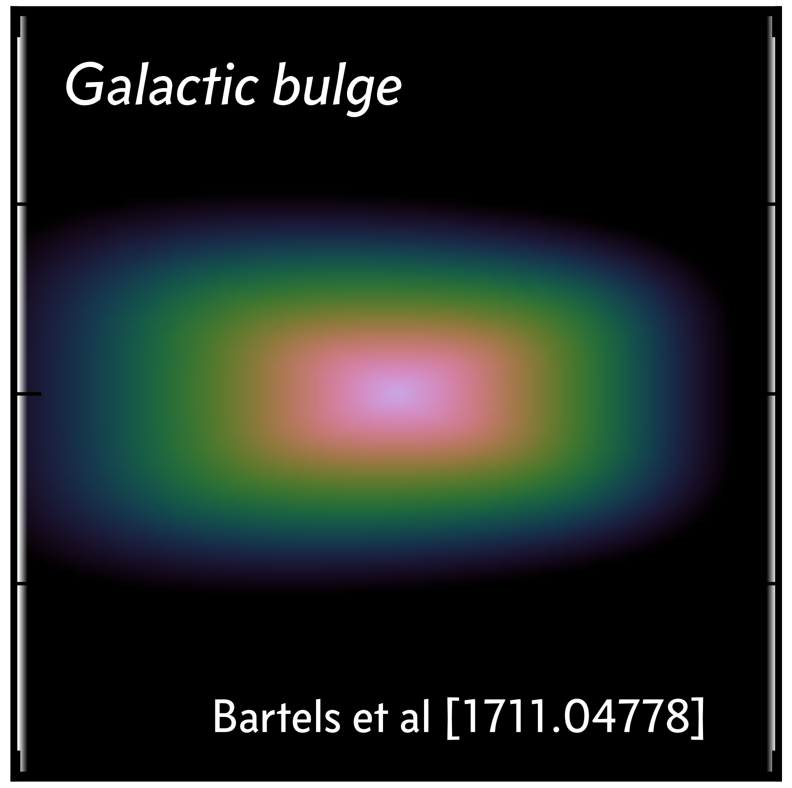
Astrophysics

Spectrum roughly consistent with MSP expectation



Brandt & Kocsis [1507.05616]

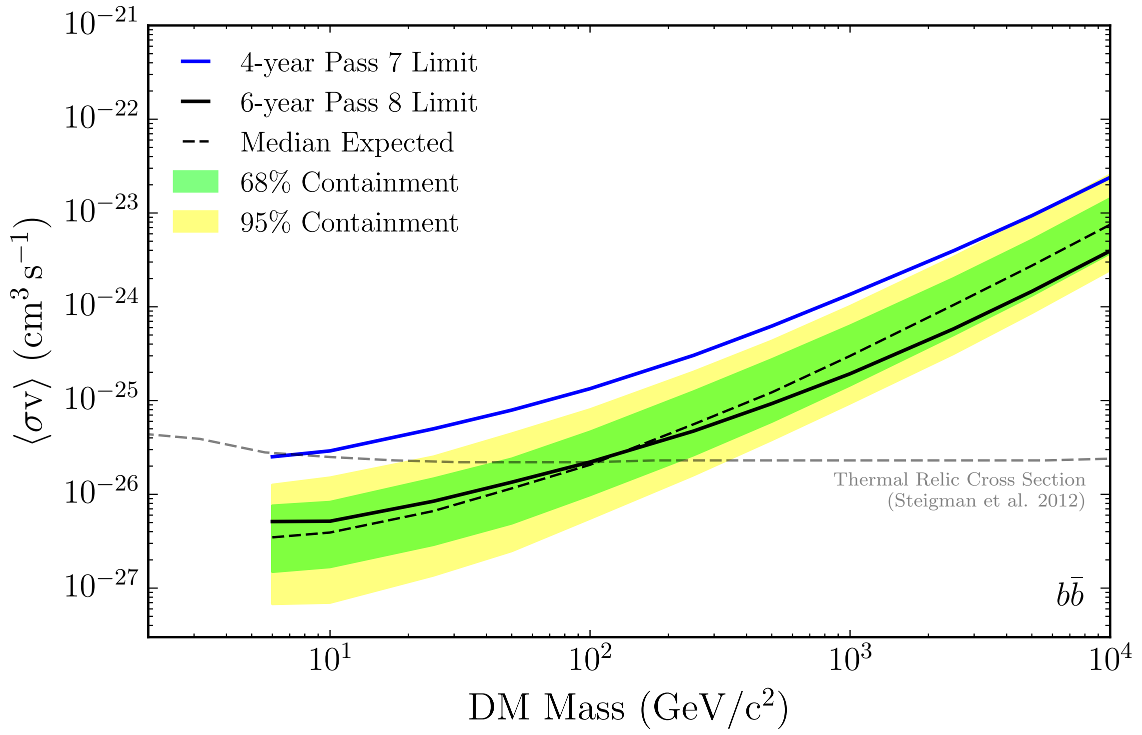
Spatial morphology consistent with stellar distribution?



Bartels et al [1711.04778]

DM annihilation signal not seen in other targets

Fermi-LAT Collaboration [1503.02641]



Distinguishing DM from PSs

Unresolved sources

PSs degenerate with smooth emission

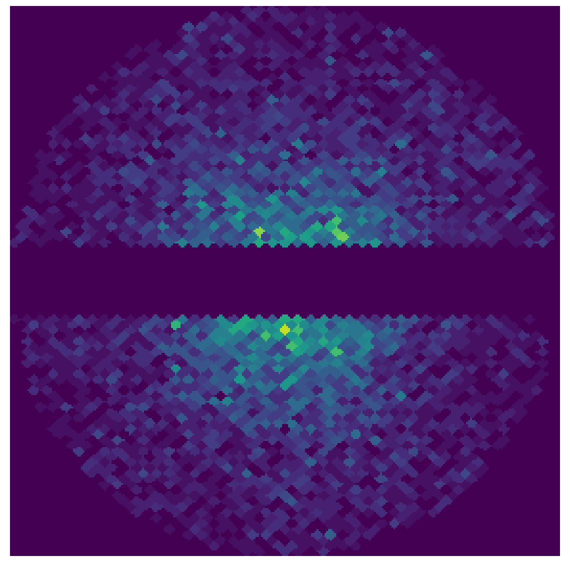
Can we infer and characterize these source populations?

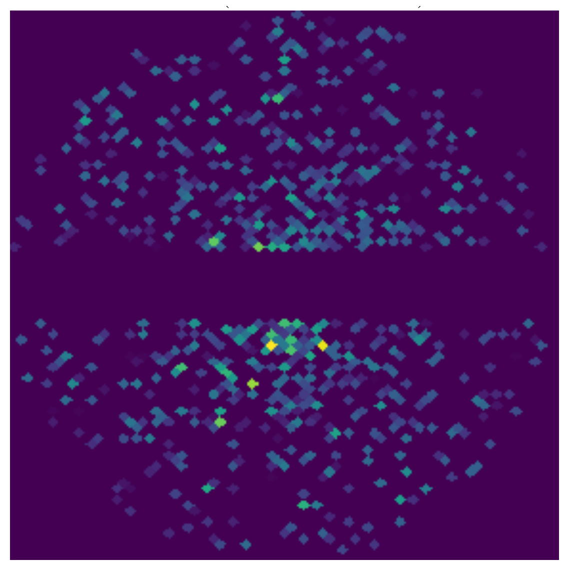
Individually-resolved PSs

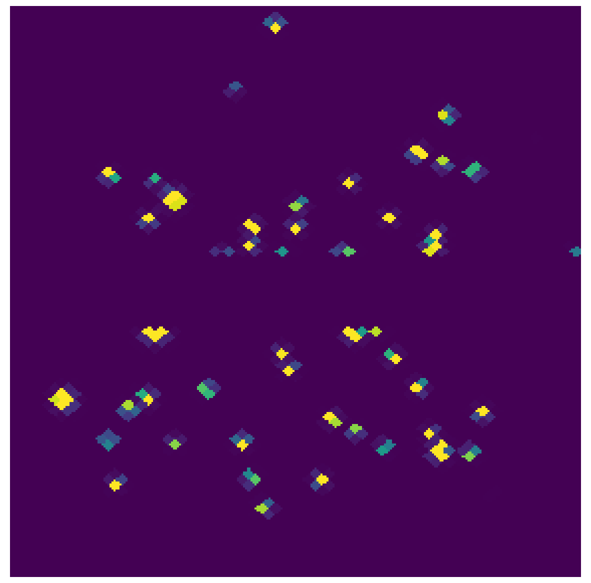
~ 1 photon

Fermi-LAT PS resolution threshold

Flux







Modeling PS populations in γ-ray data

Parameters of interest

PS population properties

$$\theta = \{A, n, S\}$$

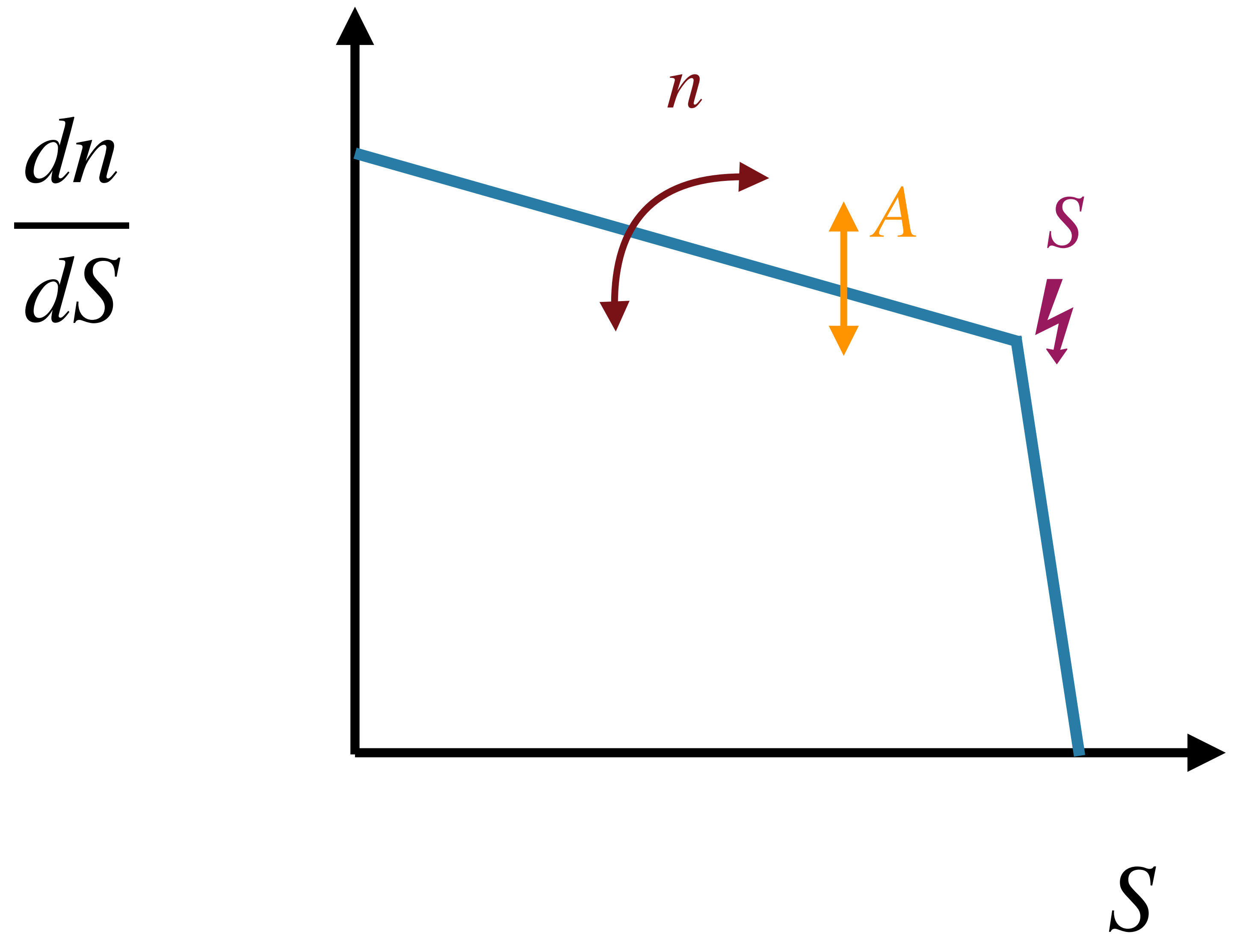
Latent variables

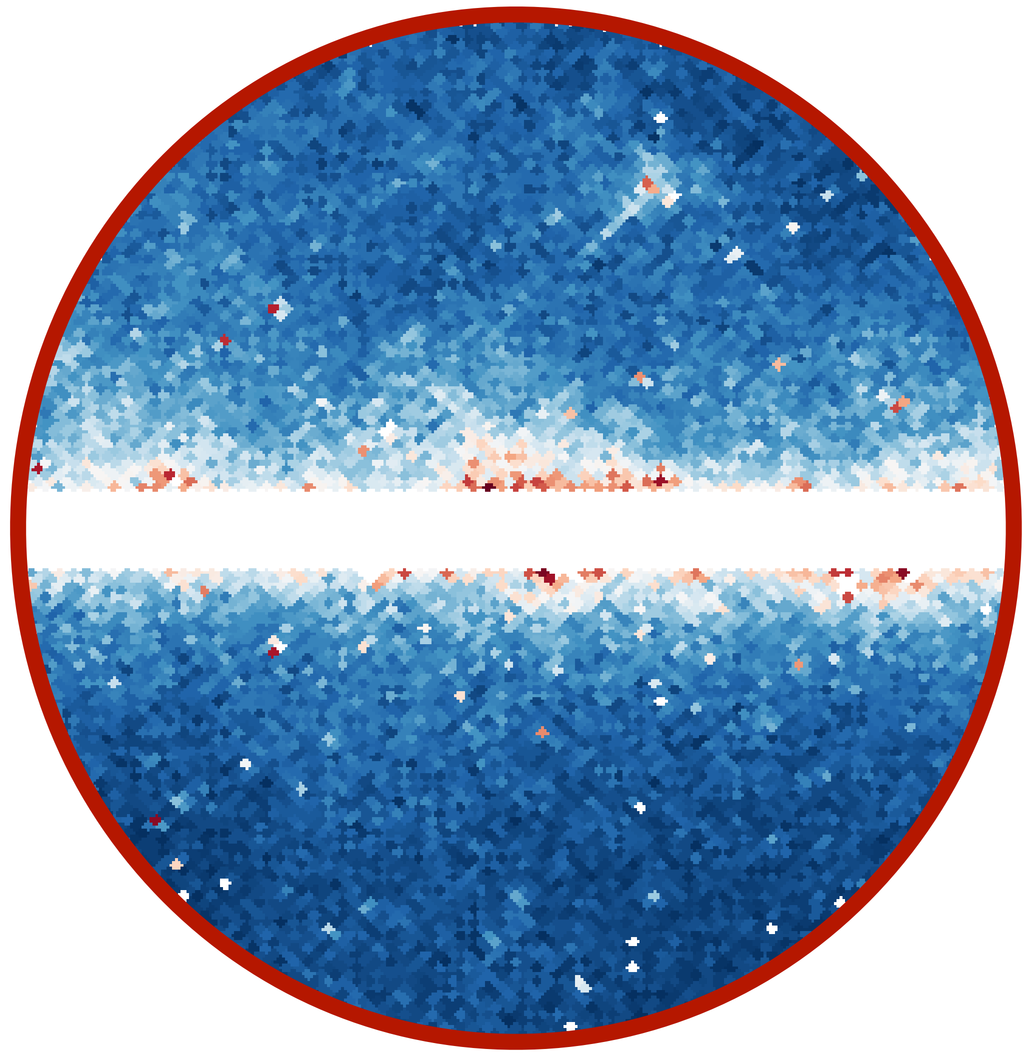
PS properties

$$\{z_{\mathrm{PS},\,i}\}$$

Observables

Gamma-ray image x



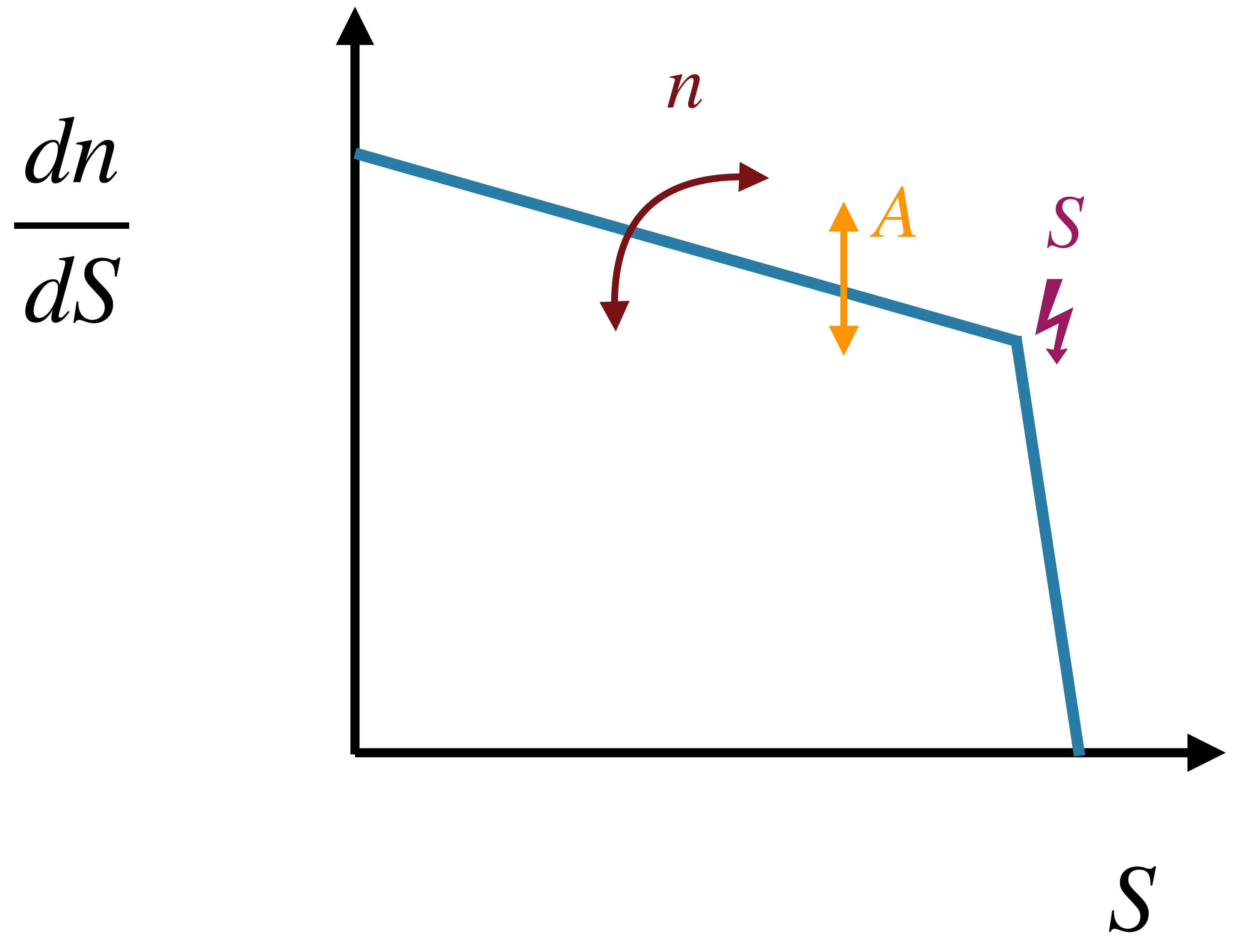


Modeling PS populations in γ-ray data

Parameters of interest

PS population properties

$\theta = \{A, n, S\}$



Latent variables

PS properties

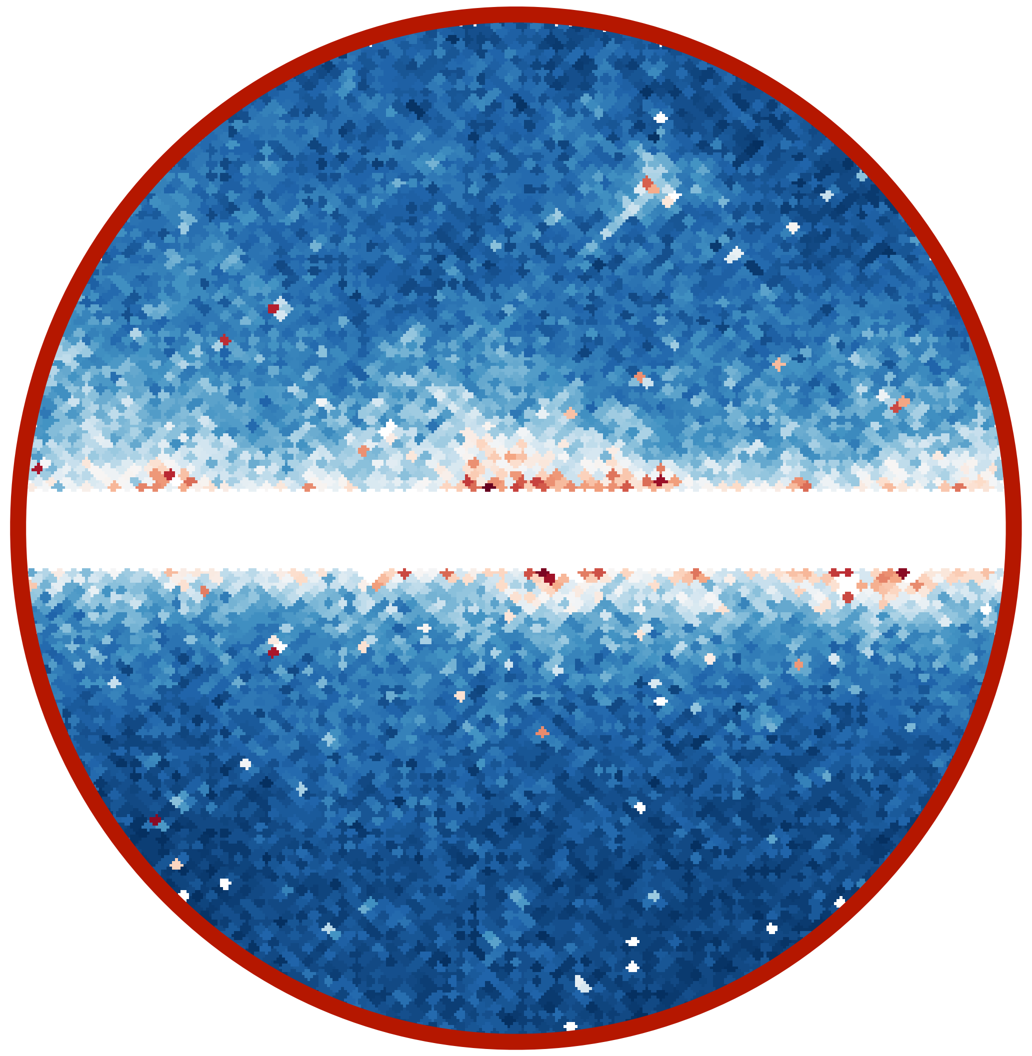
$\{z_{\mathrm{PS},\,i}\}$

Prediction (Simulation)

$p(x, z \,|\, \theta)$

Observables

Gamma-ray image x

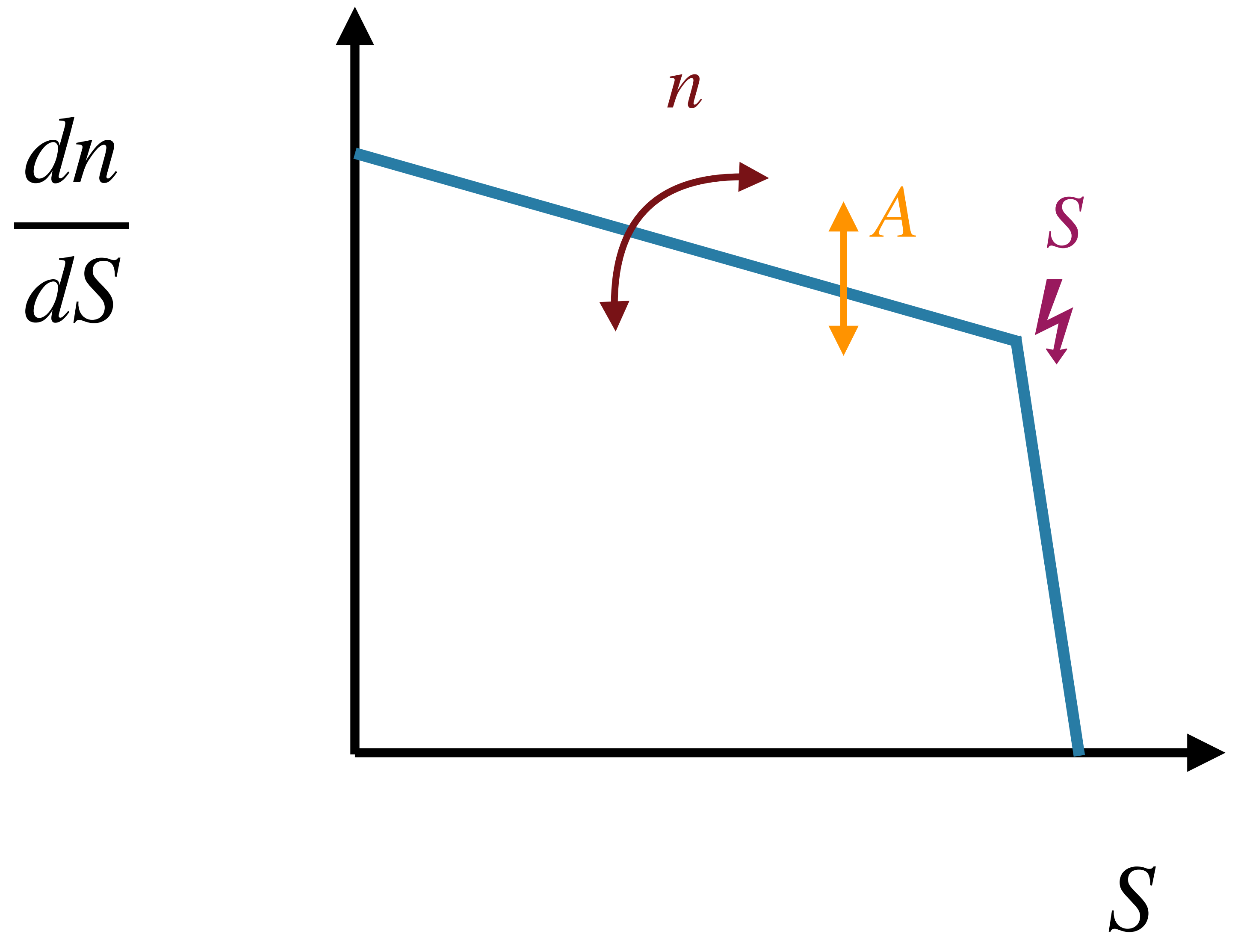


Modeling PS populations in γ-ray data

Parameters of interest

PS population properties

$\theta = \{A, n, S\}$



Latent variables

PS properties

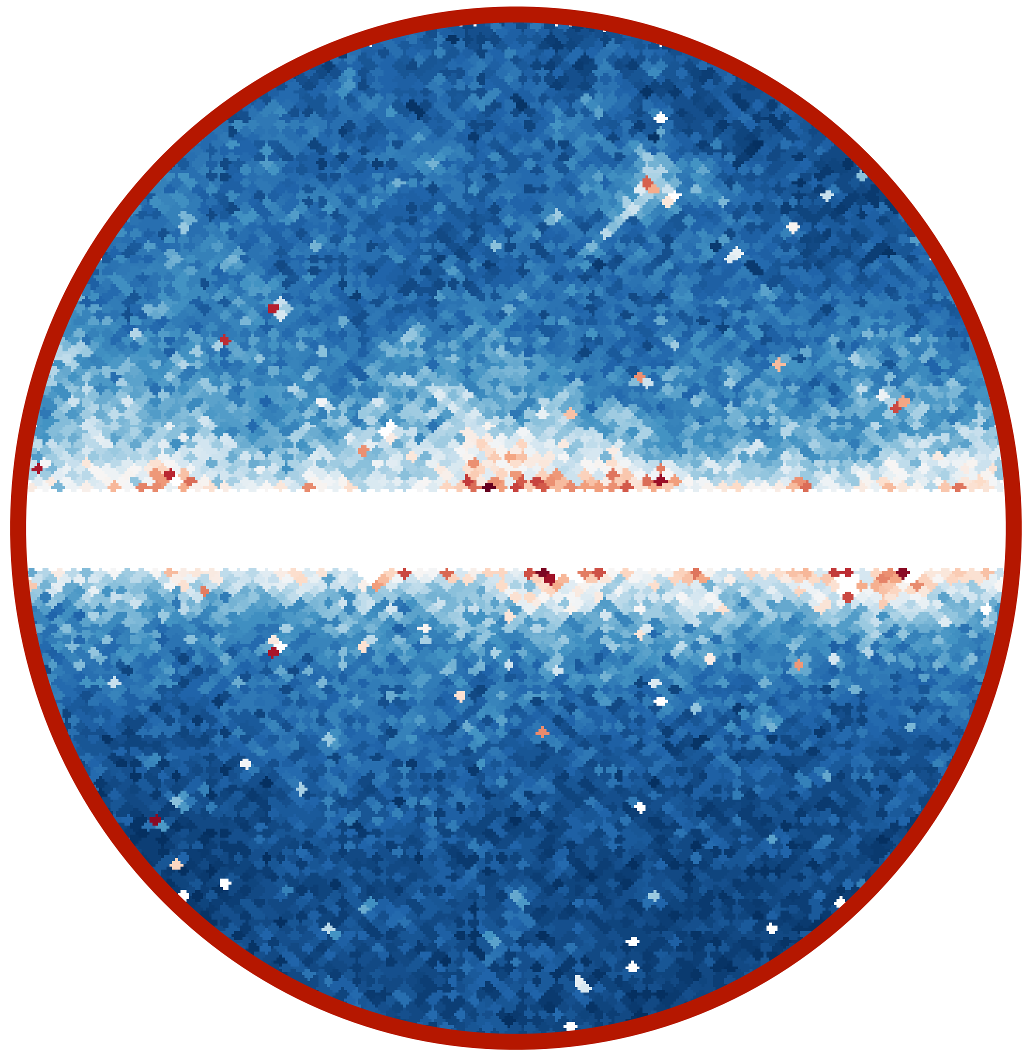
$\{z_{\mathrm{PS},\,i}\}$

Inference

$$p(x\,|\,\theta) = \int \mathrm{d}z\, p(x, z\,|\,\theta)$$

Observables

Gamma-ray image x

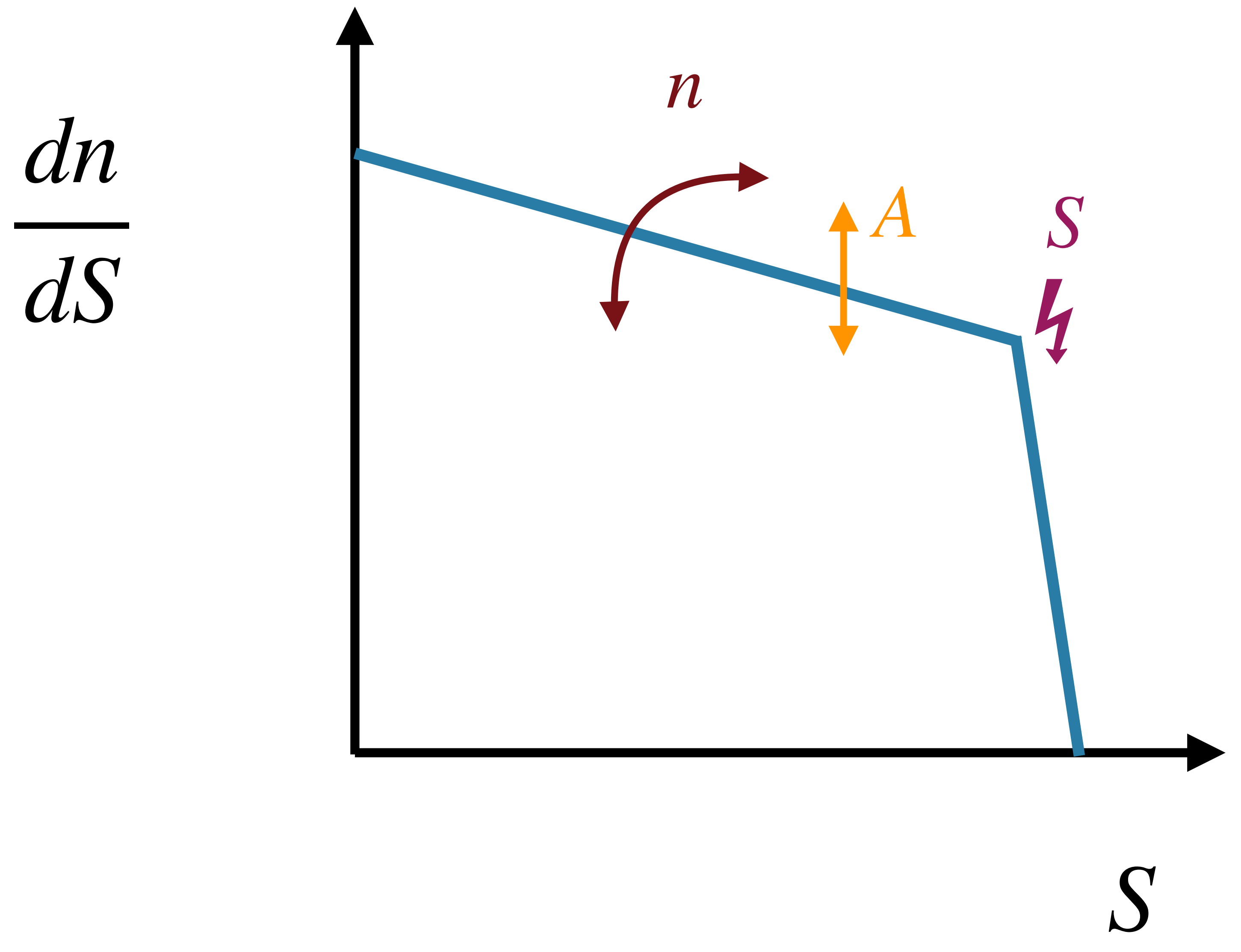


Modeling PS populations in γ-ray data

Parameters of interest

PS population properties

$$\theta = \{A, n, S\}$$



Latent variables

PS properties

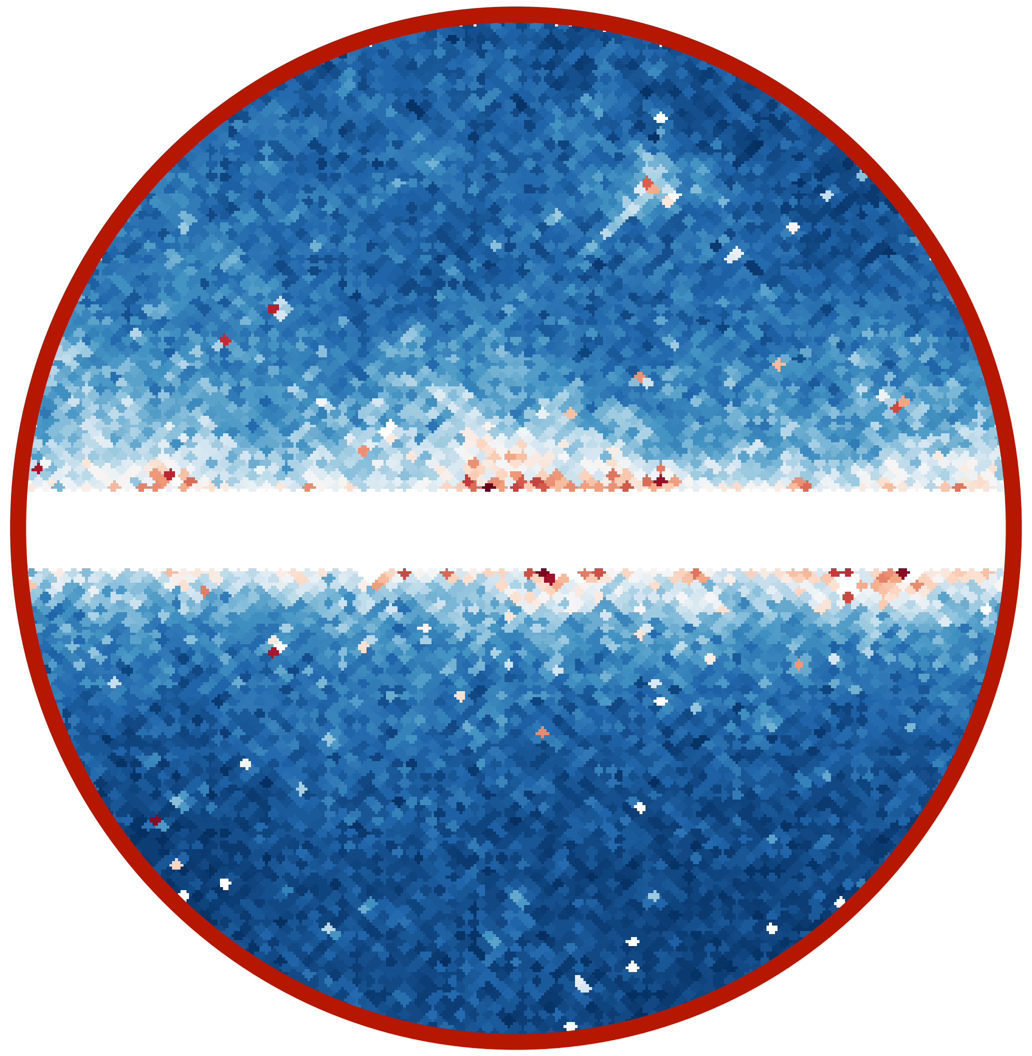
$$\{z_{\mathrm{PS},\,i}\}$$

Inference

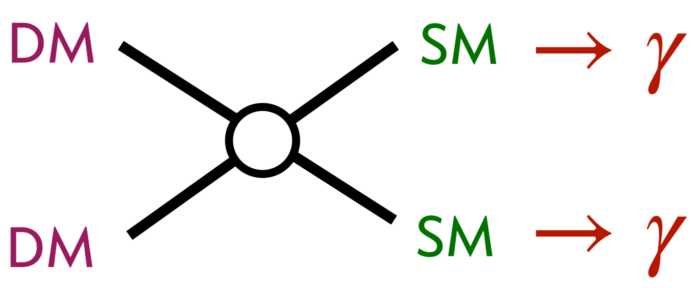
$$p(x \,|\, \theta) = \boxed{\int \mathrm{d}z}\, p(x, z \,|\, \theta)$$

Observables

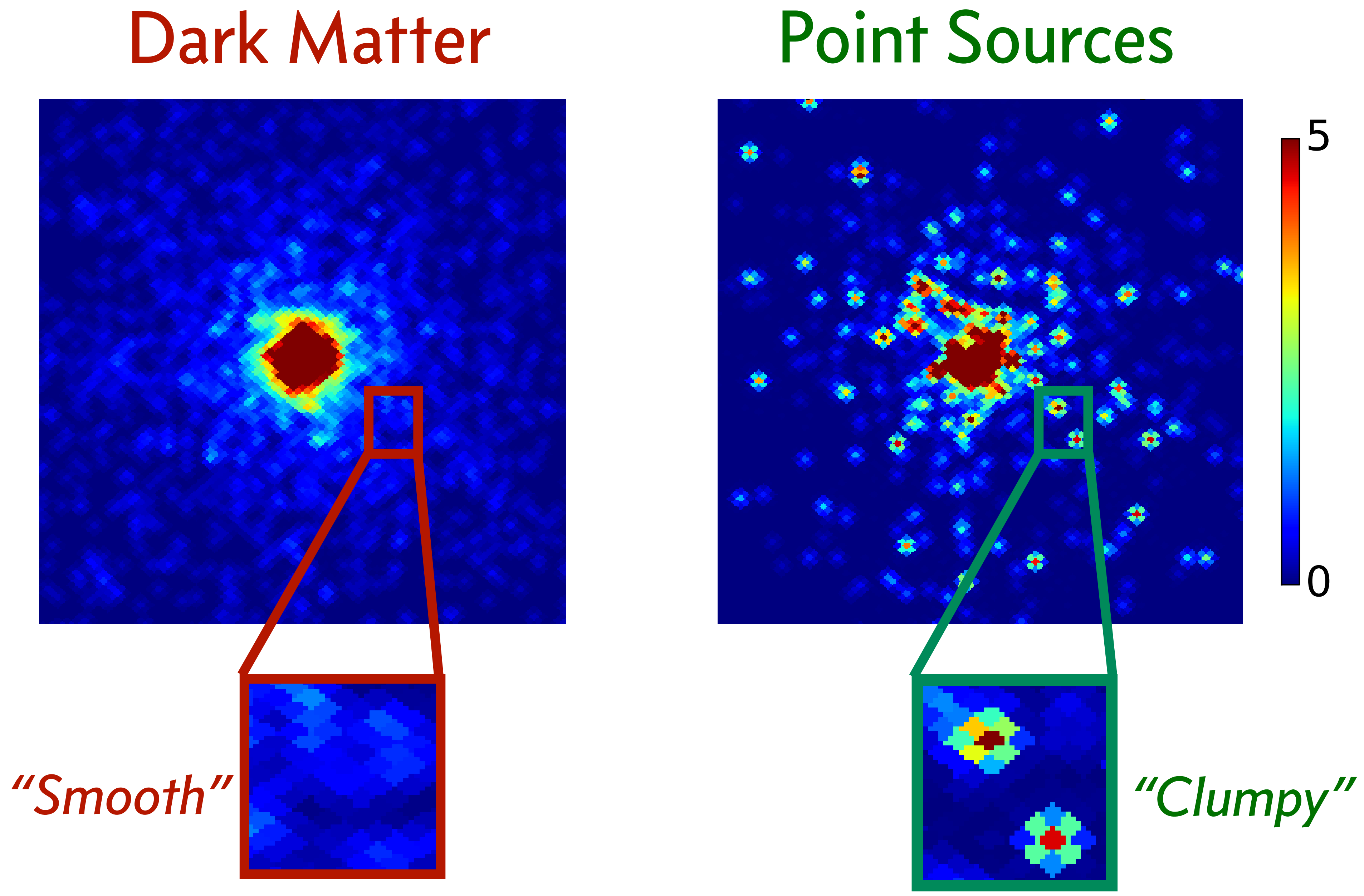
Gamma-ray image x

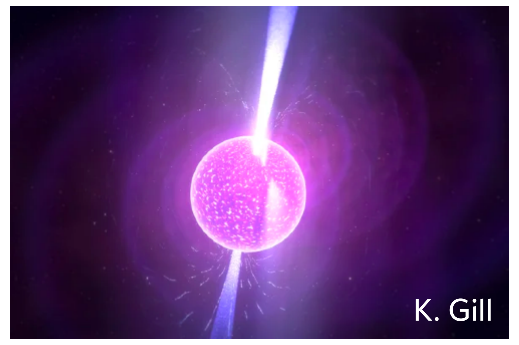


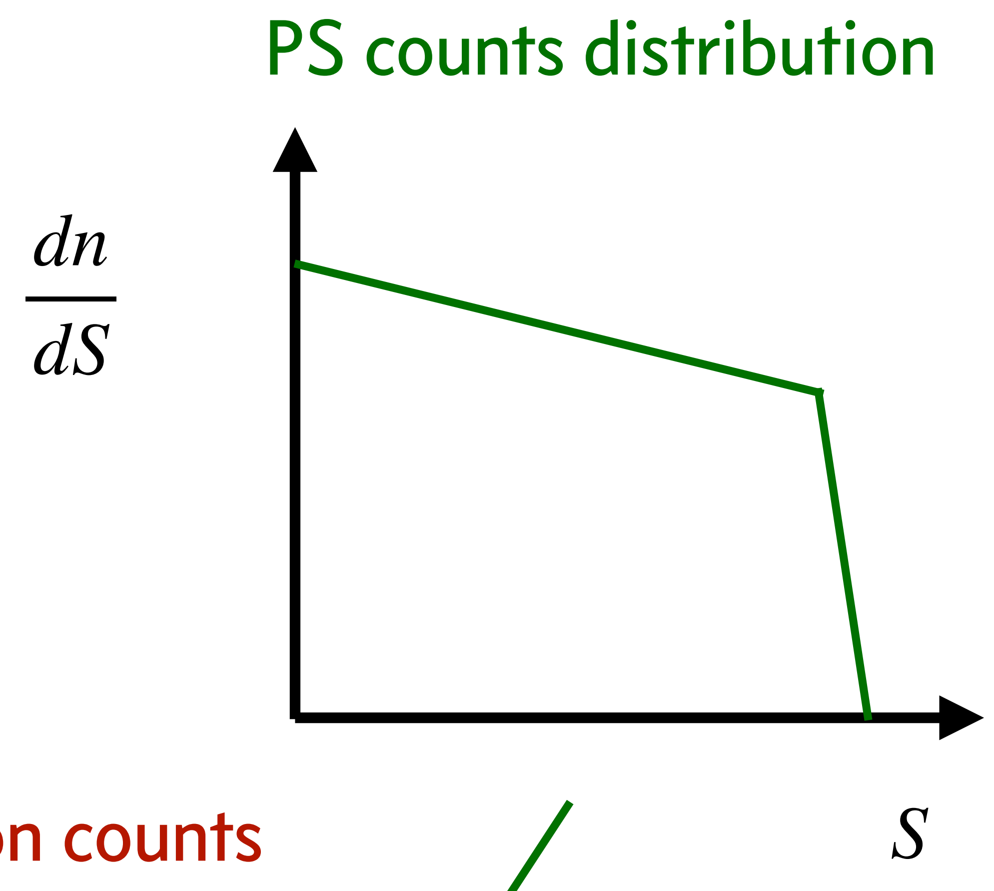
Distinguishing DM from PS with the 1-point PDF

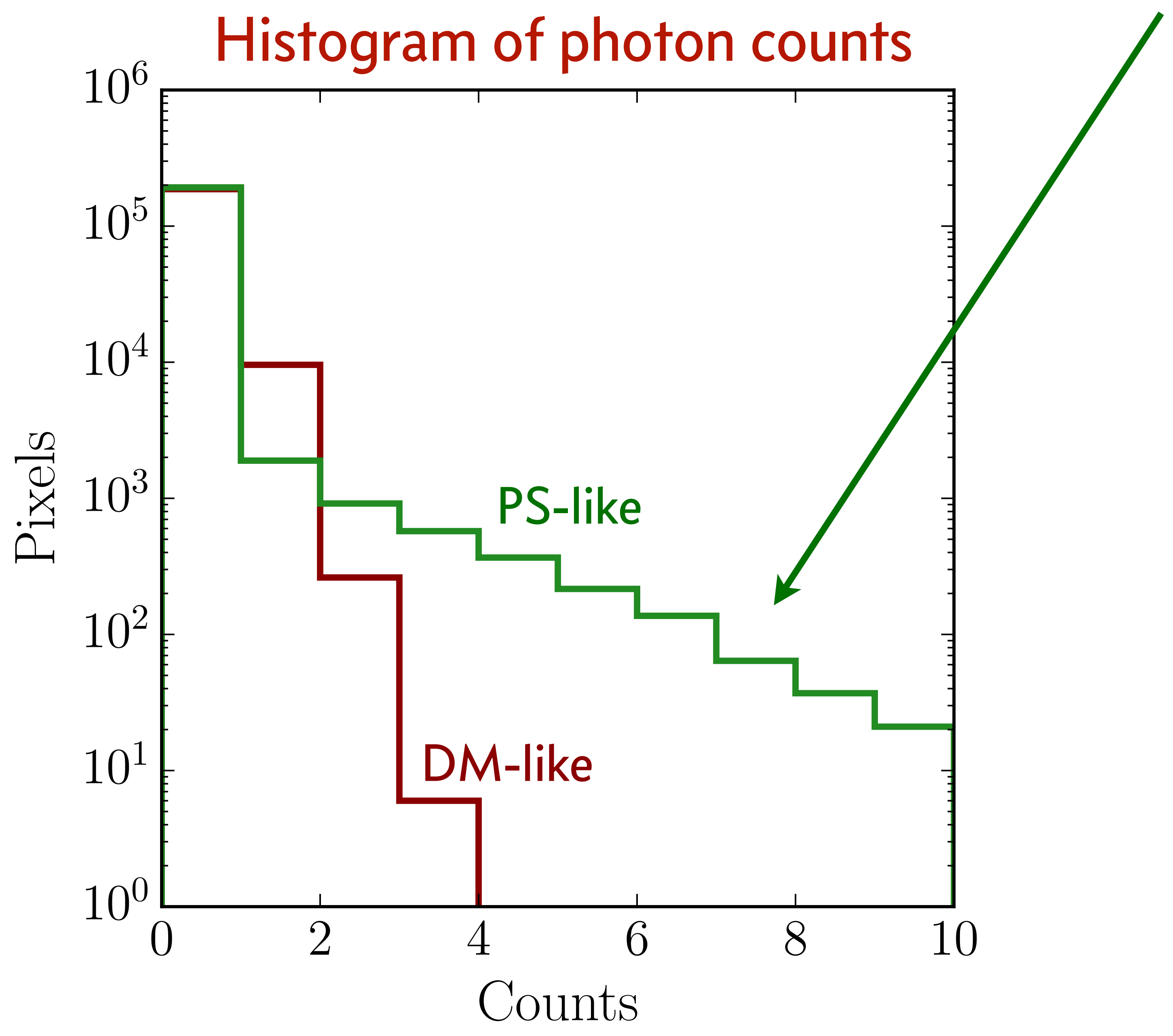


Lee et al [JCAP 2015]

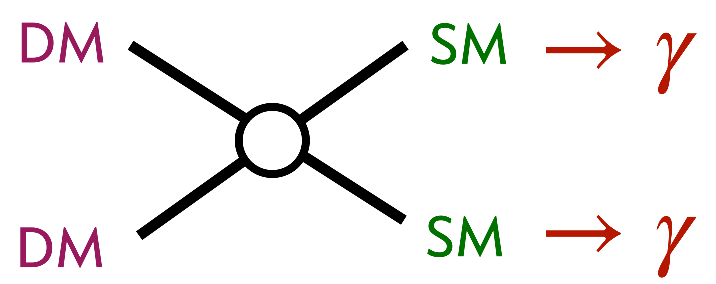




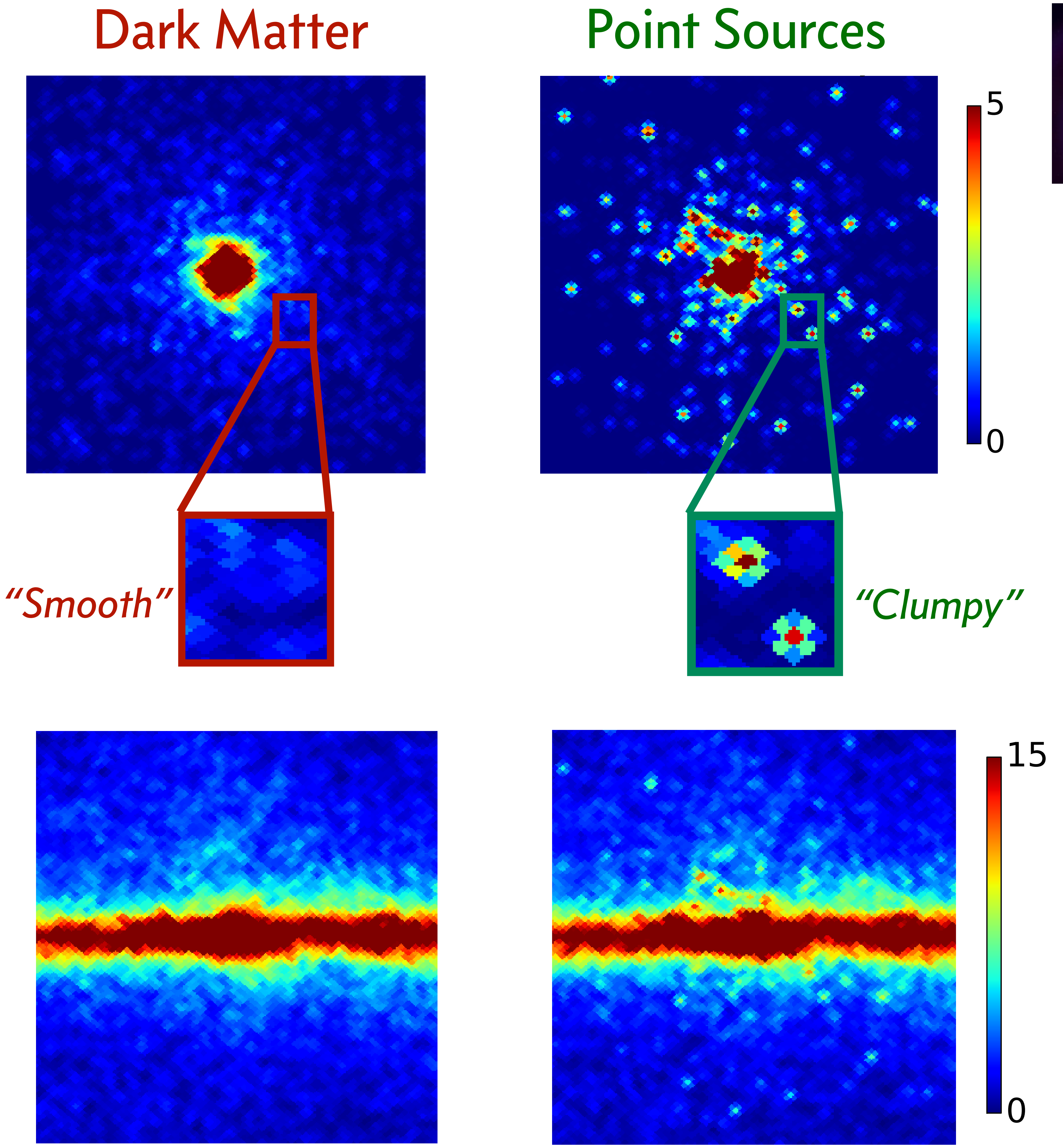


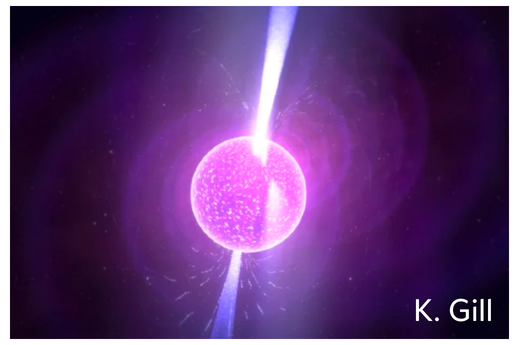


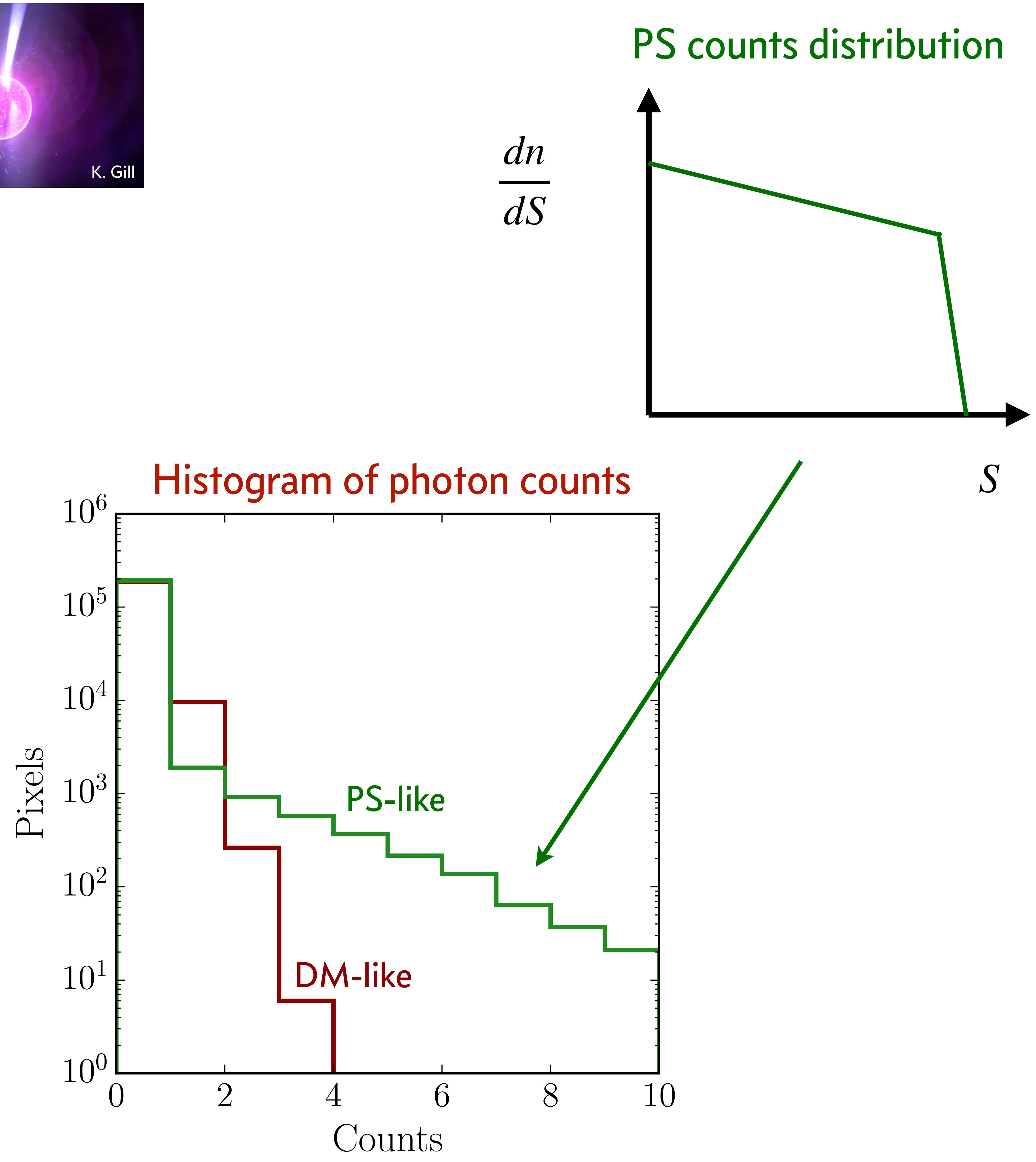
Distinguishing DM from PS with the 1-point PDF



Lee et al [JCAP 2015]

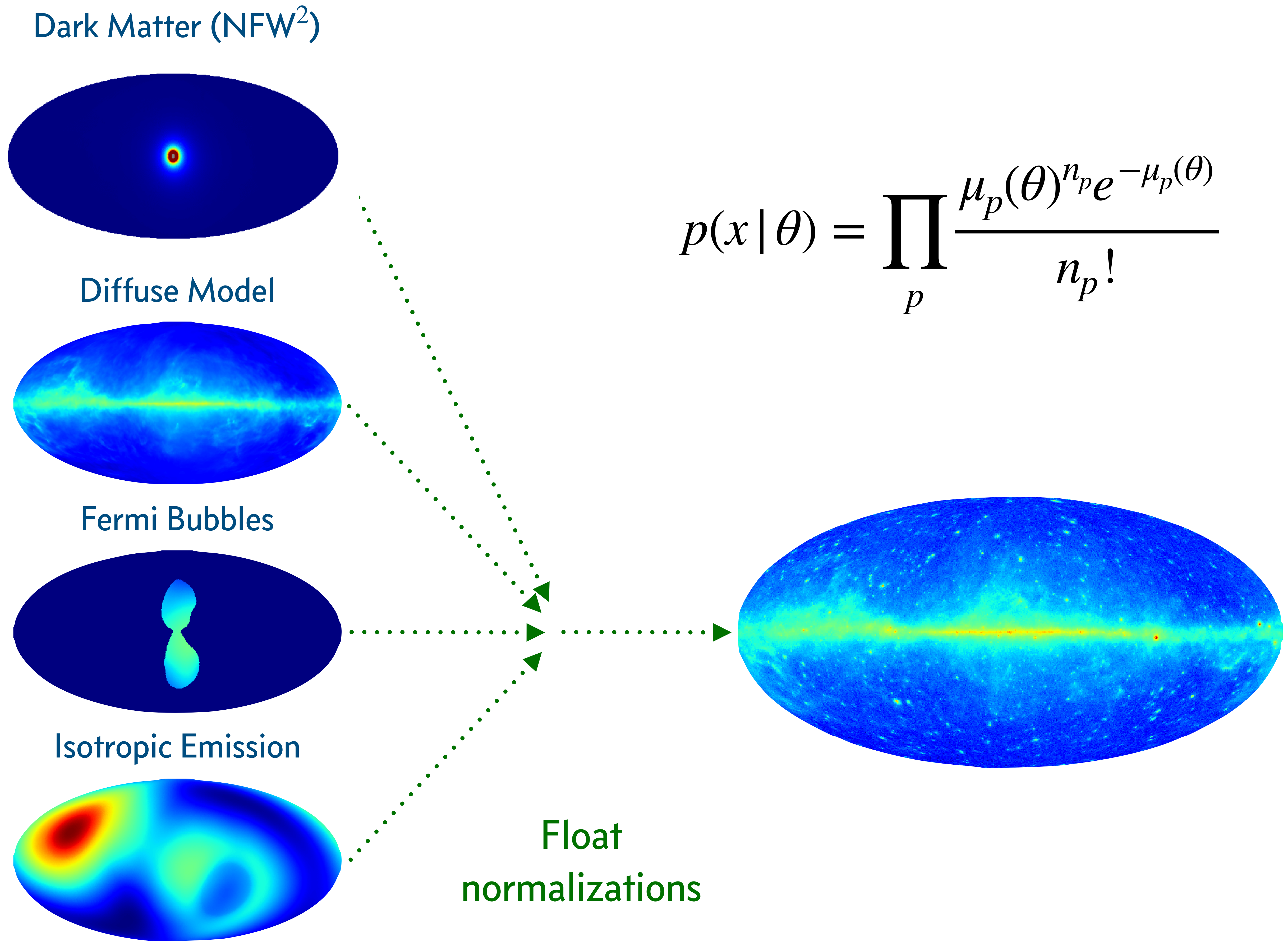




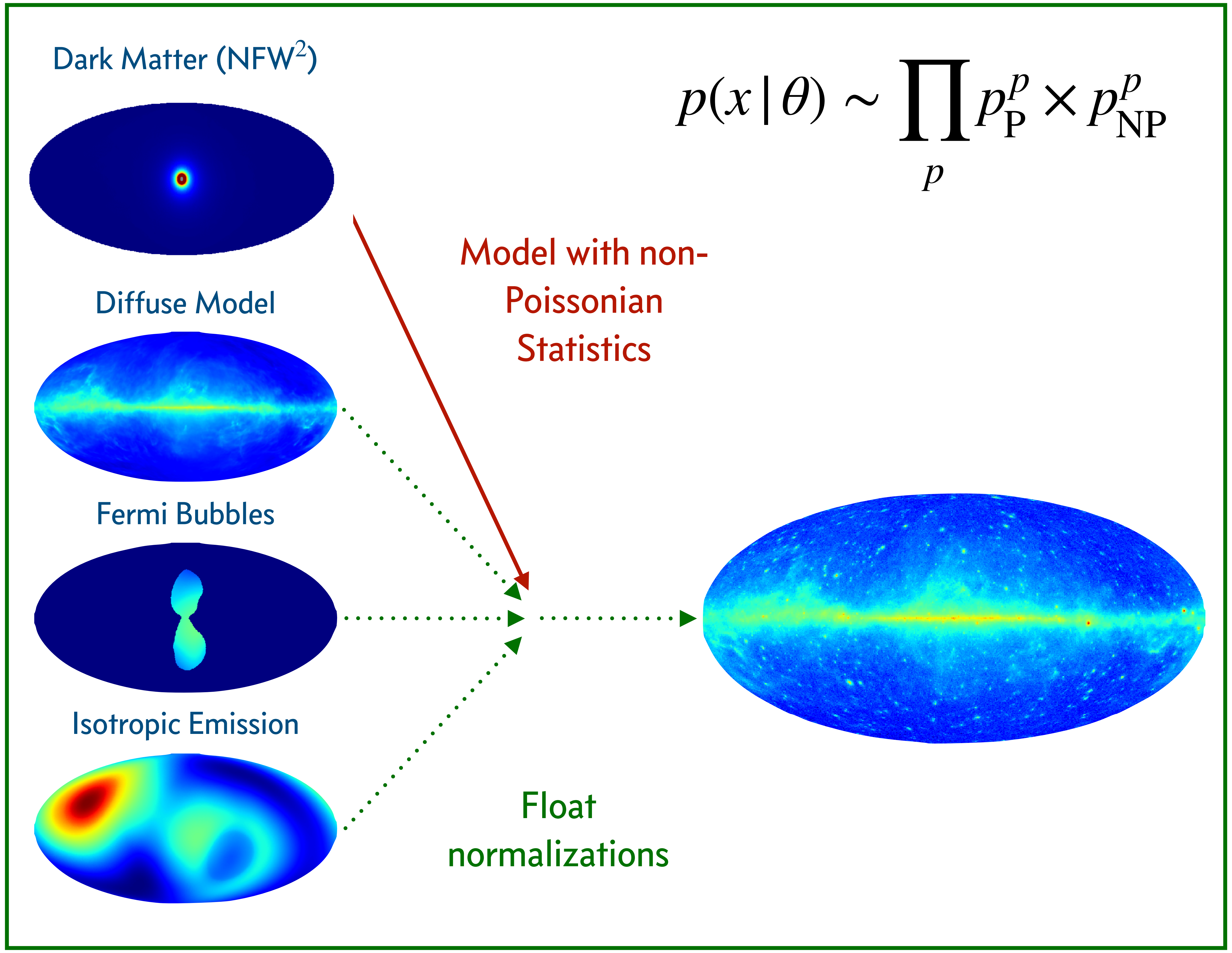


Poissonian template fitting

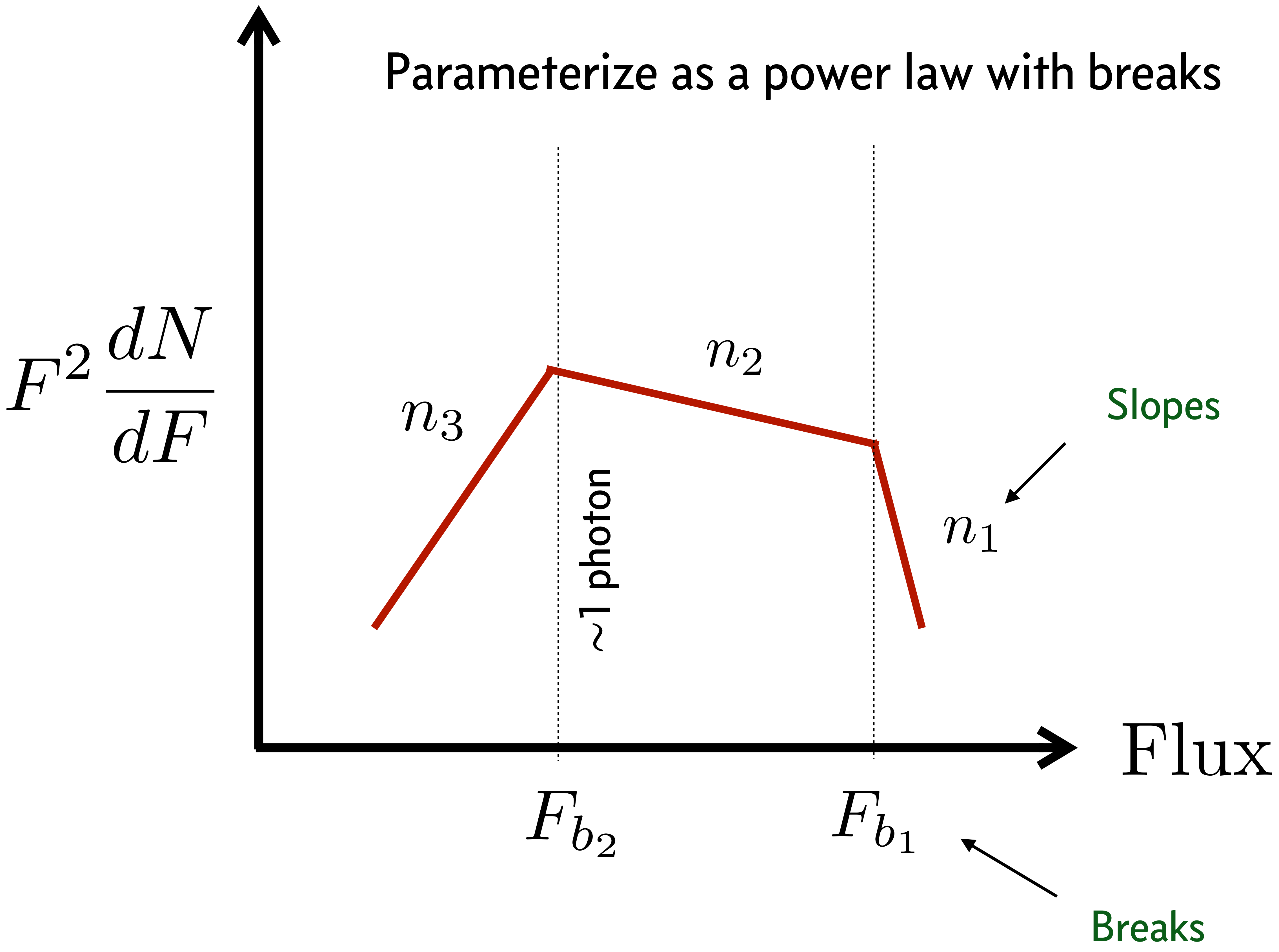
$$p(x\,|\,\theta) = \prod_p \frac{\mu_p(\theta)^{n_p} e^{-\mu_p(\theta)}}{n_p!}$$



Non-Poissonian Template Fitting (NPTF)

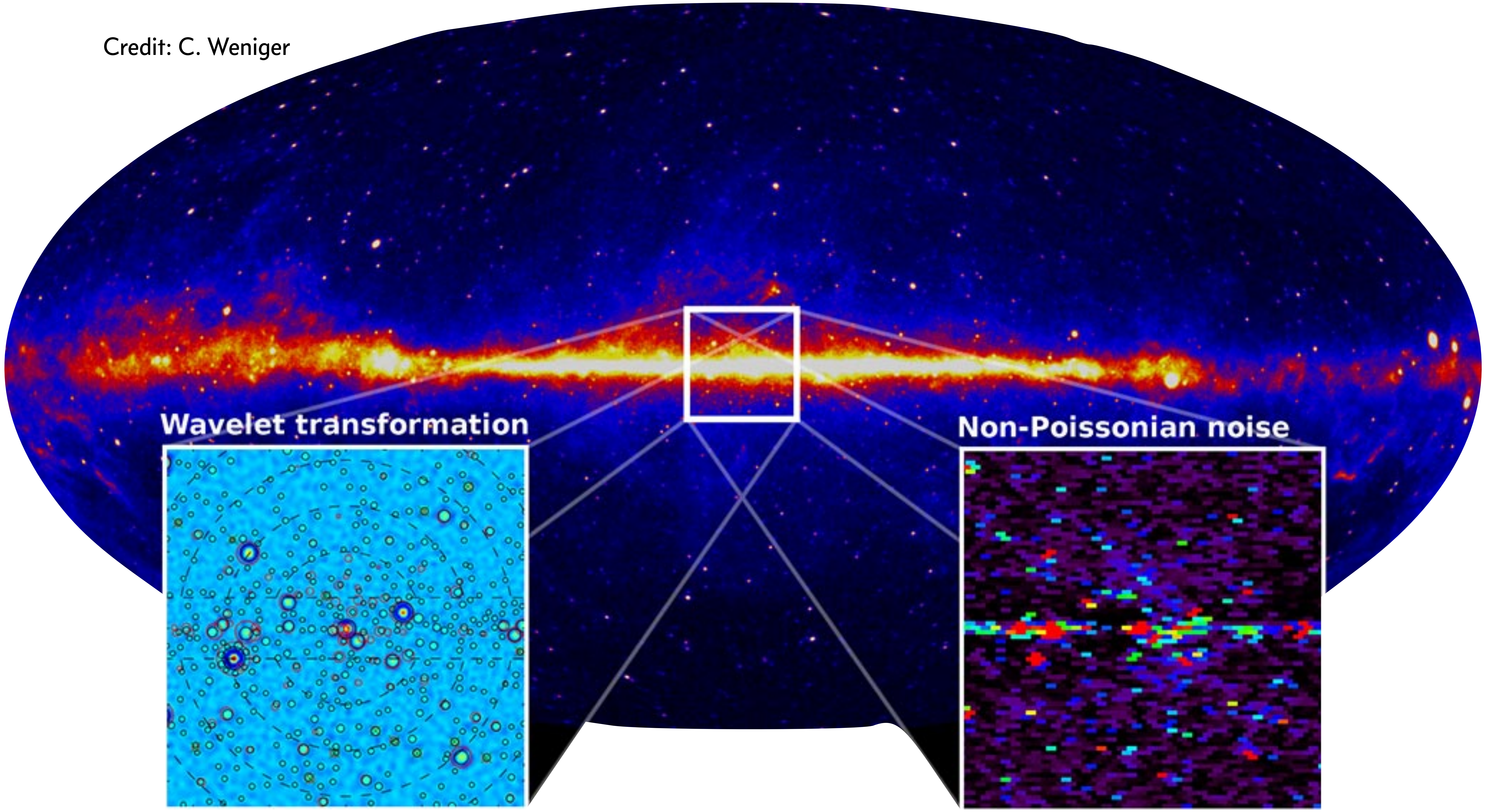


Source-count distribution gives number of sources in a given pixel with a flux between F and F+dF



Status *c.*2015: Evidence for unresolved point sources

Credit: C. Weniger

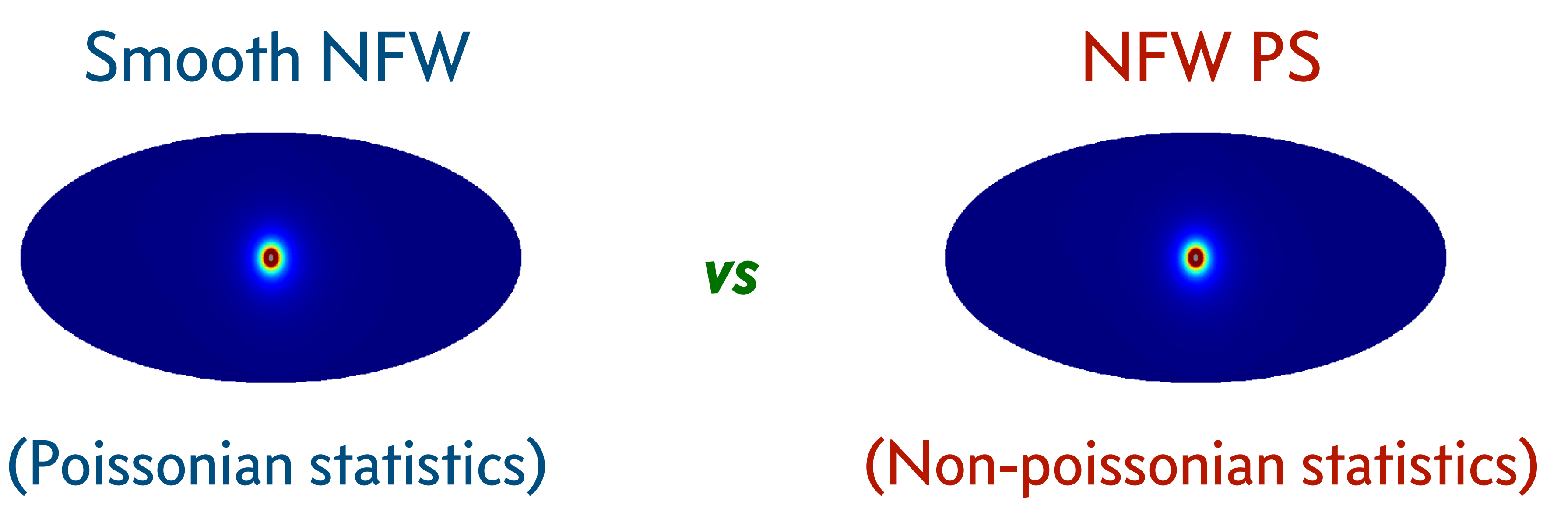


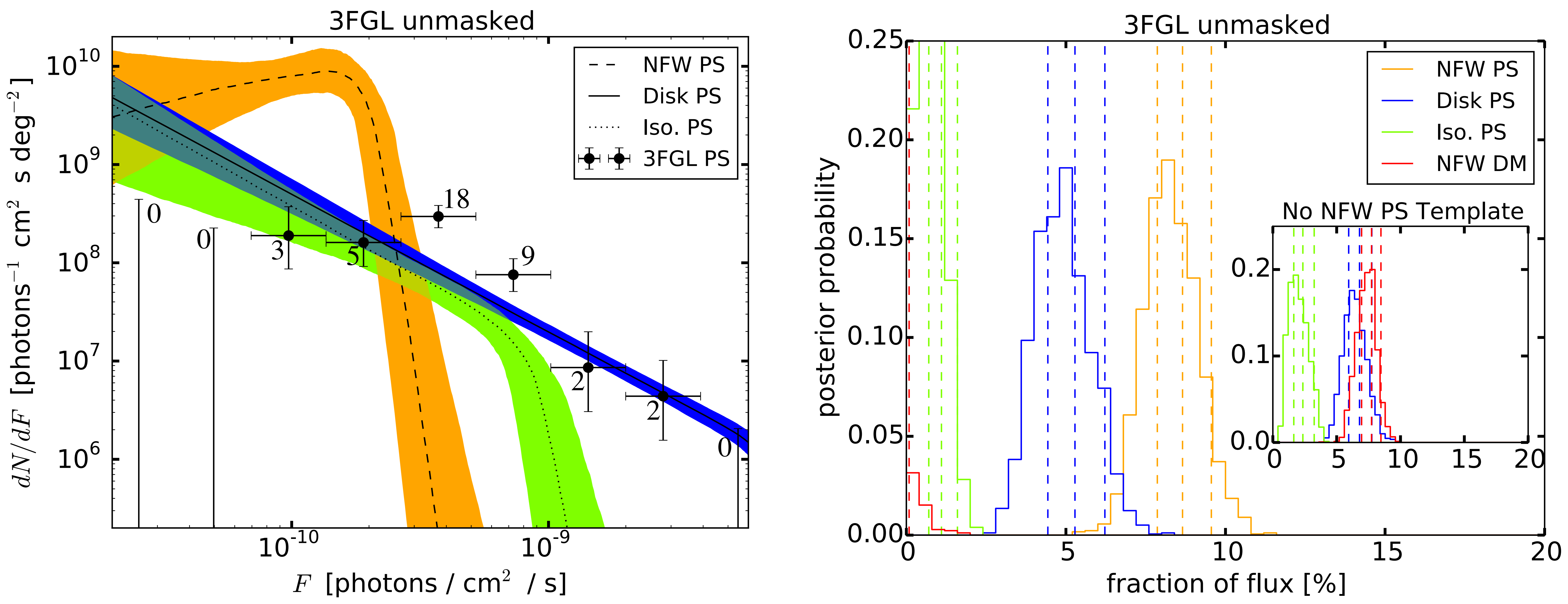
Bartels et al [PRL 2016]

Lee et al [PRL 2016]

NPTF analysis of Lee et al (c. 2015)

Which model provides a better fit to the data?





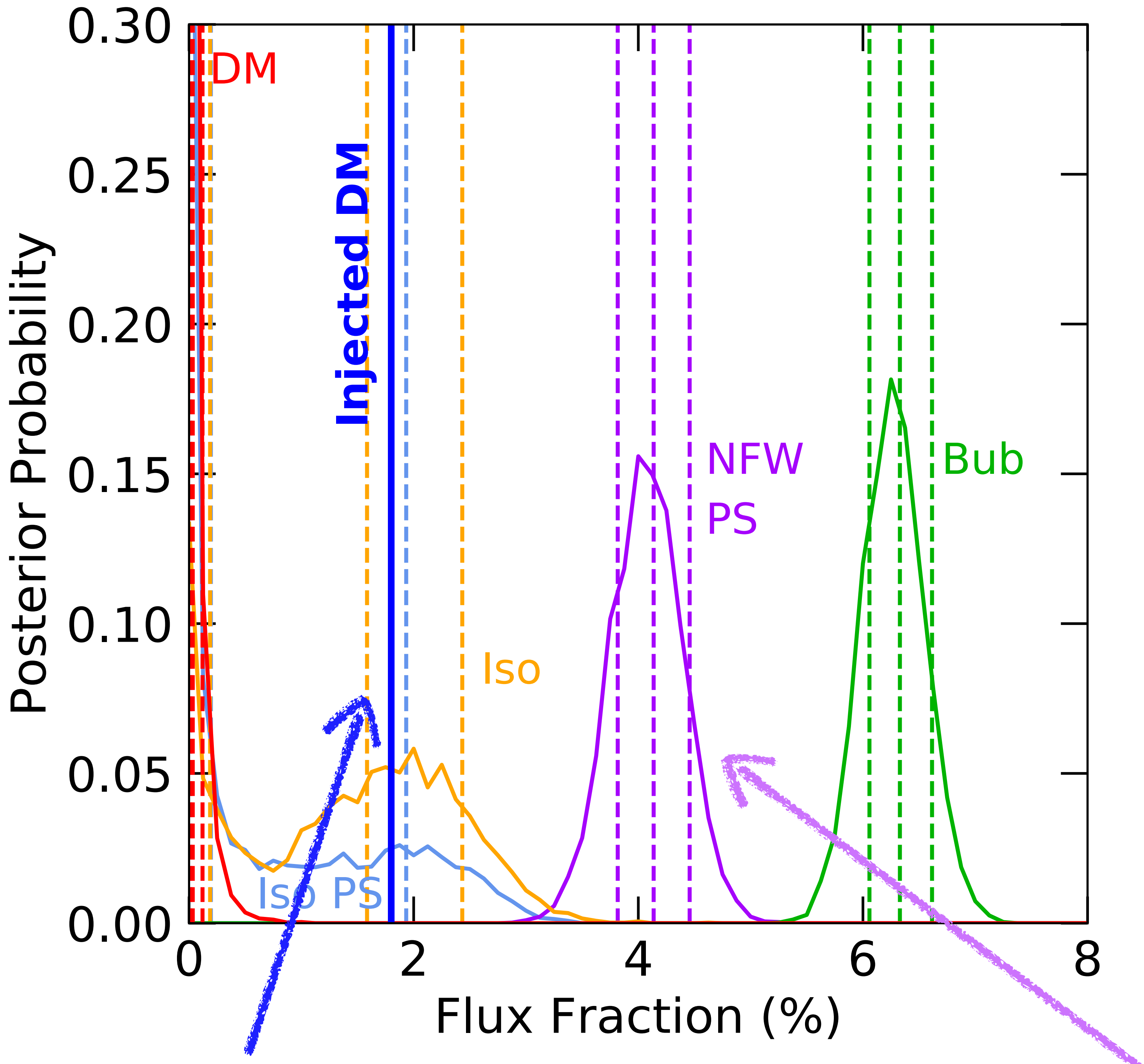
- *Excess flux is entirely accounted for by the NFW PS template*
- *Bayes factor in preference for NFW point sources is $\sim 10^7$*

Lee et al [PRL 2016]

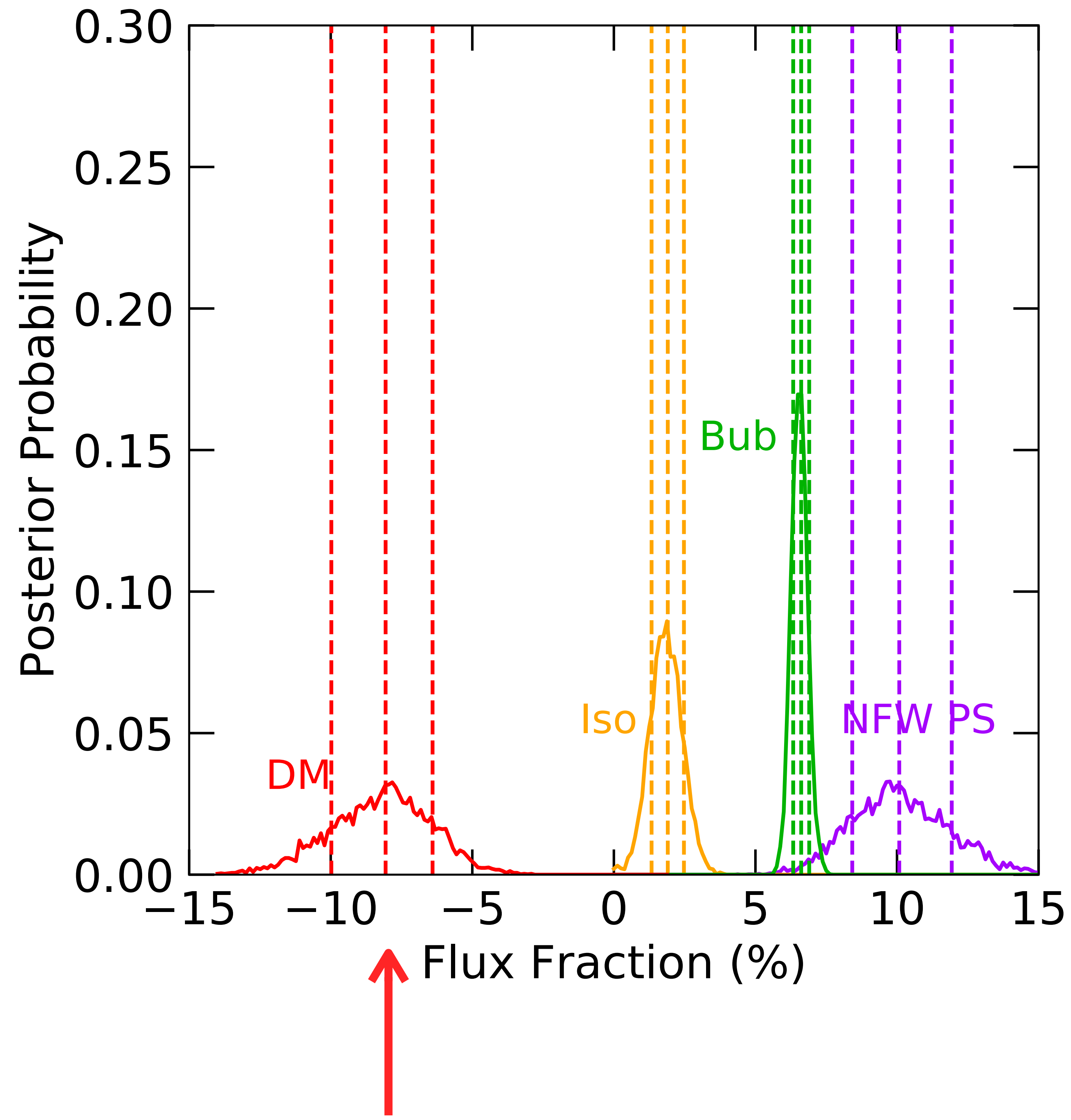
Dark matter strikes back? (c. 2019)

Leane & Slatyer [PRL 2019]
+ Leane & Slatyer [PRL 2020, PRD 2020]

Performed a closure test:
Inject a DM signal onto the real data, then try to recover it with the NPTF pipeline



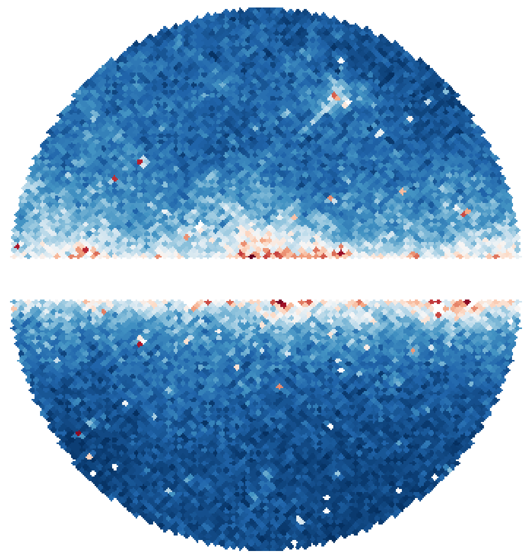
Injected DM flux gets reconstructed as PSs



DM flux can go negative if allowed to

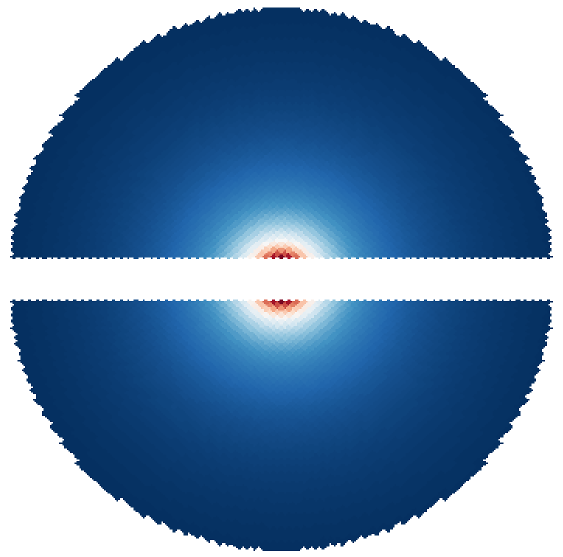
What could go wrong?

Fermi γ-ray data



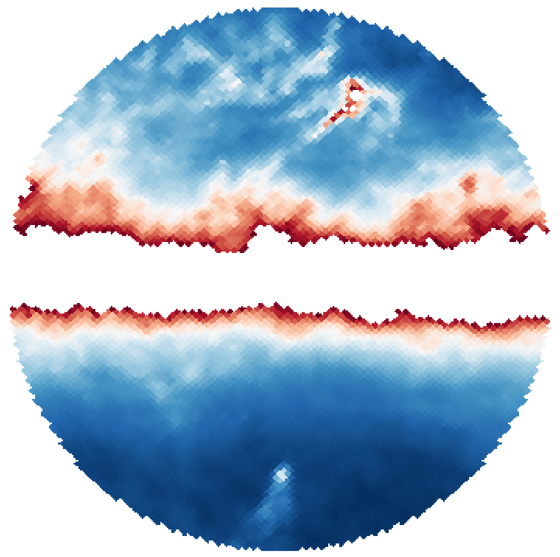
=

GCE signal (~10%)

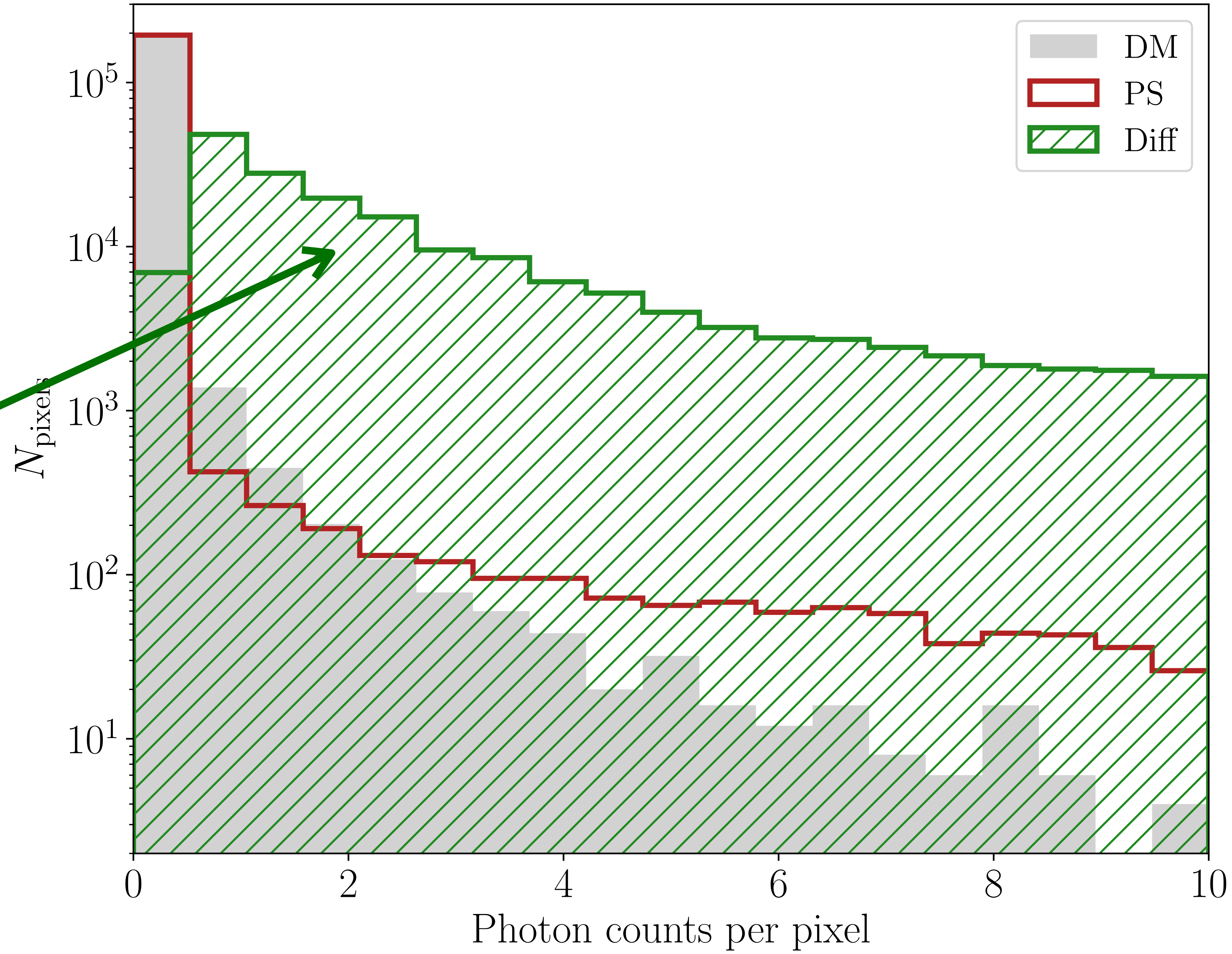


+

Diffuse background (~80%)



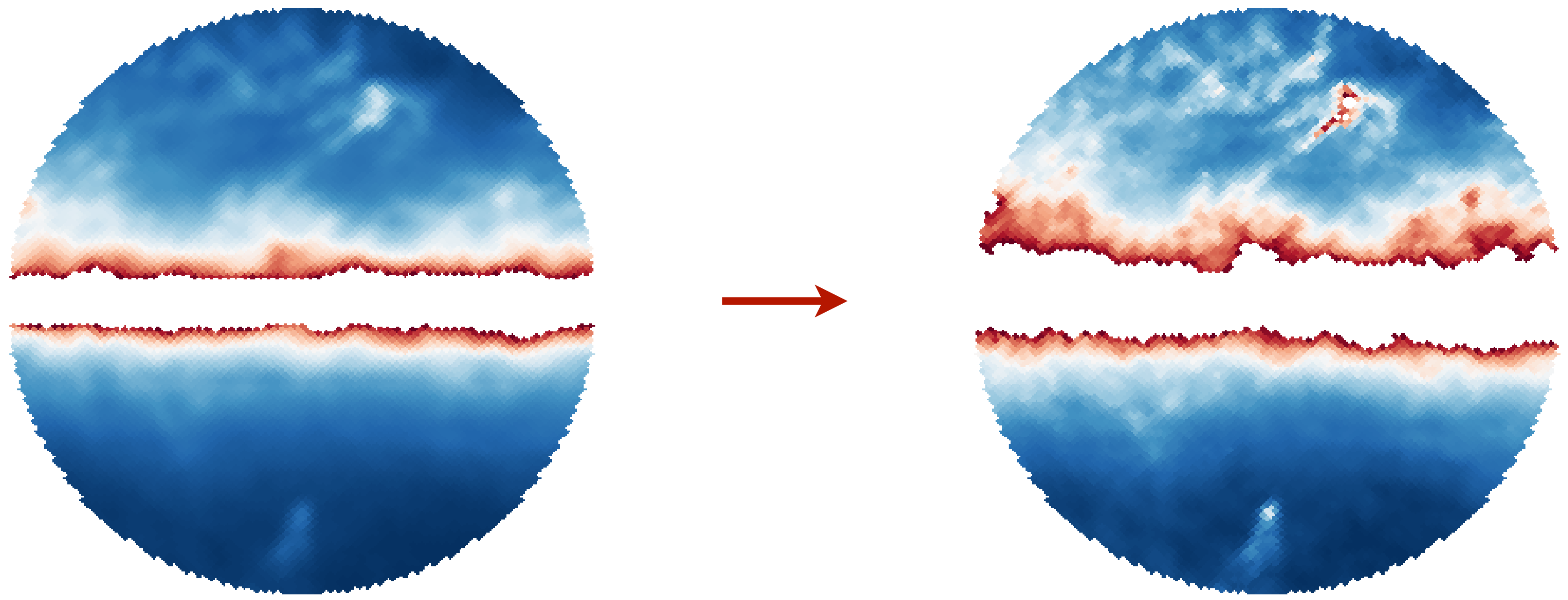
+ ...



Diffuse foregrounds make up most of the observed emission in the Galactic Center

Not an error in the NPTF method!

Better diffuse models



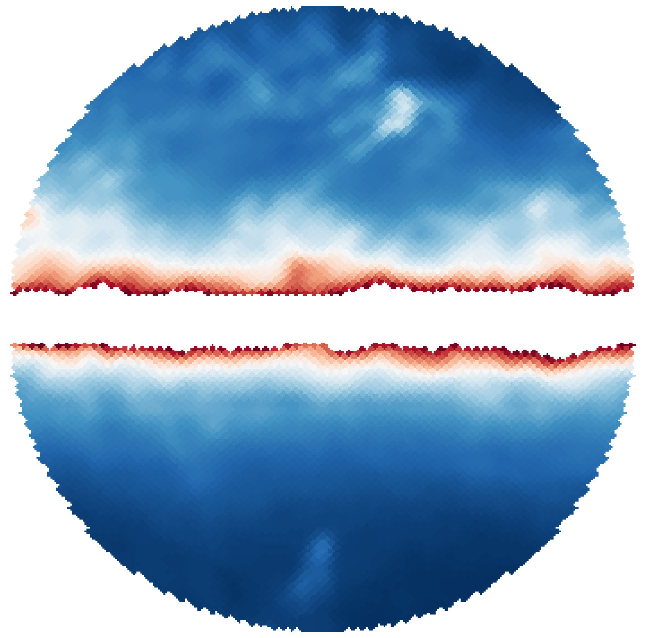
Background
Model p6v11

Background
"Model O"

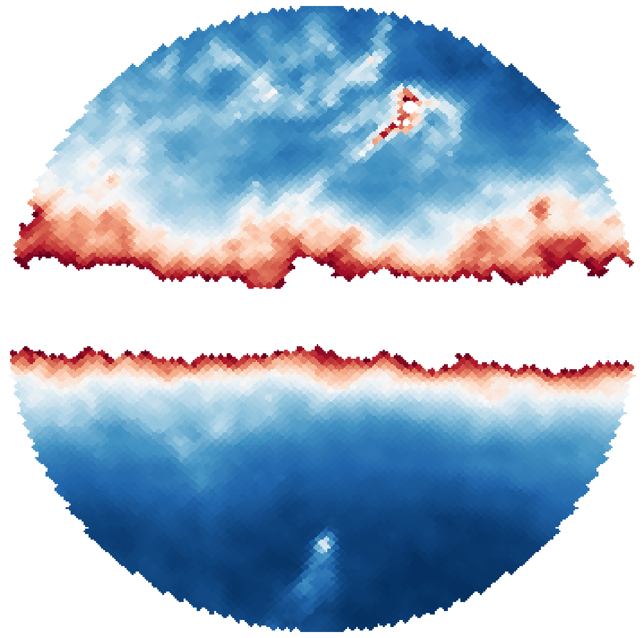
- *Updated gas tracers*
- *3D radiation field for IC*
- *Components fit in several Galactocentric rings*

Macias et al [JCAP 2019]
Macias et al [Nat. Ast. 2017]

Better diffuse models



Background
Model p6v11



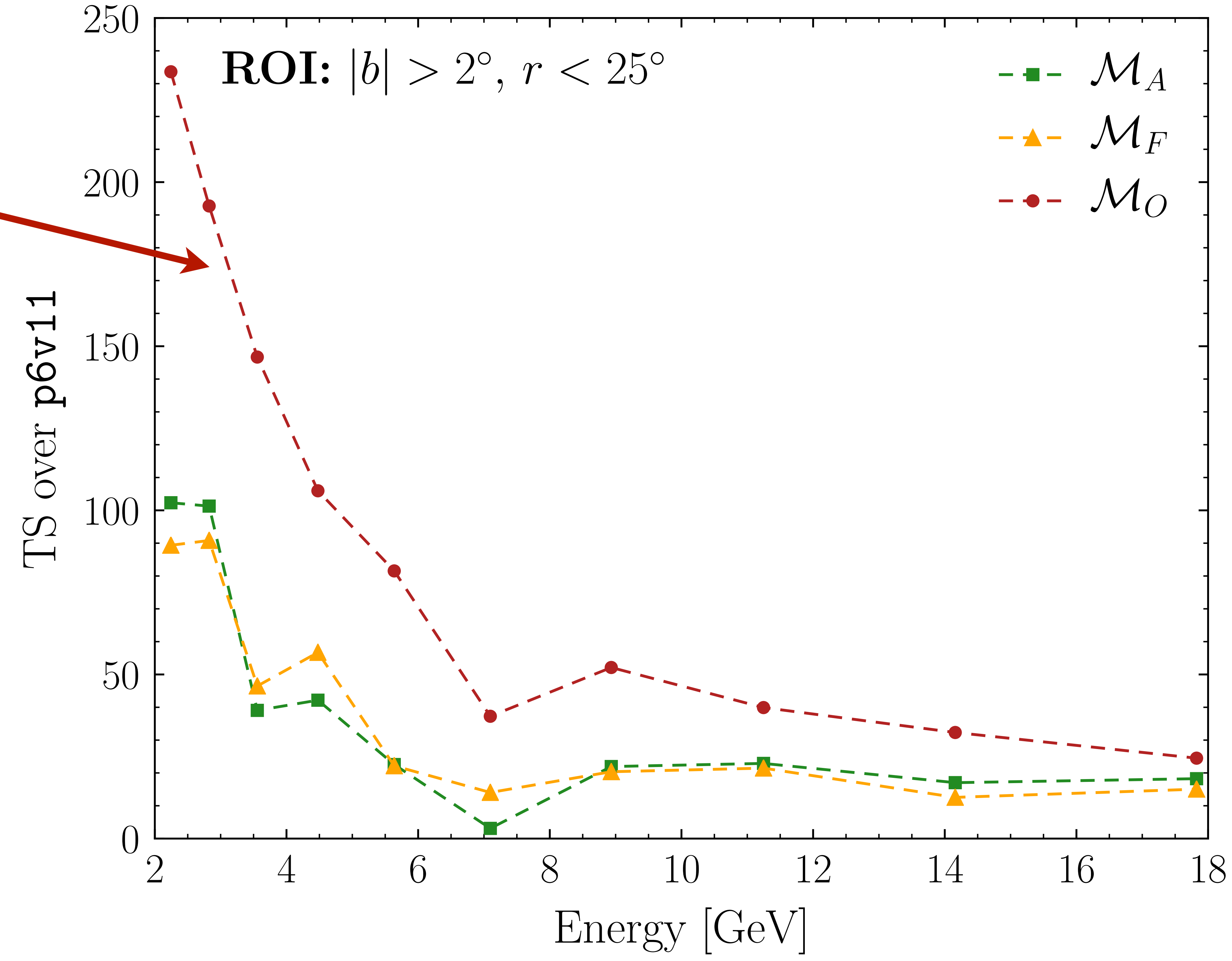
Background
"Model O"

- *Updated gas tracers*
- *3D radiation field for IC*
- *Components fit in several Galactocentric rings*

Macias et al [JCAP 2019]

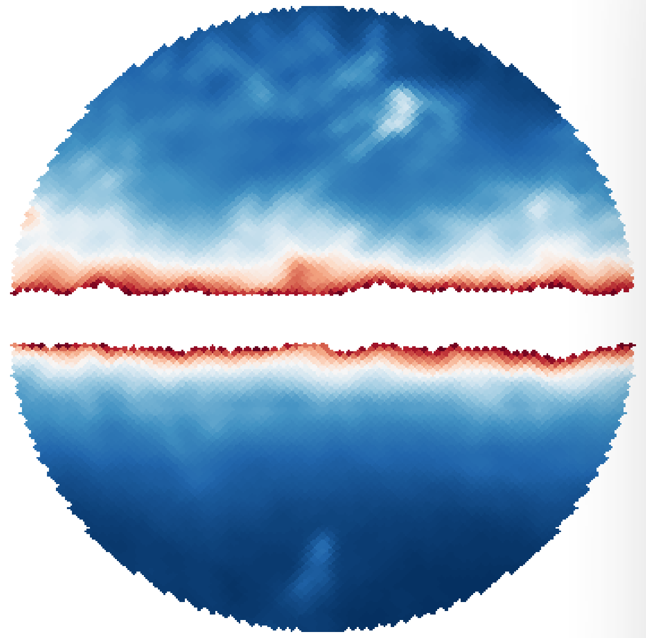
Macias et al [Nat. Ast. 2017]

TS improvement over previous models



Buschmann et al incl. SM [PRD 2020]

Better diffuse models



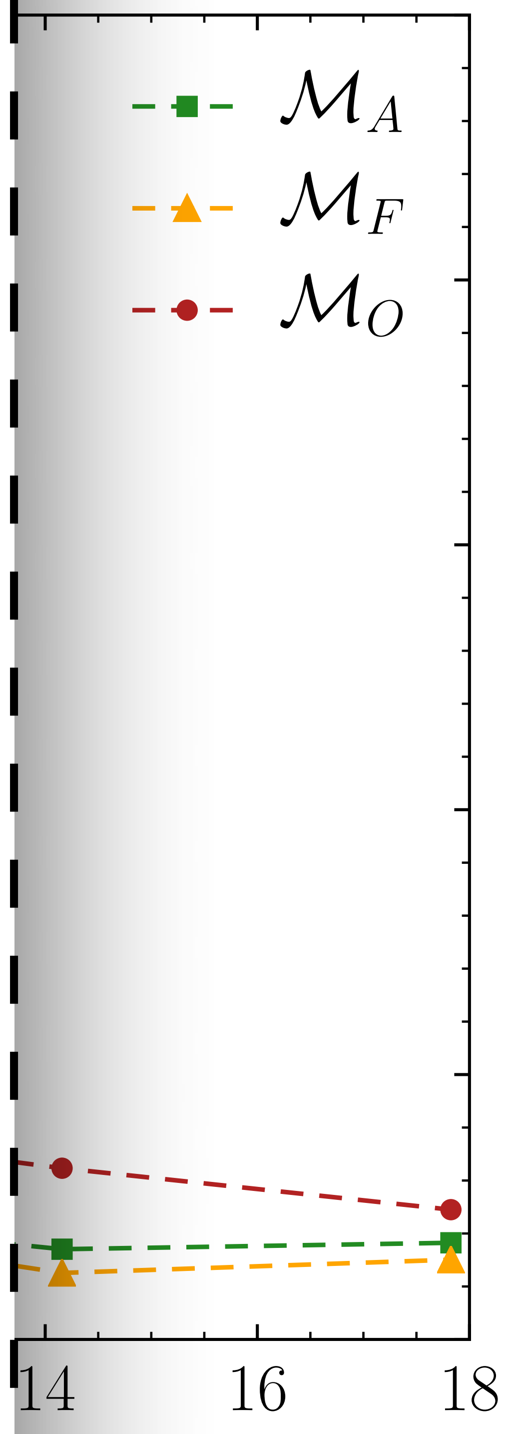
Background
Model p6v11

- *Updated gas tracers*
- *3D radiation field for IC*
- *Components fit in severa*

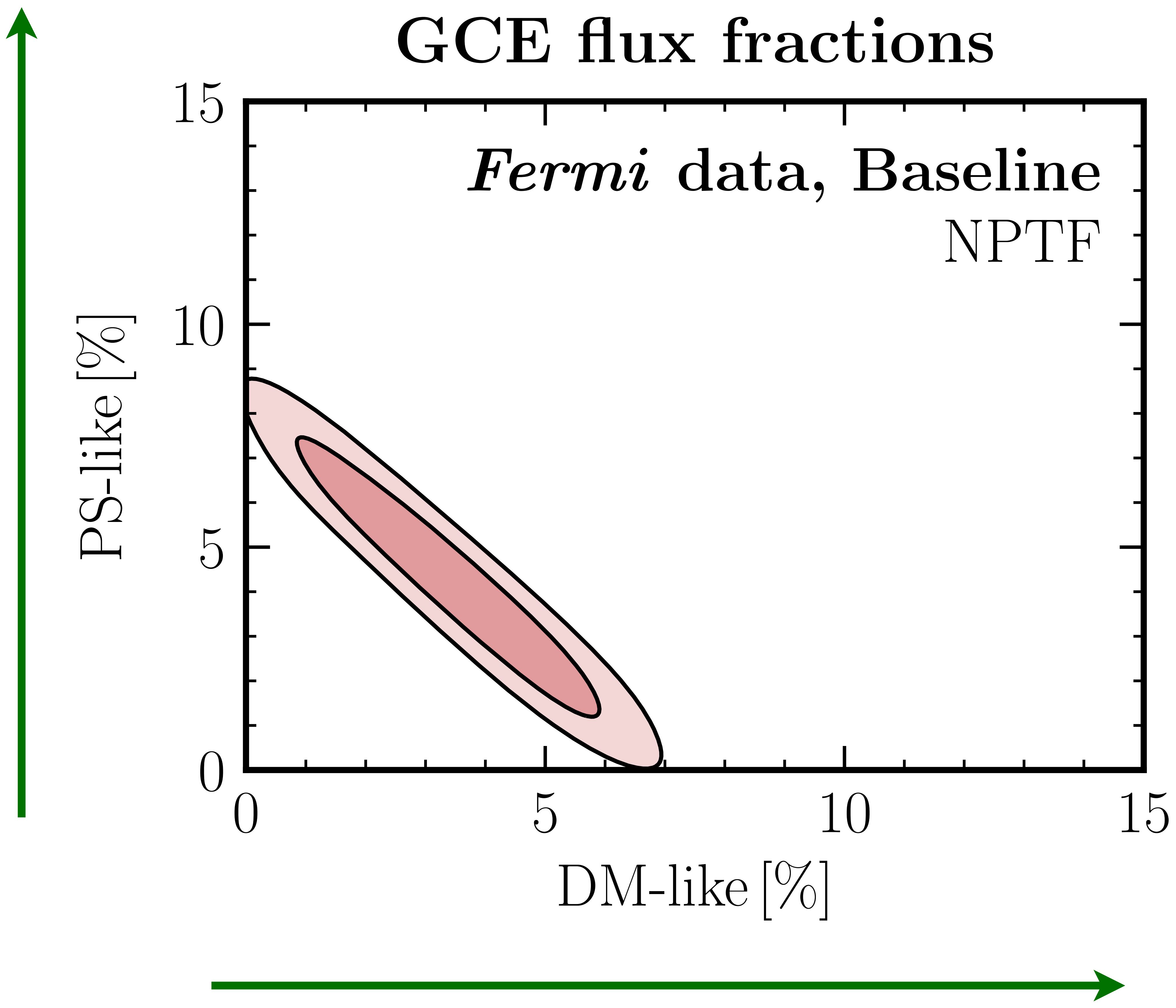
Macias et al [JCAP 2019]

Macias et al [Nat. Ast. 2017]

s models



D 2020]



The use of updated diffuse models shows a <u>reduced</u> but <u>still significant</u> preference for unresolved points sources

Giving the diffuse background mode freedom

Less freedom

More freedom

Less conservative

More information about signal

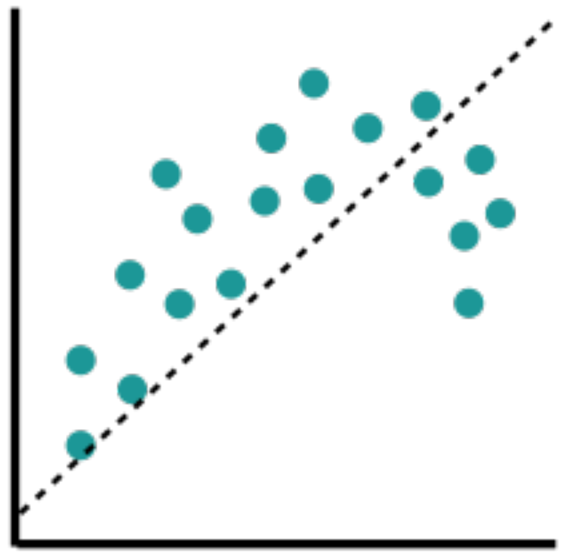
Contingent on background model

More conservative

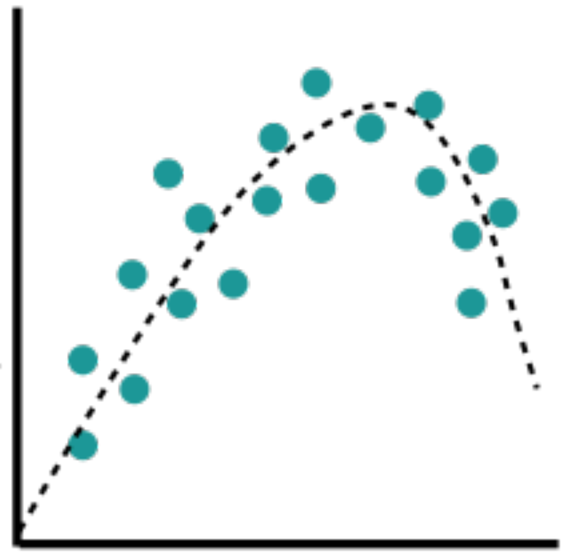
Less extractable information about signal

"Background model-independent"

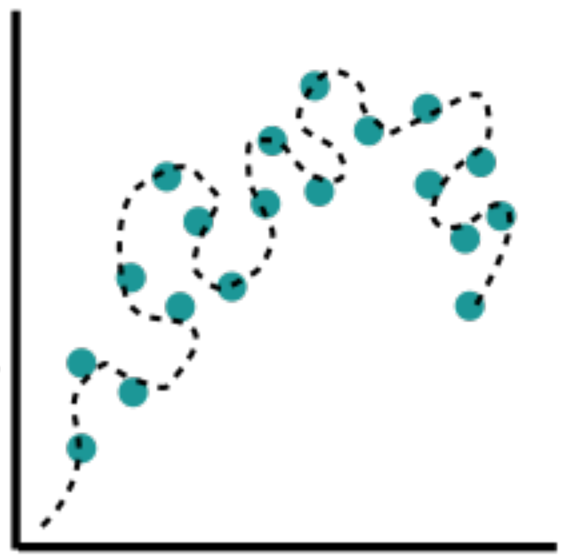
Underfit



Balanced

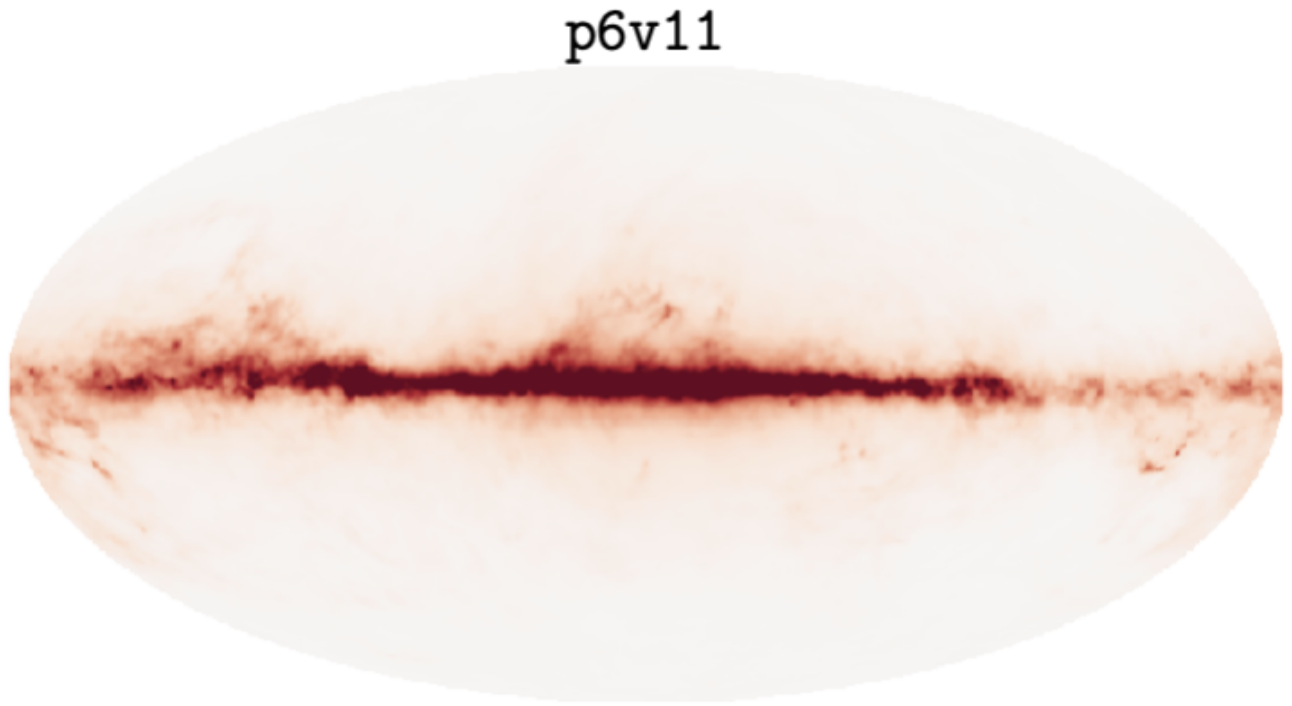


Overfit

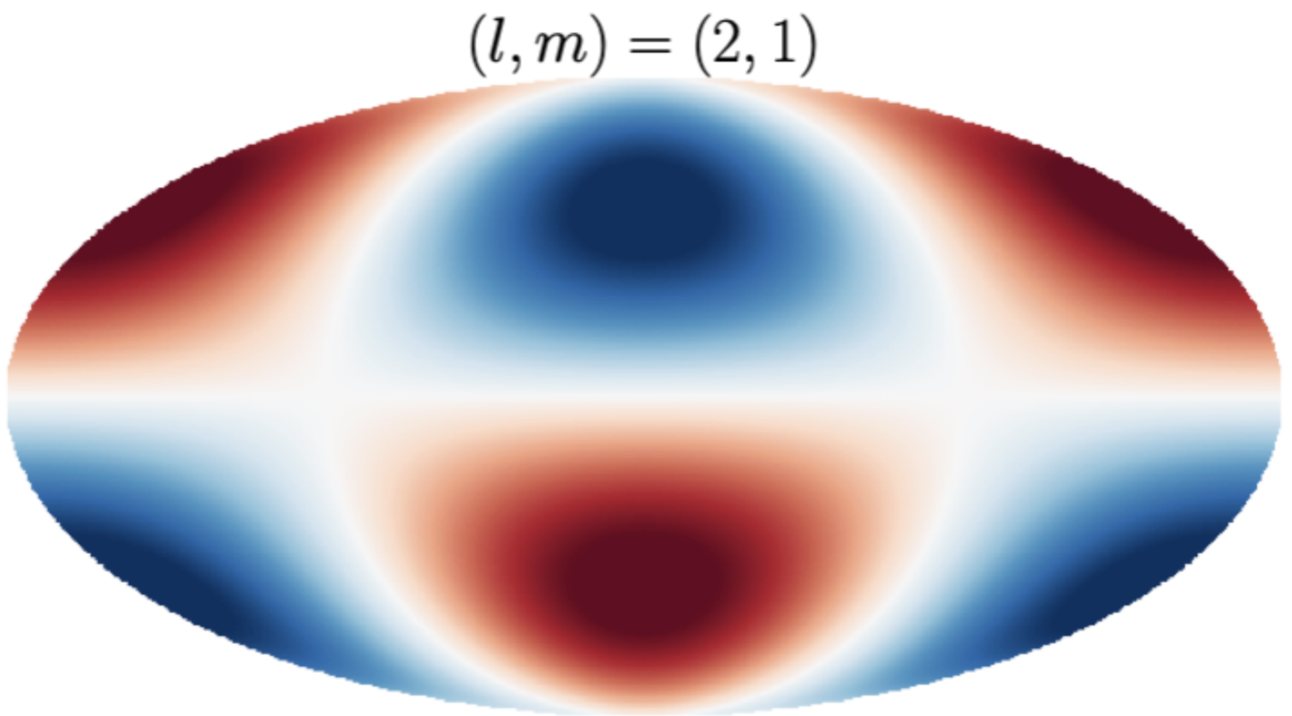


(Large-scale) harmonic marginalization

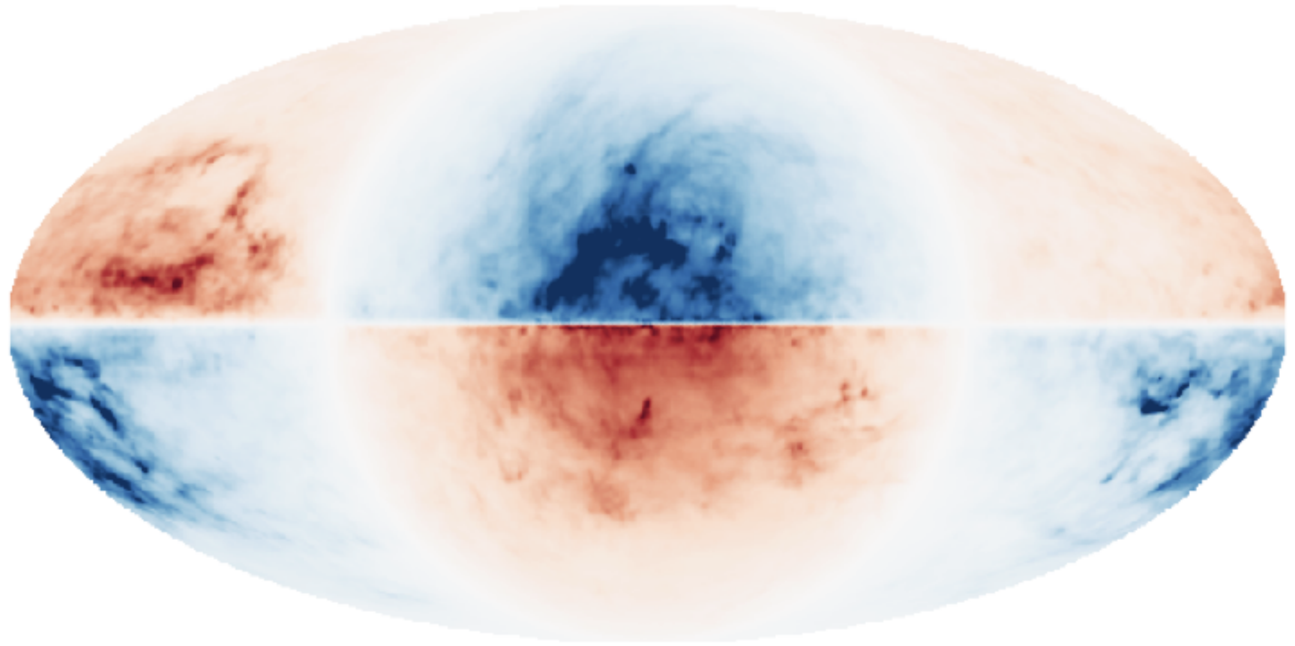
Extract large-scale harmonic components of diffuse model



$\times$



$=$



T_p^{diff} $\qquad$ $Y_p^{\ell,m}$

$$T_p^{\mathrm{harm}} \propto \underbrace{\hat{A}_{\mathrm{diff}} T_p^{\mathrm{diff}}}_{\text{Base model}} + \underbrace{\sum_{\ell,m} \hat{A}_{\ell,m} Y_p^{\ell,m} T_p^{\mathrm{diff}}}_{\text{Modulation of large scales}}$$

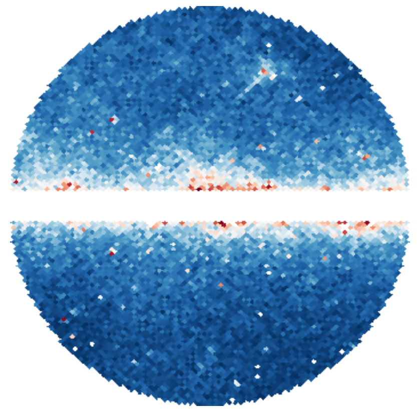
Give each large-scale component an independent degree of freedom

Buschmann et al incl. SM [PRD 2020]

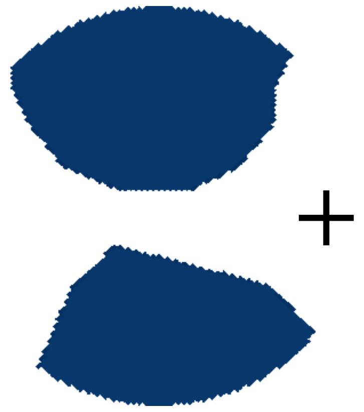
Gaussian process-augmented diffuse models

Traditional template fitting

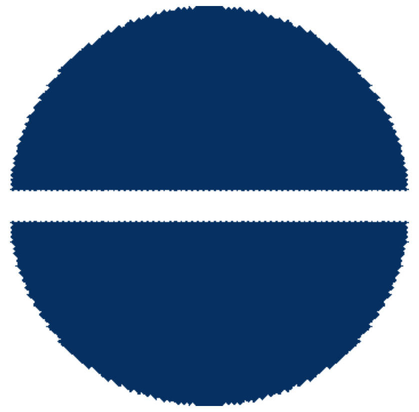
Augment background template with a GP



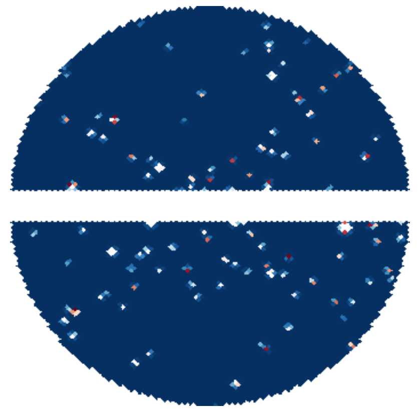
$= \mathrm{Pois}\Big(A_{\mathrm{bub}} \times$



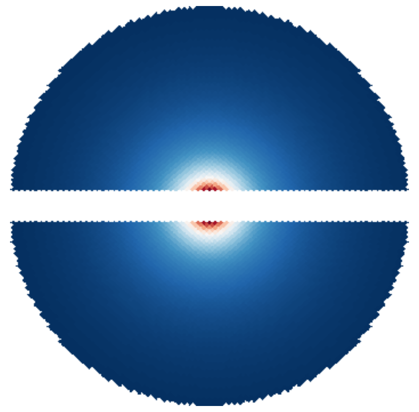
$+ A_{\mathrm{iso}} \times$



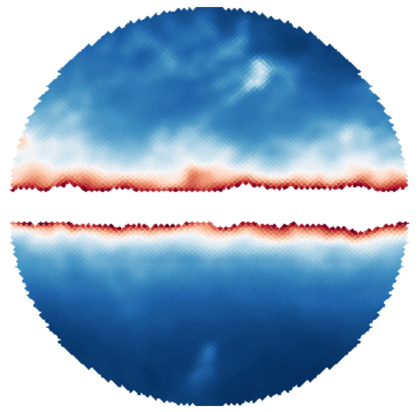
$+ A_{\mathrm{PS}} \times$



$+ A_{\mathrm{DM}} \times$



$+ \exp(\mathrm{GP}) \times$



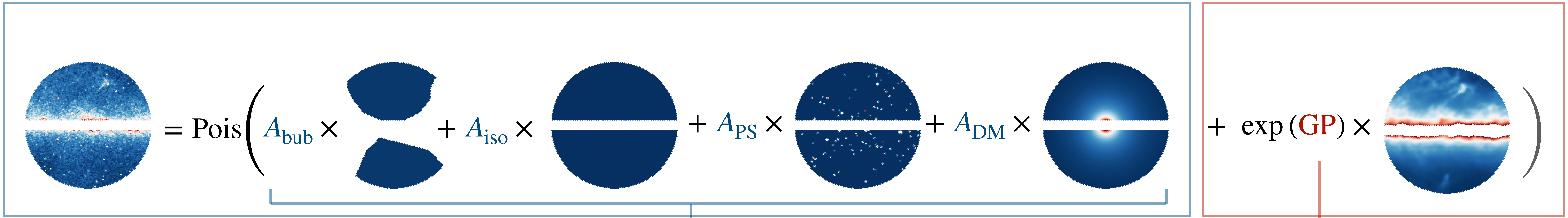
$\Big)$

SM, Cranmer [NeurIPS ML4PS Workshop 2020]

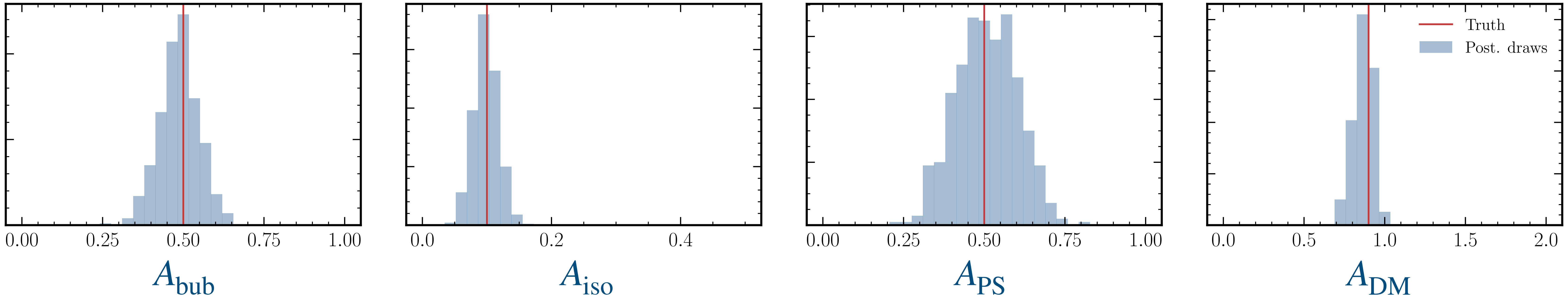
Gaussian process-augmented diffuse models

Traditional template fitting

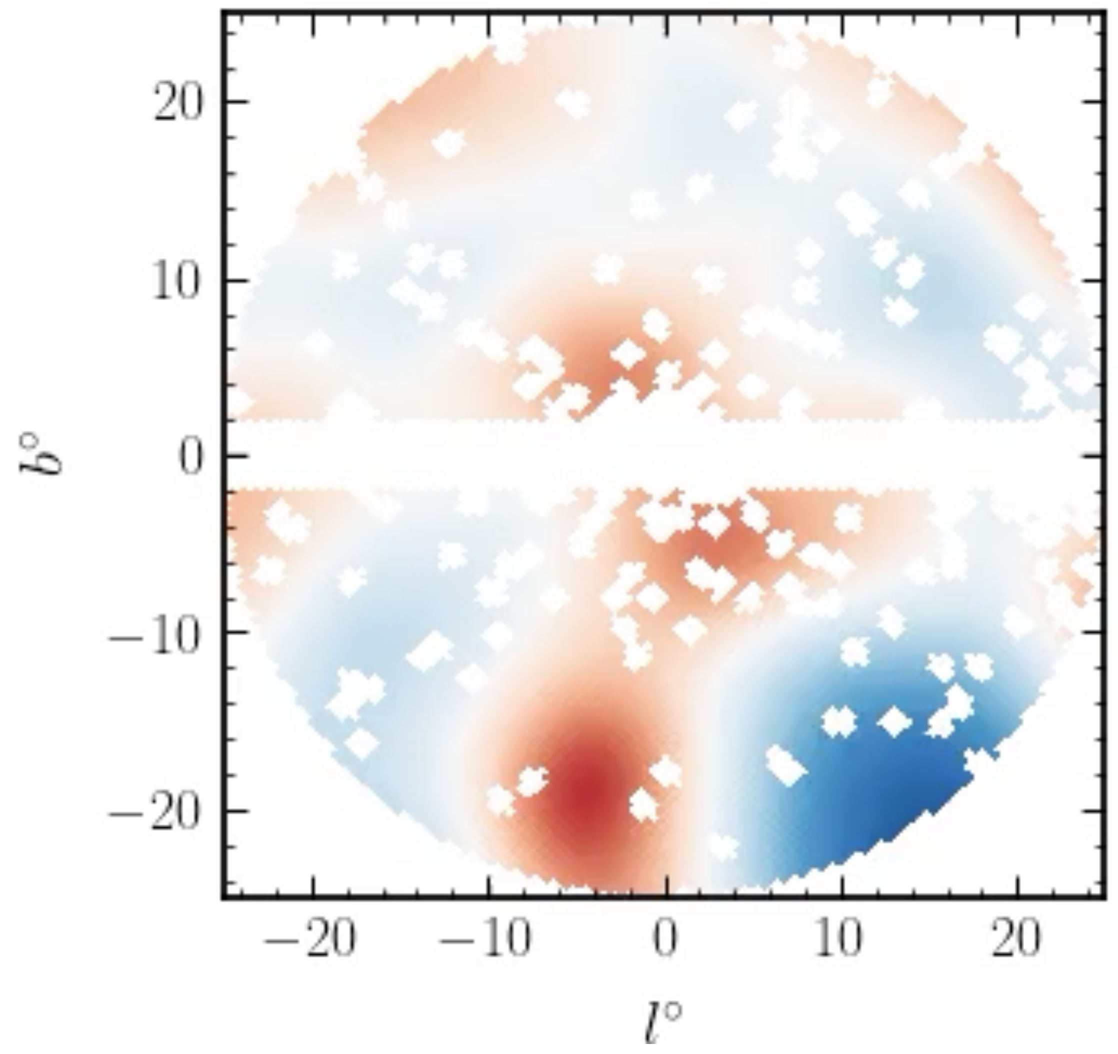
Augment background template with a GP



Component norm posteriors



GP posterior

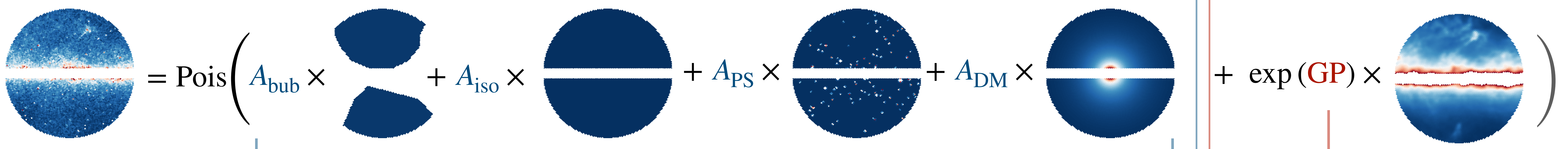


SM, Cranmer [NeurIPS ML4PS Workshop 2020]

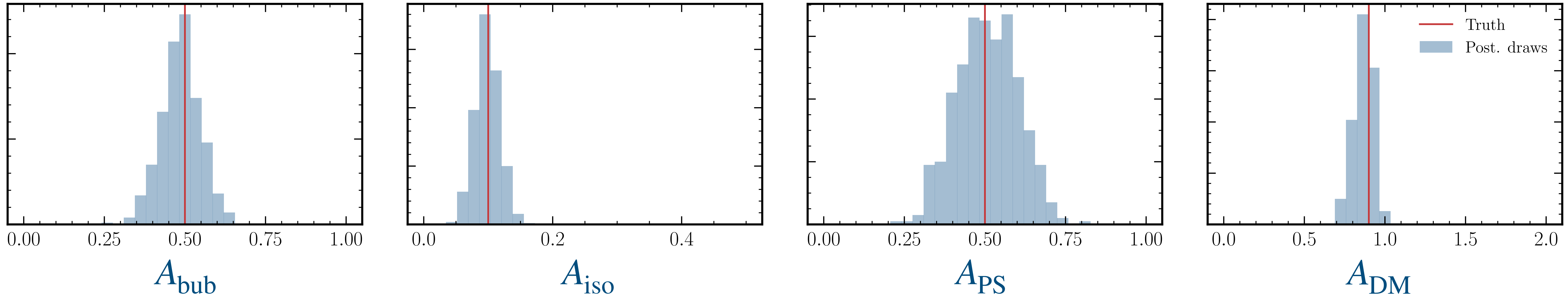
Gaussian process-augmented diffuse models

Traditional template fitting

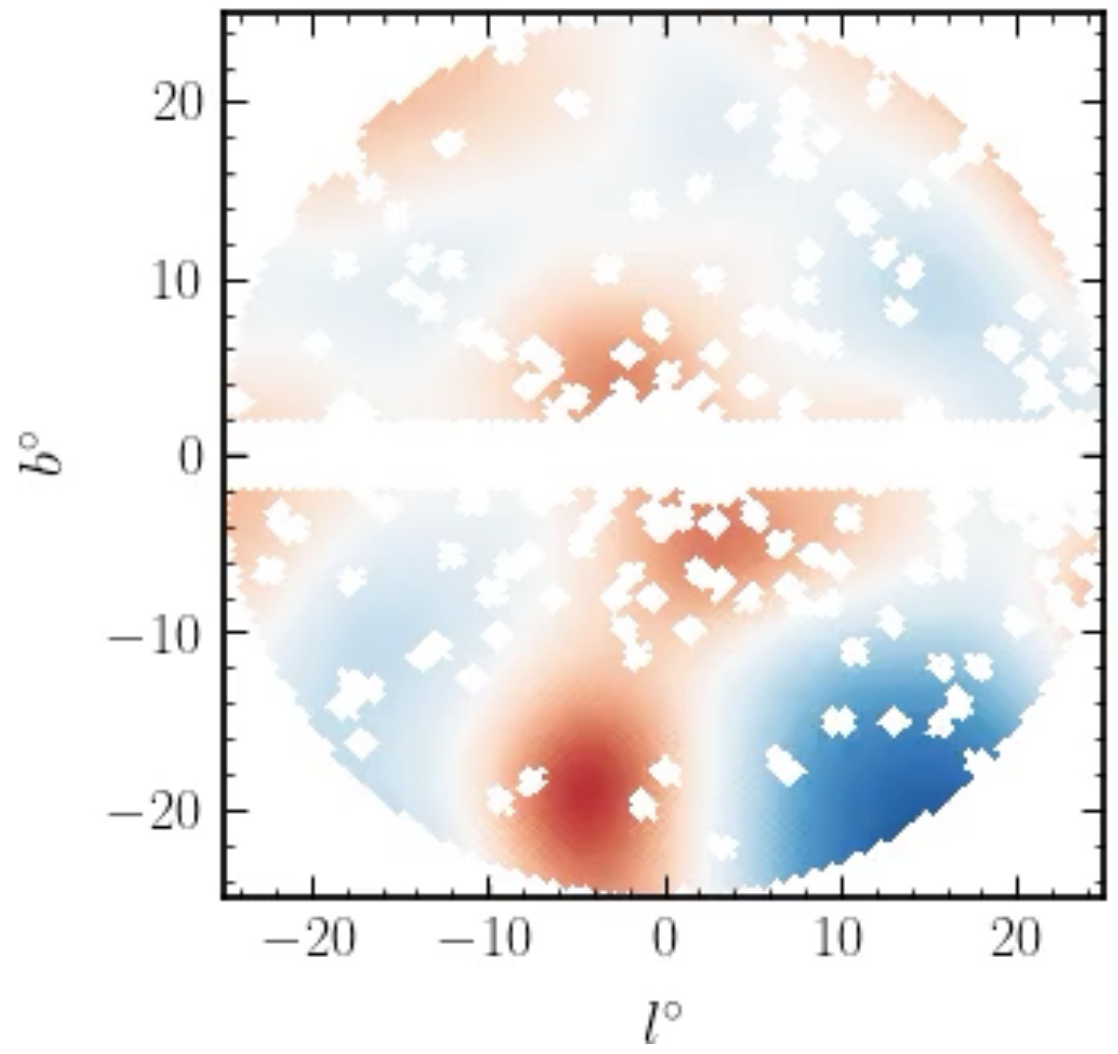
Augment background template with a GP



Component norm posteriors



GP posterior

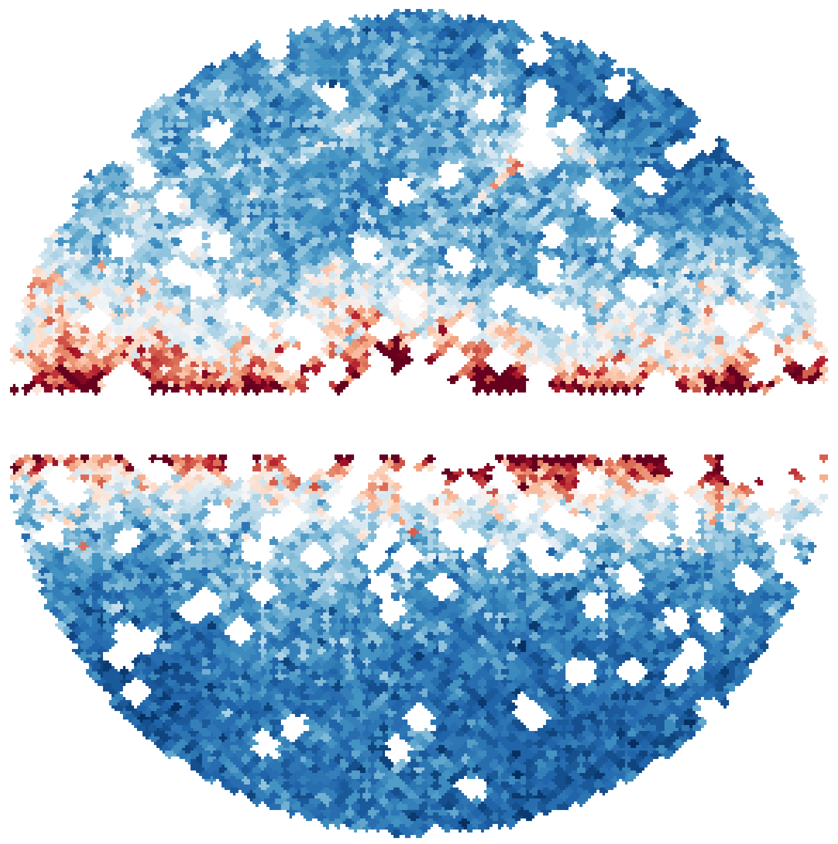


Sparse variational GPs in GPyTorch

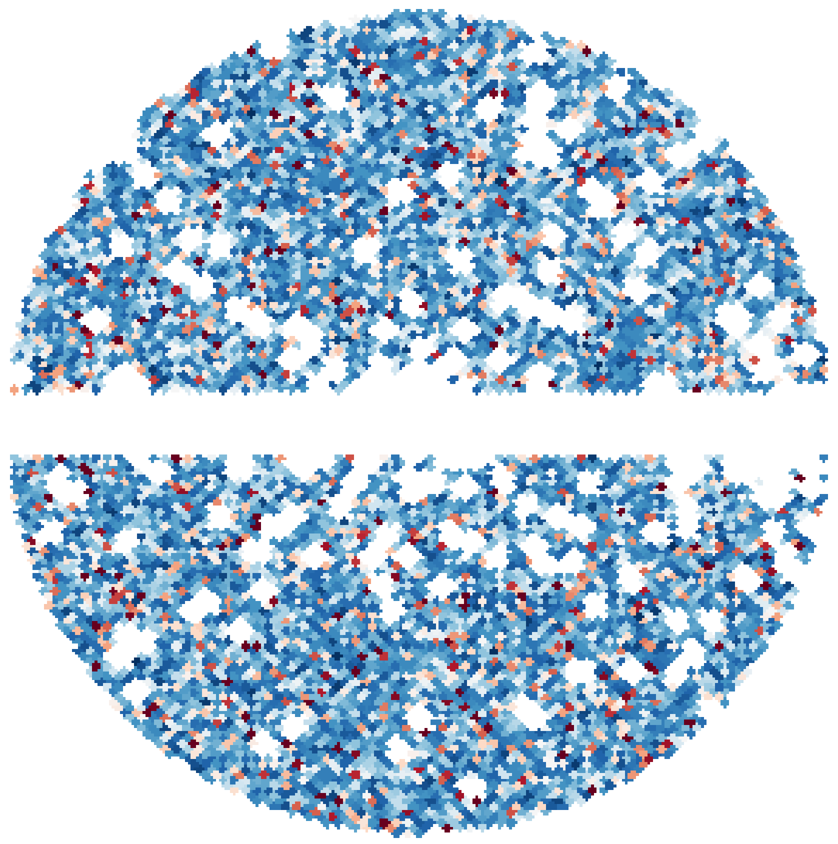
SM, Cranmer [NeurIPS ML4PS Workshop 2020]

Information content in the 1-point PDF

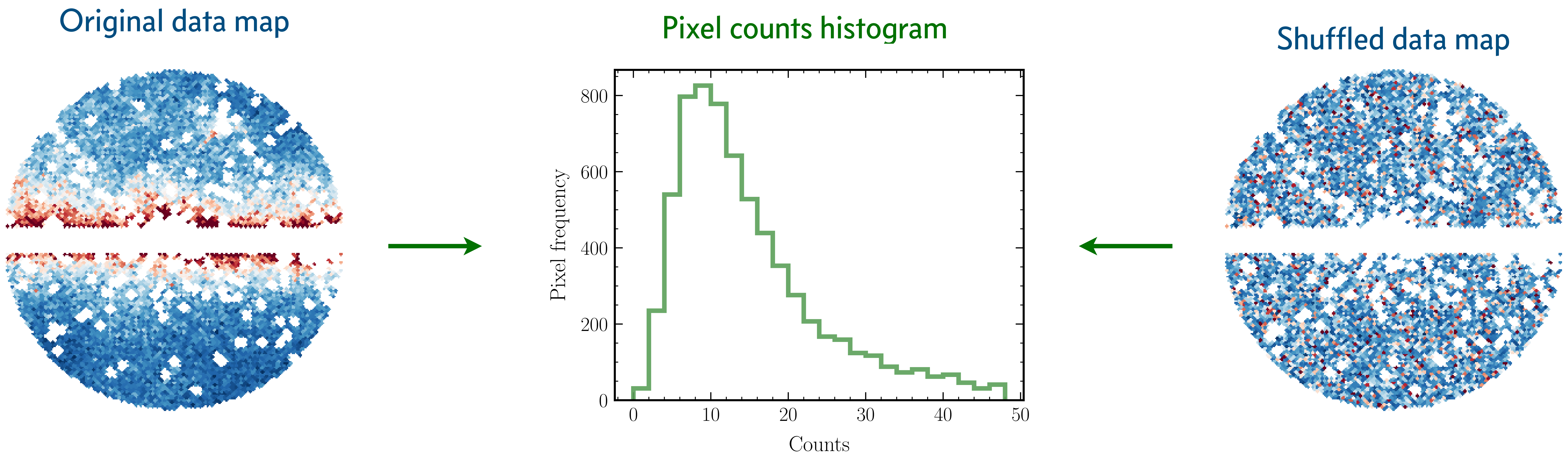
Original data map



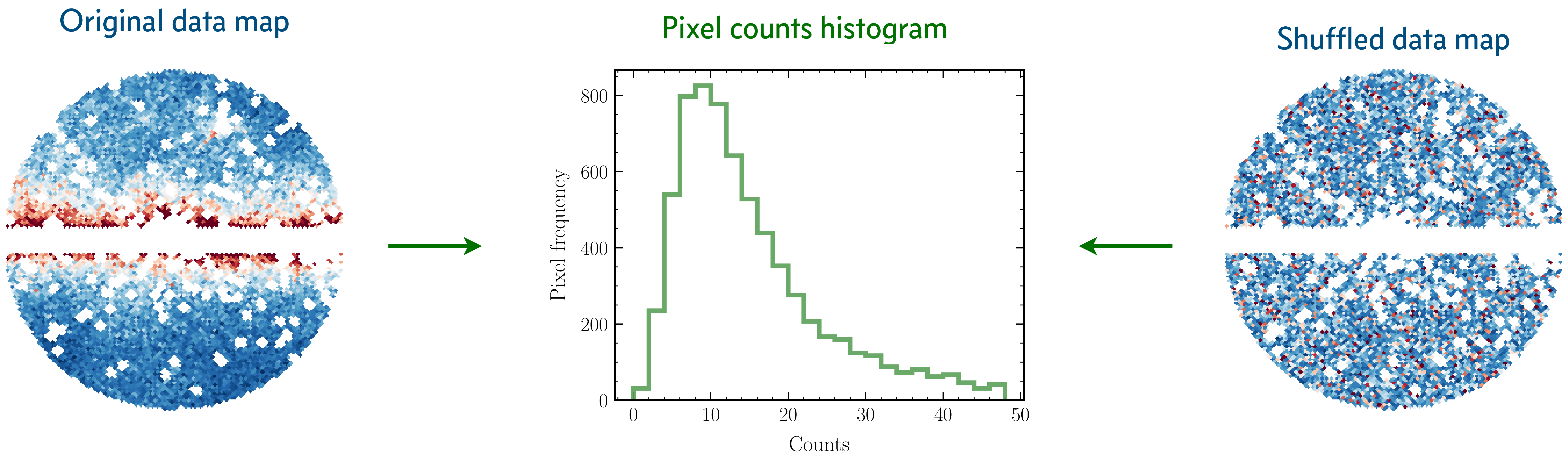
Shuffled data map



Information content in the 1-point PDF



Information content in the 1-point PDF



Can we model the full likelihood $p(x \mid \theta)$ instead of the per-pixel likelihood $\prod_p p(x^p \mid \theta)$?

Exploiting higher-order statistics can "regularize" issues associated with model misspecification

Modeling the posterior with normalizing flows

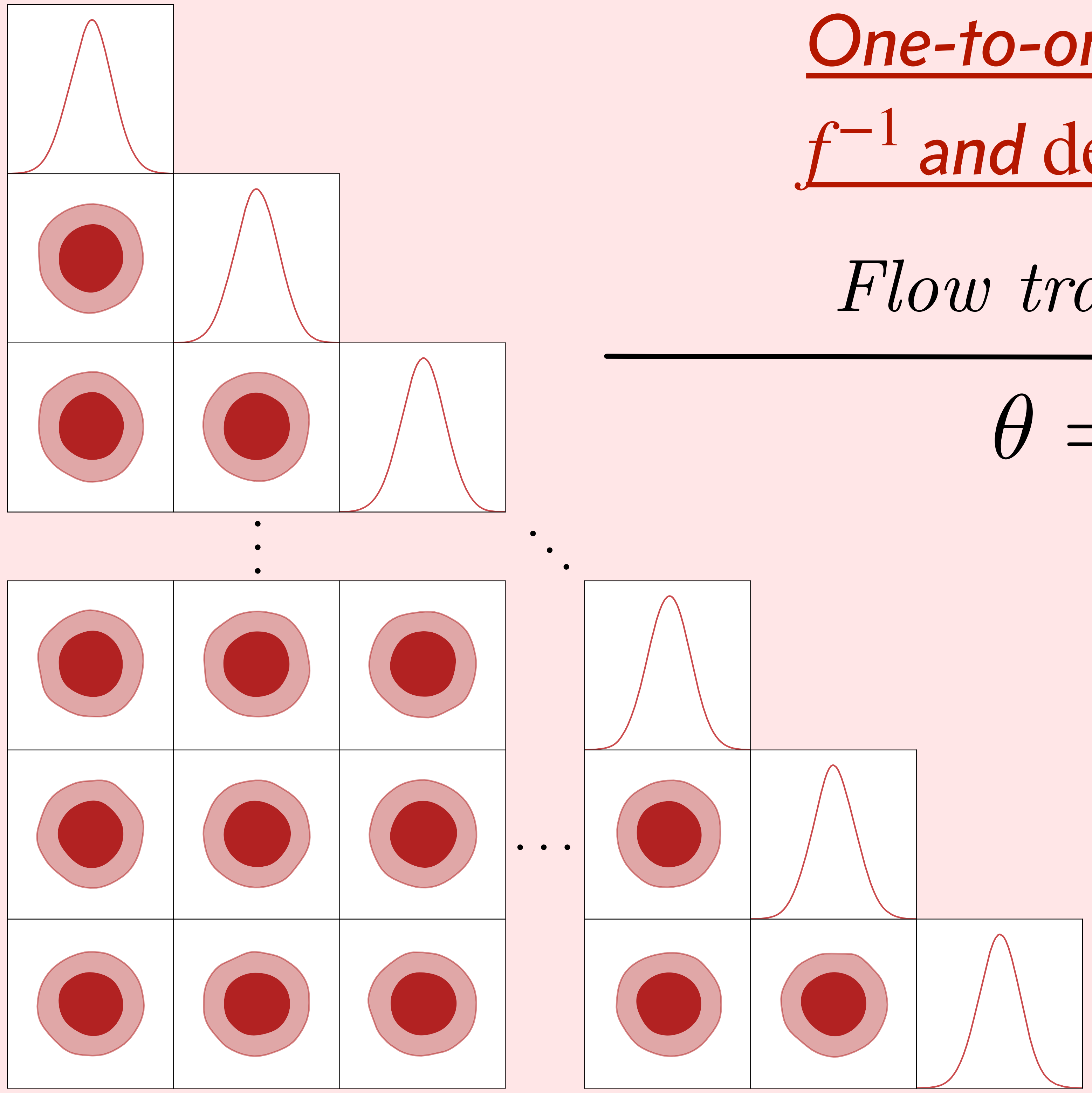
Normalizing flow

Base distribution

$u \sim \pi(u) = \mathcal{N}(u; 0, \mathbb{1})$

Posterior distribution

$\theta \sim p(\theta) = \pi\left(f^{-1}(\theta)\right) \left|\det \nabla f\right|^{-1}$

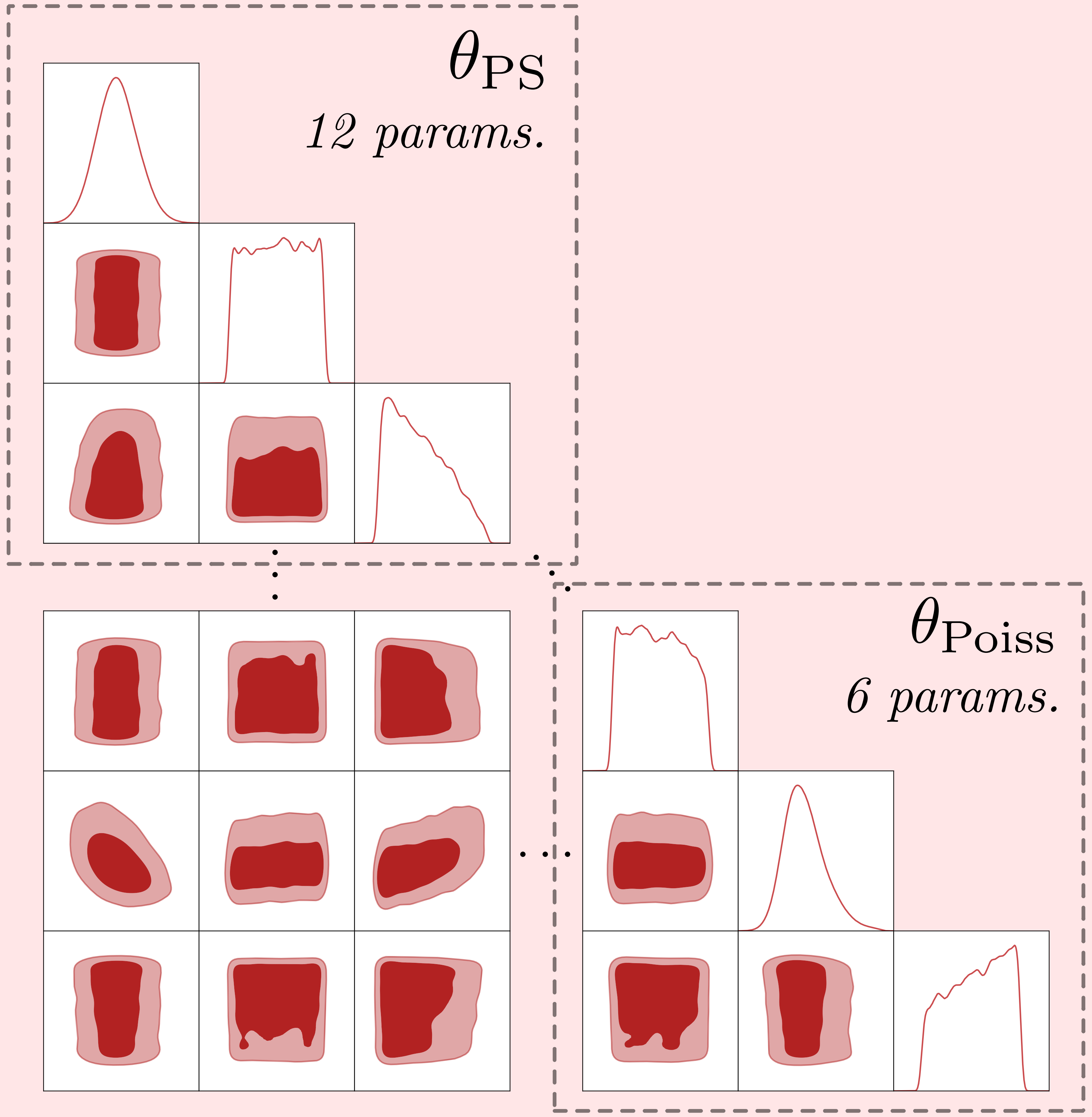


<u>*One-to-one transformation*</u>

<u>f^{-1} *and* $\det \nabla f$ *are tractable*</u>

Flow transformations

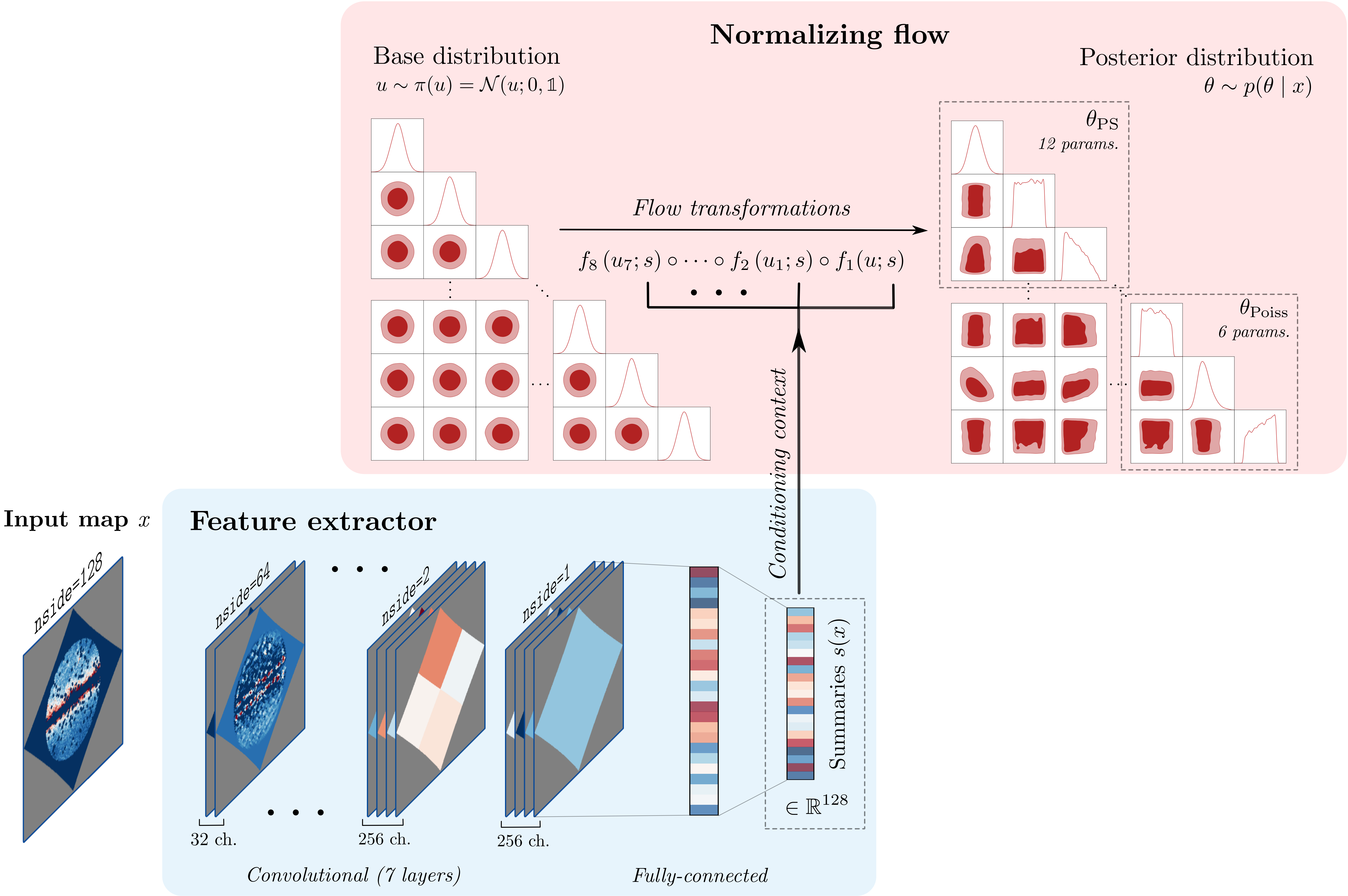
$\theta = f(u)$



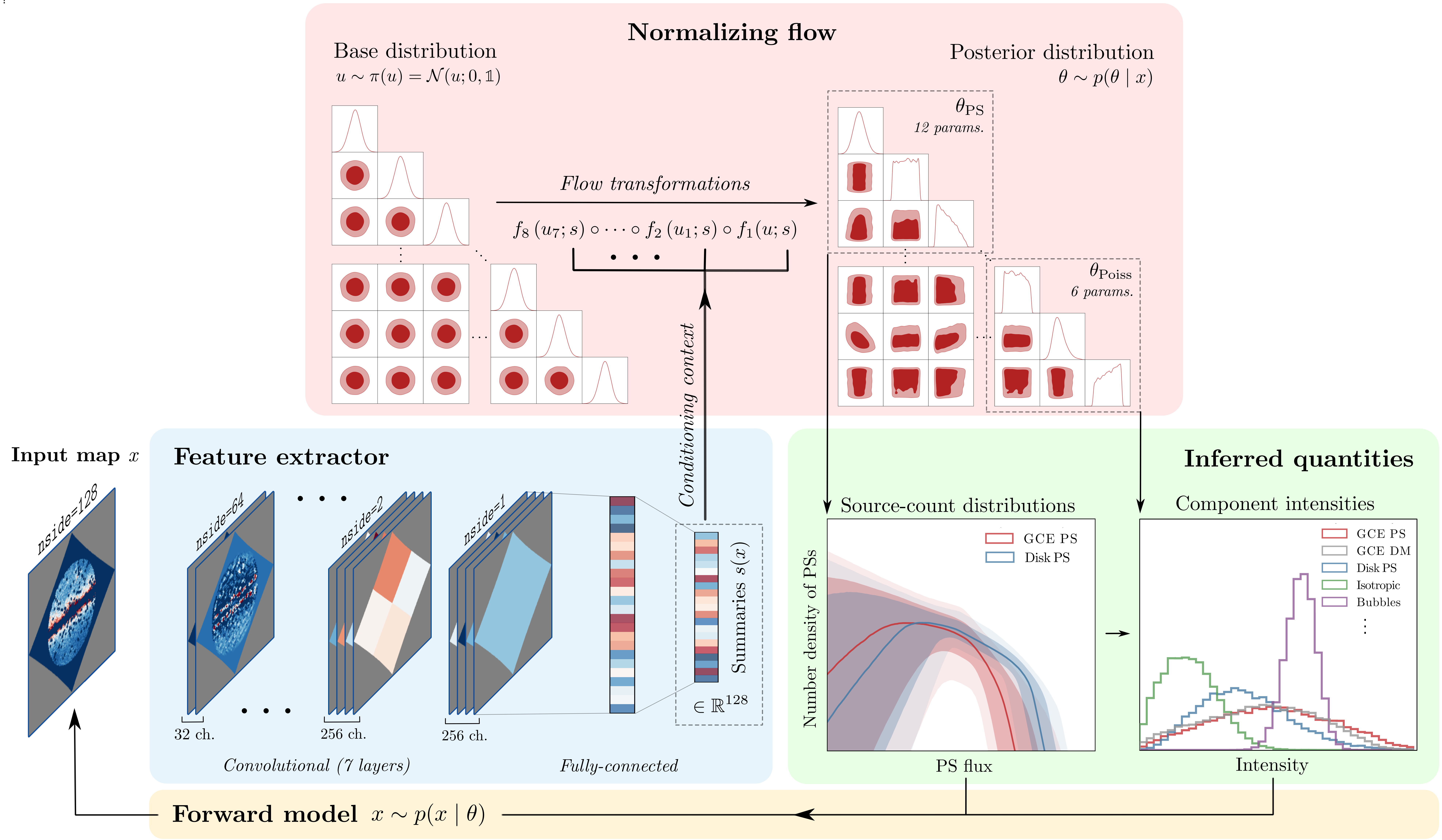
Review: Papamakarios et al [arXiv:1912.02762]

SM, Cranmer [PRD 2022]

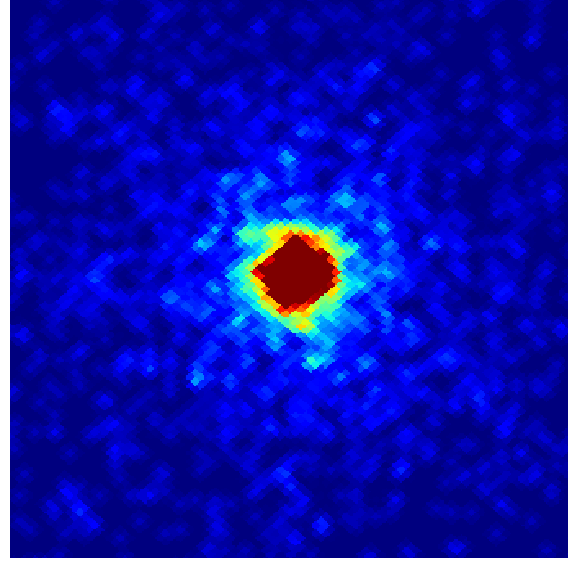
Going beyond the 1-point PDF with SBI

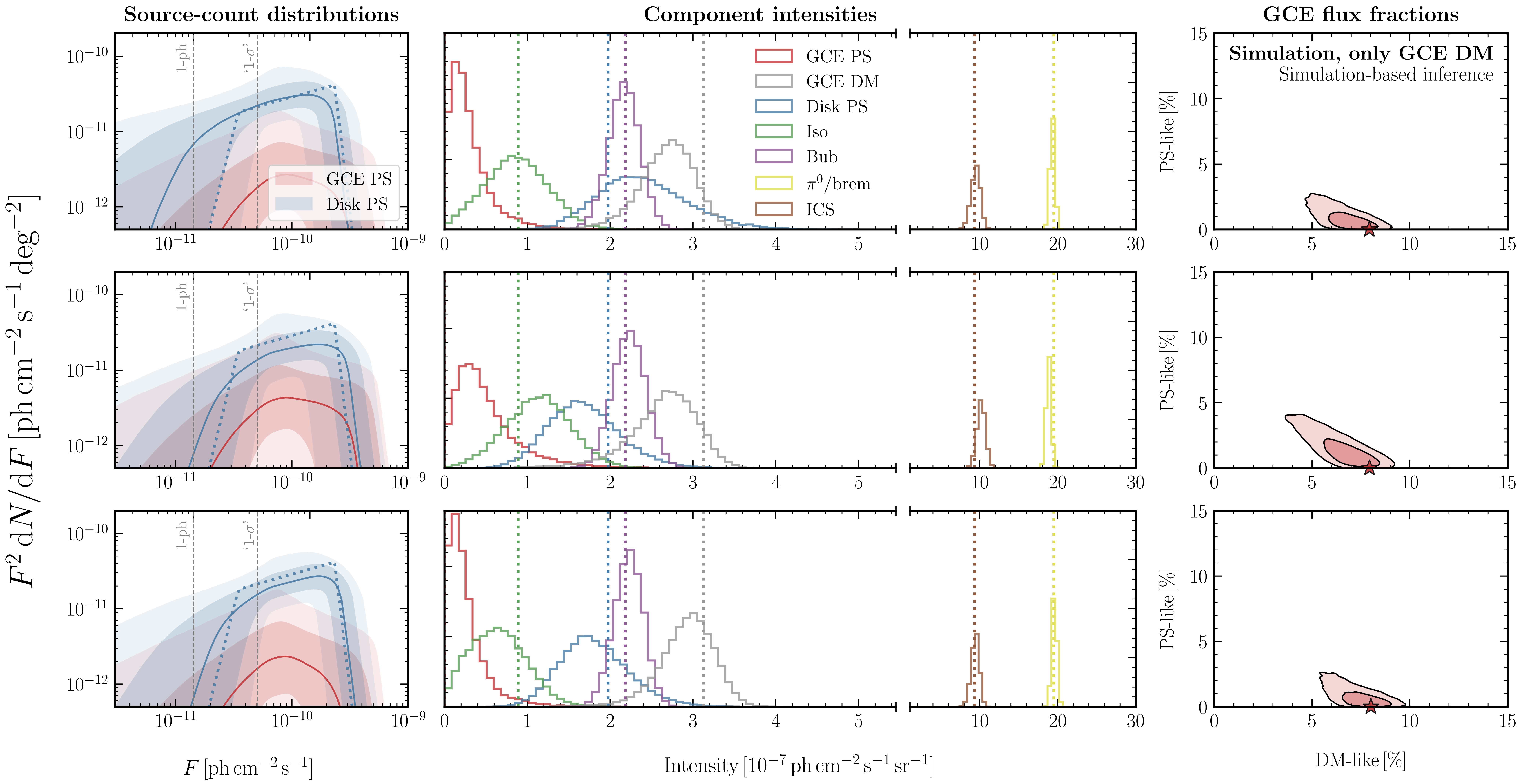


Going beyond the 1-point PDF with SBI

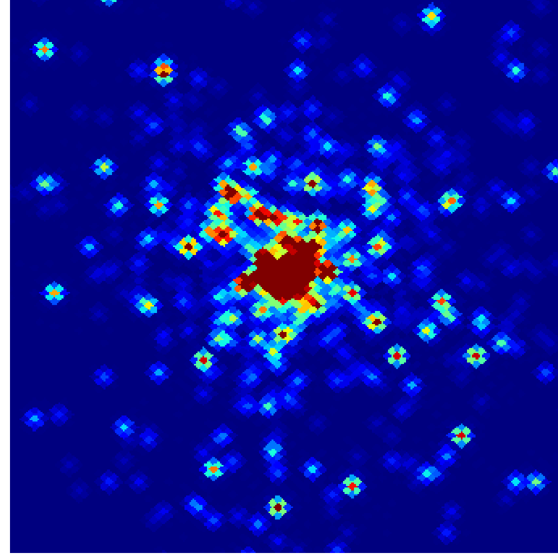


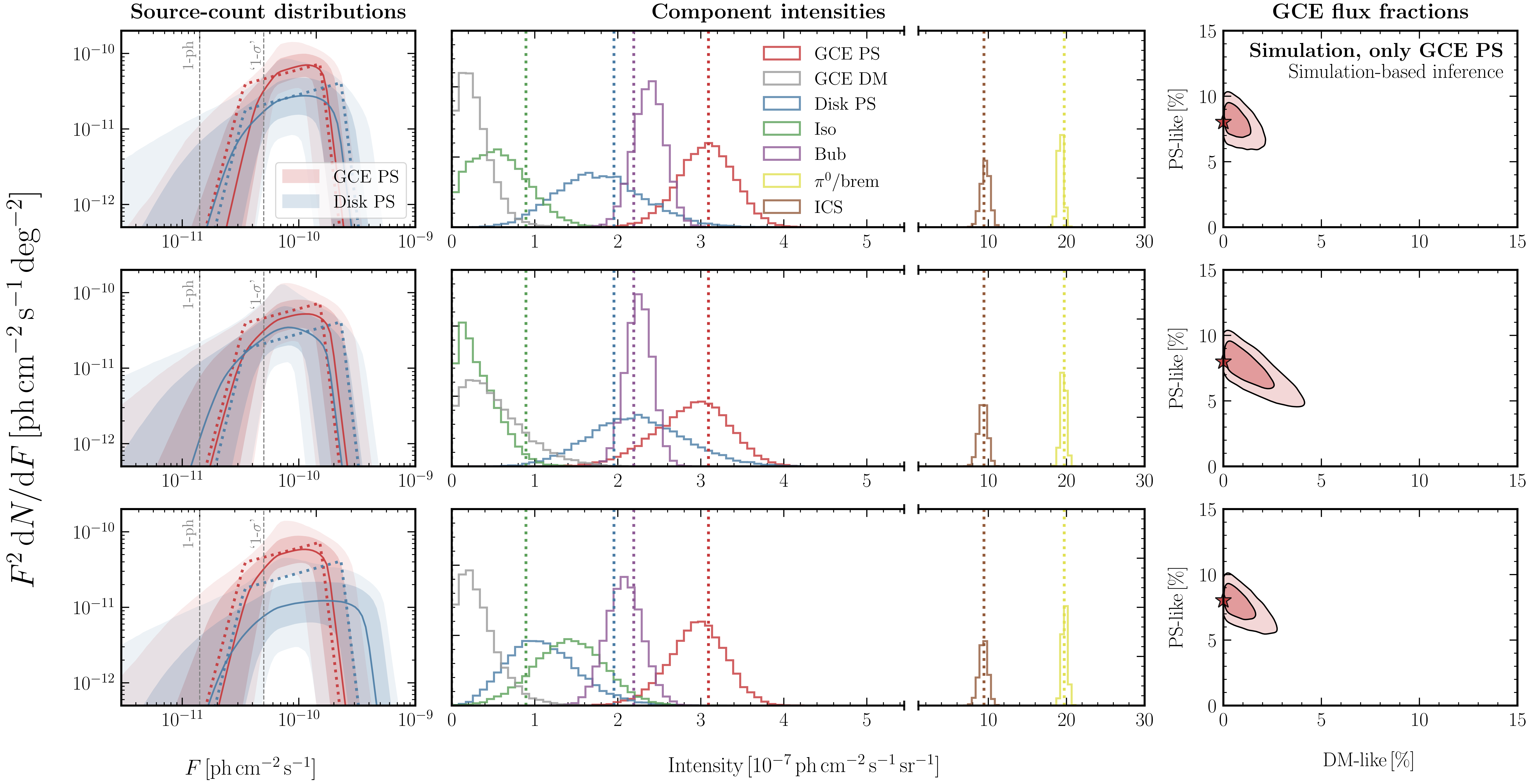
Tests on simulations: DM only





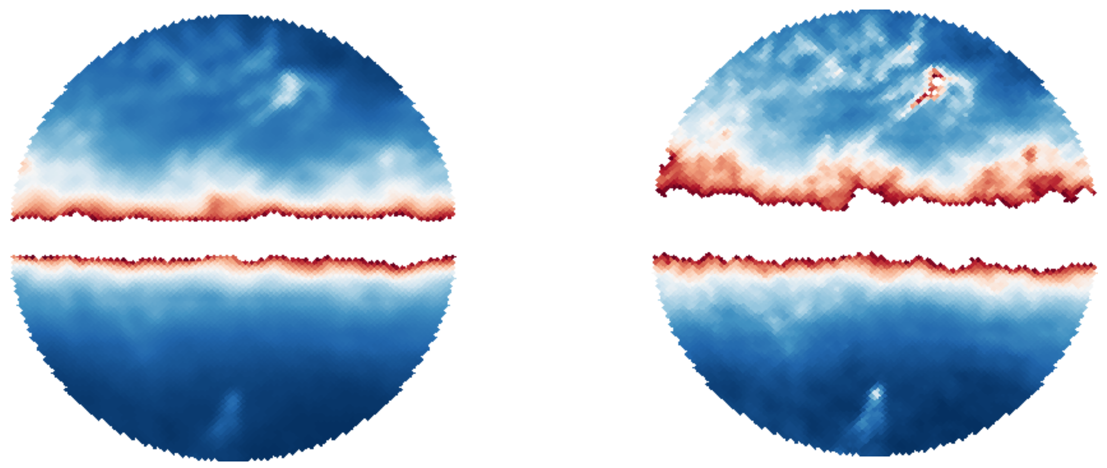
Tests on simulations: PS only



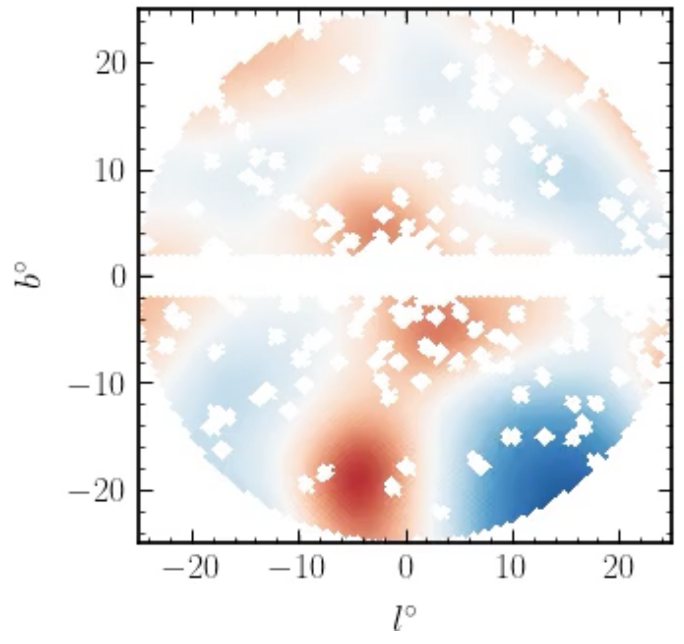


Robustness test: mismodeling with DM

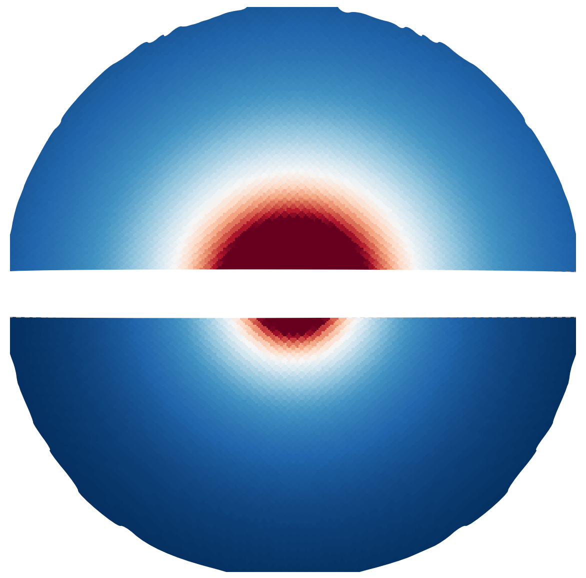
Train with one diffuse model / test with another



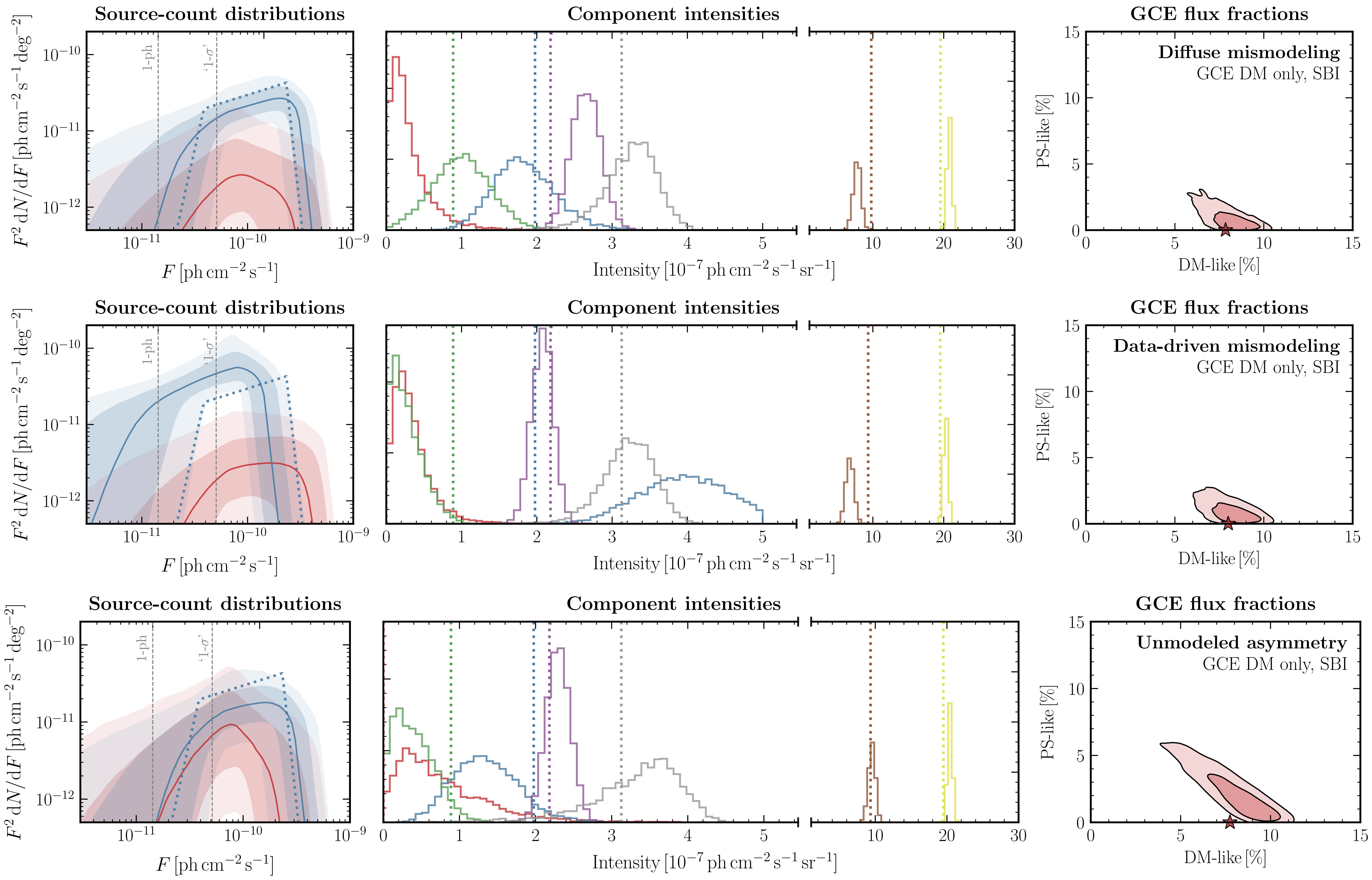
Test on GP-modulated background



Test on Substantially-asymmetric signal



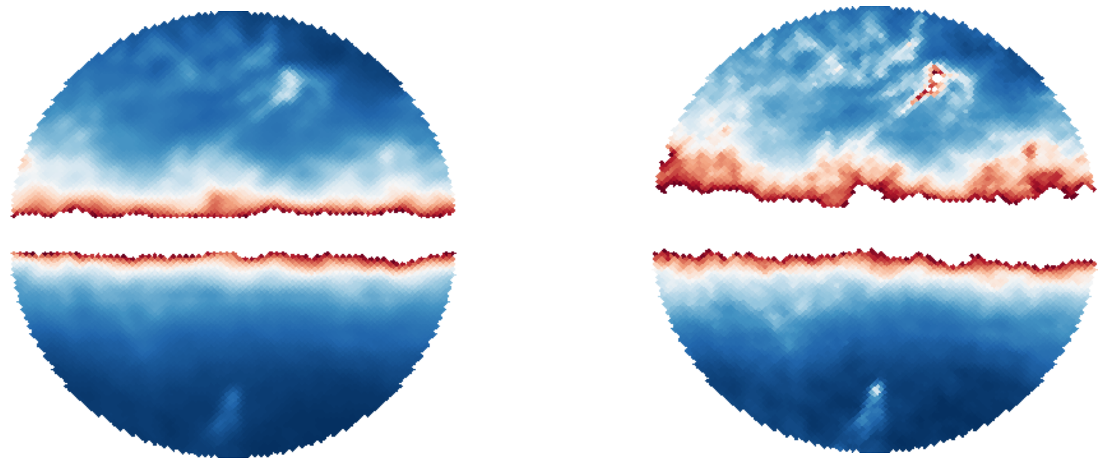
(cf. Leane & Slatyer [PRL 2020, PRD 2020])



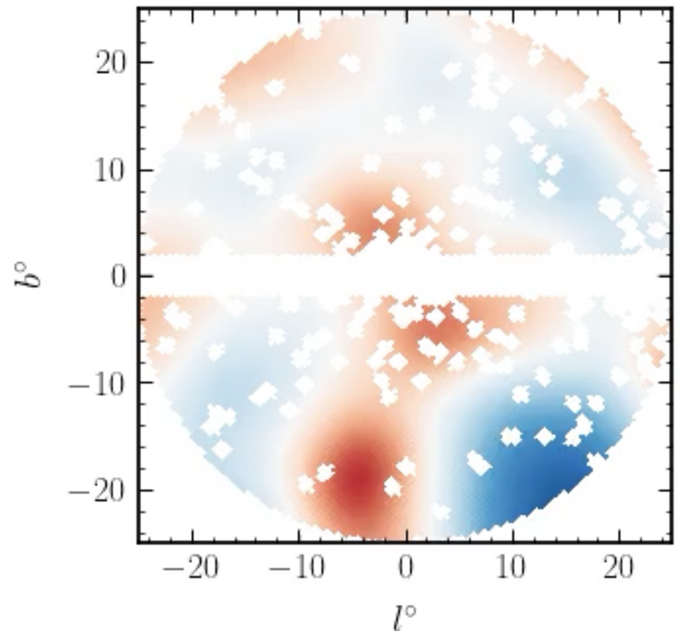
Generally well-behaved under known forms of systematic mismodeling

Robustness test: mismodeling with DM

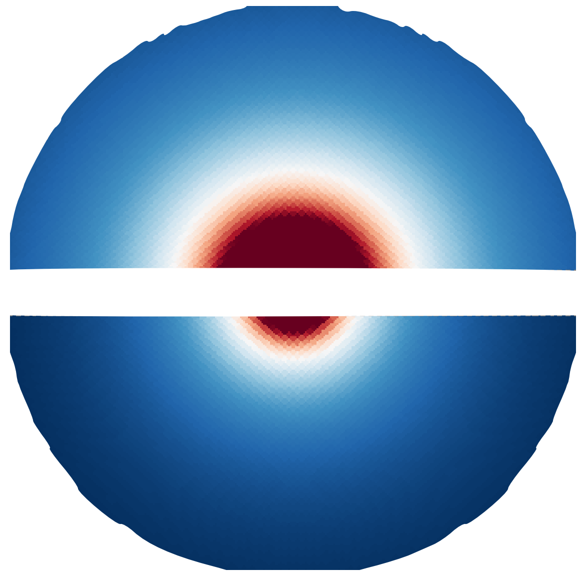
Train with one diffuse model / test with another



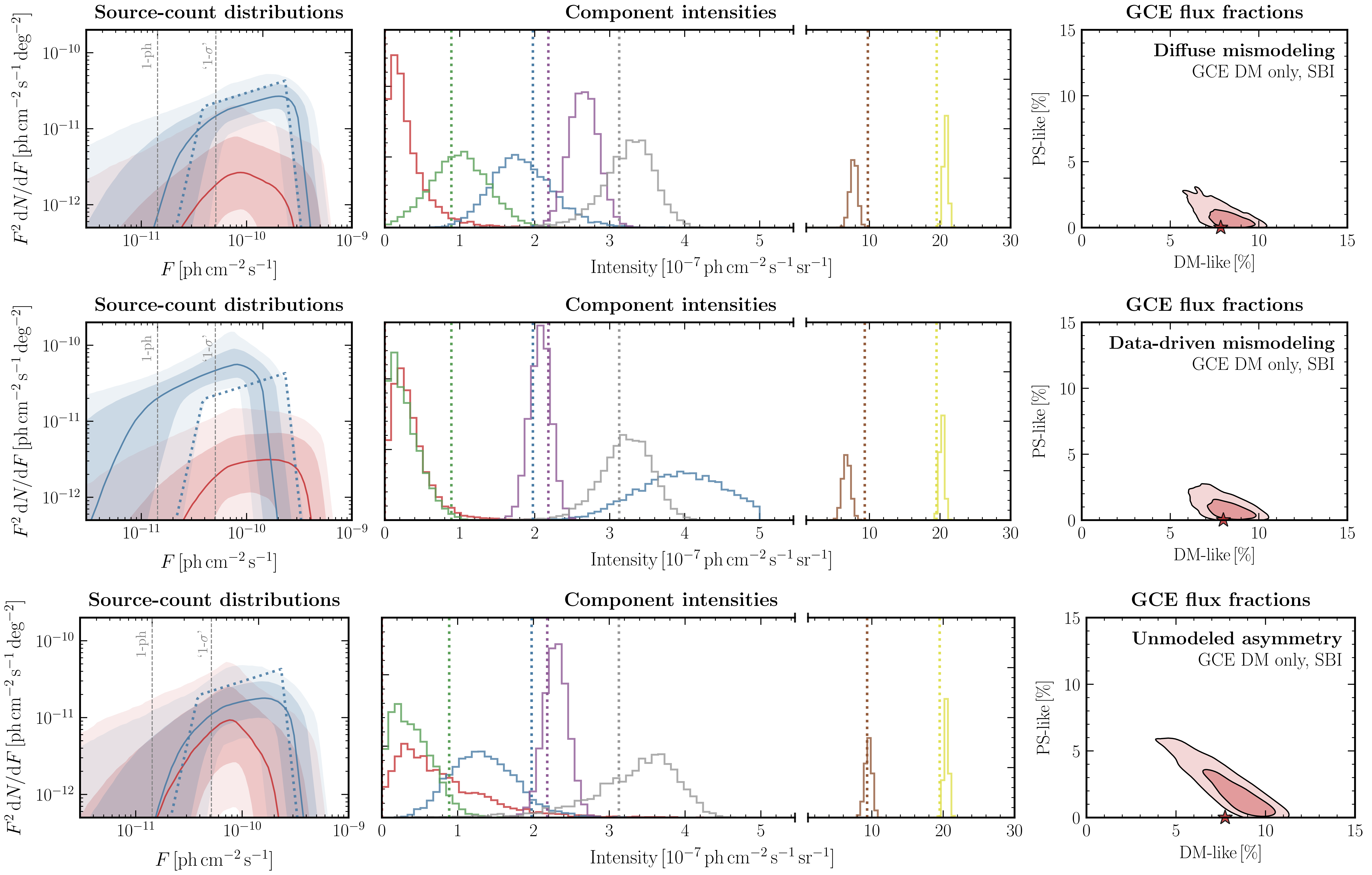
Test on GP-modulated background



Test on Substantially-asymmetric signal



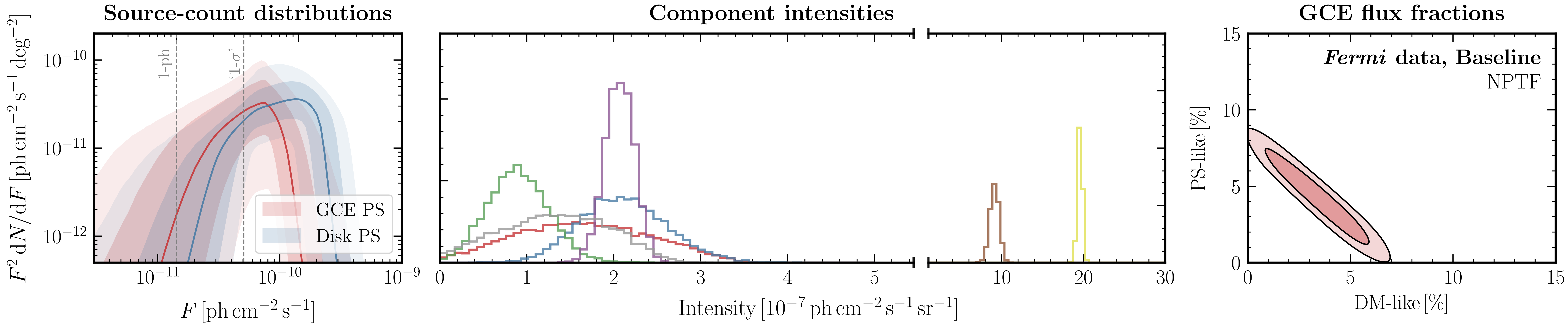
(cf. Leane & Slatyer [PRL 2020, PRD 2020])



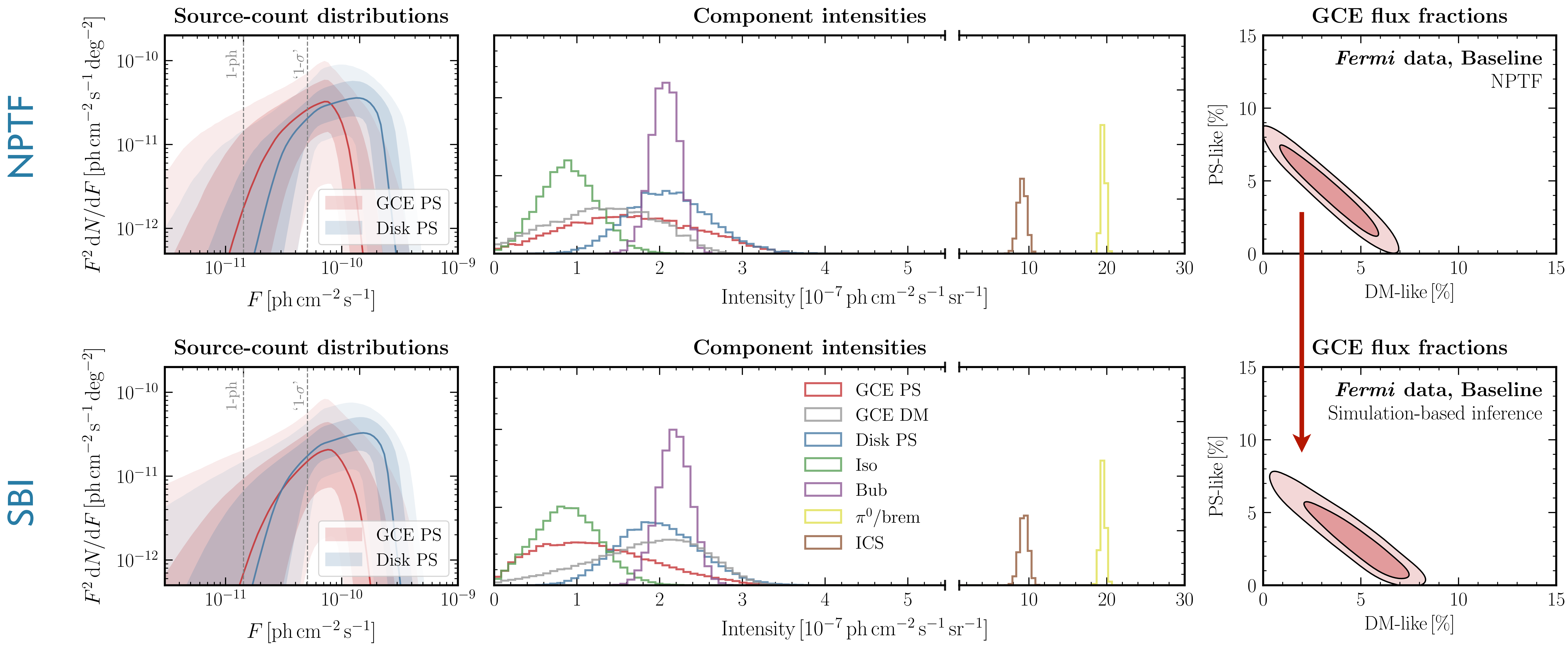
Generally well-behaved under known forms of systematic mismodeling

Application to *Fermi* γ-ray data

NPTF

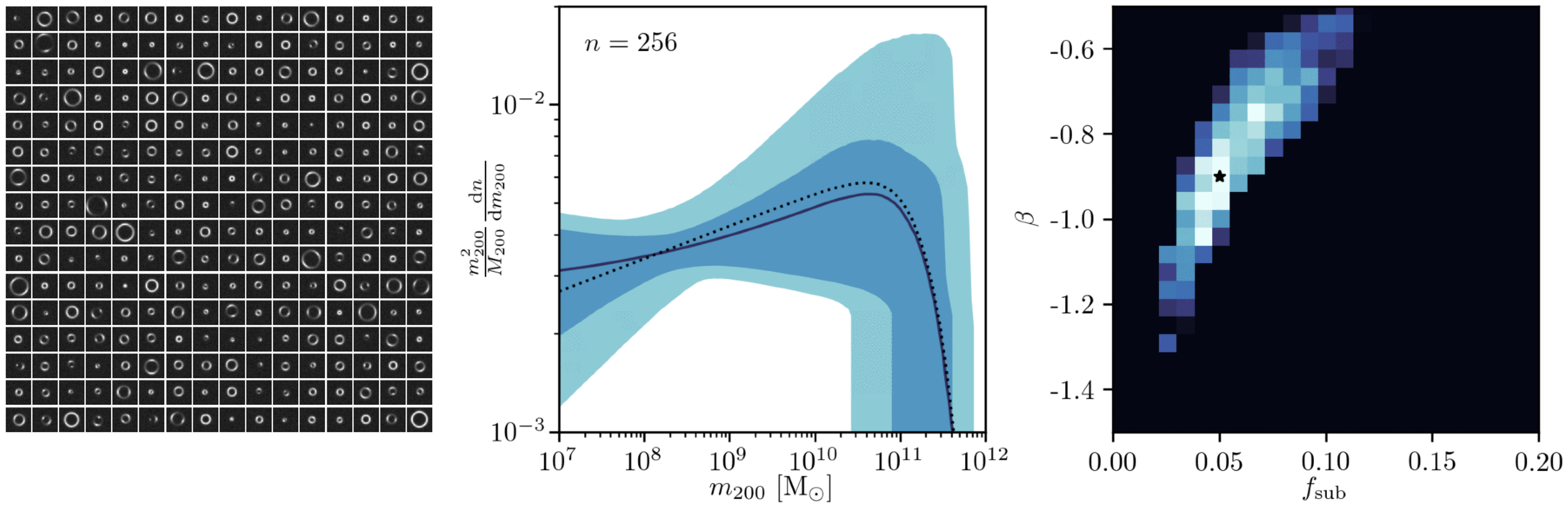


Application to *Fermi* γ-ray data

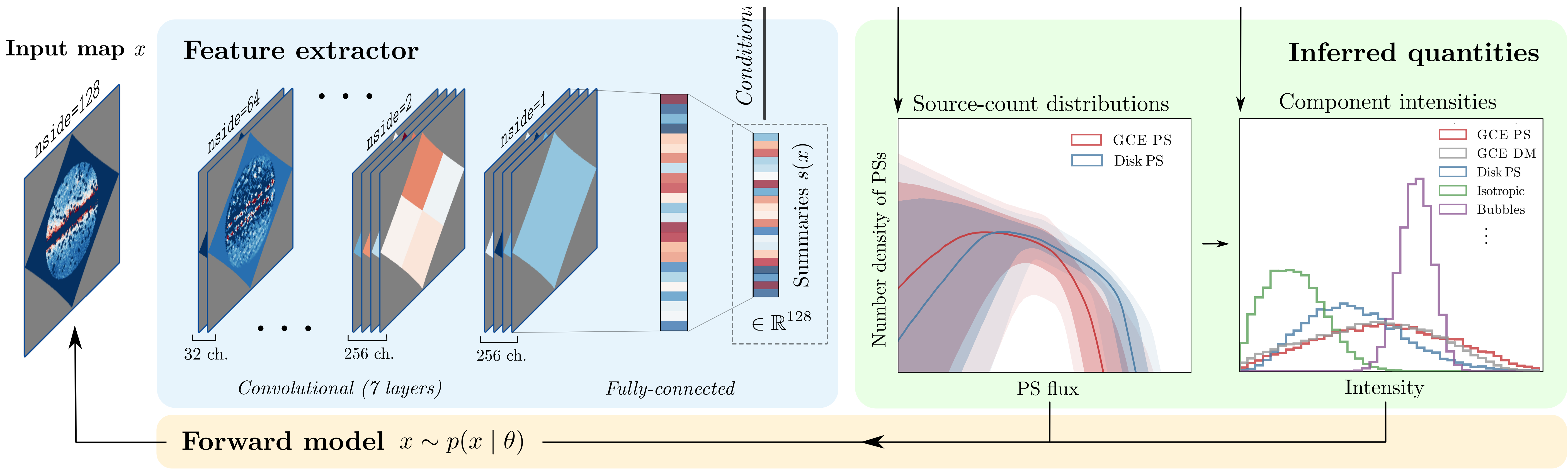


Exploiting more information in the γ-ray maps results in <u>smaller</u>, but <u>still significant PS-like component</u>

Conclusions



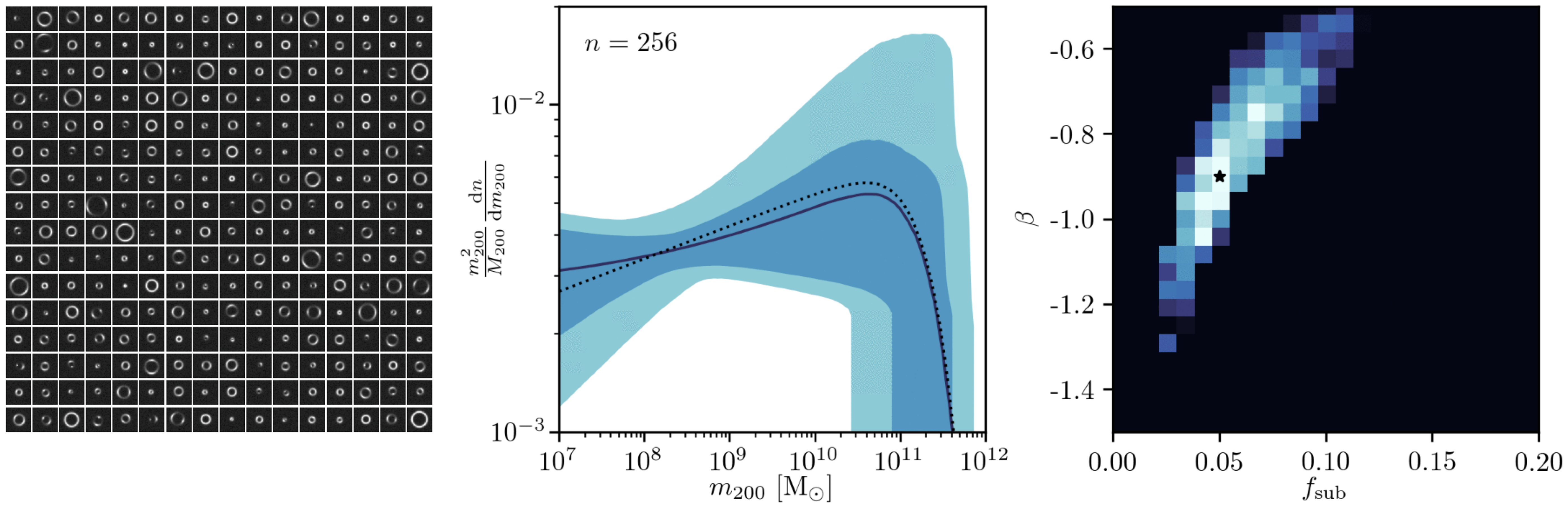
Enable fast and scalable inference from large samples of data: *crucial for next-generation of surveys*



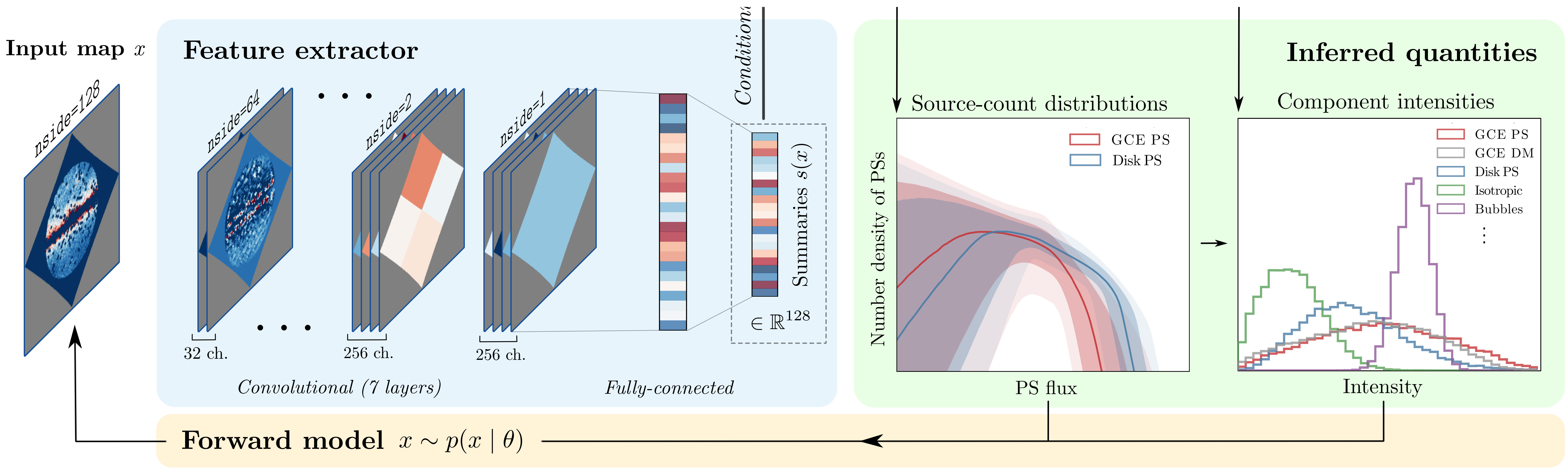
Can capture more information from models with intractable likelihood without resorting to simplifications

github.com/smsharma/mining-for-substructure-lens
github.com/smsharma/fermi-gce-flows

Conclusions



Enable fast and scalable inference from large samples of data: *crucial for next-generation of surveys*



Can capture more information from models with intractable likelihood without resorting to simplifications

github.com/smsharma/mining-for-substructure-lens
github.com/smsharma/fermi-gce-flows