Inferring Dark Matter Substructure with Weak and Strong and Gravitational Lensing

Siddharth Mishra-Sharma

SM, Ken Van Tilburg, and Neal Weiner [20xx.xxxxx]
Johann Brehmer, SM, Joeri Hermans, Gilles Louppe, and Kyle Cranmer [1909.02005]

NYU Center for Cosmology and Particle Physics

University of Michigan High Energy Theory Brown Bag

January 28, 2020

Inferring Dark Matter Substructure with Weak and Strong and Gravitational Lensing

Siddharth Mishra-Sharma

SM, Ken Van Tilburg, and Neal Weiner [20xx.xxxxx]
Johann Brehmer, SM, Joeri Hermans, Gilles Louppe, and Kyle Cranmer [1909.02005]

NYU | Center for Cosmology and Particle Physics

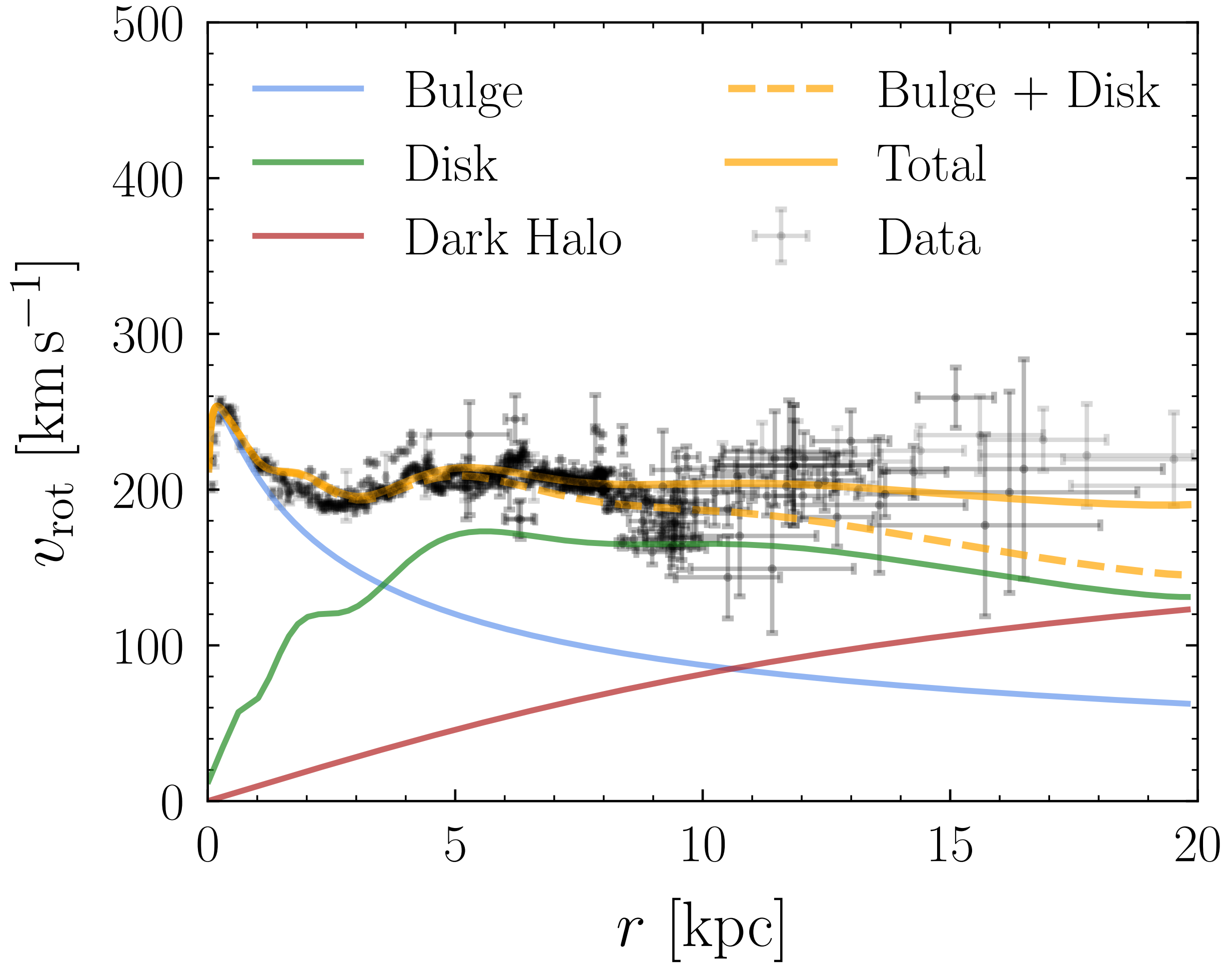
University of Michigan High Energy Theory Brown Bag

January 28, 2020

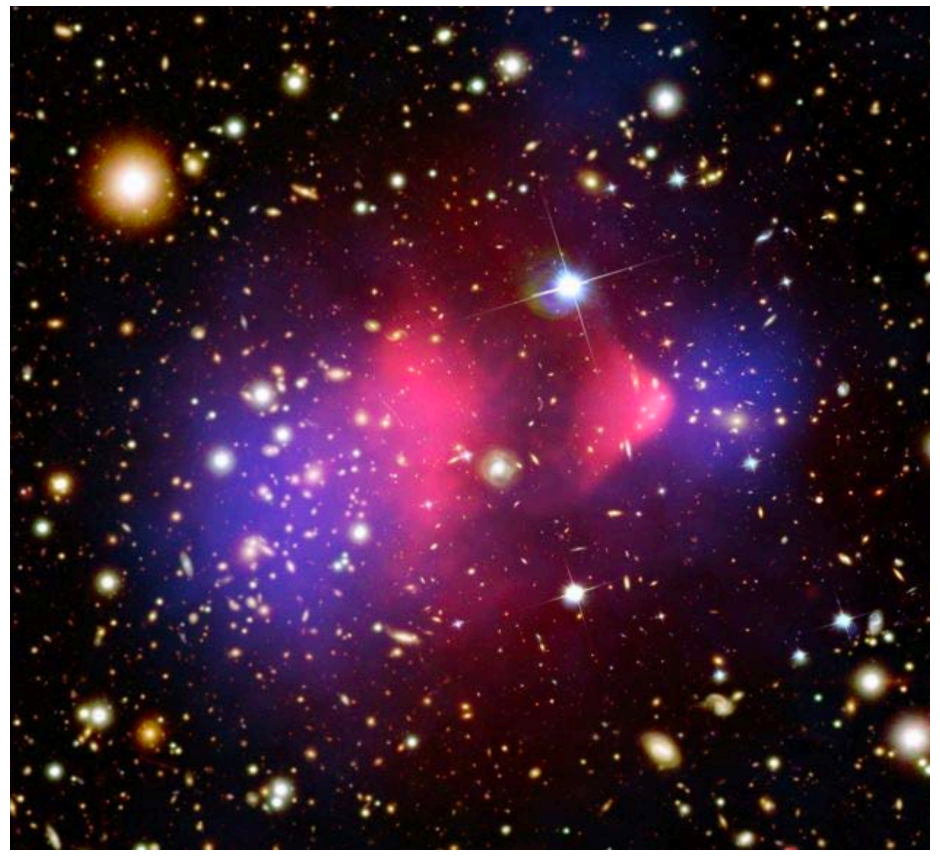
Evidence for dark matter...

...exists over a diversity of scales and physical systems

Galaxy rotation curves

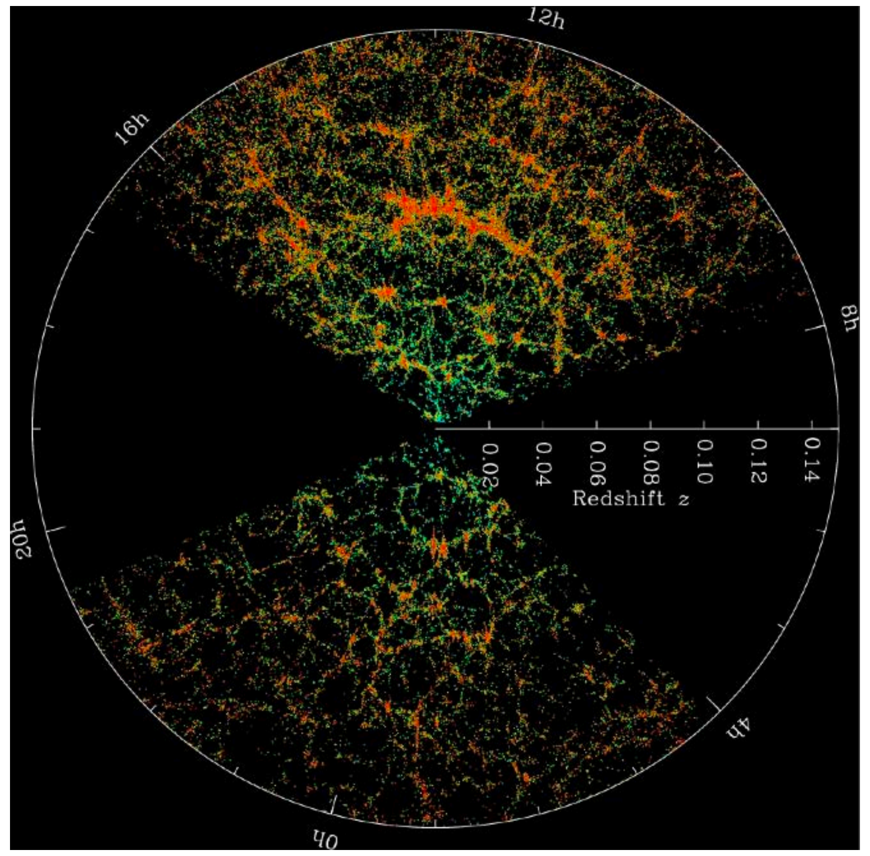


Galaxy clusters



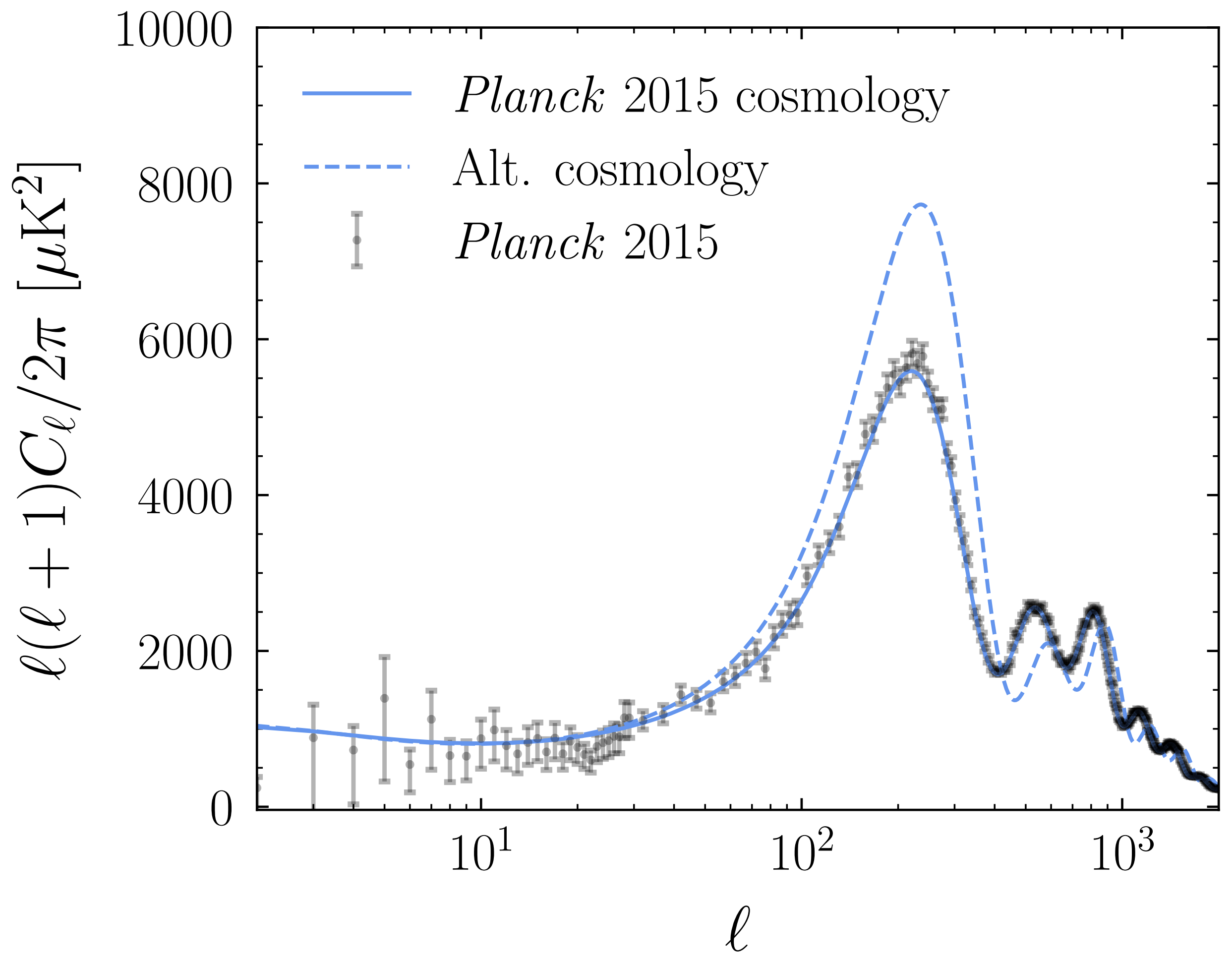
Bullet Cluster. Credits: NASA

Large-scale structure



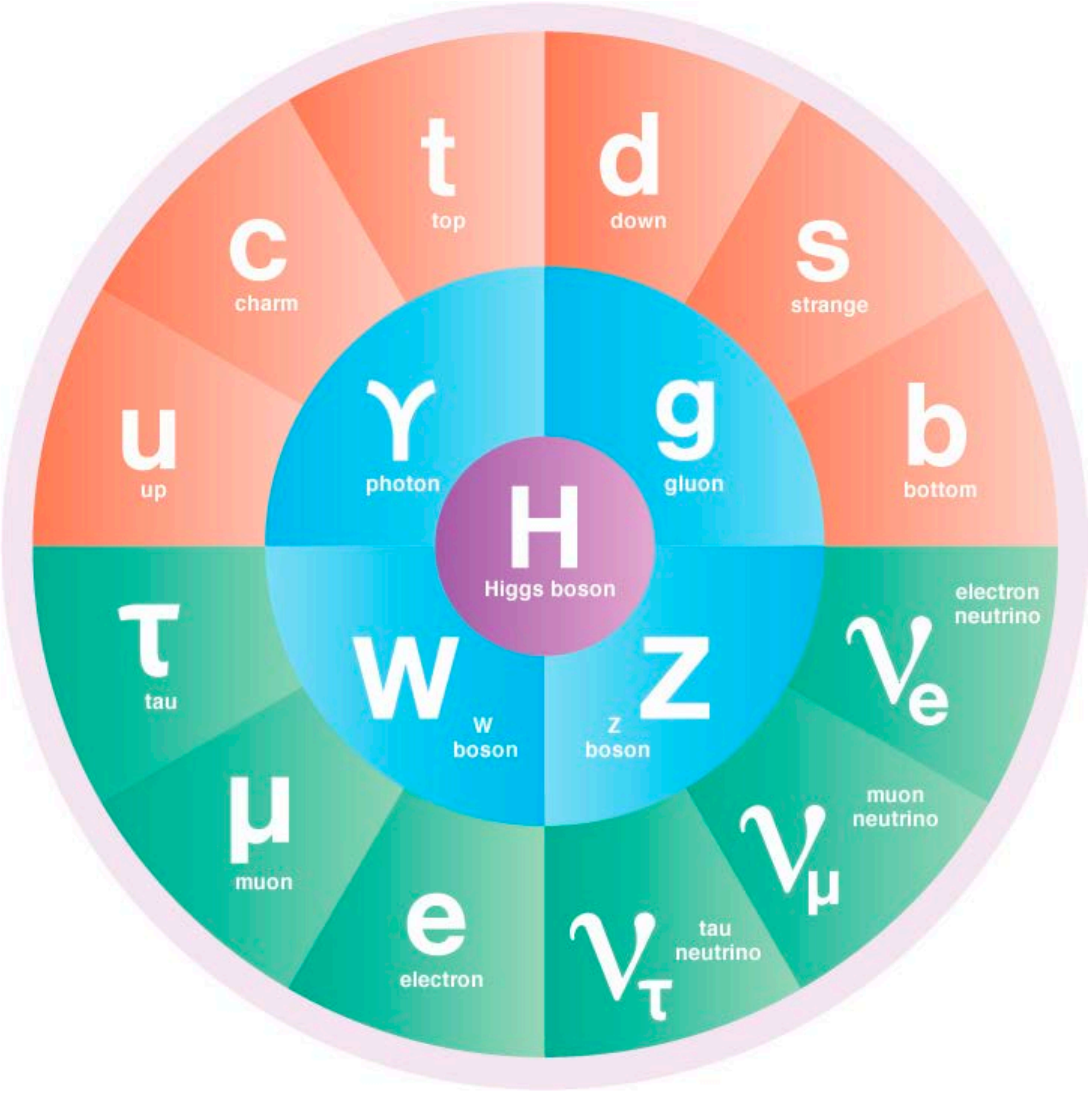
Credits: SDSS

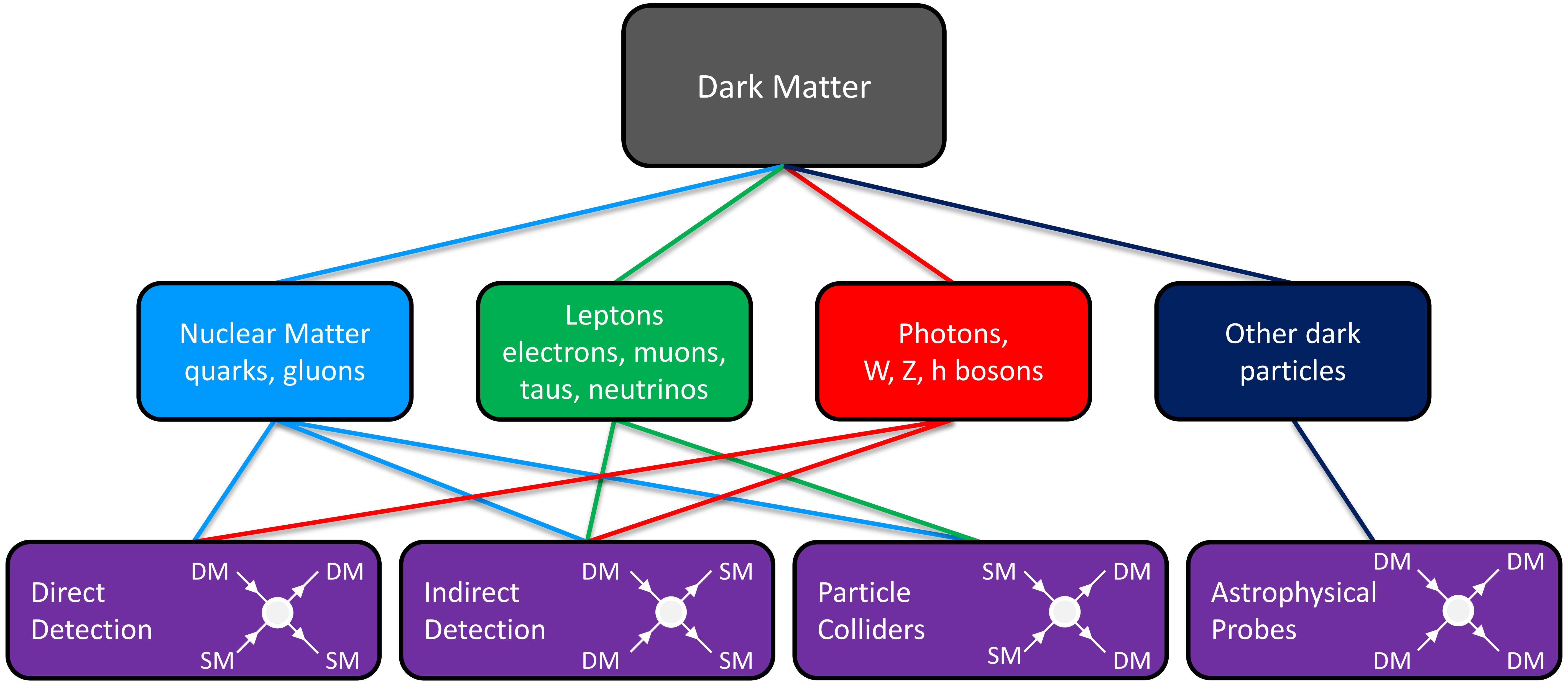
CMB



Increasing scale

Pinning down dark matter microphysics...





Snowmass CF4 report (Bauer et al, 2015)

…through macroscopic effects

Underlying particle physics can be manifest by understanding macroscopic distribution of dark matter on small scales

Model	Probe	Parameter	Value
Warm Dark Matter	Halo Mass	Particle Mass	$m \sim 18\,\mathrm{keV}$
Self-Interacting Dark Matter	Halo Profile	Cross Section	$\sigma_{\mathrm{SIDM}}/m_\chi \sim 0.1$–$10\,\mathrm{cm}^2/\,\mathrm{g}$
Baryon-Scattering Dark Matter	Halo Mass	Cross Section	$\sigma \sim 10^{-30}\,\mathrm{cm}^2$
Axion-Like Particles	Energy Loss	Coupling Strength	$g_{\phi e} \sim 10^{-13}$
Fuzzy Dark Matter	Halo Mass	Particle Mass	$m \sim 10^{-20}\,\mathrm{eV}$
Primordial Black Holes	Compact Objects	Object Mass	$M > 10^{-4} M_\odot$
Weakly Interacting Massive Particles	Indirect Detection	Cross Section	$\langle \sigma v \rangle \sim 10^{-27}\,\mathrm{cm}^3/\,\mathrm{s}$
Light Relics	Large-Scale Structure	Relativistic Species	$N_{\mathrm{eff}} \sim 0.1$

LSST Dark Matter White Paper (Drlica-Wagner et al, 2019)

Microphysics from Macrophysics

(Self-interactions, scalar field DM…)

(Subhalo mass function, subhalo profiles…)

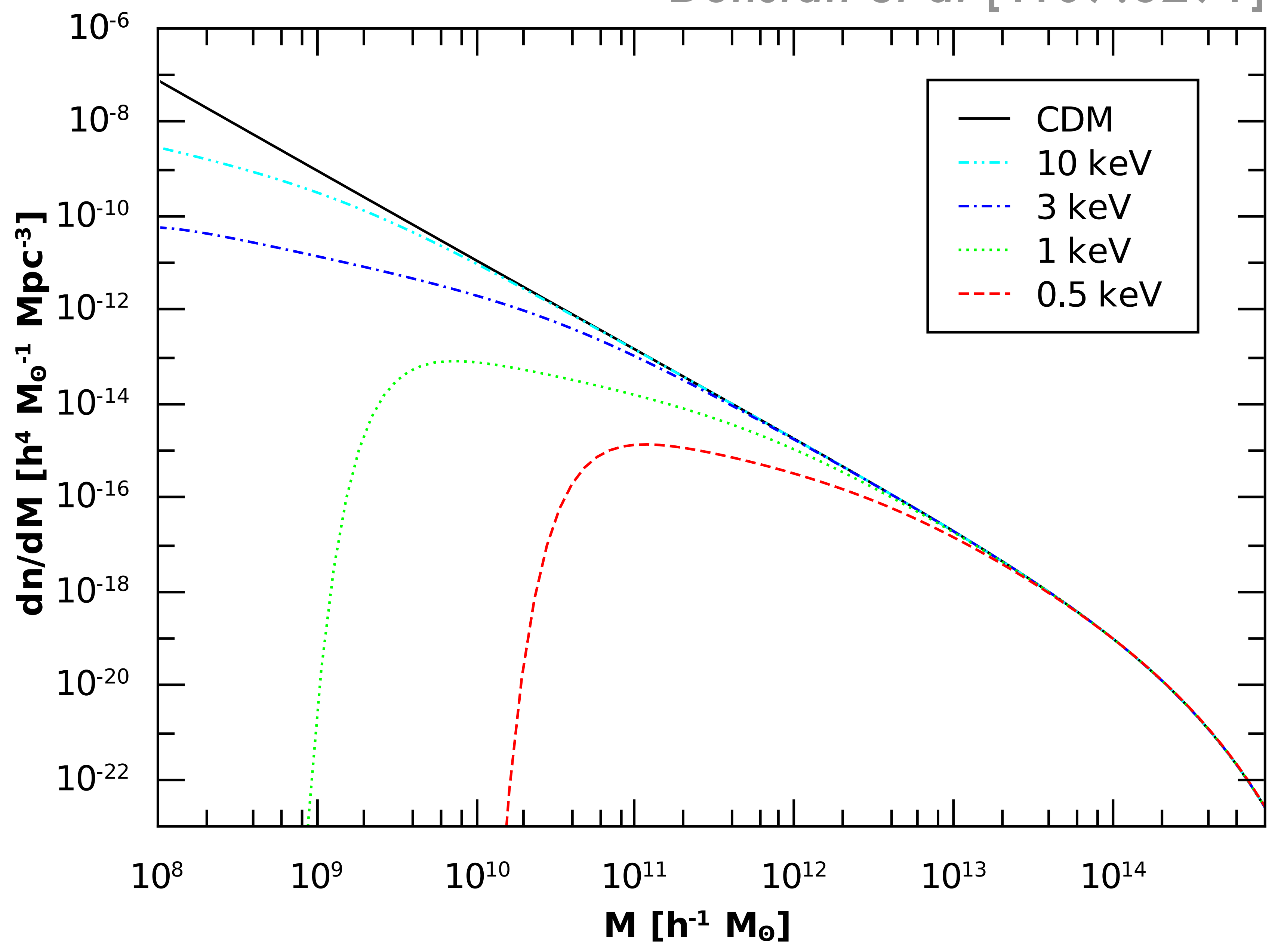
...through macroscopic effects

Underlying particle physics can be manifest by understanding macroscopic distribution of dark matter on small scales

Model	Probe	Parameter	Value
Warm Dark Matter	Halo Mass	Particle Mass	$m \sim 18\,\mathrm{keV}$
Self-Interacting Dark Matter	Halo Profile	Cross Section	$\sigma_{\mathrm{SIDM}}/m_\chi \sim 0.1\text{–}10\,\mathrm{cm}^2/\,\mathrm{g}$
Baryon-Scattering Dark Matter	Halo Mass	Cross Section	$\sigma \sim 10^{-30}\,\mathrm{cm}^2$
Axion-Like Particles	Energy Loss	Coupling Strength	$g_{\phi e} \sim 10^{-13}$
Fuzzy Dark Matter	Halo Mass	Particle Mass	$m \sim 10^{-20}\,\mathrm{eV}$
Primordial Black Holes	Compact Objects	Object Mass	$M > 10^{-4} M_\odot$
Weakly Interacting Massive Particles	Indirect Detection	Cross Section	$\langle \sigma v \rangle \sim 10^{-27}\,\mathrm{cm}^3/\,\mathrm{s}$
Light Relics	Large-Scale Structure	Relativistic Species	$N_{\mathrm{eff}} \sim 0.1$

LSST Dark Matter White Paper (Drlica-Wagner et al, 2019)

Dunstan et al [1109.6291]



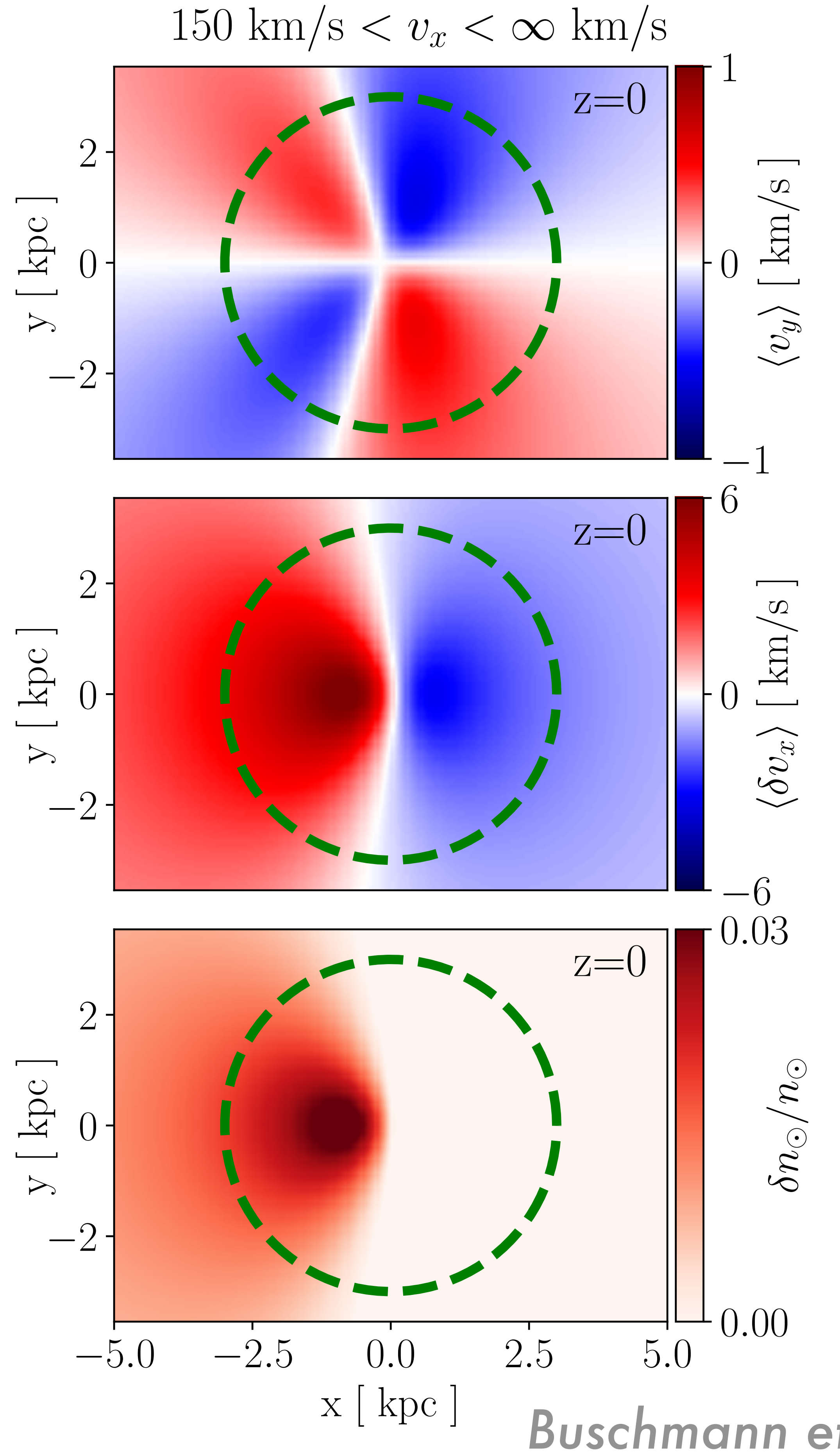
Microphysics from **Macrophysics**

(Self-interactions, scalar field DM...)

(Subhalo mass function, subhalo profiles...)

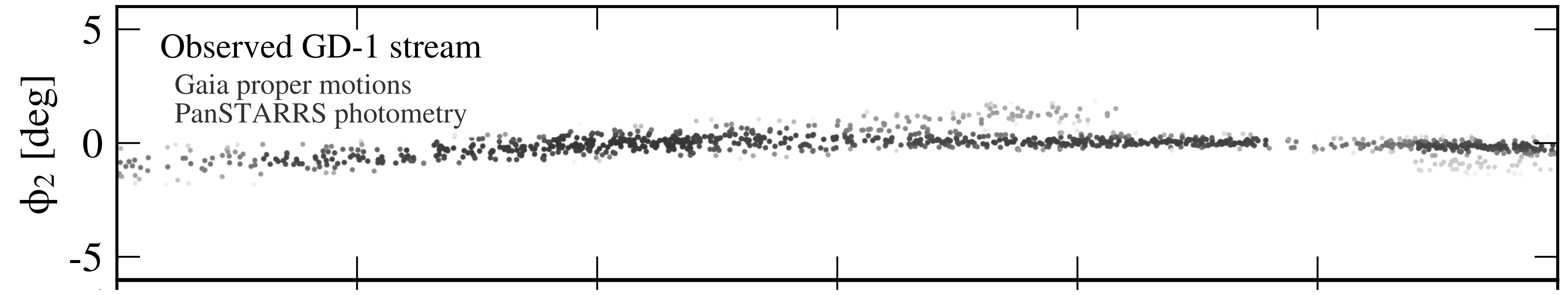
Gravitational probes of dark Galactic substructure

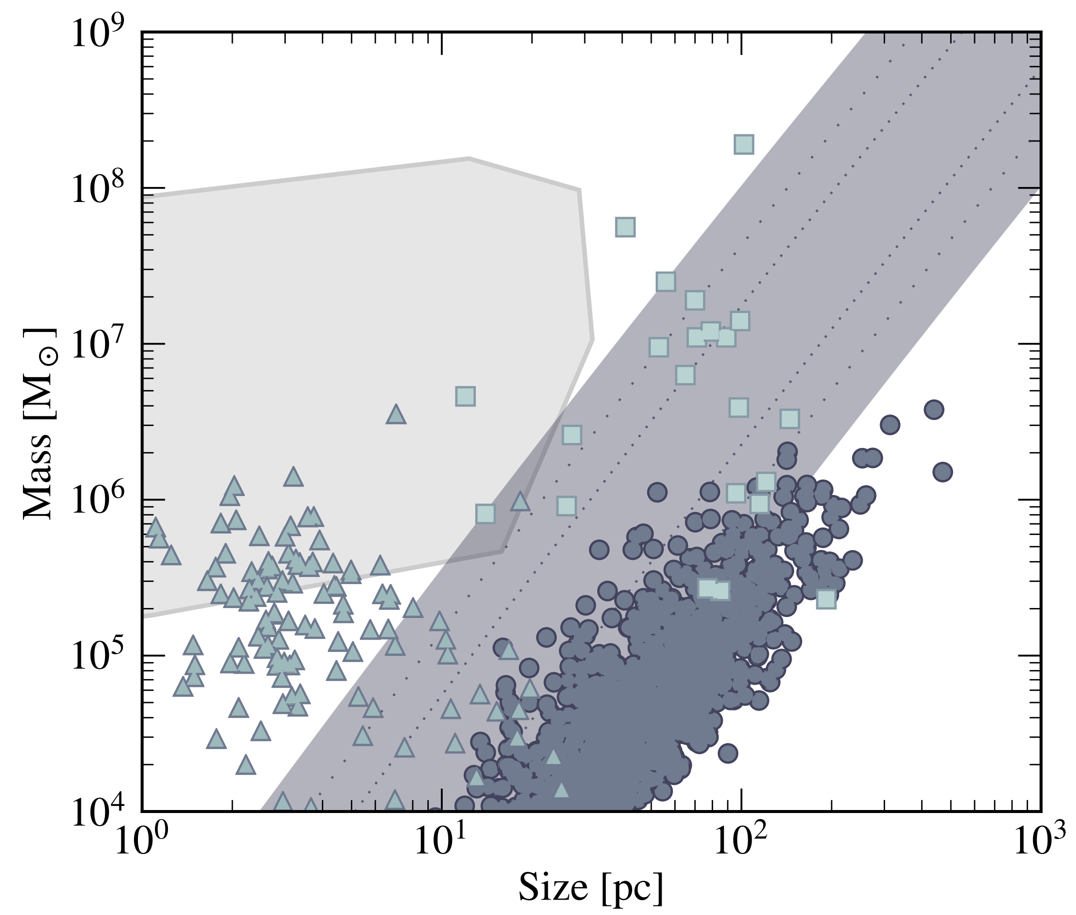
Phase-space perturbation of Milky Way stars



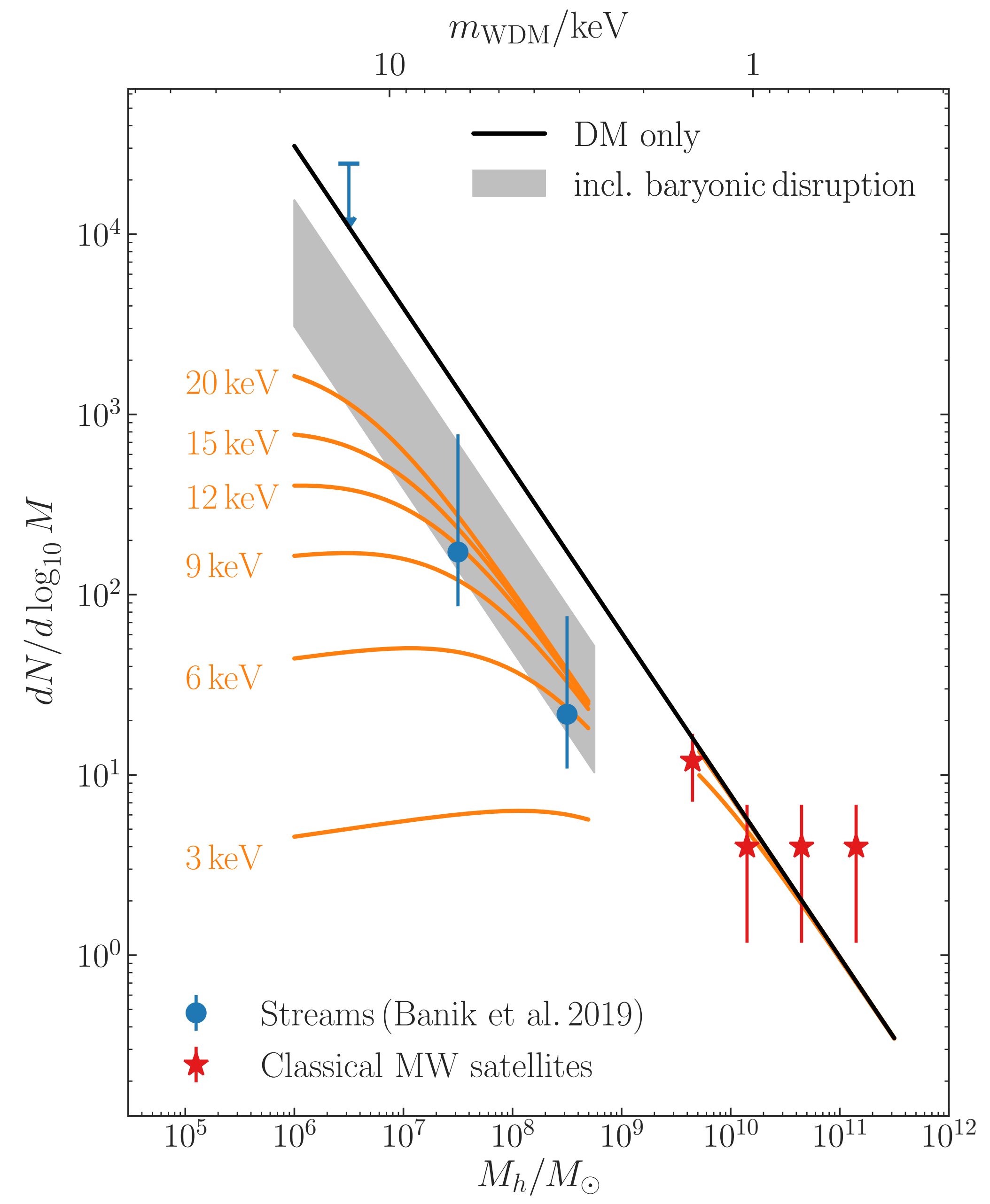
Buschmann et al [1711.03554]

Phase-space perturbation of cold stellar steams



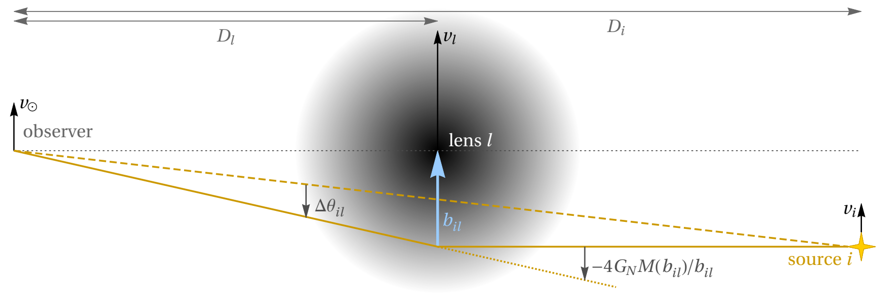


Bonaca et al [1811.03631]



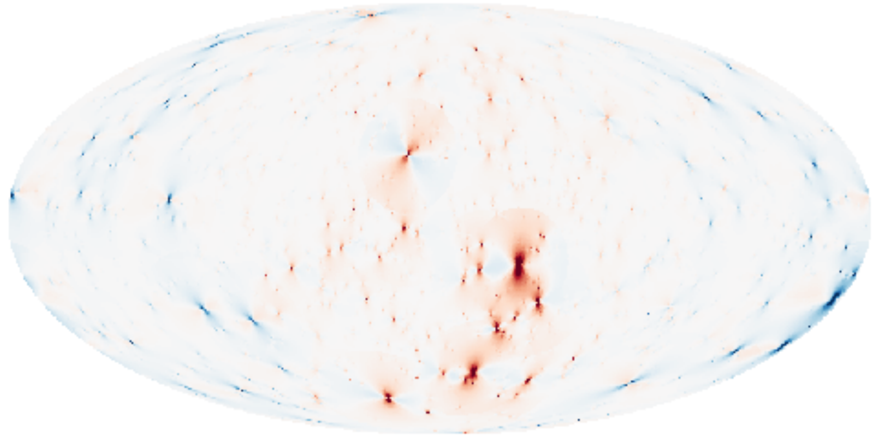
Banik et al [1911.02663]

Outline



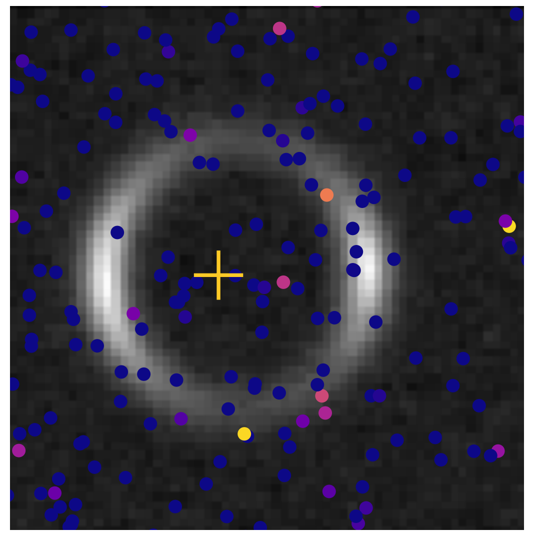
Gravitational Lensing

A Brief Primer



Inferring Galactic Substructure

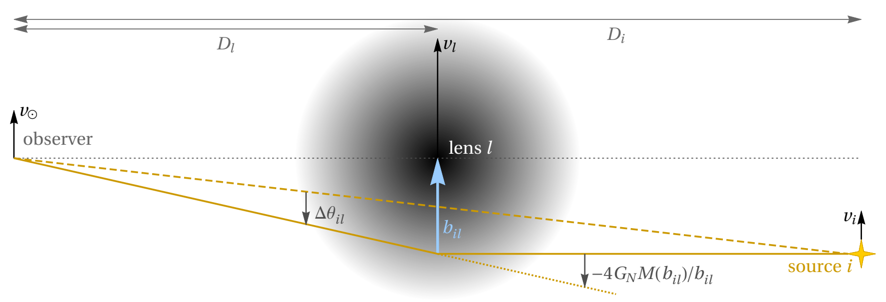
With Astrometry & Weak Lensing



Inferring Extragalactic Substructure

With Likelihood-free Inference & Strong Lensing

Outline



Gravitational Lensing

A Brief Primer

Inferring Galactic Substructure

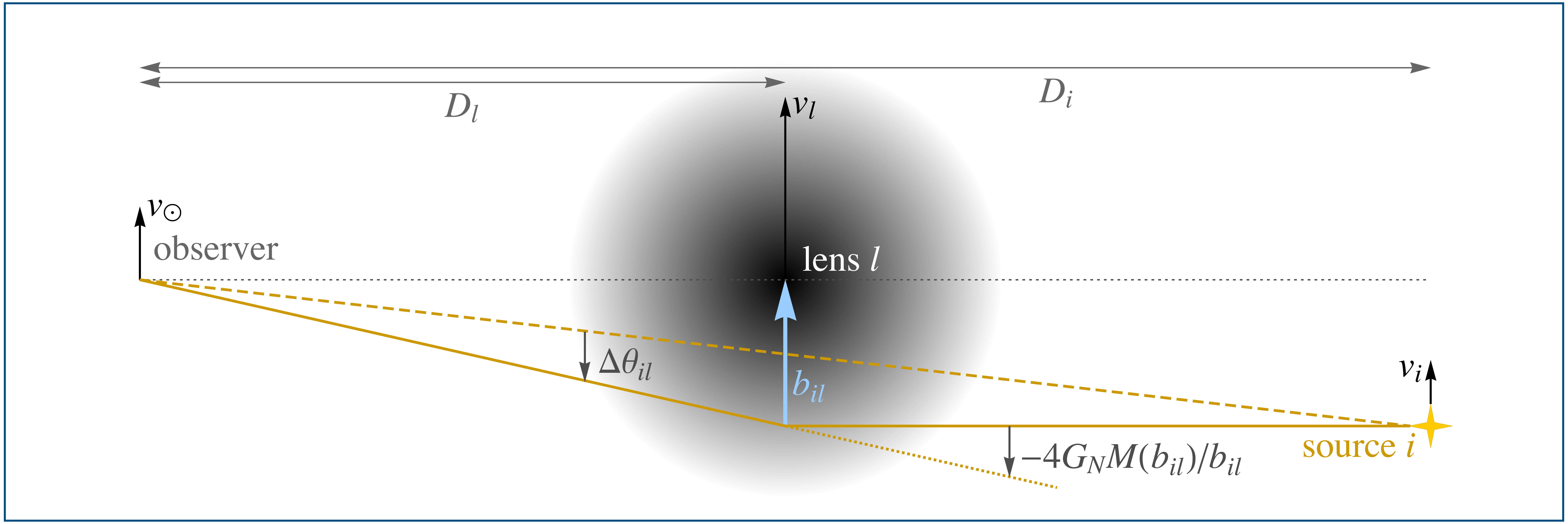
With Astrometry & Weak Lensing

Inferring Extragalactic Substructure

With Likelihood-free Inference & Strong Lensing

Gravitational lensing

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

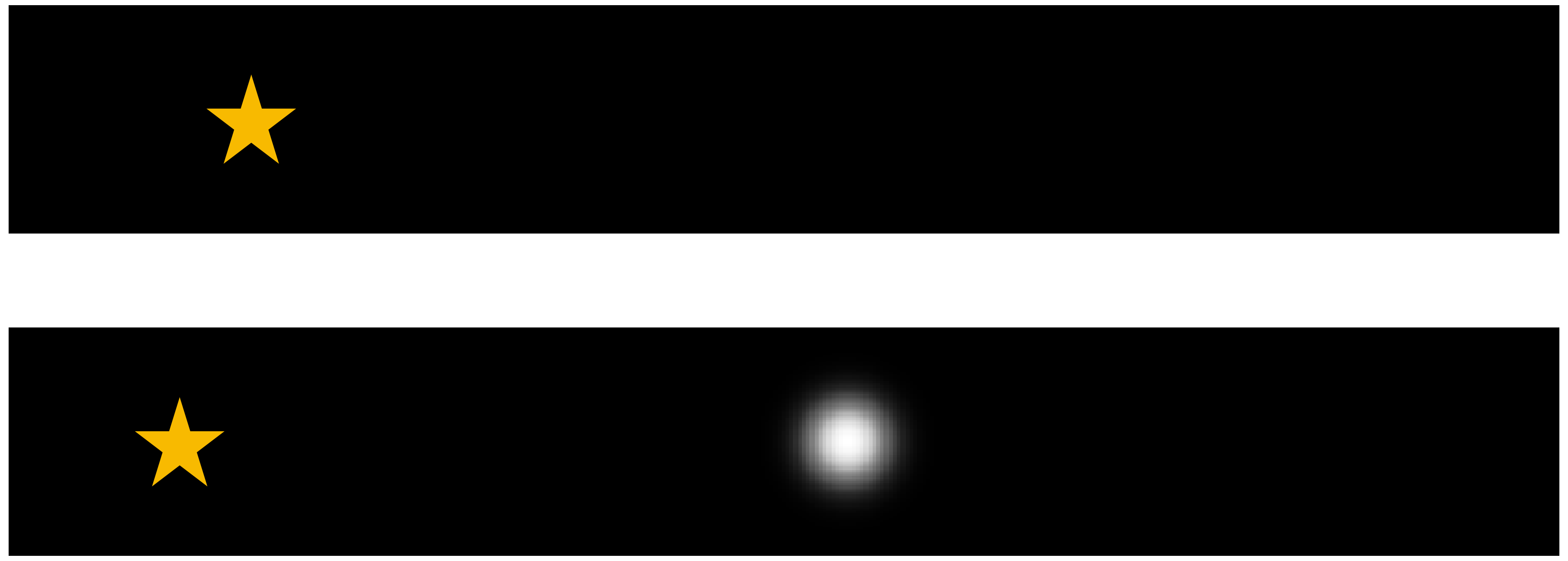


Van Tilburg et al, 2018

$$\overrightarrow{\Delta\theta} = \frac{2}{D_l}\vec{\nabla}_\theta \int \mathrm{d}z\, \Psi_G(\vec{r})$$

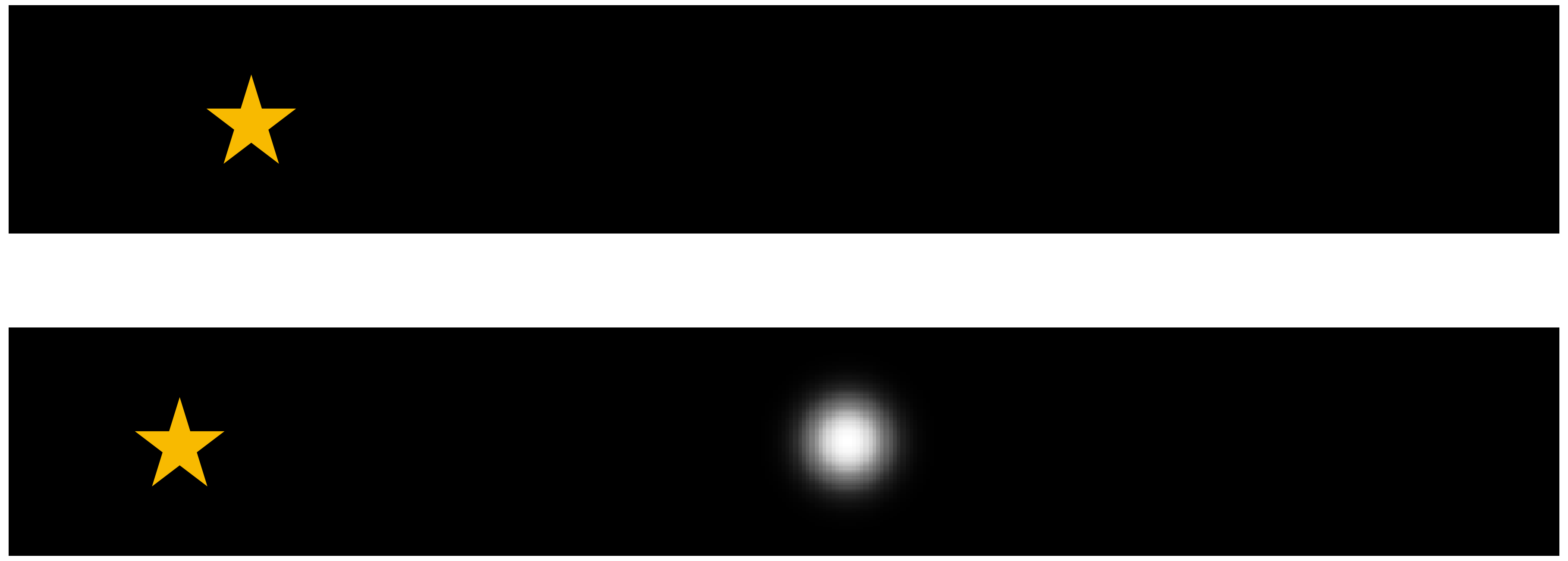
Different lensing regimes

Weak lensing: *Single image*

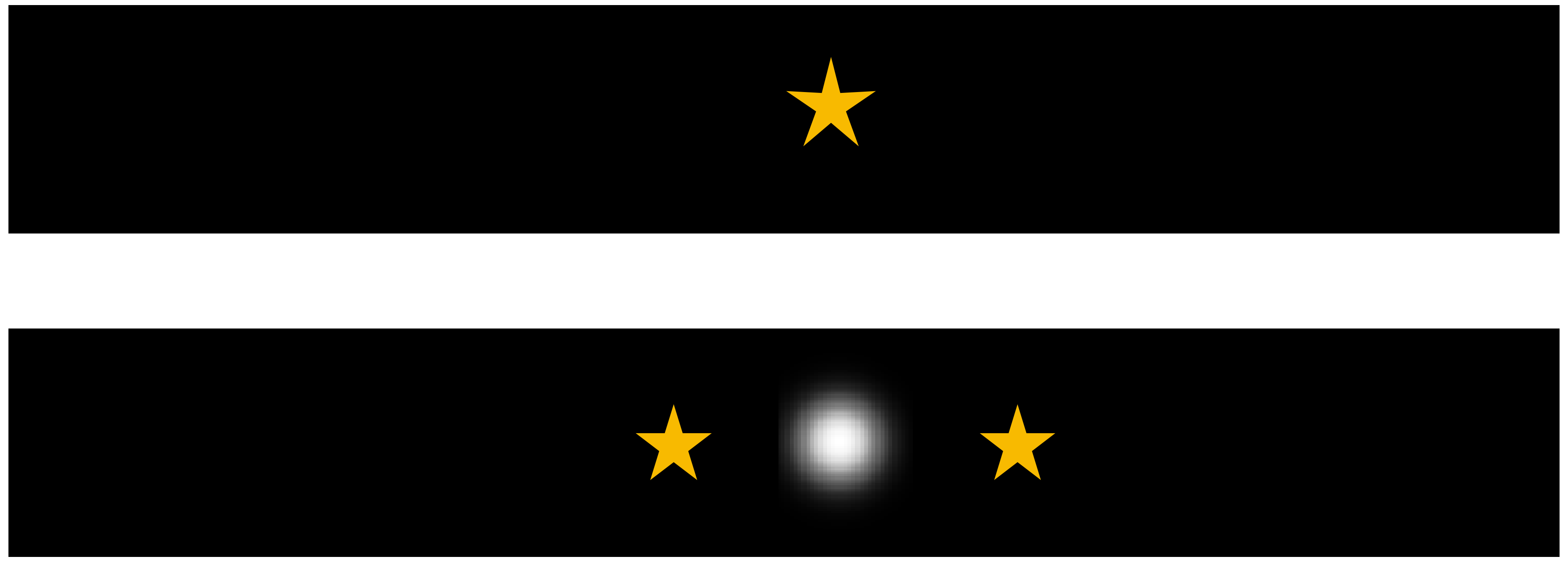


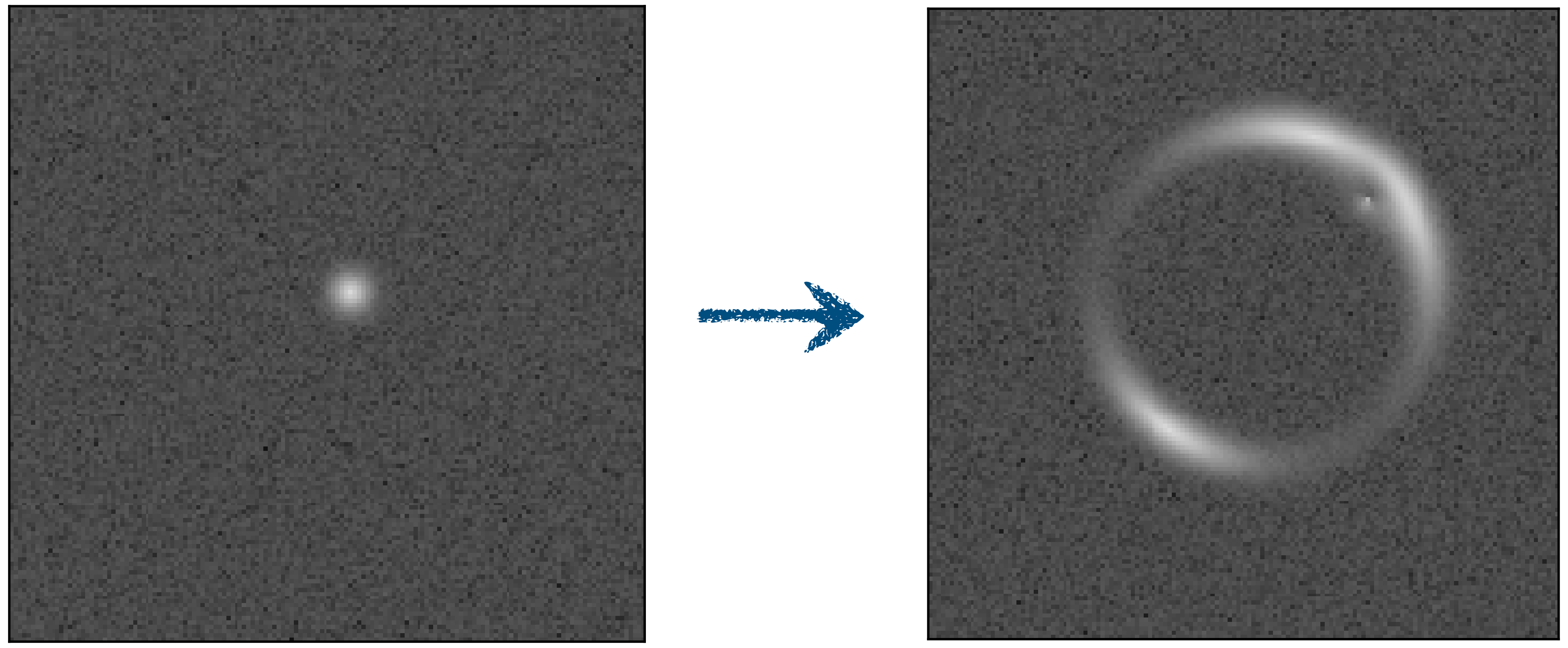
Different lensing regimes

Weak lensing: *Single image*



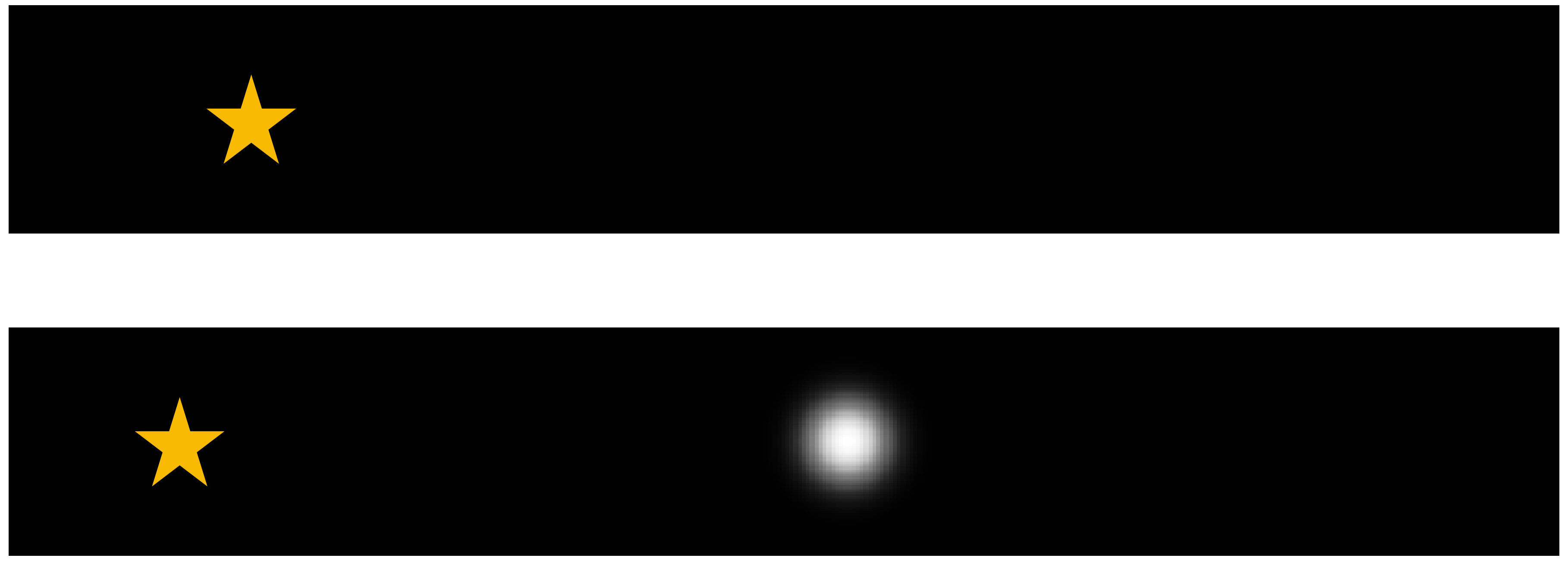
Strong lensing: *Multiple/extended images*



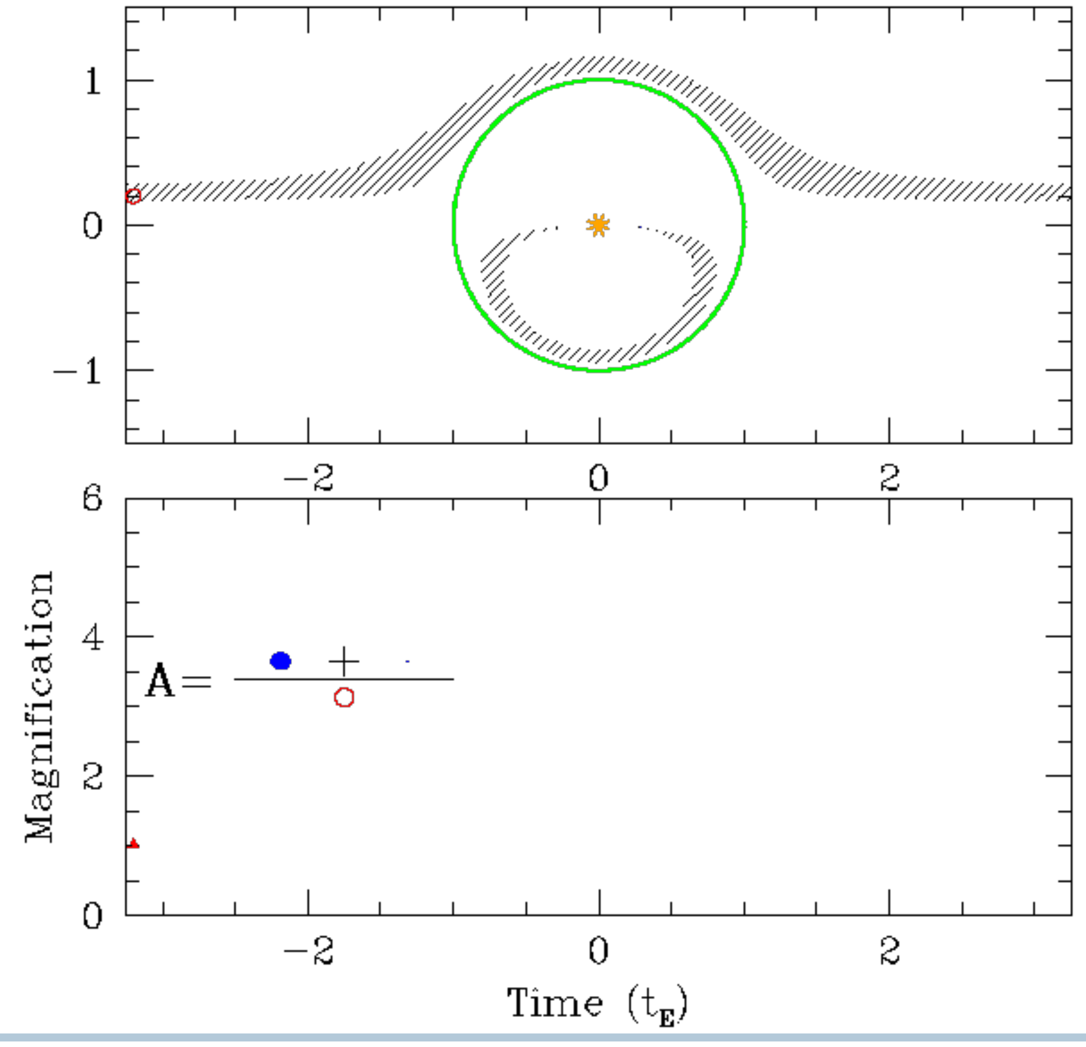


Different lensing regimes

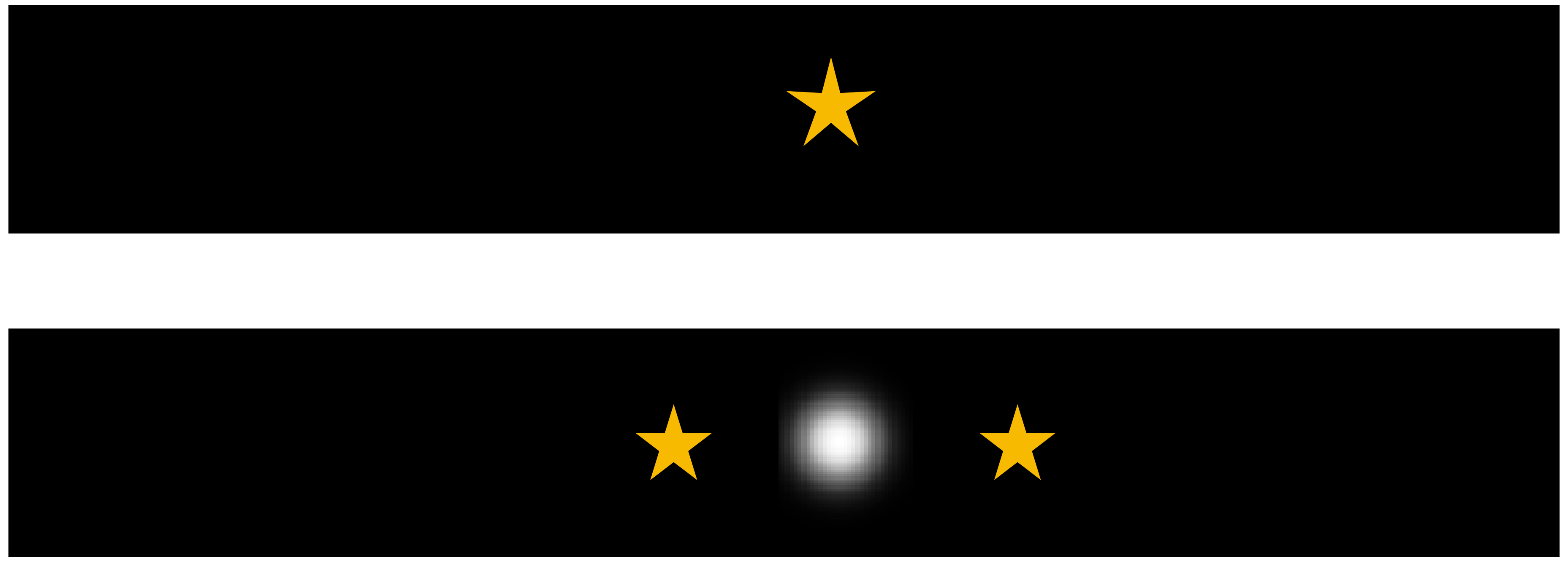
Weak lensing: *Single image*

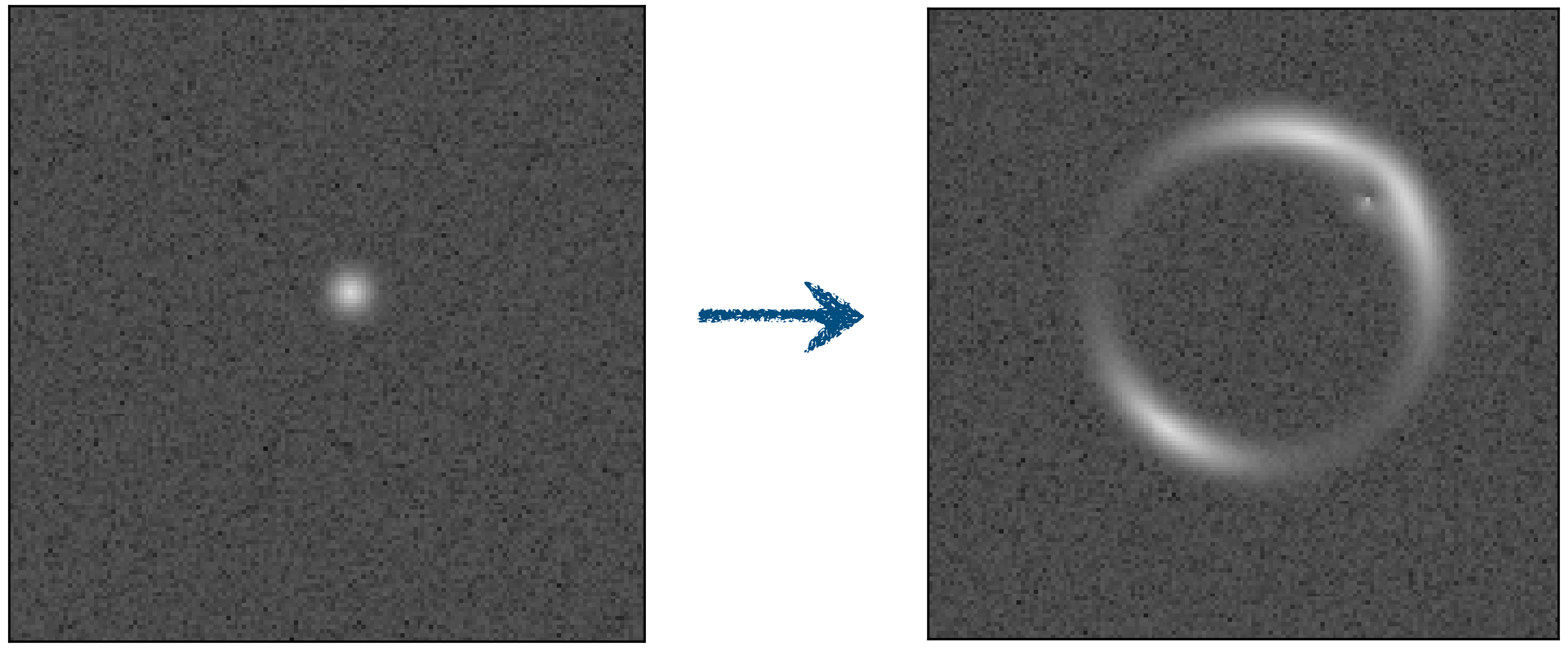


Microlensing: *Change in brightness(t)*



Strong lensing: *Multiple/extended images*

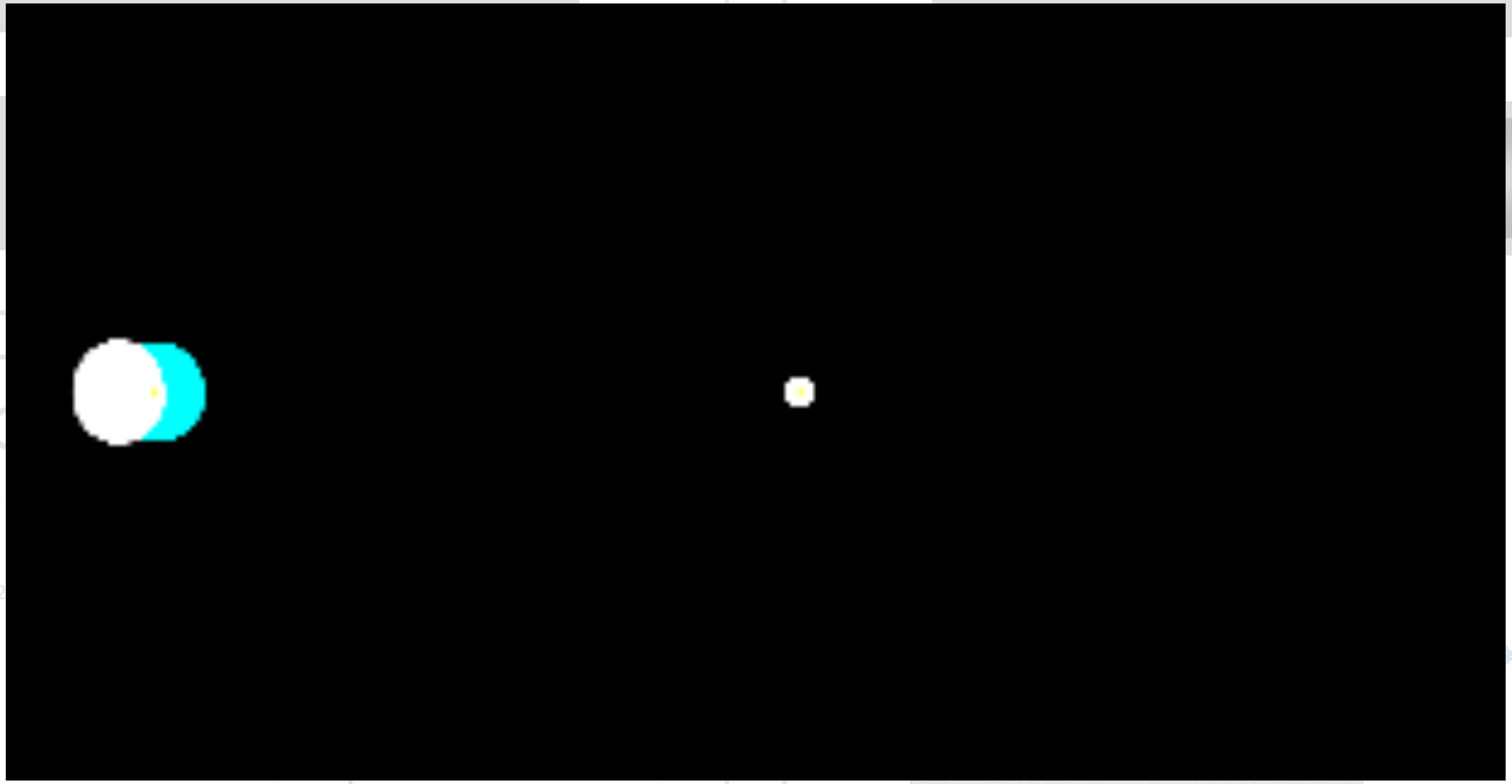




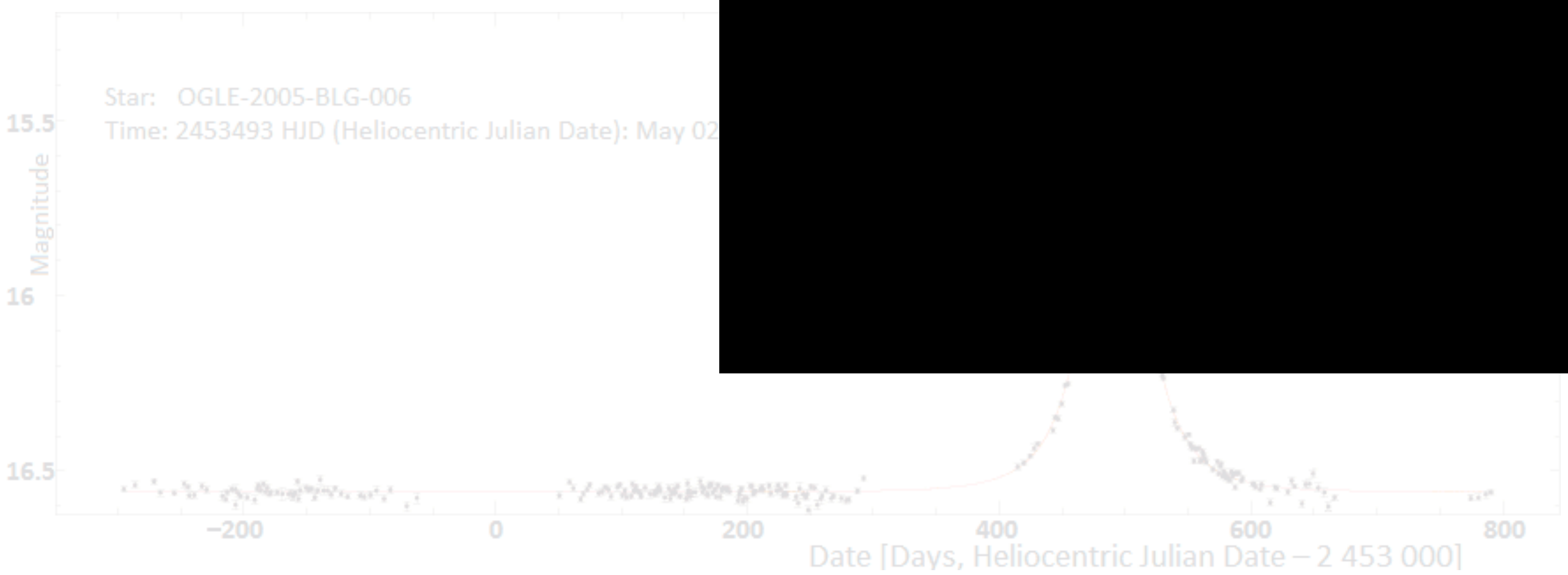
Different lensing regimes

Weak lensing: *Single image*

Strong lensing: *Multiple/extended images*



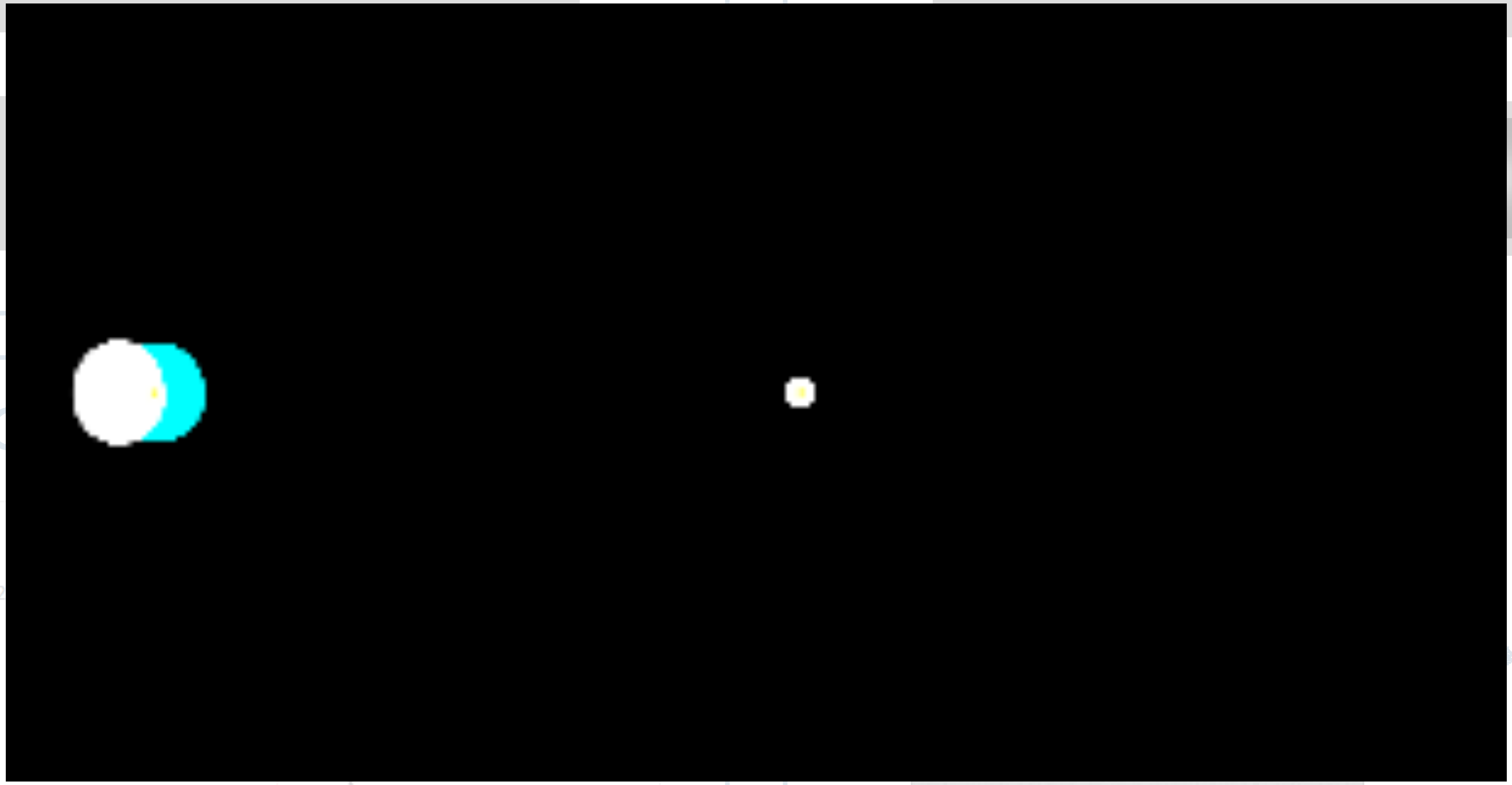
Microlensing:



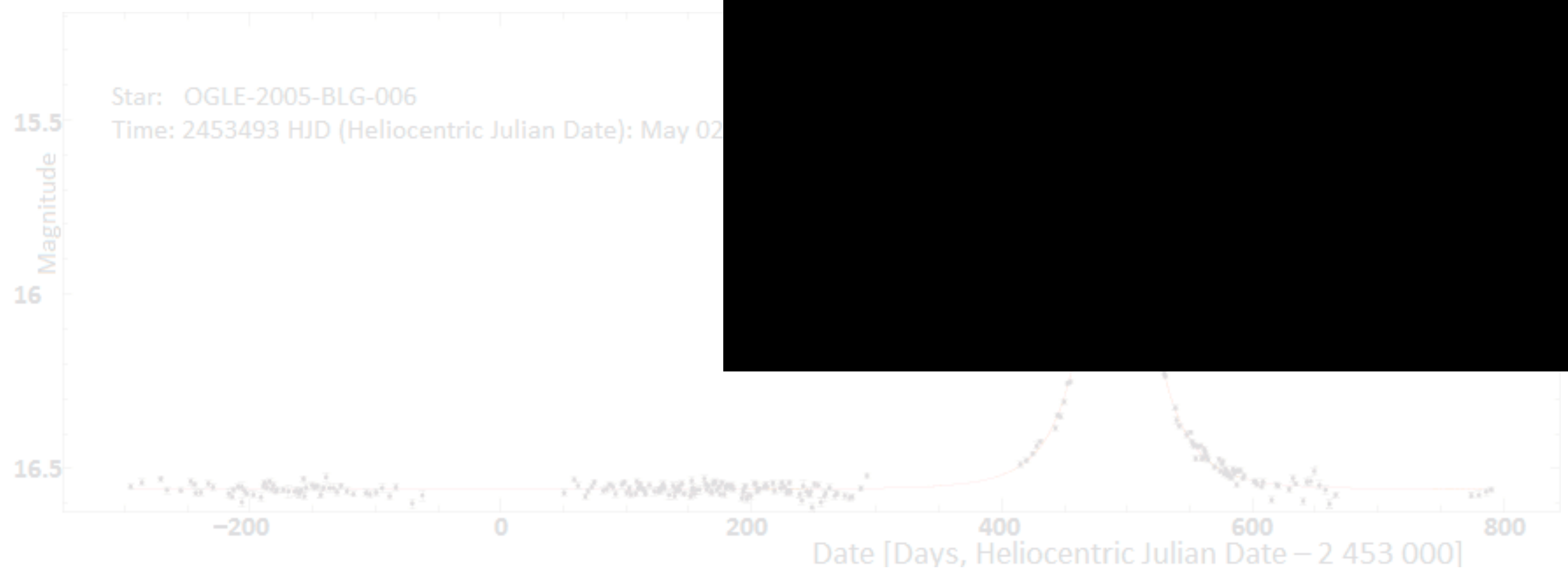
Different lensing regimes

Weak lensing: *Single image*

Strong lensing: *Multiple/extended images*



Microlensing:

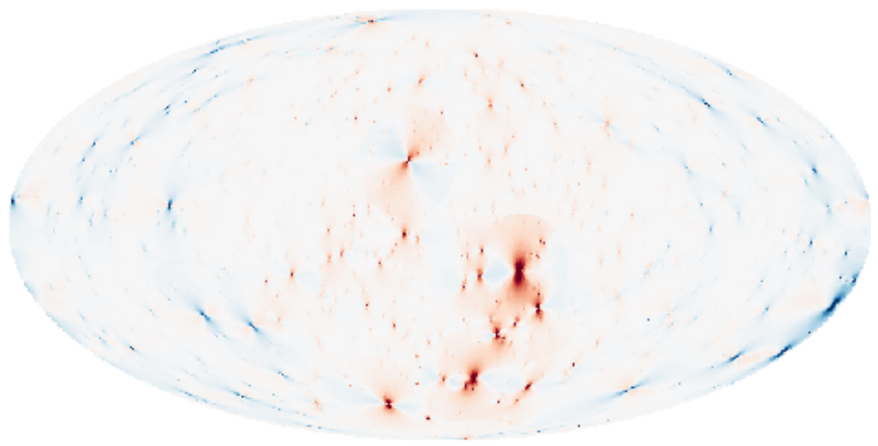


Outline

Van Tilburg Weiner

Gravitational Lensing

A Brief Primer



Inferring Galactic Substructure

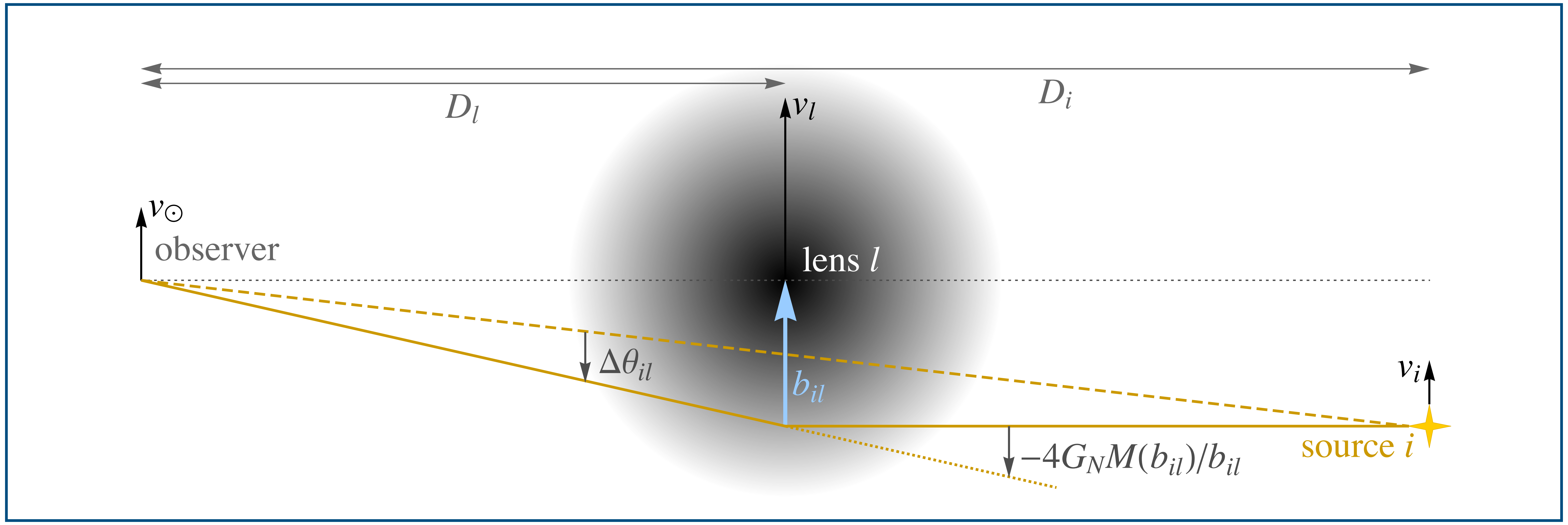
With Astrometry & Weak Lensing

Inferring Extragalactic Substructure

With Likelihood-free Inference & Strong Lensing

Gravitational lensing

Intervening mass causes a shift in the *apparent position* of luminous sources



Van Tilburg et al, 2018

Magnitude of the shift for Galactic subhalo lenses

$$\Delta\boldsymbol{\theta}_{il} = -\left(1 - \frac{D_l}{D_i}\right)\frac{4G_N M(b_{il})}{b_{il}}\hat{\mathbf{b}}_{il} \approx 400\ \mu\text{as}\left(\frac{M(b_{il})}{10^6\ M_\odot}\right)\left(\frac{10^2\ \text{pc}}{b_{il}}\right)$$

$1\ \mu\text{as} \approx 5 \times 10^{-12}\ \text{rad}$

$M(b)$: projected enclosed mass

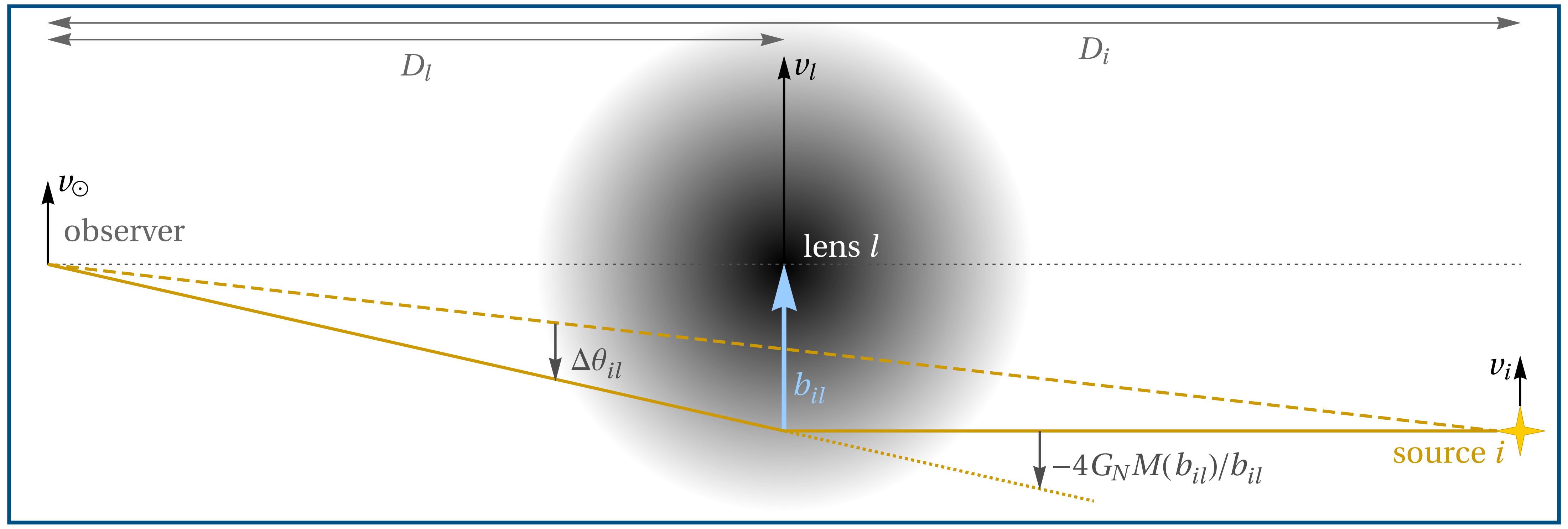
Problem: true position of background sources unknown


Shifts in position smaller than typical angular density variations of sources
Potential solution: go into the time-domain

Moving lenses induce motions in background sources

$$\mathbf{v}_{il} \equiv \dot{\mathbf{b}}_{il} = \mathbf{v}_l - \left(1 - \frac{D_l}{D_i}\right)\mathbf{v}_\odot - \frac{D_l}{D_i}\mathbf{v}_i$$

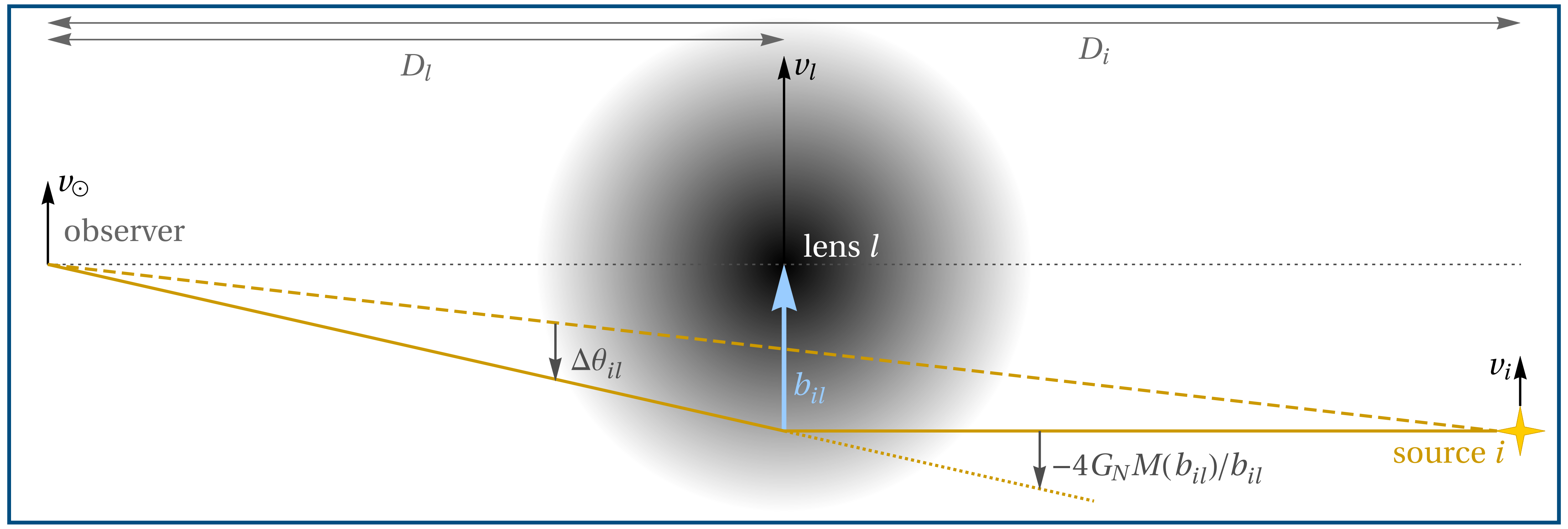
$$\boldsymbol{\mu}(\mathbf{b}) = 4G\left\{\frac{M(b)}{b^2}\left[2\hat{\mathbf{b}}(\hat{\mathbf{b}}\cdot\mathbf{v}_l) - \mathbf{v}_l\right] - \frac{M'(b)}{b}\hat{\mathbf{b}}(\hat{\mathbf{b}}\cdot\mathbf{v}_l)\right\}$$



Moving lenses induce motions in background sources

$$\mathbf{v}_{il} \equiv \dot{\mathbf{b}}_{il} = \mathbf{v}_l - \left(1 - \frac{D_l}{D_i}\right)\mathbf{v}_\odot - \frac{D_l}{D_i}\mathbf{v}_i$$

$$\boldsymbol{\mu}(\mathbf{b}) = 4G\left\{\frac{M(b)}{b^2}\left[2\hat{\mathbf{b}}(\hat{\mathbf{b}}\cdot\mathbf{v}_l) - \mathbf{v}_l\right] - \frac{M'(b)}{b}\hat{\mathbf{b}}(\hat{\mathbf{b}}\cdot\mathbf{v}_l)\right\}$$



Typical size of time-domain effects for Galactic lenses

Angular velocity shift: $$\Delta\dot{\theta}_{il} \sim \frac{4G_N M(b_{il}) v_{il}}{b_{il}^2} \sim 10^{-3}\ \mu\text{as y}^{-1}\left(\frac{M(b_{il})}{10^6\ M_\odot}\right)\left(\frac{10^2\ \text{pc}}{b_{il}}\right)^2$$

Angular acceleration shift: $$\Delta\ddot{\theta}_{il} \sim \frac{4G_N M(b_{il}) v_{il}^2}{b_{il}^3} \sim 4\times10^{-3}\ \mu\text{as y}^{-2}\left(\frac{M(b_{il})}{M_\odot}\right)\left(\frac{10^{-2}\ \text{pc}}{b_{il}}\right)^3$$

Smaller than current or anticipated astrometric precision

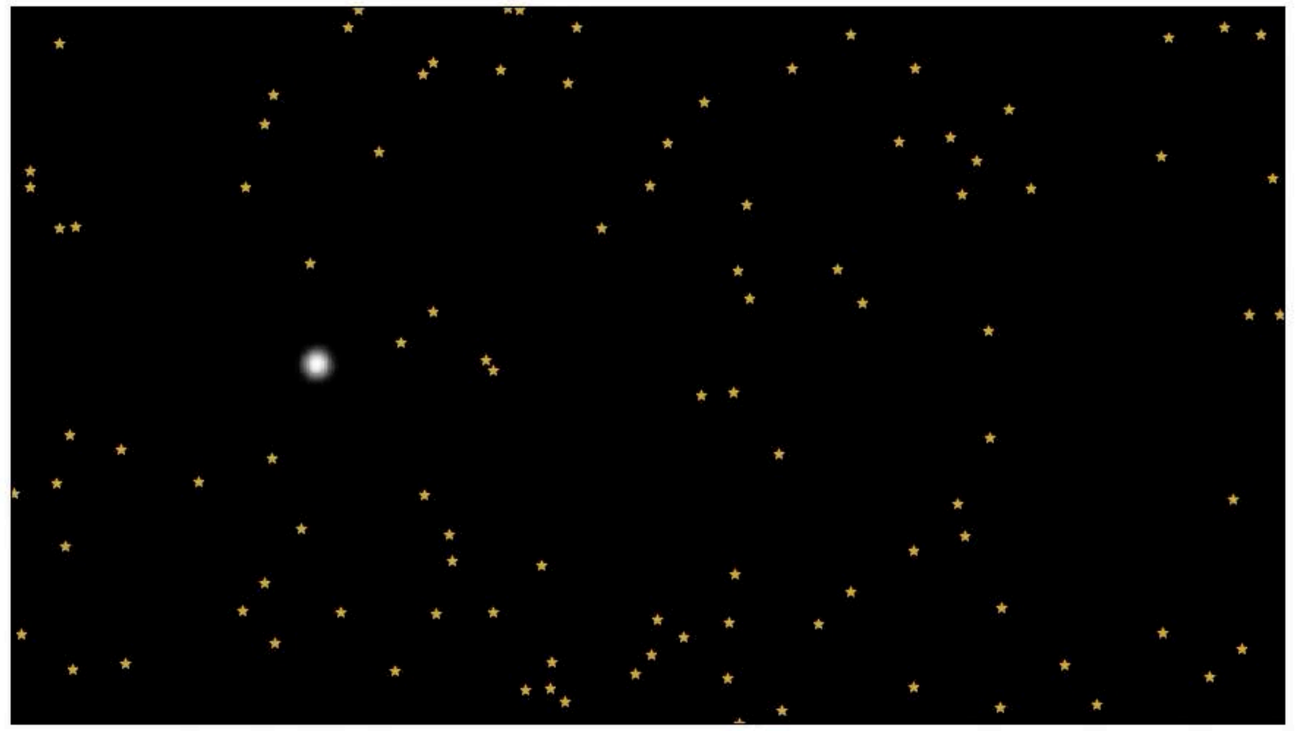
Blips and outlier motions

Blips and outlier motions

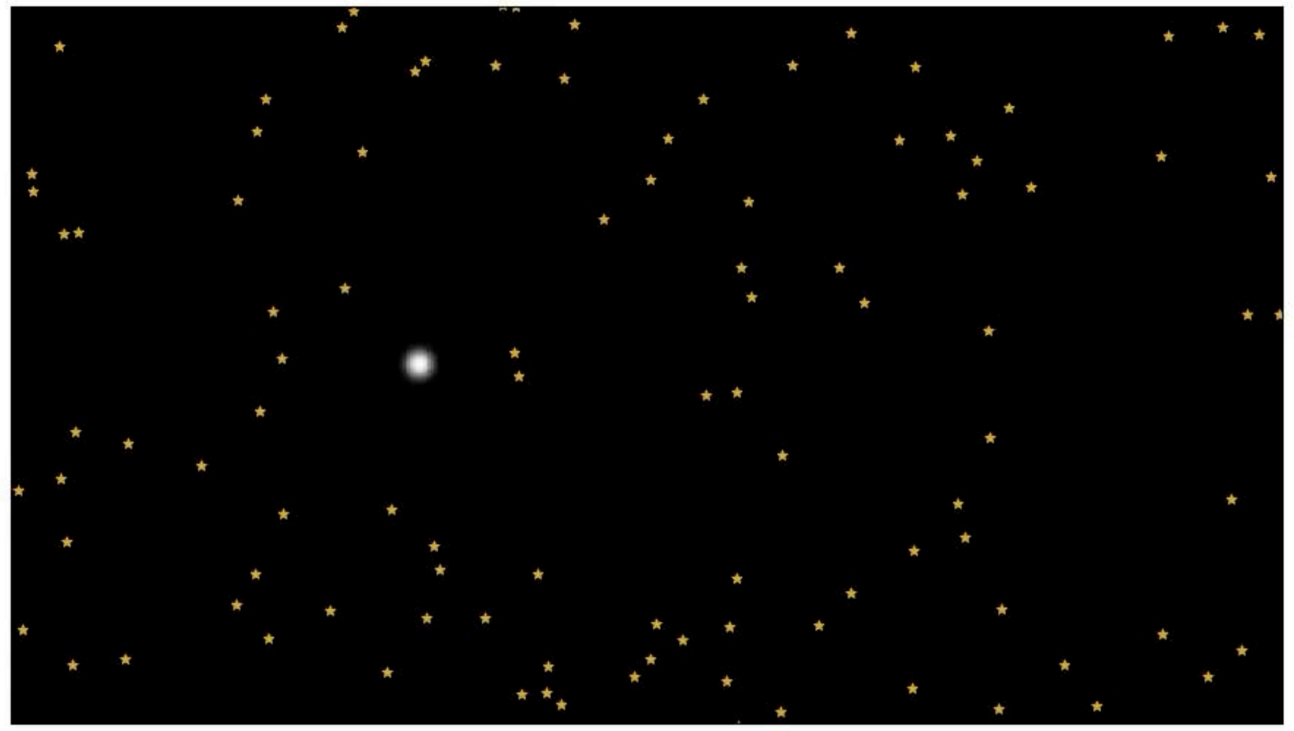
Blips and outlier motions

Compact lenses can induce *dramatic variations* in source positions

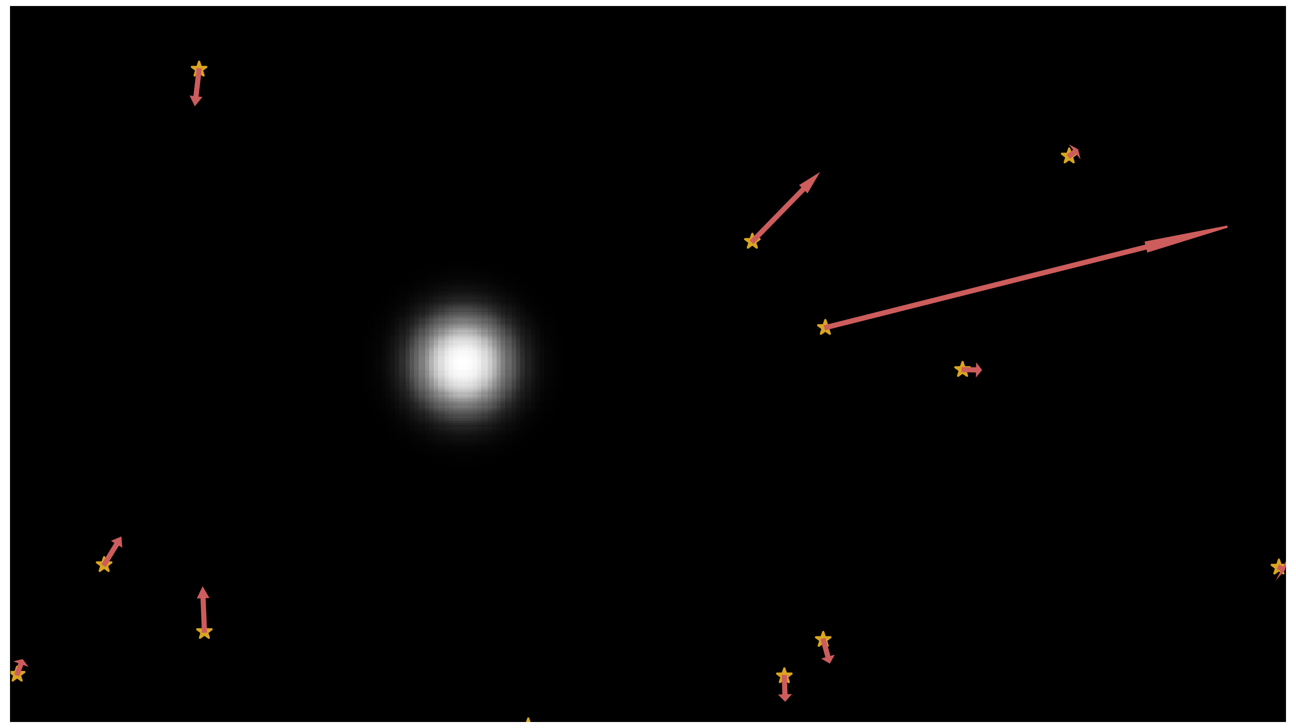
Mono-blip



Multi-blip



Outlier velocities/accelerations



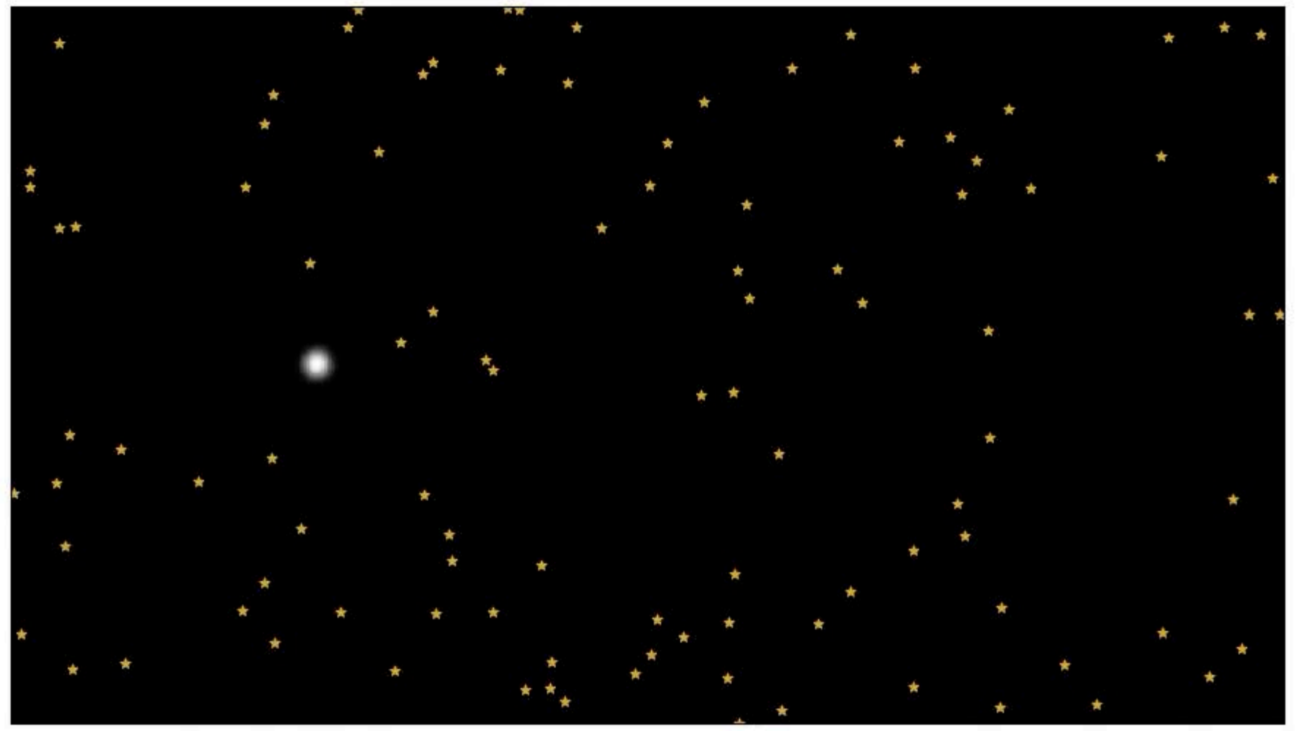
Observables systematically studied in

Van Tilburg, Taki, Weiner, *"Halometry from Astrometry"*, [1804.01991]

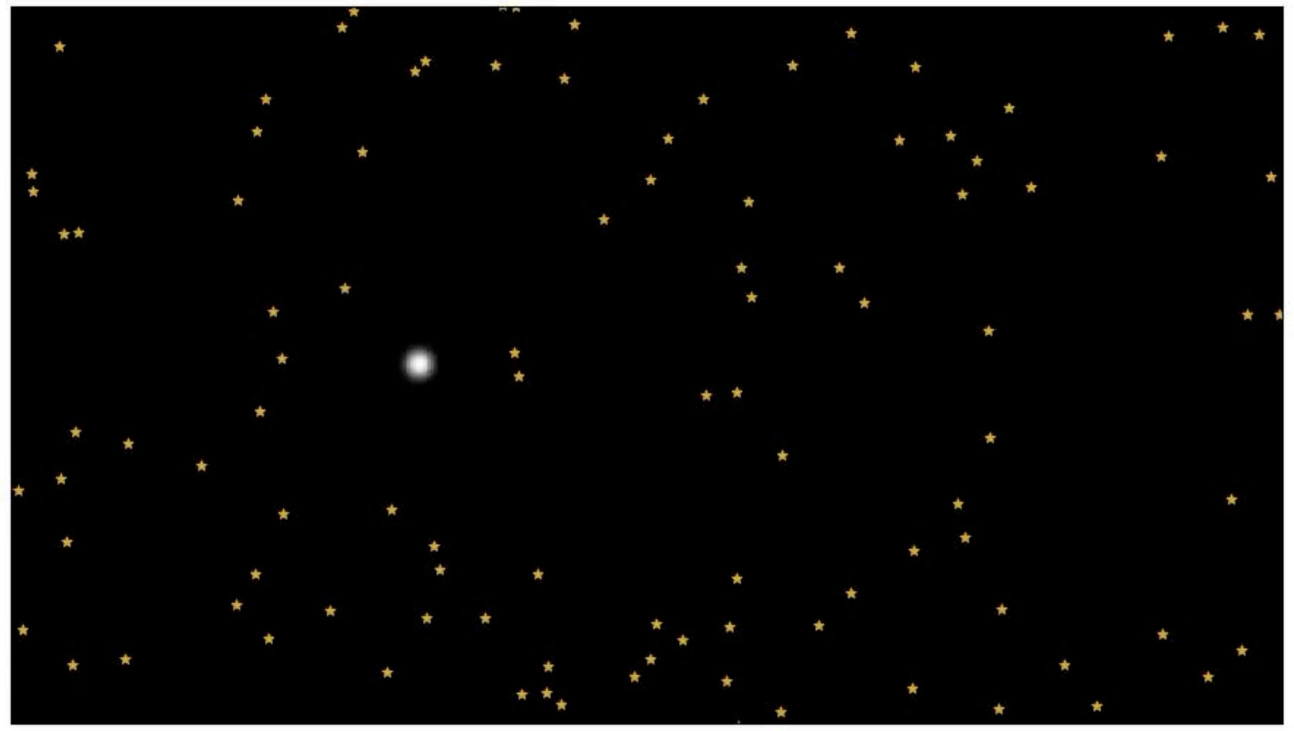
Blips and outlier motions

Compact lenses can induce *dramatic variations* in source positions

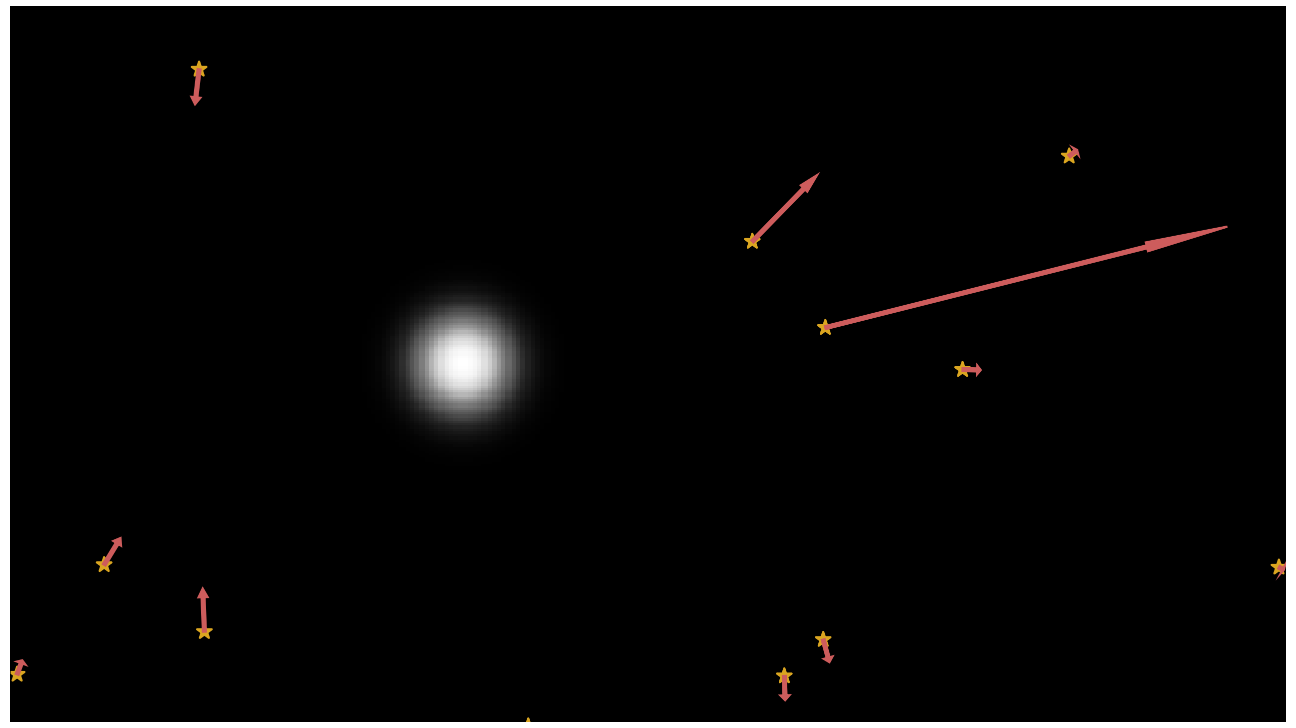
Mono-blip



Multi-blip



Outlier velocities/accelerations



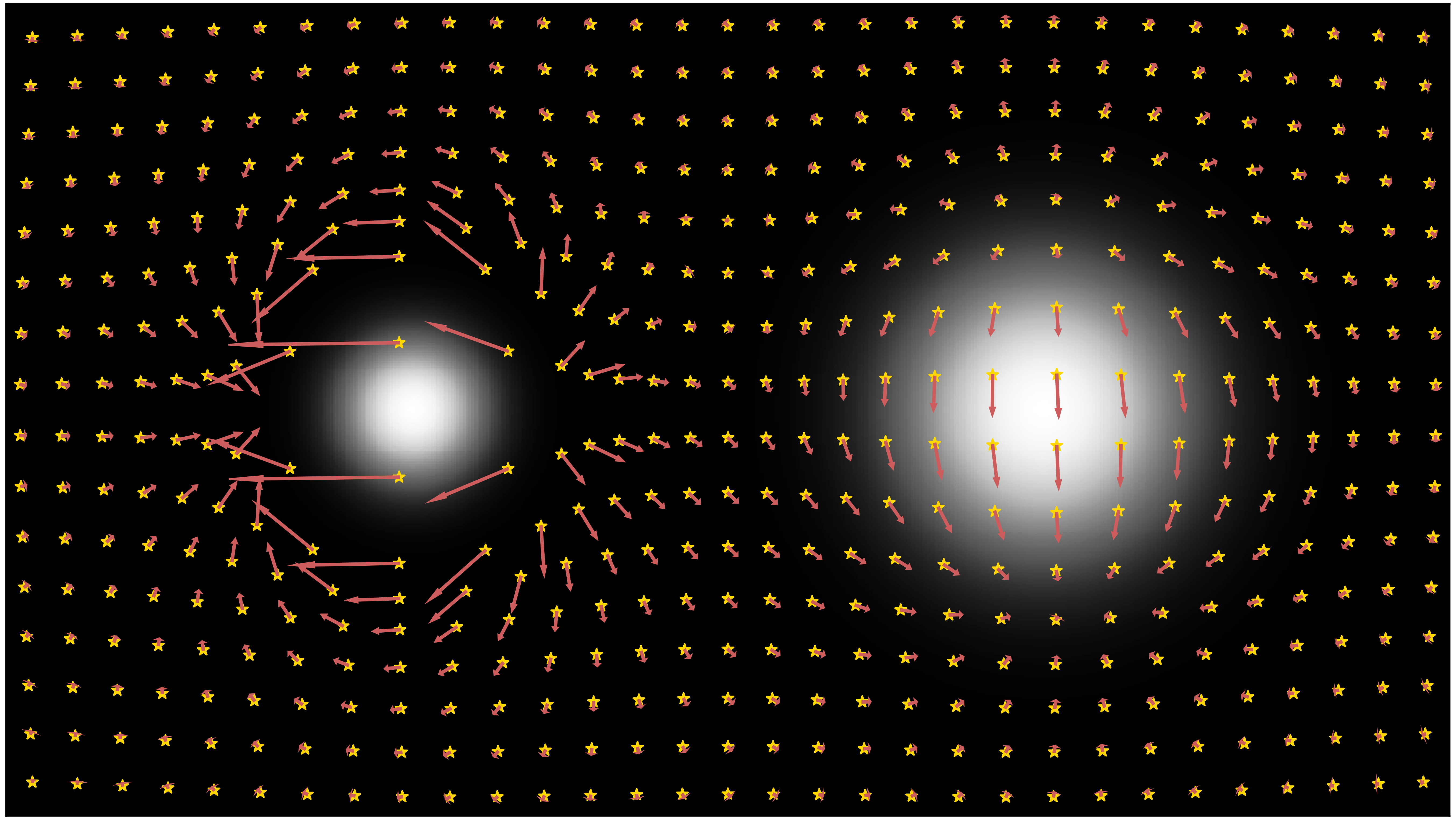
Observables systematically studied in

Van Tilburg, Taki, Weiner, *"Halometry from Astrometry"*, [1804.01991]

Extended lenses

Problem: extended lenses strongly suppress lensing effects

$$\Delta\dot{\theta}_{il} \sim \frac{4G_N M(b_{il}) v_{il}}{b_{il}^2} \sim 10^{-3}\ \mu\text{as y}^{-1} \left(\frac{M(b_{il})}{10^6\ M_\odot}\right)\left(\frac{10^2\ \text{pc}}{b_{il}}\right)^2$$

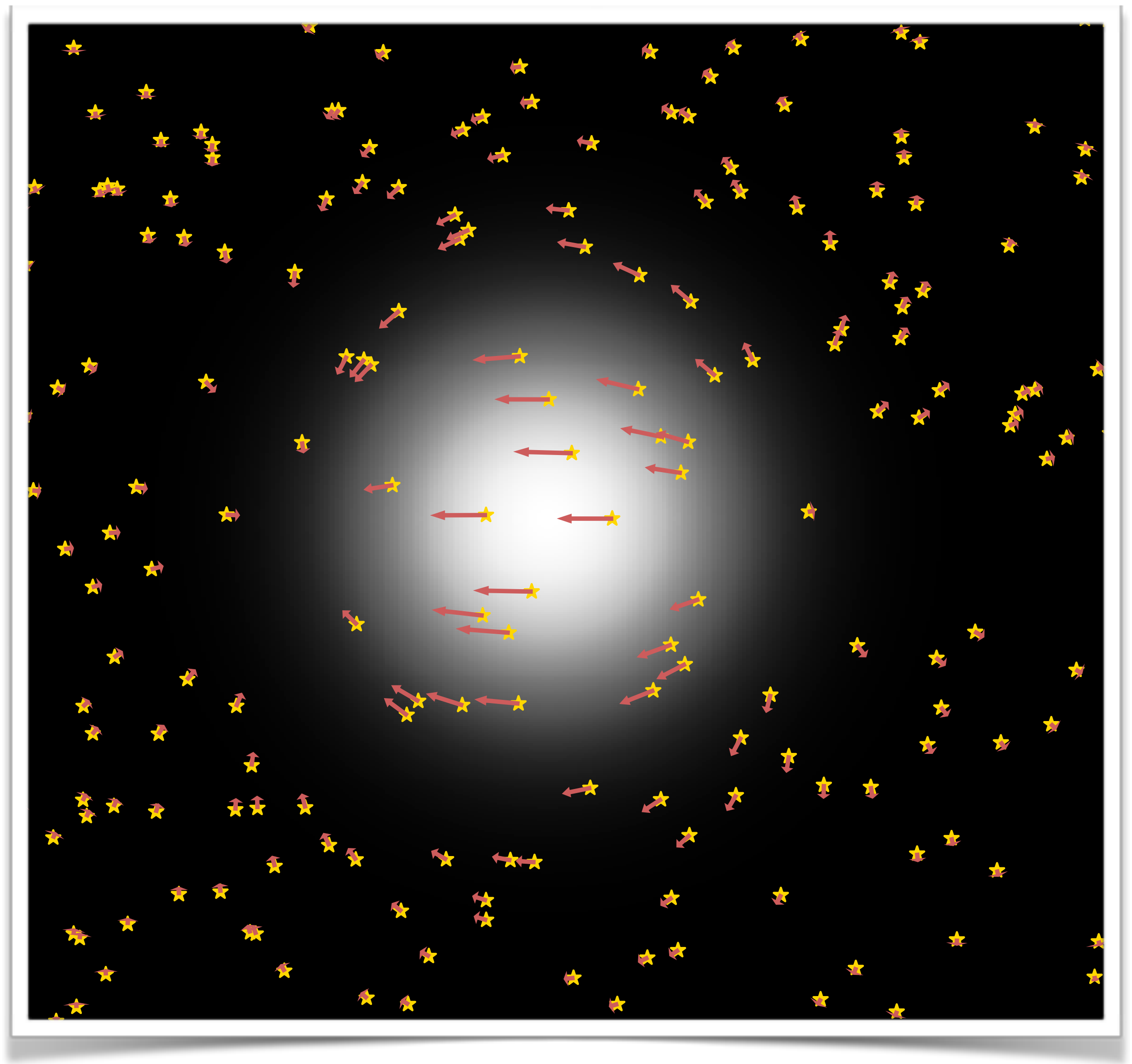


Potential solution: study correlated motions

Induced velocity dipole pattern

Correlated lens-induced motions: local templates

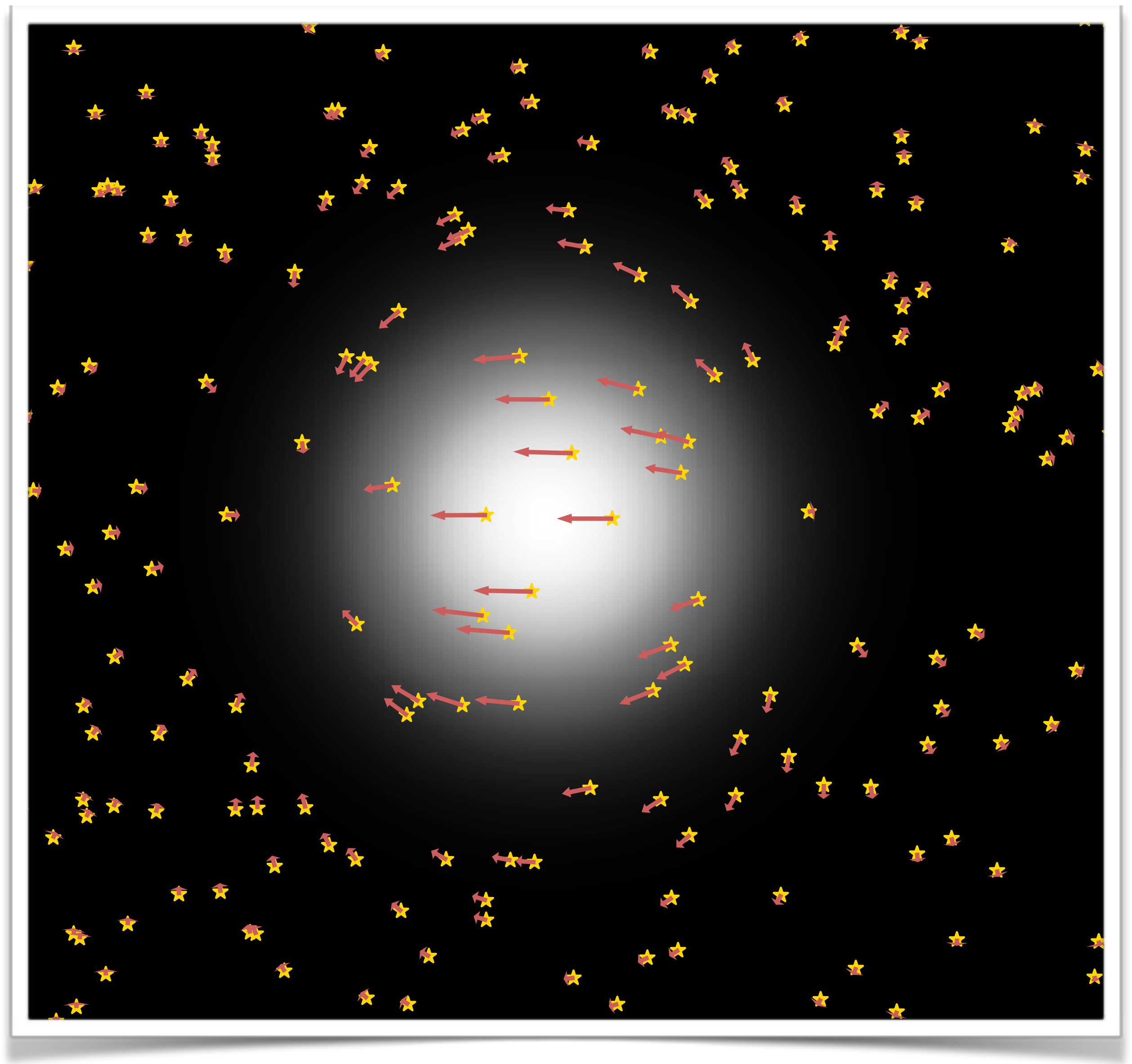
Can use templates for expected induced motions to look for substructure lenses in dense source regions



Prospects studied in
Van Tilburg, Taki, Weiner,
"Halometry from Astrometry", JCAP [1804.01991]

Correlated lens-induced motions: local templates

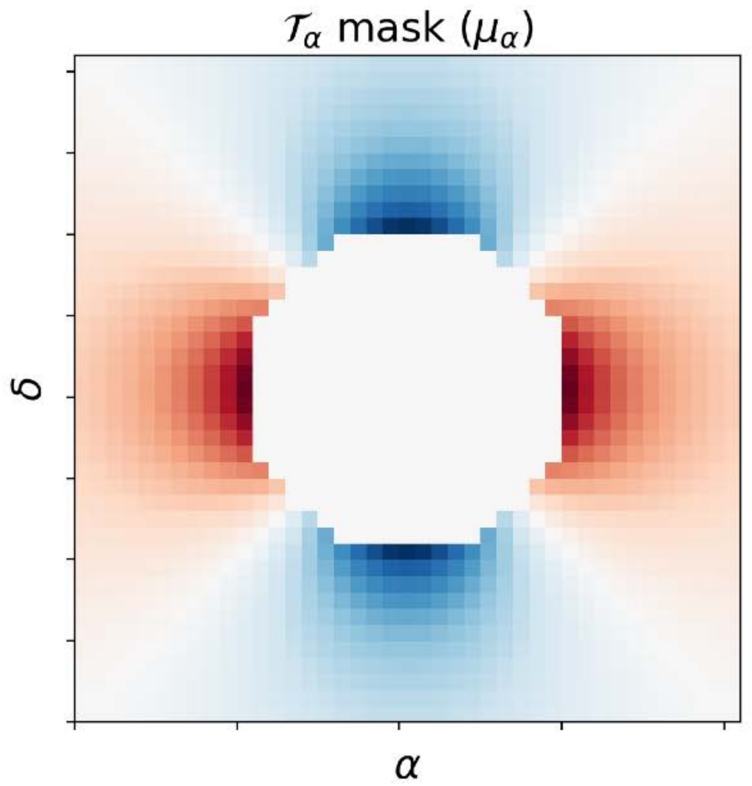
Can use templates for expected induced motions to look for substructure lenses in dense source regions



Prospects studied in
Van Tilburg, Taki, Weiner,
"Halometry from Astrometry", JCAP [1804.01991]

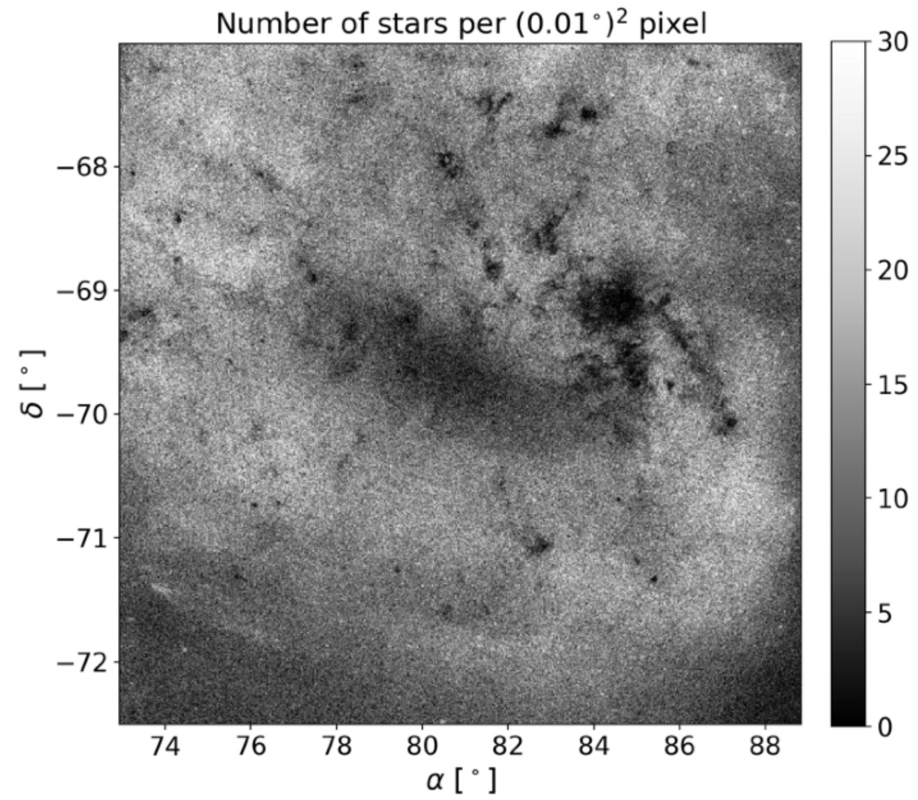
Ongoing analysis

Velocity templates

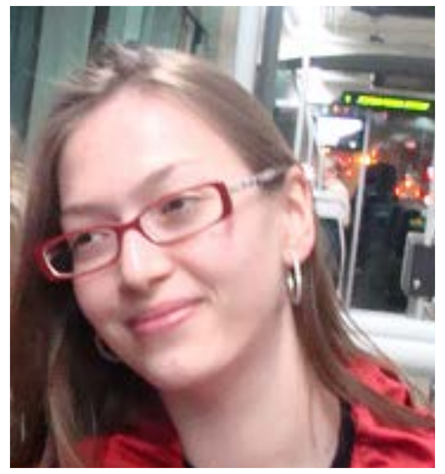


\+

LMC field of view

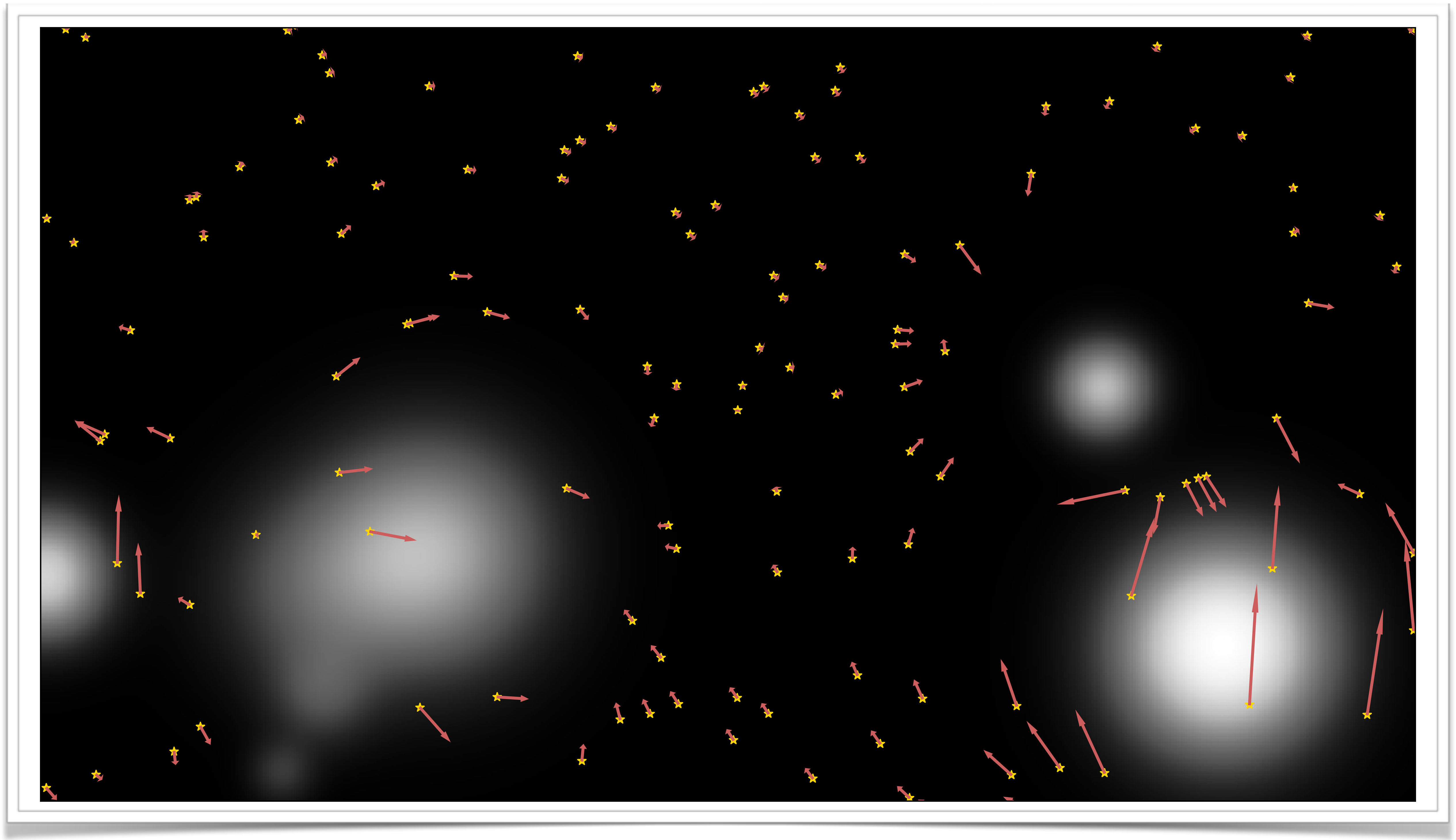


Mondino *Taki* *Van Tilburg* *Weiner*

Global, correlated lens-induced motions

Alternatively, can look for global patterns in induced motion of sources due to a population of lenses

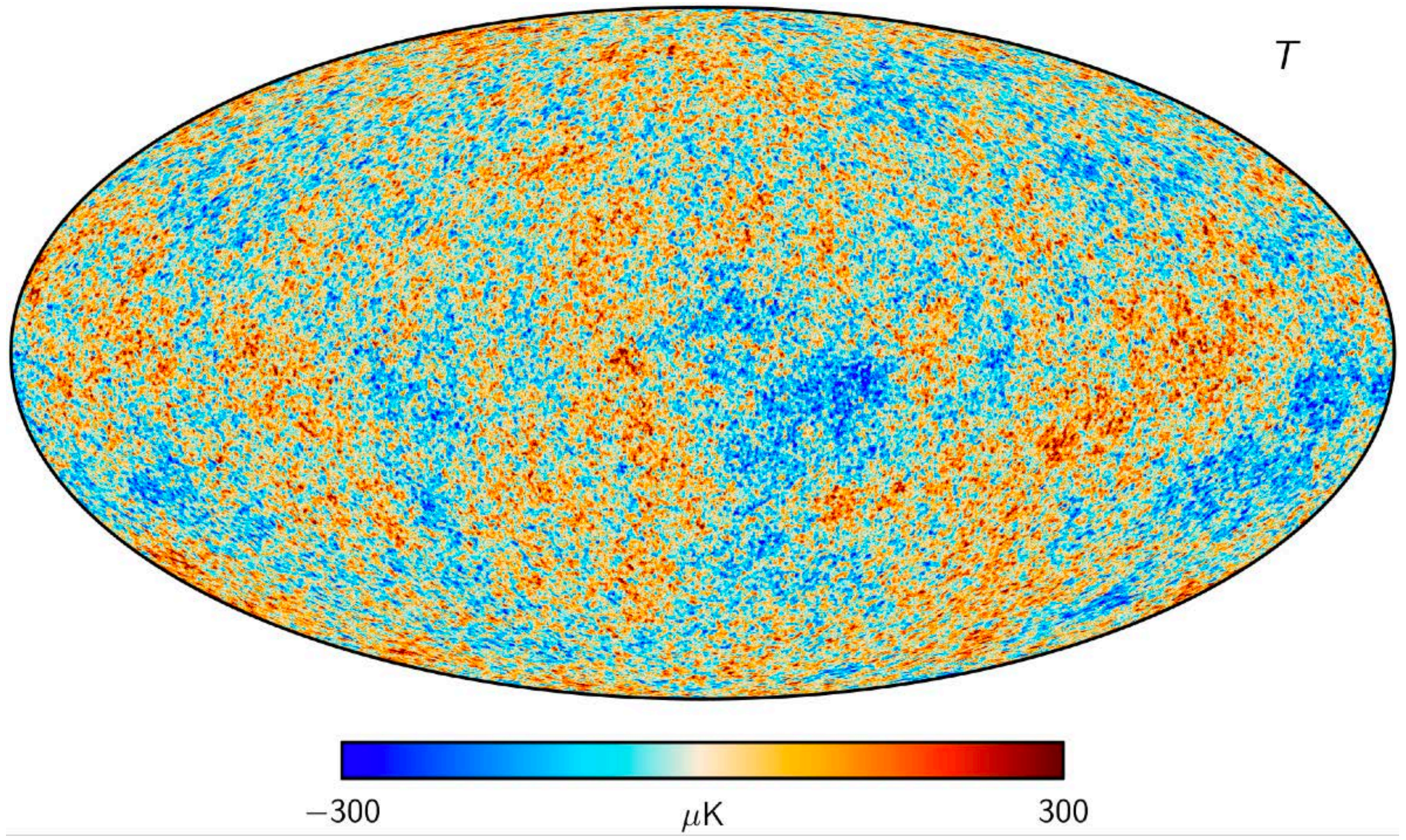


One tool to do this: *angular correlation functions*

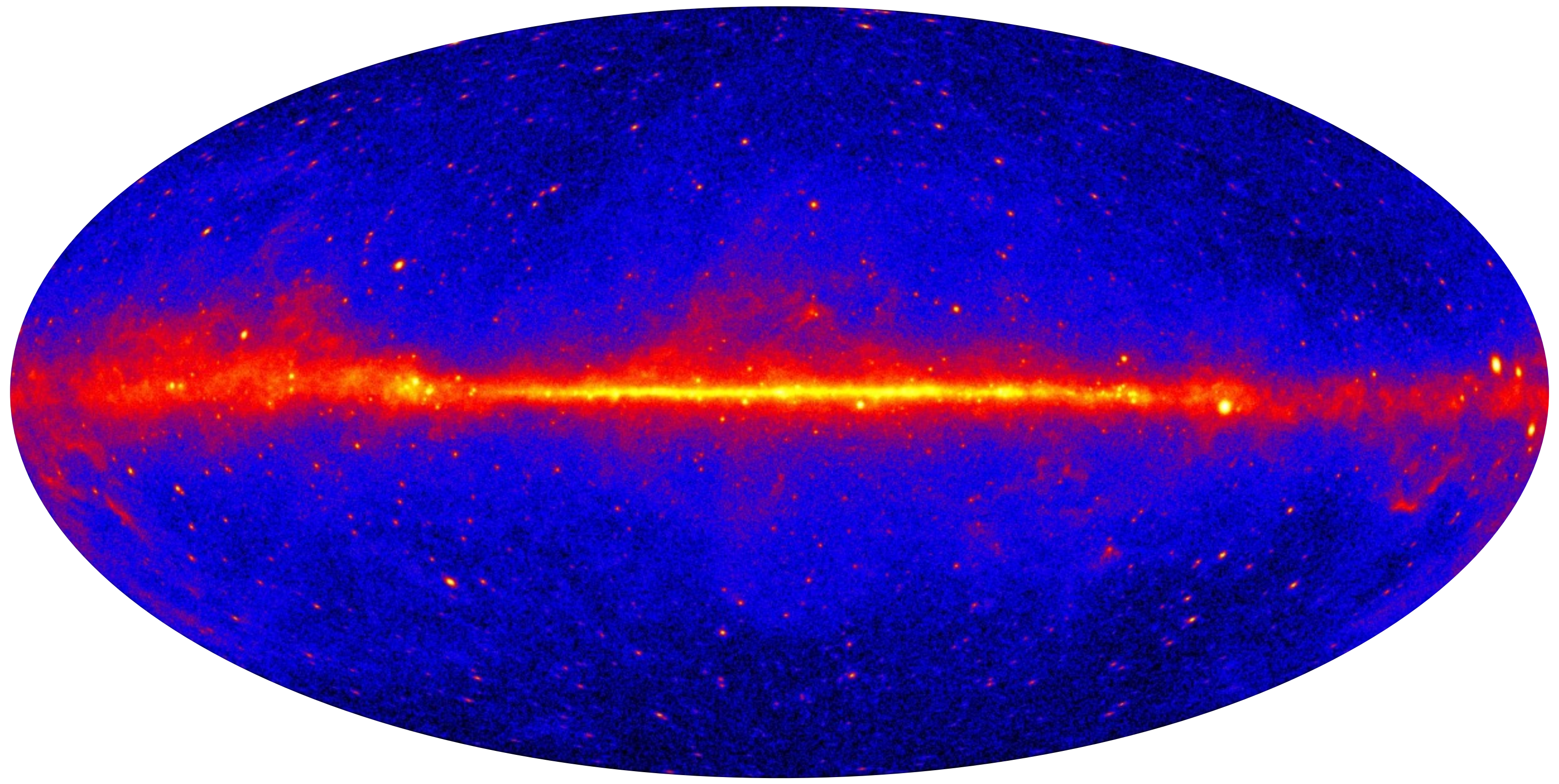
Statistics for analyzing large-scale structure

A lot of cosmology boils down to looking for patterns at different scales...

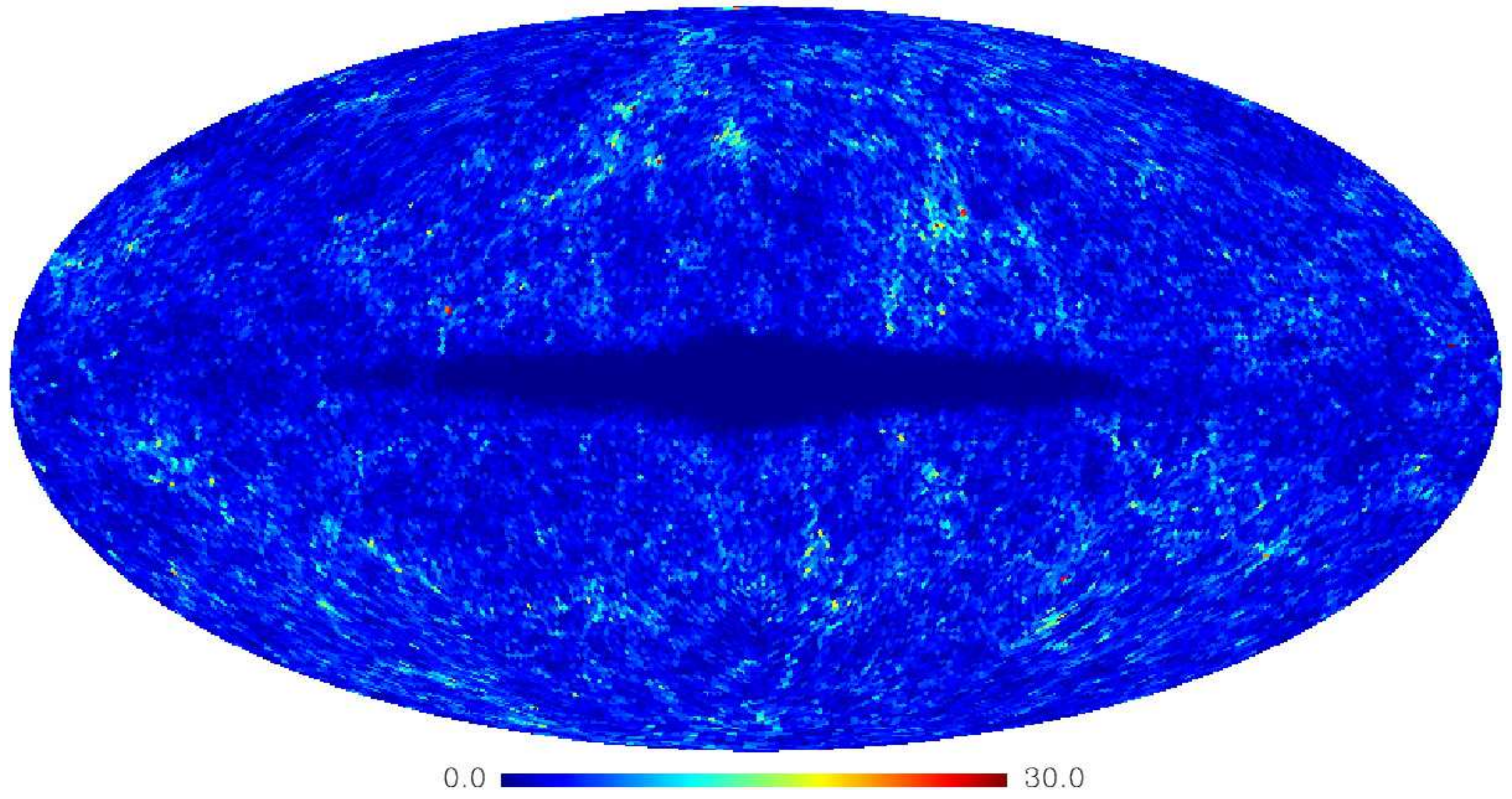
Planck CMB



Fermi gamma-rays



2MASS galaxies

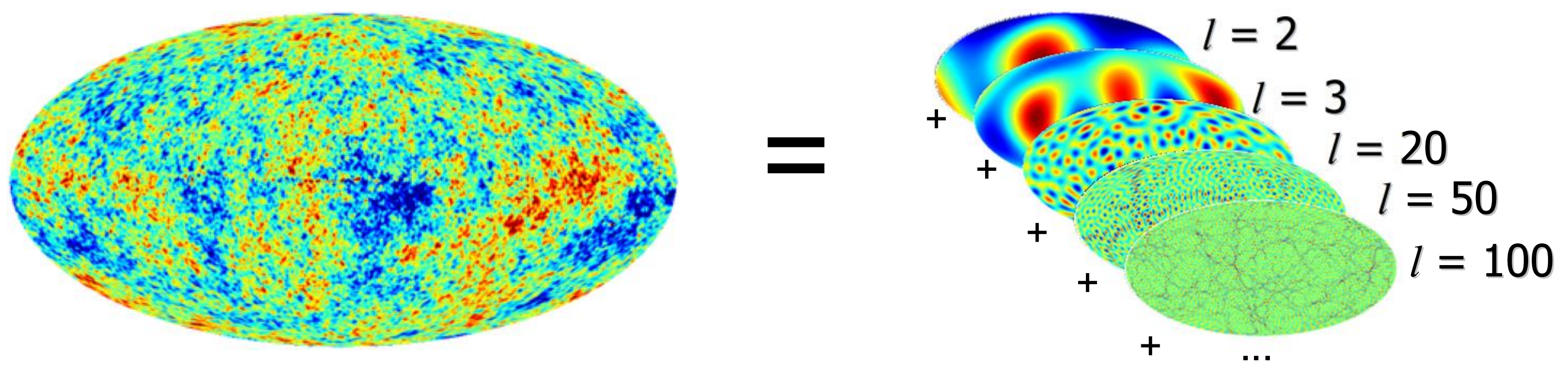


The 2-point function has been one of the most important tool for studying large-scale structure

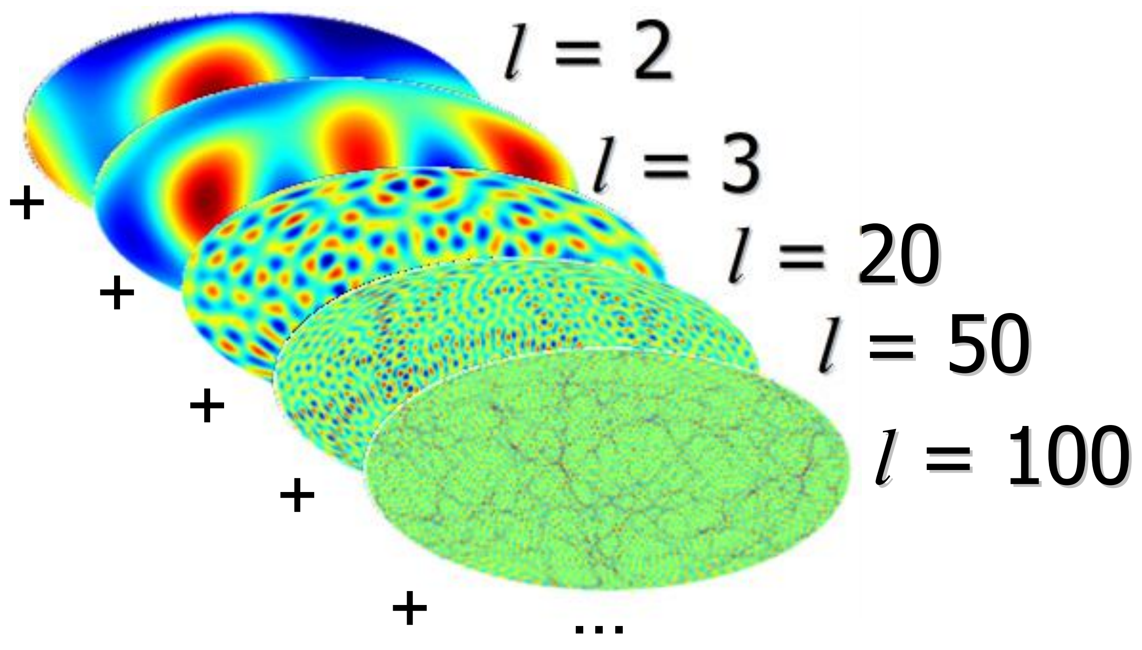
Angular Power Spectra 101

A scalar field $T(\hat{n})$ on a sphere can be expressed as a linear superposition of $Y_{\ell m}(\hat{n})$ spherical harmonics

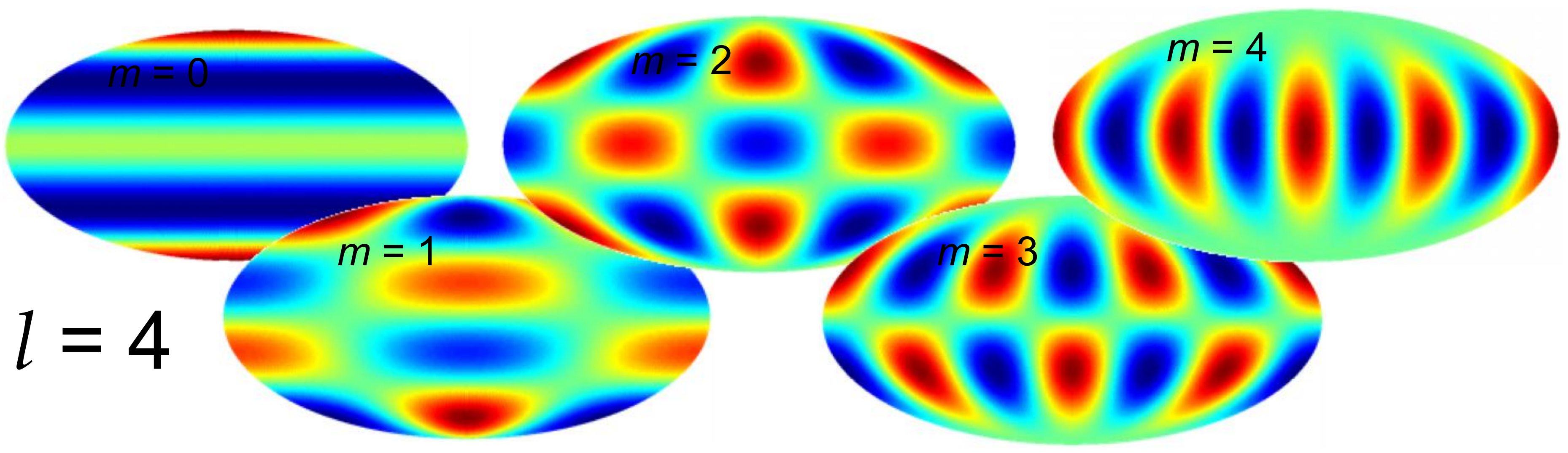
$$T(\hat{n}) = \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n})$$



ℓ encodes angular scale



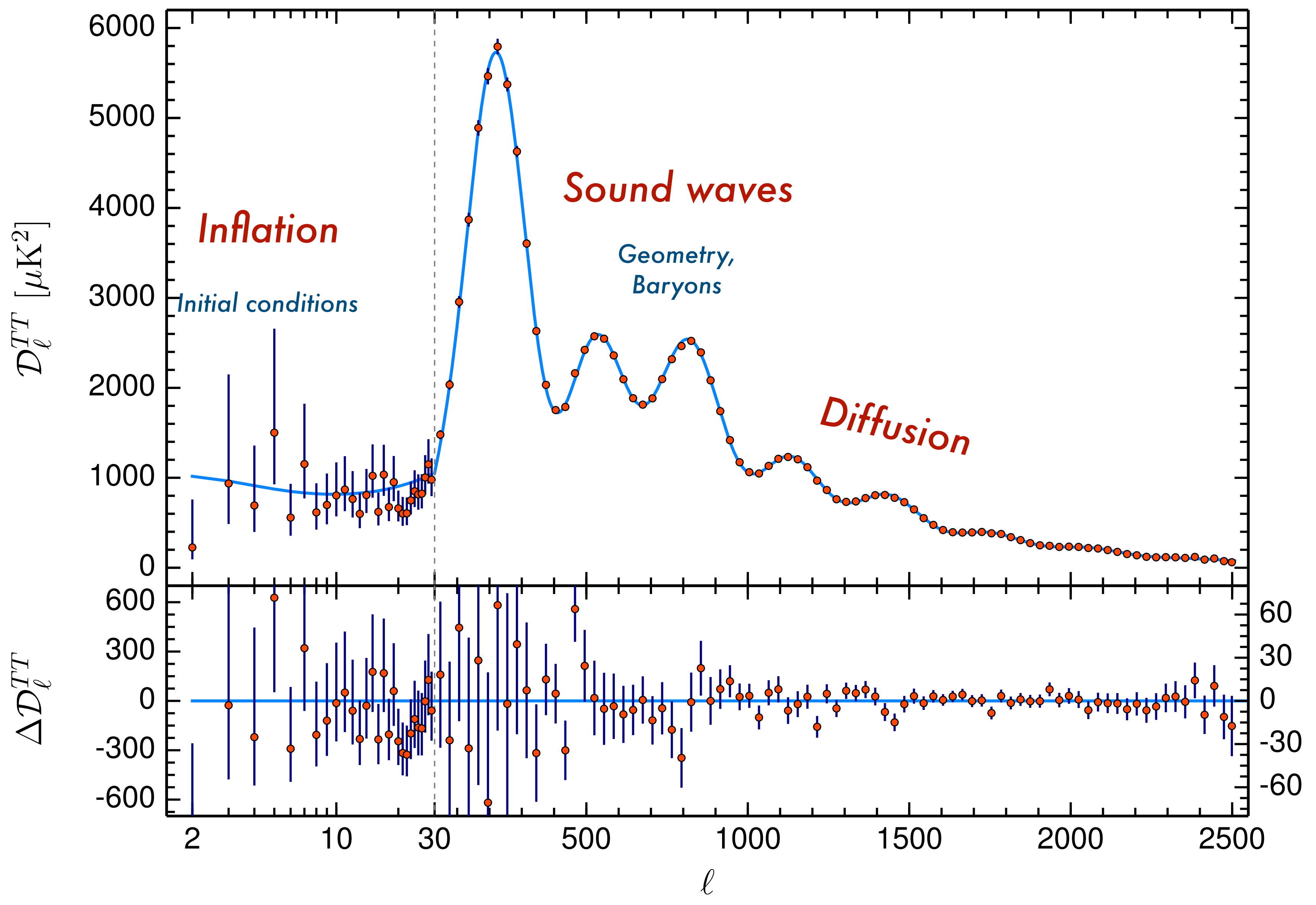
m encodes orientation



Angular Power Spectra 101

The spherical harmonic coefficient can be determined through a convolution

$$a_{\ell m} = \int_{4\pi} T(\hat{n}) Y^*_{\ell m}(\hat{n}) d\Omega \qquad C_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$



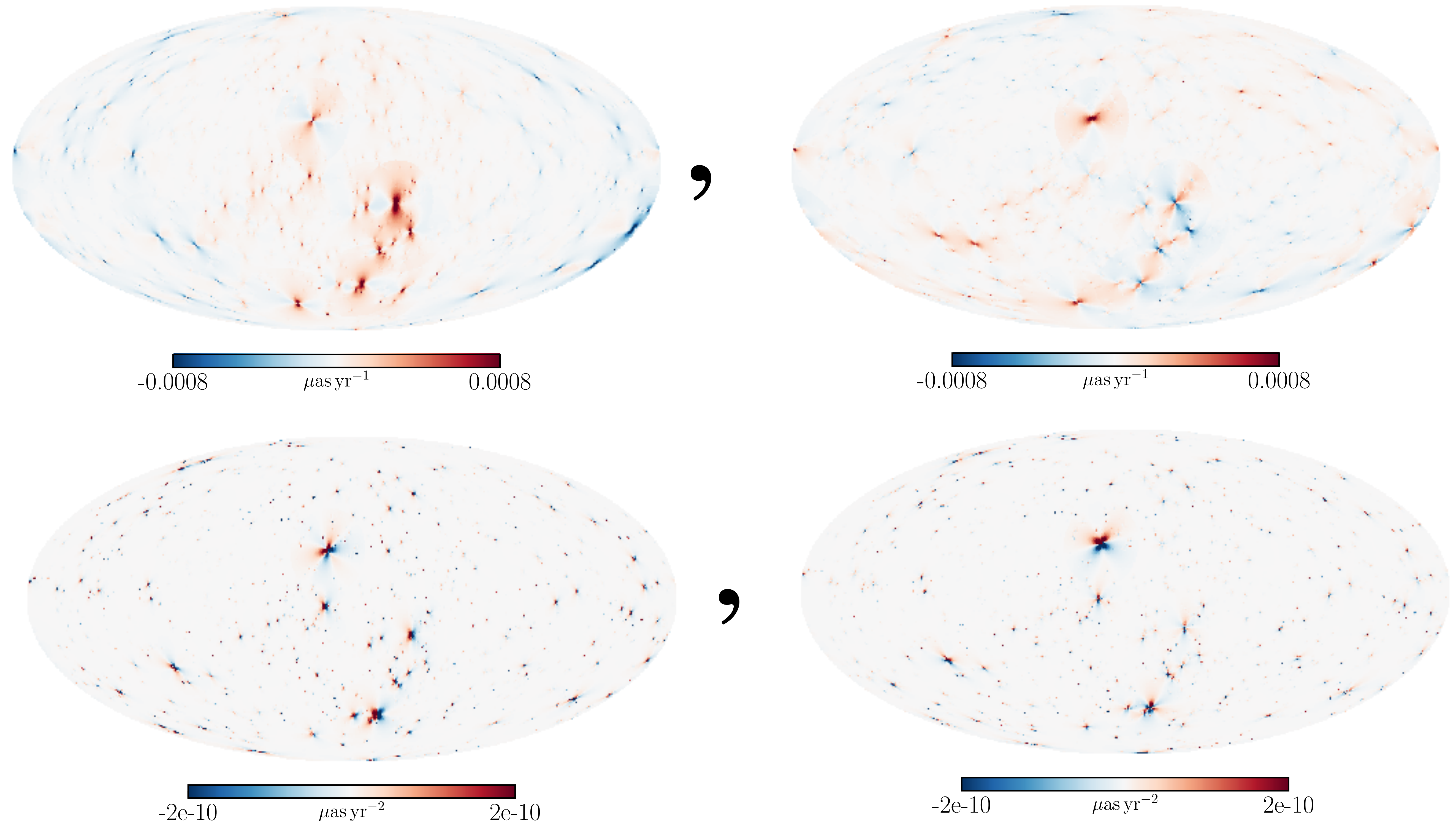
Shape of power spectrum contains a wealth of information about underlying physics

Power spectra of lens-induced motions?

Can map of induced velocities/accelerations due to Galactic dark matter subhalos be analyzed the same way?

$$\vec{\mu} = \{ \quad , \quad \}$$

$$\vec{\alpha} = \{ \quad , \quad \}$$



These are vector fields

How to measure and interpret a power spectrum?

Angular Power Spectra 201: vector fields

Any vector field $\vec{\mu}(\hat{n})$ on a sphere can be expressed as a linear superposition of vector spherical harmonics $\vec{\Psi}_{\ell m}(\hat{n})$ and $\vec{\Phi}_{\ell m}(\hat{n})$

$$\vec{\mu}(\hat{n}) = \sum_{\ell m} \underbrace{\mu^{(1)}_{\ell m}\vec{\Psi}_{\ell m}(\hat{n})}_{(\nabla \times) = 0} + \underbrace{\mu^{(2)}_{\ell m}\vec{\Phi}_{\ell m}(\hat{n})}_{(\nabla \cdot) = 0}$$

$$\vec{\Psi}_{\ell m} = \nabla Y_{\ell m}$$
$$\vec{\Phi}_{\ell m} = \hat{n} \times \nabla Y_{\ell m}$$

Physically, corresponds to decomposing vector field in a curl-free and divergence-free part (*Helmholtz-Hodge decomposition*)

The spherical harmonic coefficient can again be determined through convolutions

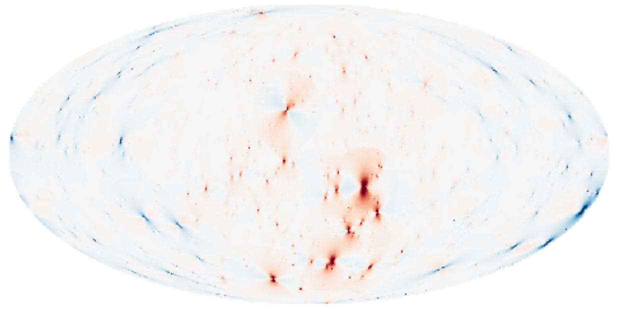
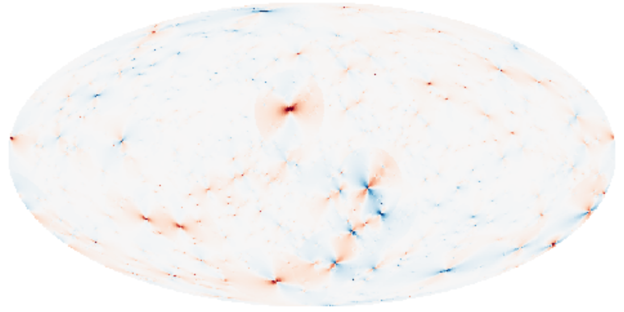
$$\mu^{(1)}_{\ell m} = \int \mathrm{d}\Omega\, \vec{\mu} \cdot \vec{\Psi}^{*}_{\ell m}; \quad \mu^{(2)}_{\ell m} = \int \mathrm{d}\Omega\, \vec{\mu} \cdot \vec{\Phi}^{*}_{\ell m}$$

Application to lens-induced motions

The lensing deflection is "sourced" from the *gradient* of the gravitational potential

$$\overrightarrow{\Delta\theta} = \frac{2}{D_l}\vec{\nabla}_\theta \int \mathrm{d}z\, \Psi_G(\vec{r})$$

Induced deflection/velocity/acceleration fields have vanishing curl

$\nabla \times \{$  ,  $\} \equiv 0$

Curl-free

$$C_\ell^{\mu(1)} \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\left|\mu_{\ell m}^{(1)}\right|^2 \simeq \sum_l \left(\frac{4G_N v}{D_l^2}\right)^2 \frac{\pi}{2}\ell(\ell+1)\left[\int_0^\infty \mathrm{d}\beta M\left(\beta D_l\right) J_1(\ell\beta)\right]^2$$

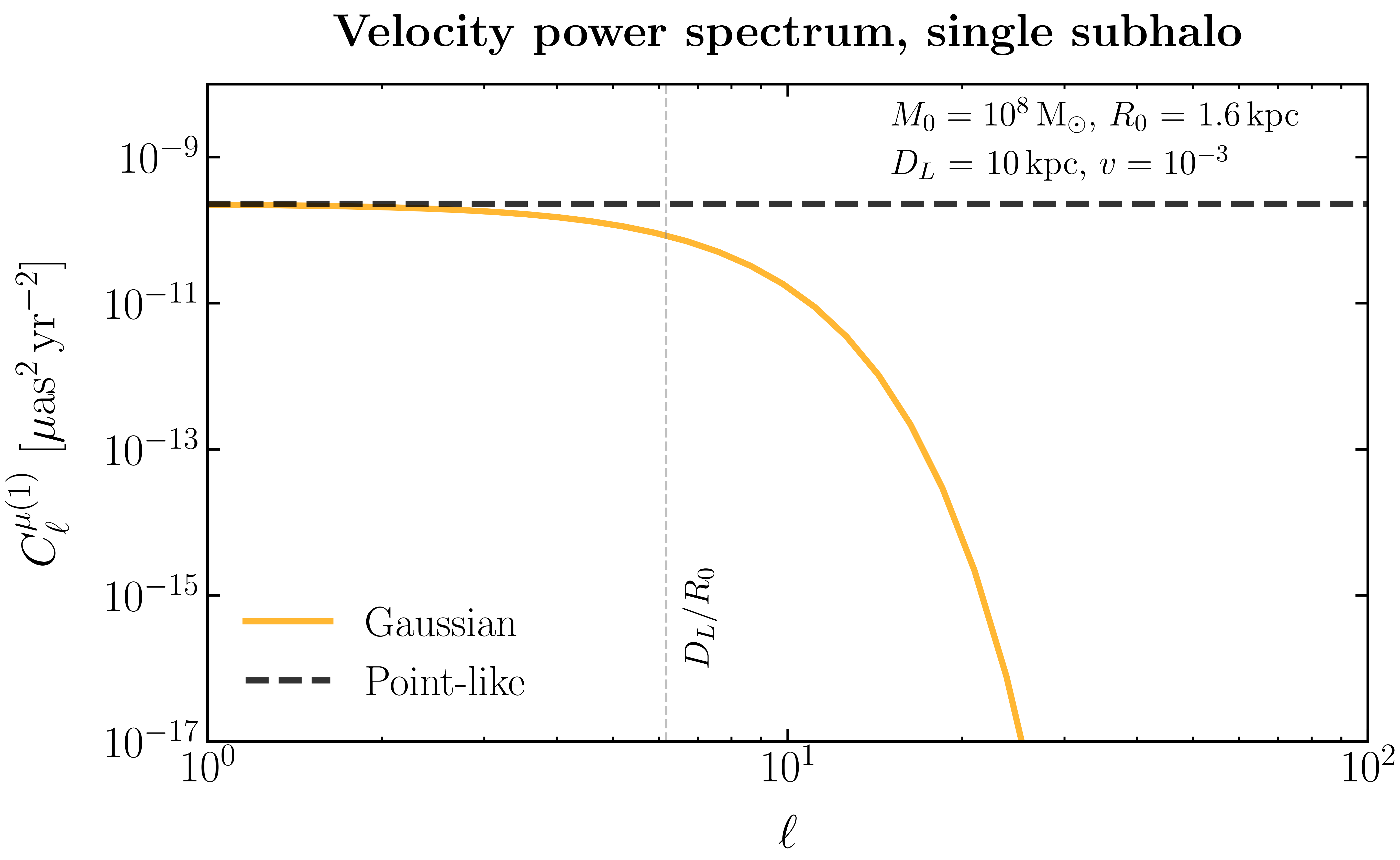
Divergence-free

$$C_\ell^{\mu(2)} \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\left|\mu_{\ell m}^{(2)}\right|^2 = 0$$

All vector lensing observables have only curl-free modes in harmonic decomposition

The lensing signal: extended lenses

$$\rho(r) = \frac{M_0}{2\sqrt{2}\pi^{3/2} R_0^3} e^{-r^2/2R_0^2}$$



$$C_\ell^{\mu(1)} \simeq \left(\frac{4G_N M_0 v}{D_l^2}\right)^2 \frac{\pi}{2}\, e^{-\ell^2 R_0^2/D_l^2}$$

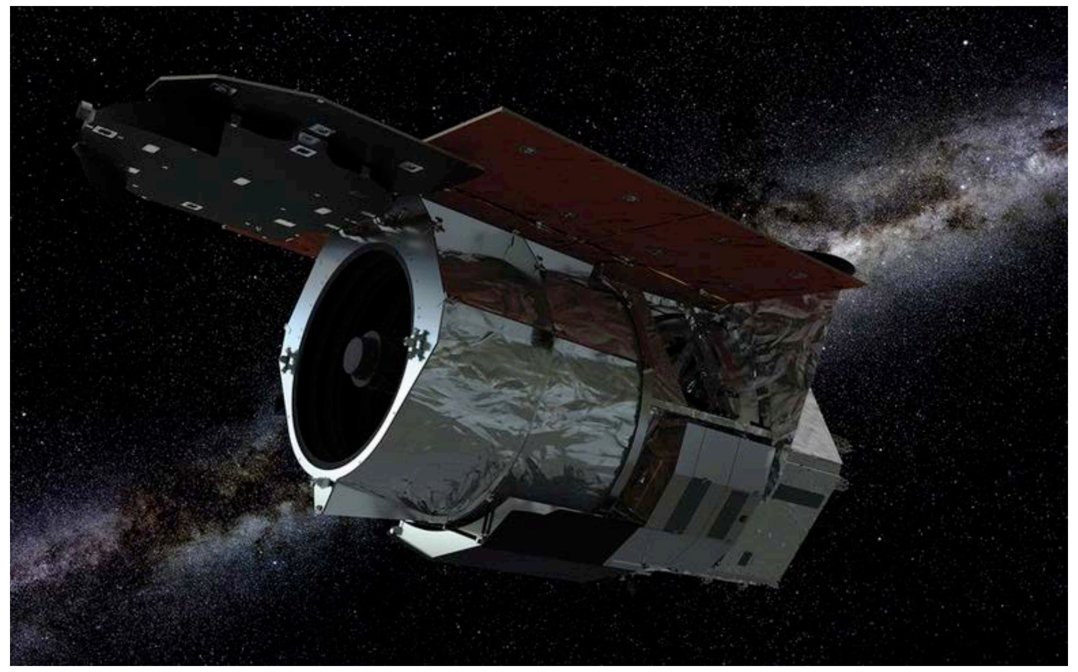
Suppression at smaller scales

Astrometric precision

Space-based, optical telescopes

Current: *Gaia, HST*

Future: *Theia*, WFIRST



Noise configuration

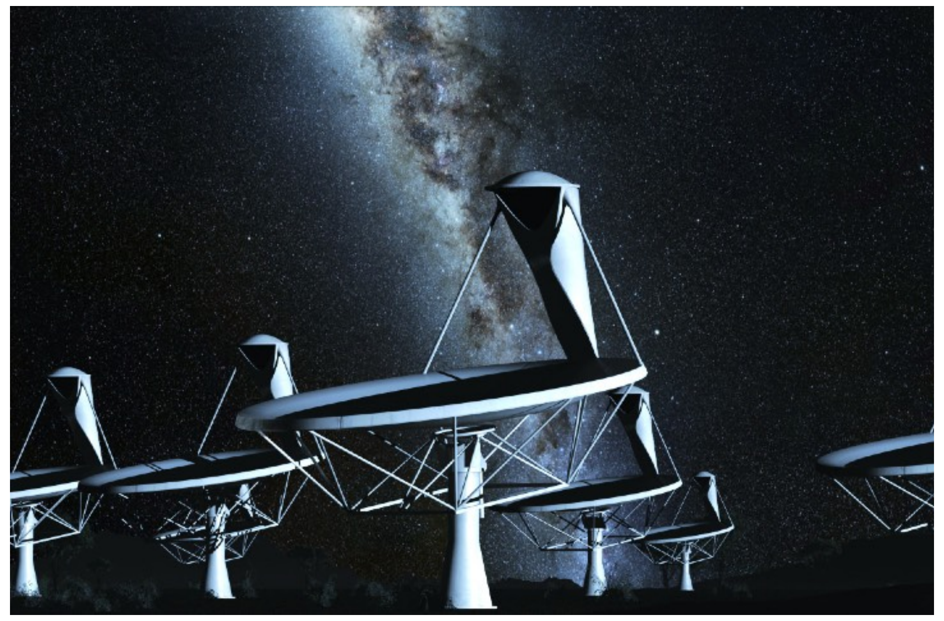
$$\sigma_\alpha = 0.1\,\mu\mathrm{as\,yr}^{-2}$$

$$N_q = 10^{11}$$

Ground-based, radio interferometry

Current: VLA (Very Large Array)

Future: SKA (Square Kilometer Array)



Noise configuration

$$\sigma_\mu = 1\,\mu\mathrm{as\,yr}^{-1}$$

$$N_q = 10^{8}$$

Cold dark matter

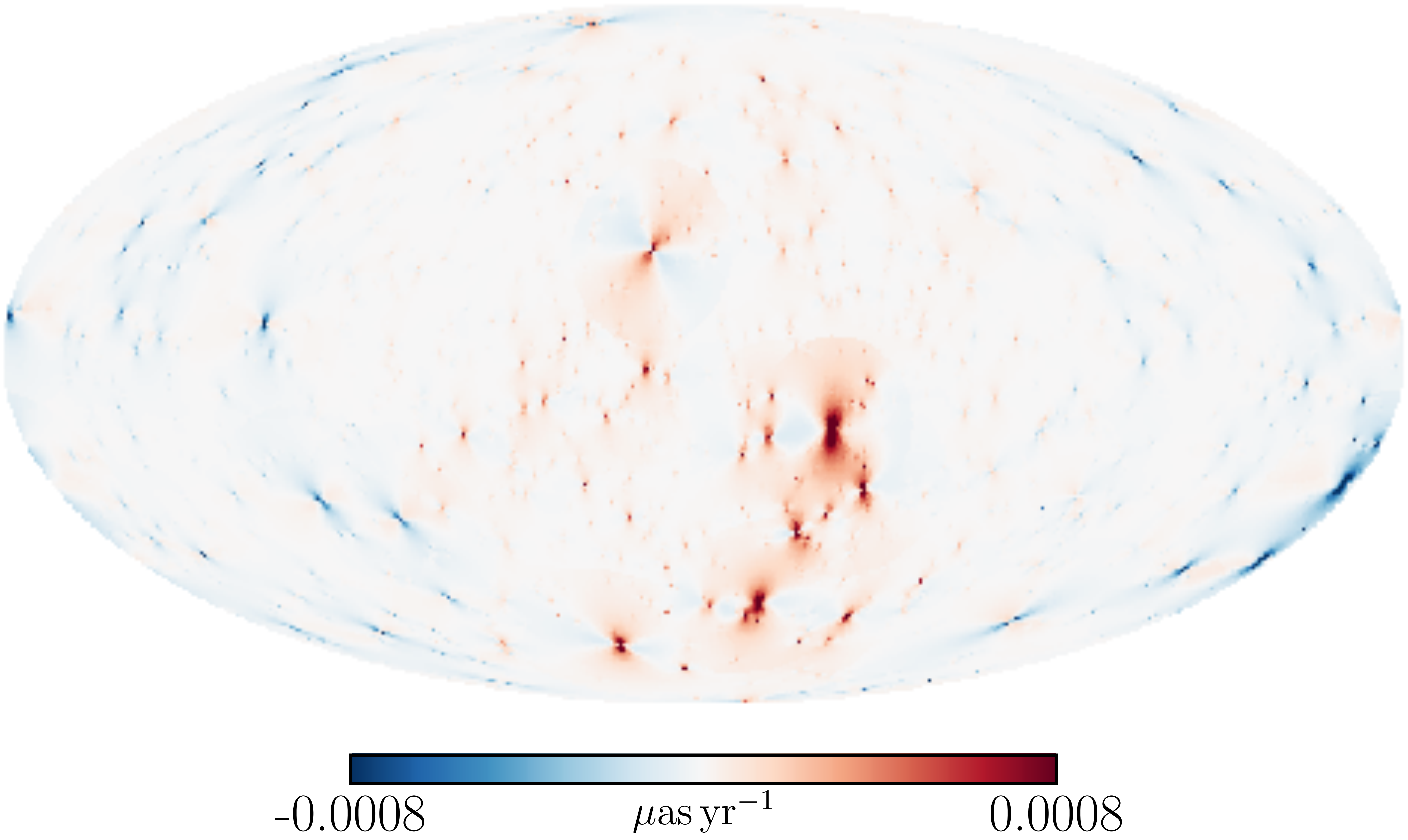
ΛCDM predicts a *broad, scale invariant* spectrum of subhalos distributed in the Milky Way

Credit: T. Brown and J.Tumlinson

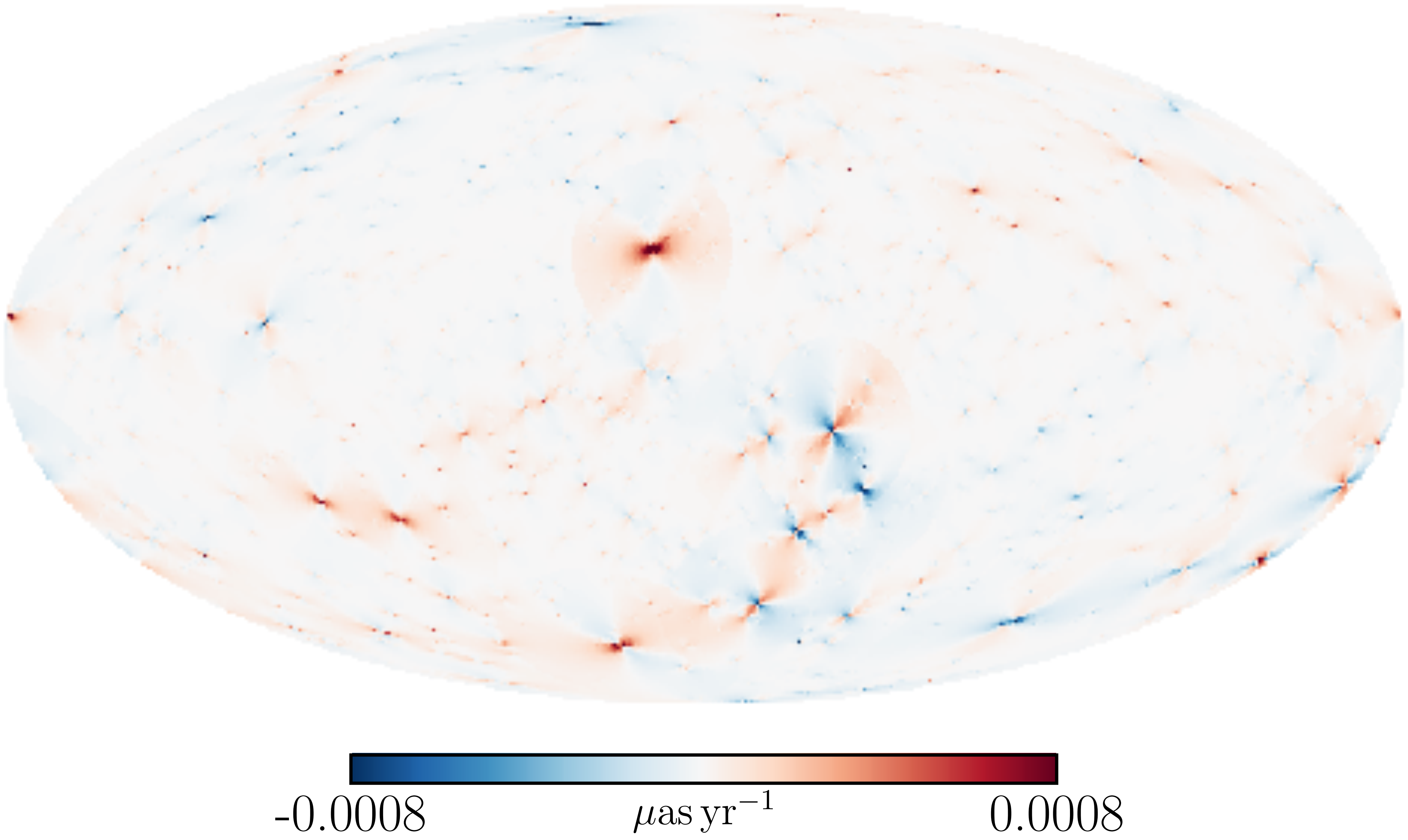
Cold dark matter

ΛCDM predicts a *broad, scale invariant* spectrum of subhalos distributed in the Milky Way

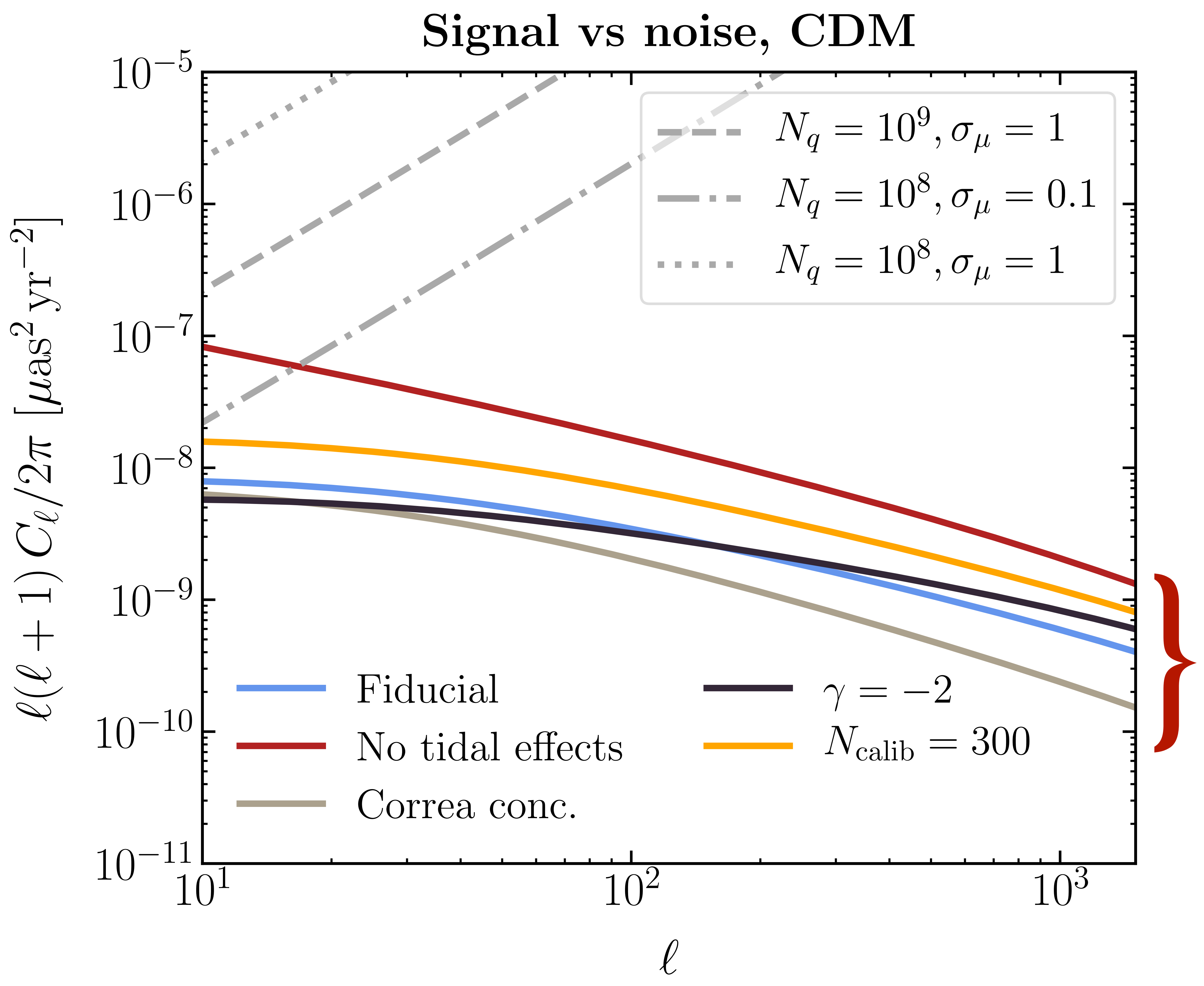
Induced longitudinal proper motions, μ_l



Induced latitudinal proper motions, μ_b



Cold dark matter: total signal and noise

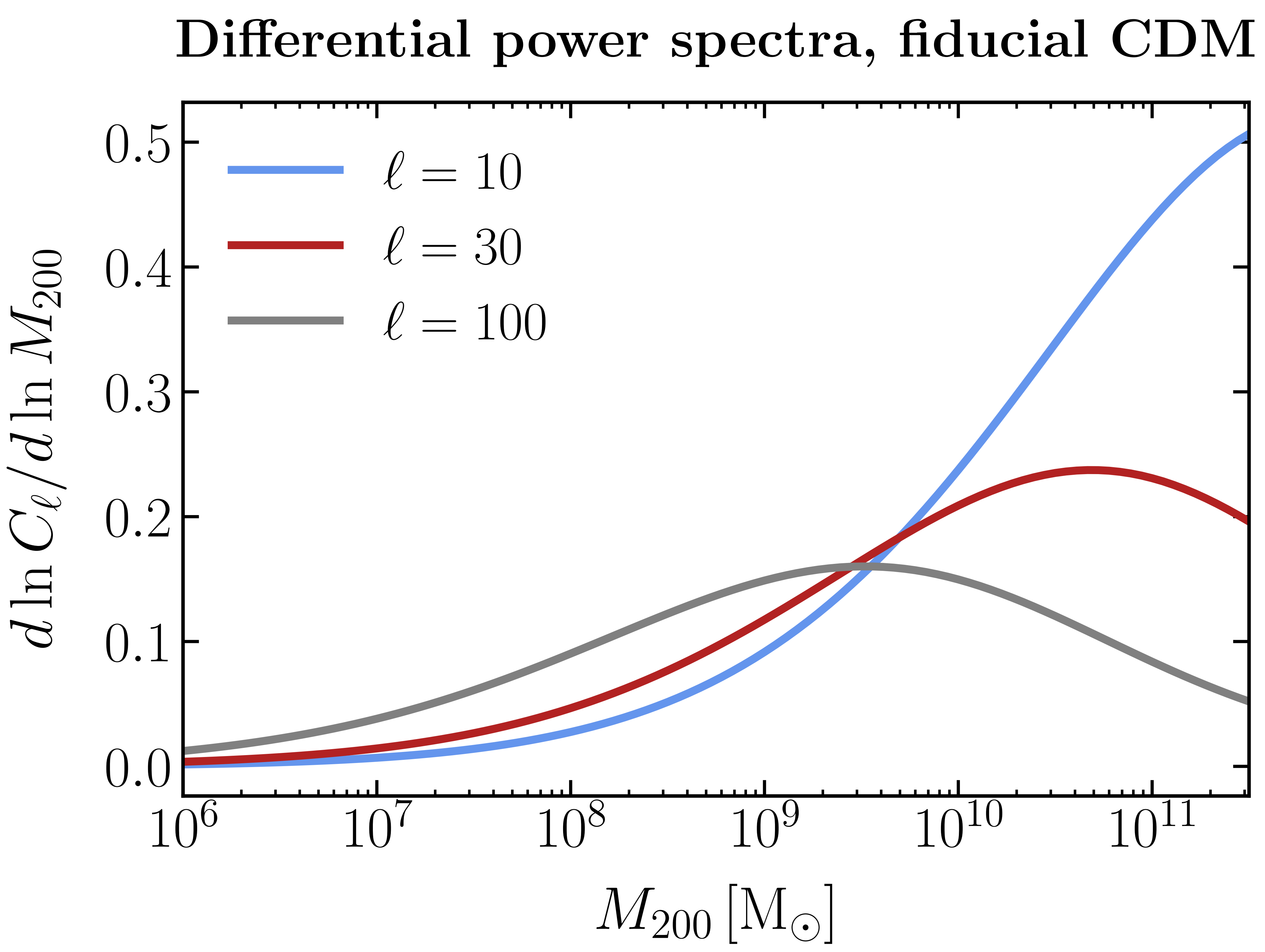


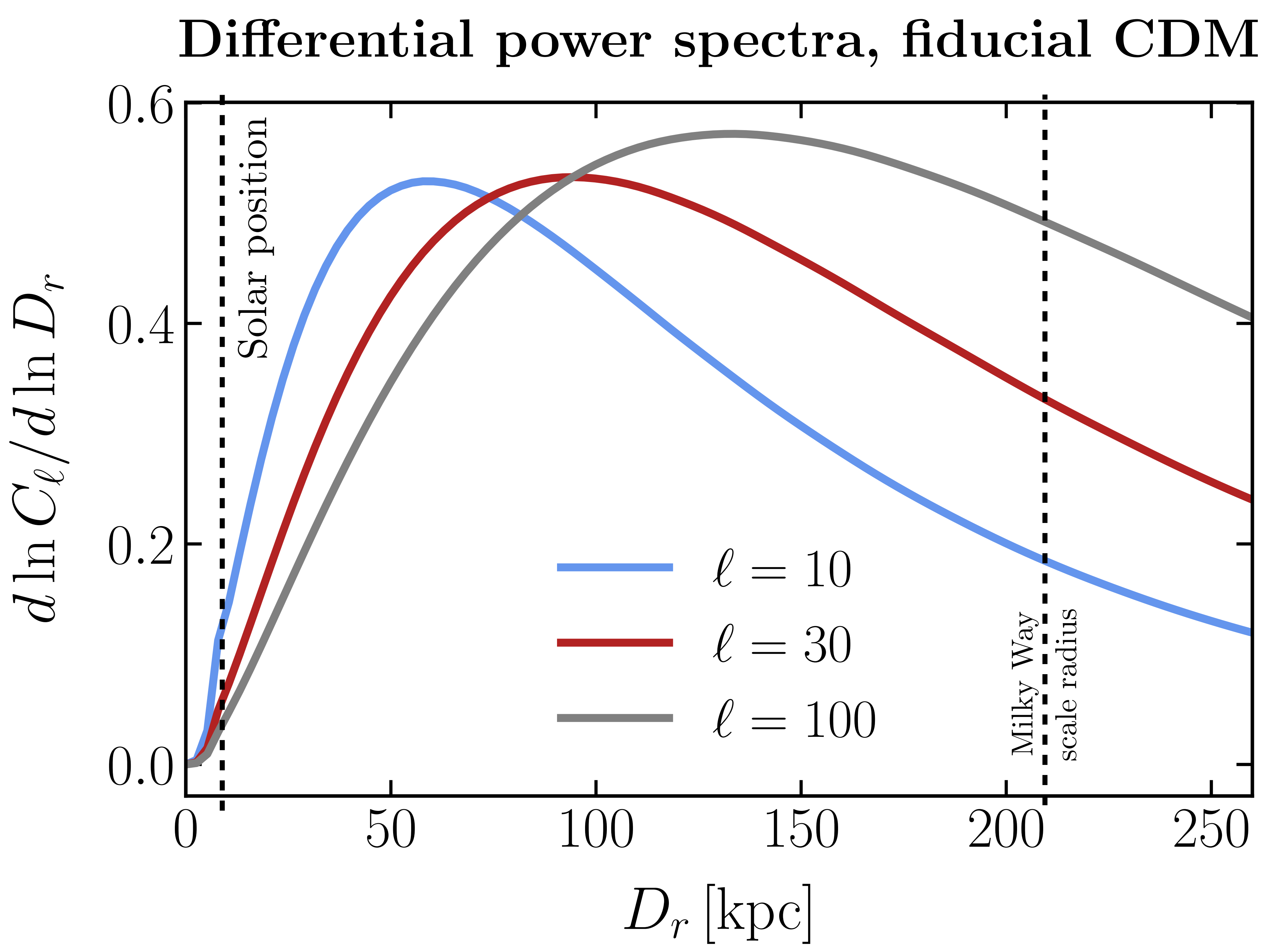
Brackets uncertainty on CDM halo properties

Get total expected signal as *convolution over subhalo distribution properties*

$$C_\ell^{\rm tot} = \int_{M,r,v} d^3v\, d^3r\, dM\, f_\oplus(\mathbf{v},t)\frac{dN}{d\mathbf{r}}\frac{dN}{dM}C_\ell\left(M,\mathbf{v},D_l(\mathbf{r}),\mathbf{r}\right)$$

Cold dark matter: mass and location in Galaxy

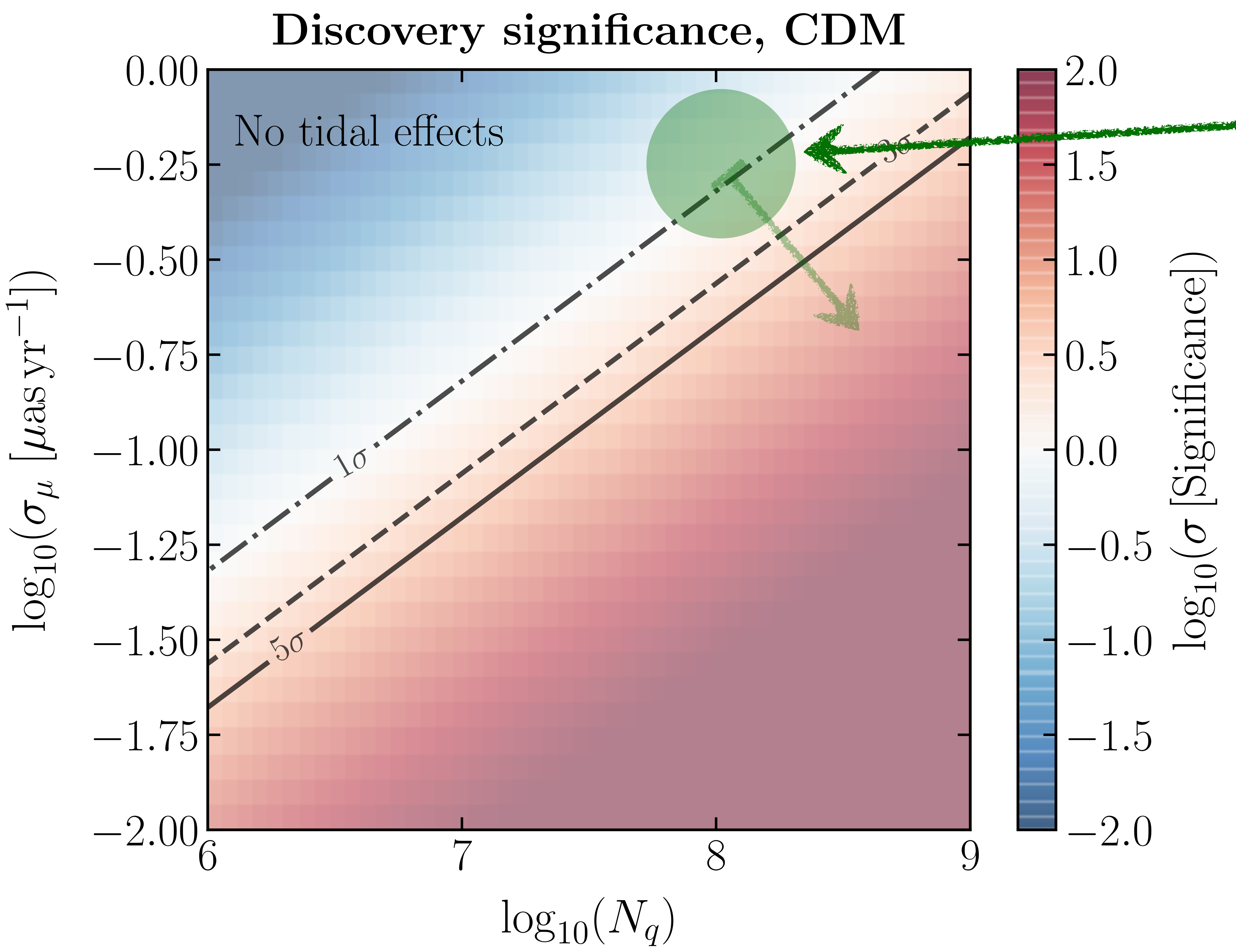




Most of the sensitivity comes from more massive halos

Sensitive to subhalo population in the bulk Milky Way halo

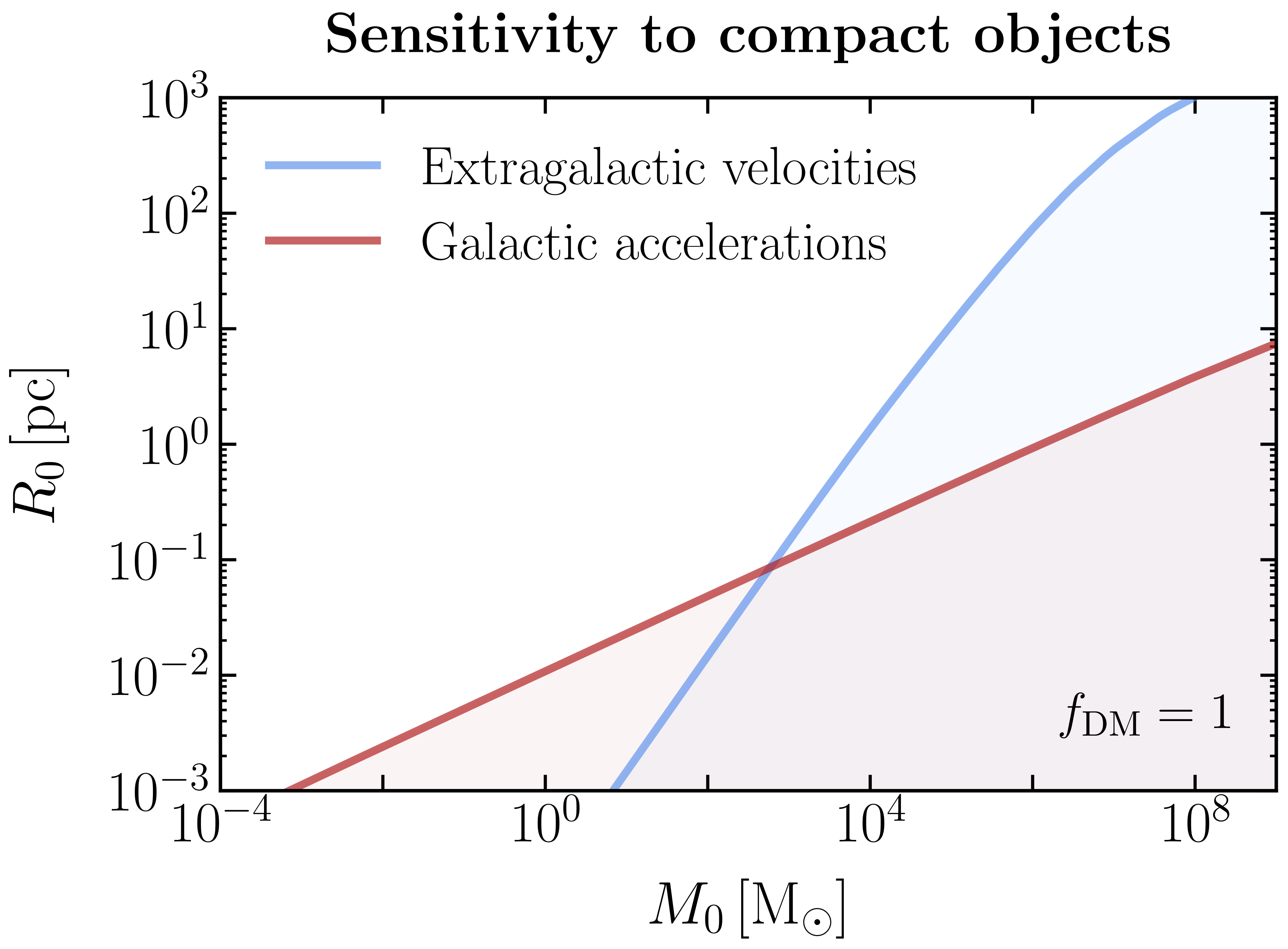
Cold dark matter: discovery potential

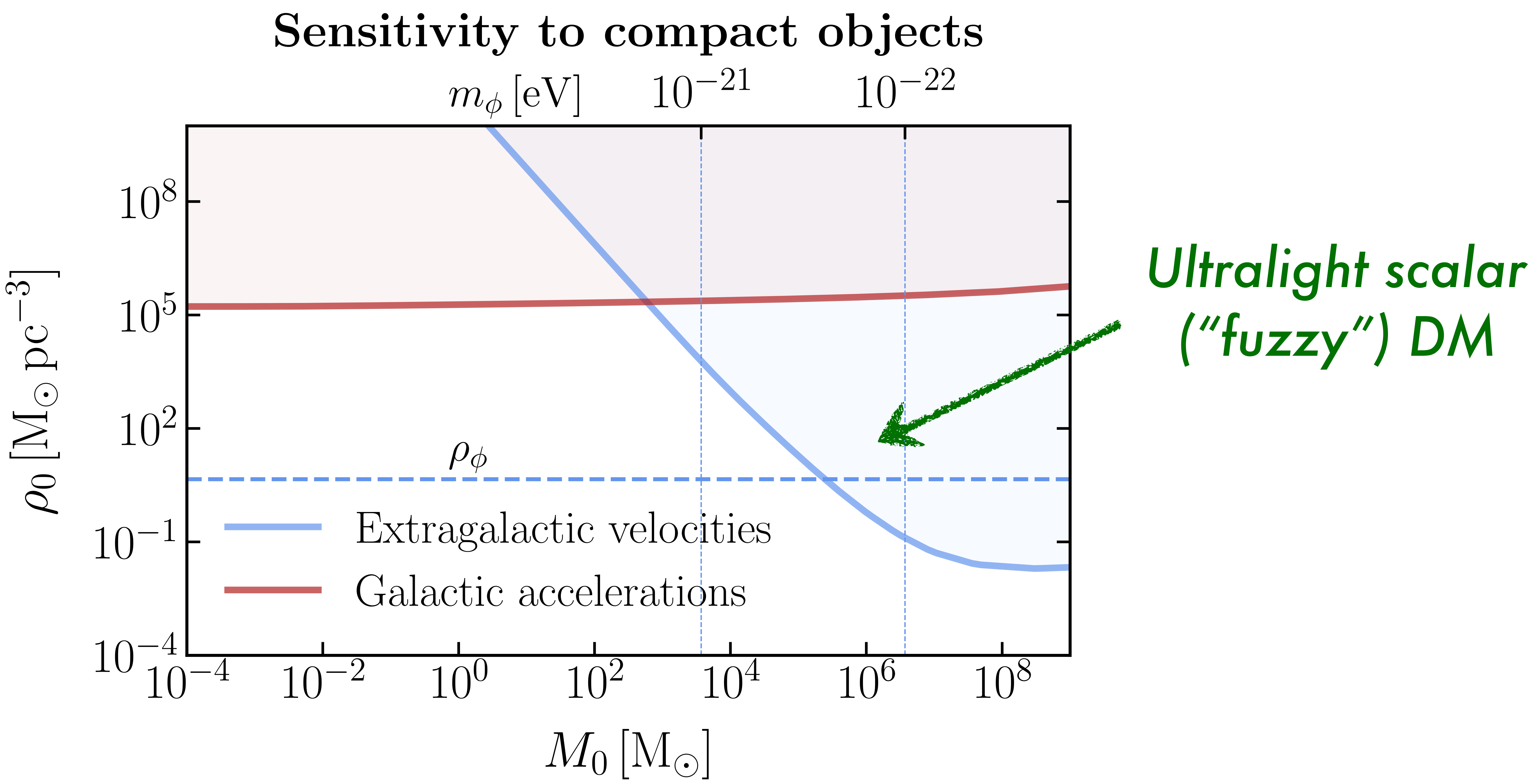


Anticipated noise levels with future surveys

A CDM subhalo population can be detected!

Compact objects in the Milky Way: sensitivity projections





Relatively extended subhalos can be detected, unlike conventional searches

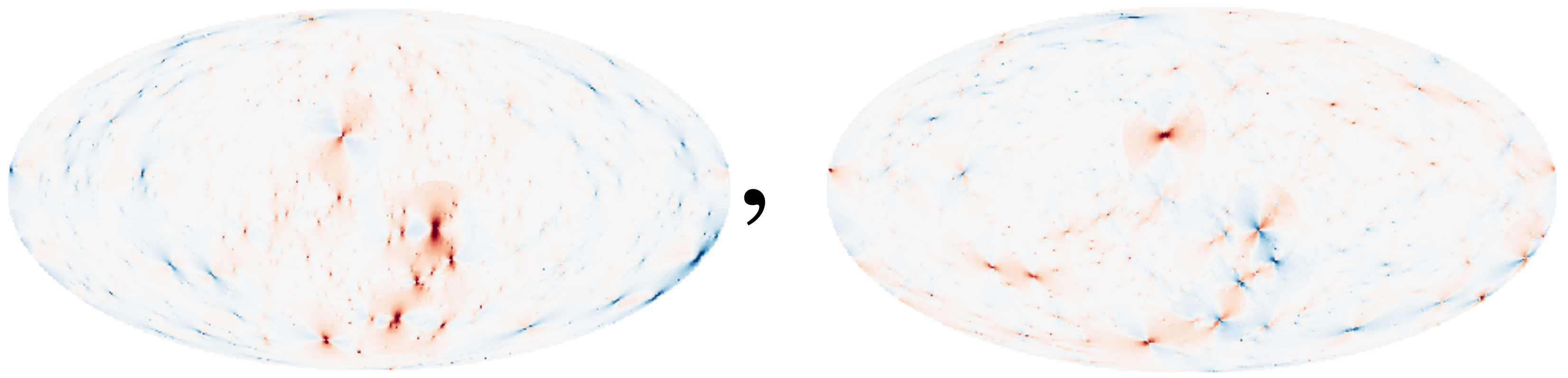
Discovery handles

Can leverage several features of the lensing signal to ensure discovery against systematic/instrumental noise

Curl of lensing signal vanishes

$$C_\ell^{\mu(1)} \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\left|\mu_{\ell m}^{(1)}\right|^2$$ = *Signal + Noise*

$$C_\ell^{\mu(2)} \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\left|\mu_{\ell m}^{(2)}\right|^2$$ = *Noise*

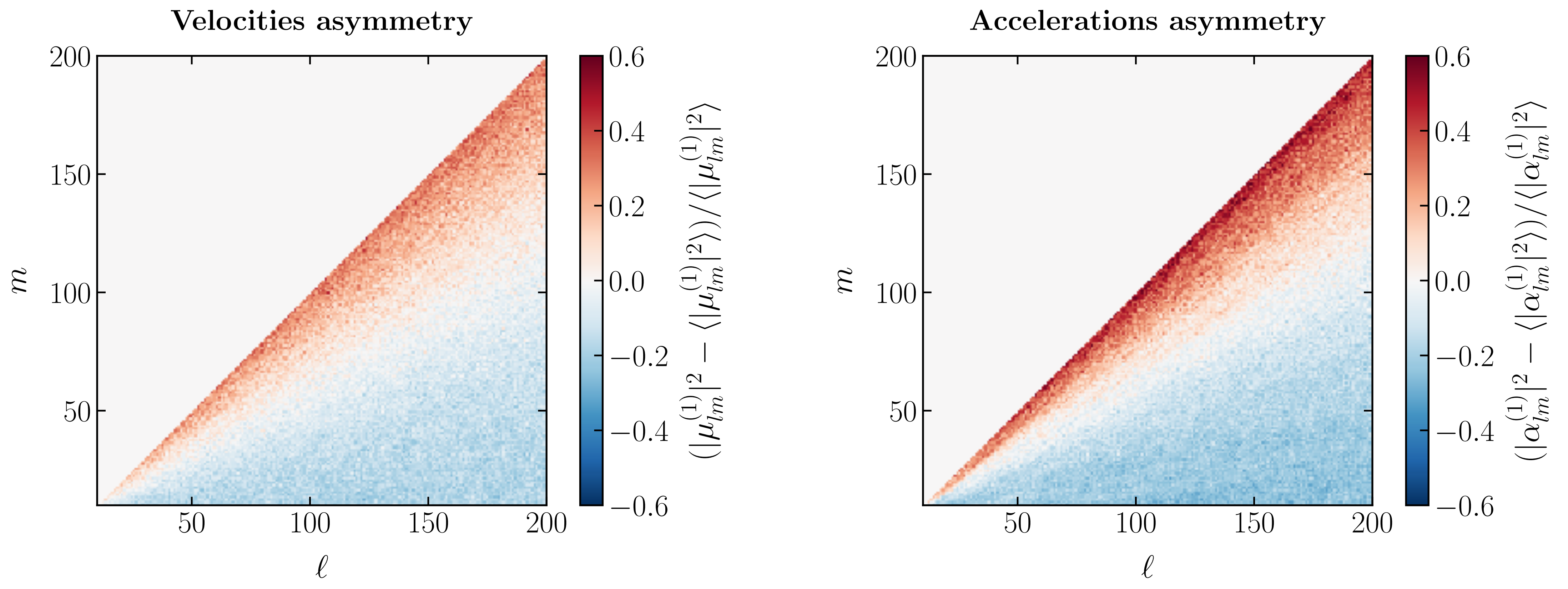
$\nabla \times \{$  $\} \equiv 0$

Use divergence-free modes as control region

Preferred velocity due to Sun's motion leads to azimuthal asymmetry

$$\mu_{\ell m}^{(1)} = -\frac{\ell(\ell+1)}{D_l}\int \mathrm{d}\Omega\, \psi(\boldsymbol{\beta})\mathbf{v}\cdot\mathbf{\Psi}_{\ell m}^*$$

Asymmetric

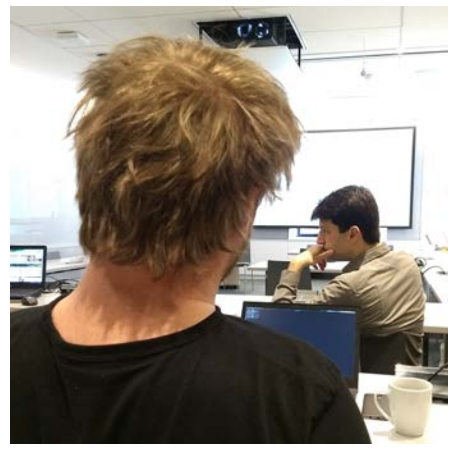


Use systematic asymmetry in m-modes as discriminant

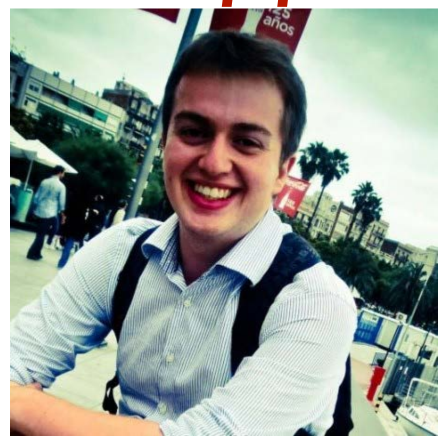
Outline

Brehmer

Hermans



Louppe



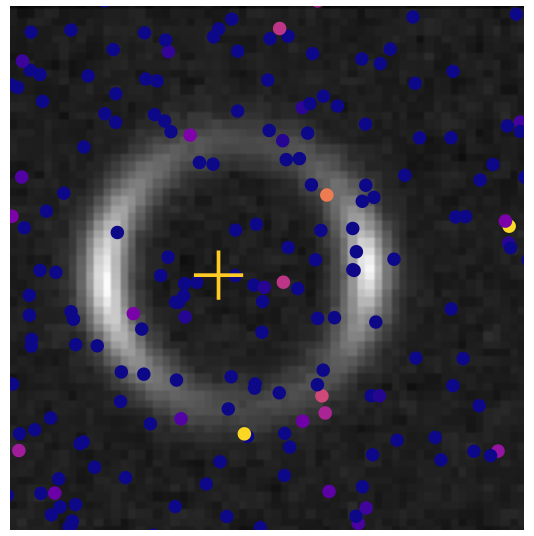
Cranmer

Gravitational Lensing

A Brief Primer

Inferring Galactic Substructure

With Astrometry & Weak Lensing

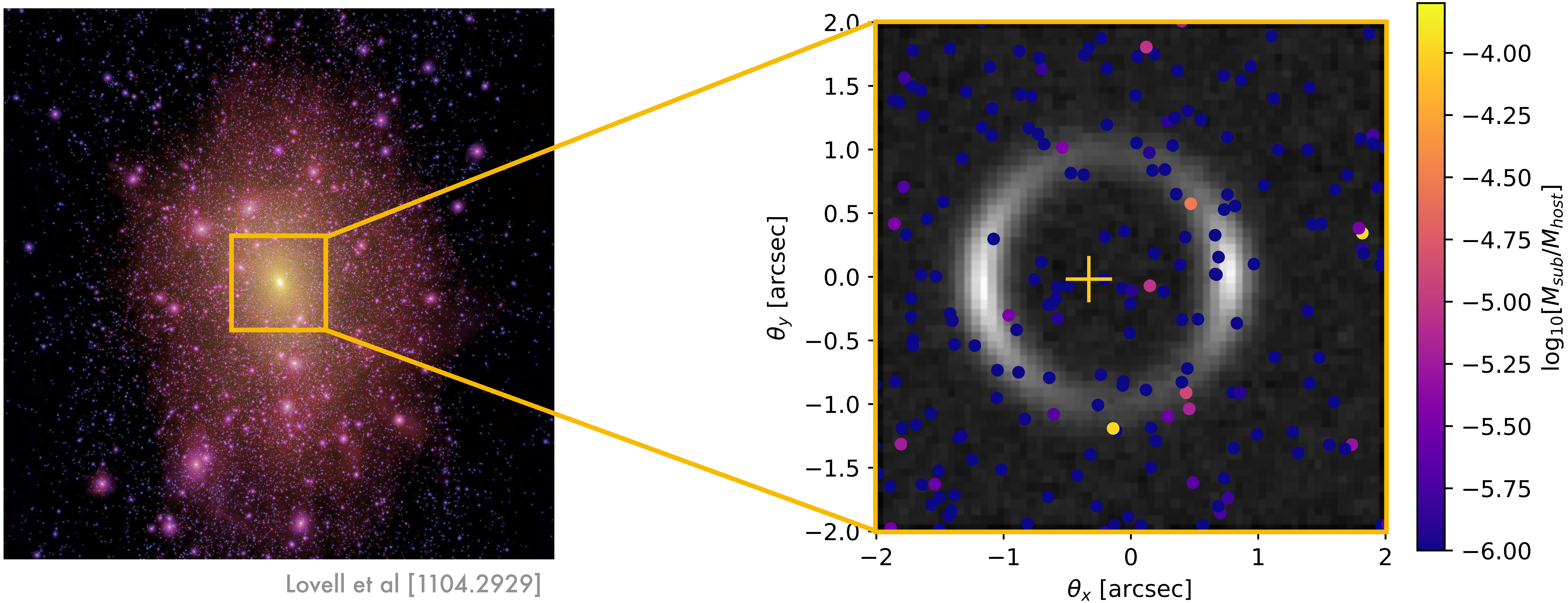


Inferring Extragalactic Substructure

With Likelihood-free Inference & Strong Lensing

Strong lensing: effect of substructure

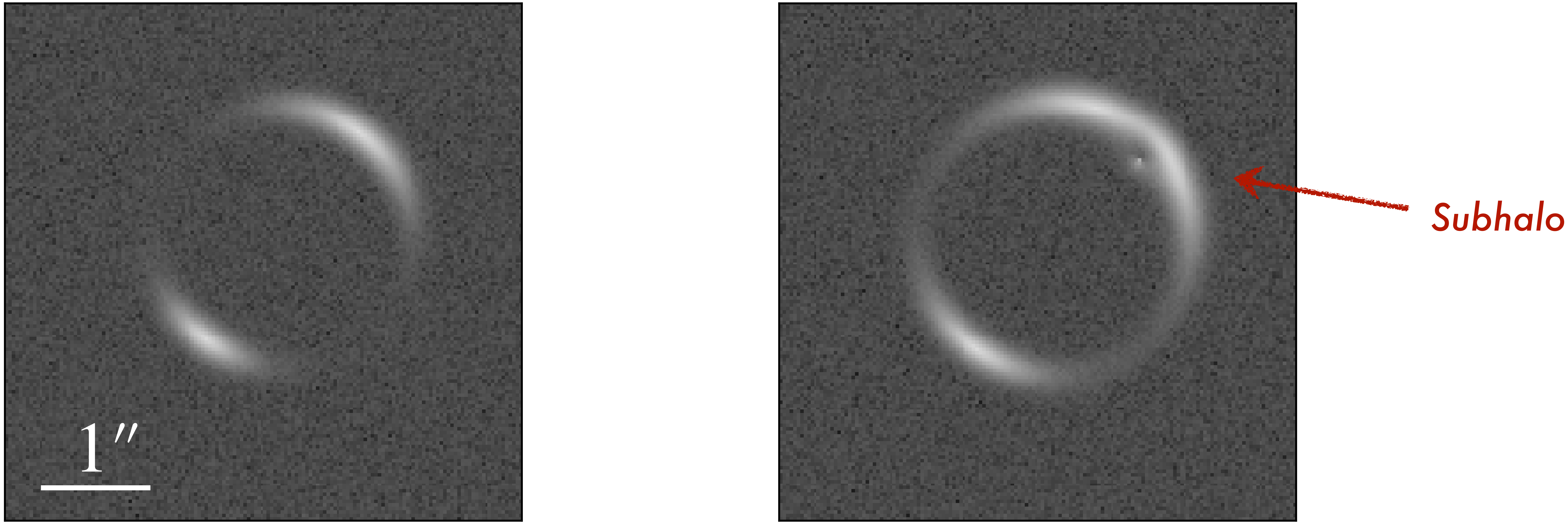
Substructure causes percent-level shifts in strongly lensed image



Lovell et al [1104.2929]

Strong lensing: effect of substructure

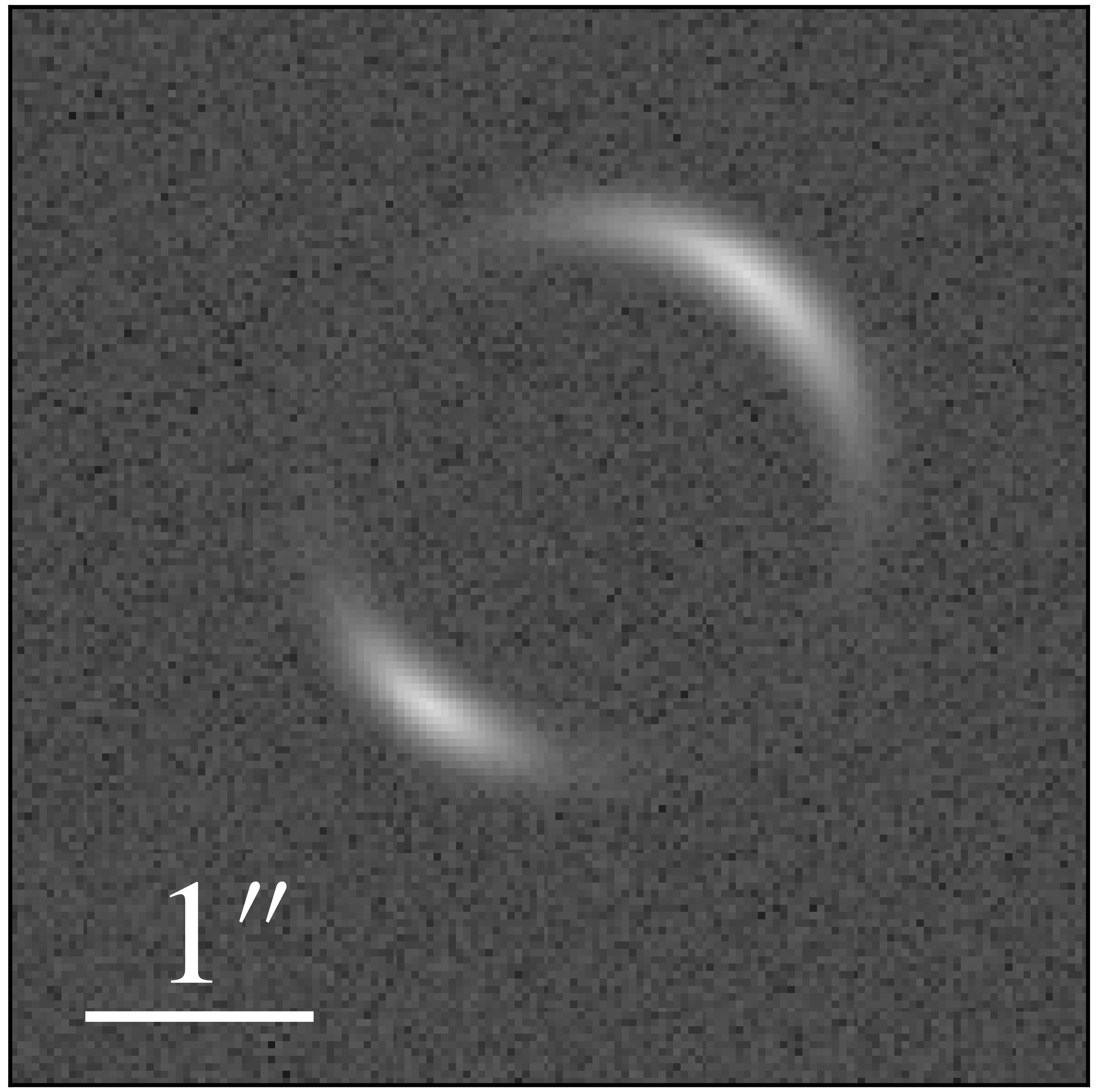
Smooth halo only | Smooth halo + subhalo



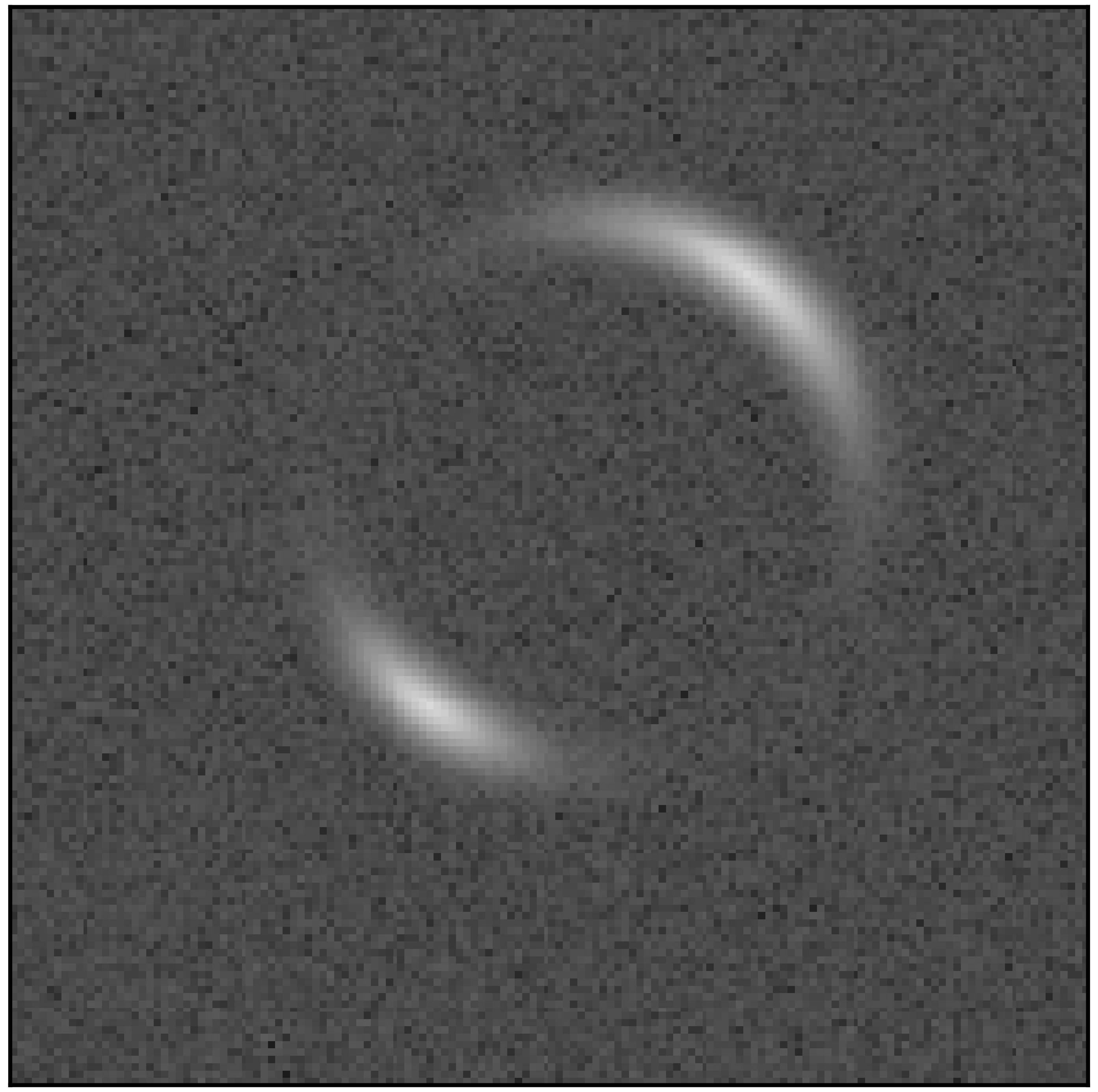
Substructure perturbs lensing rings compared to only smooth halo

Strong lensing: effect of substructure... in reality

Smooth halo only



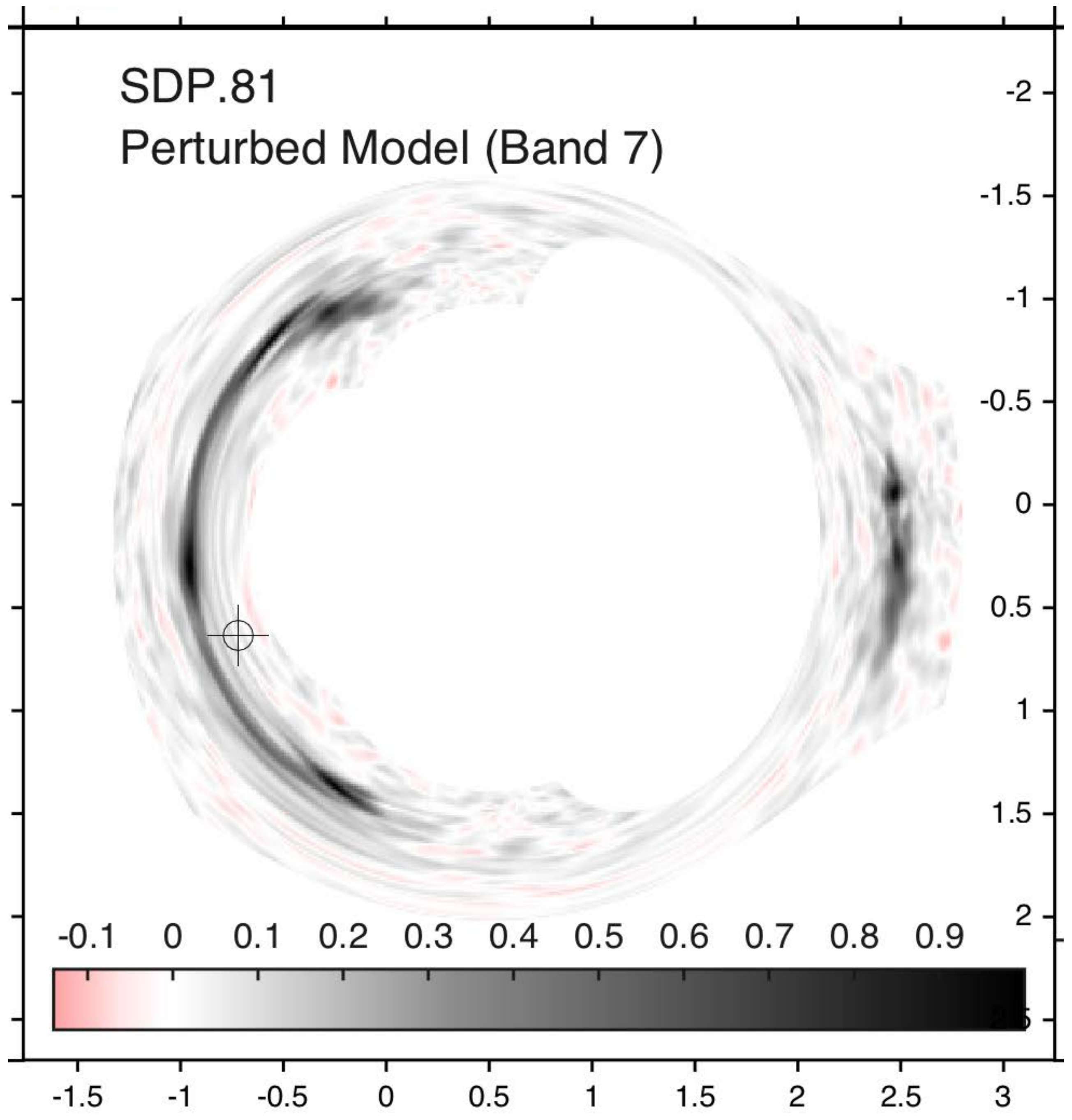
Smooth halo + subhalos

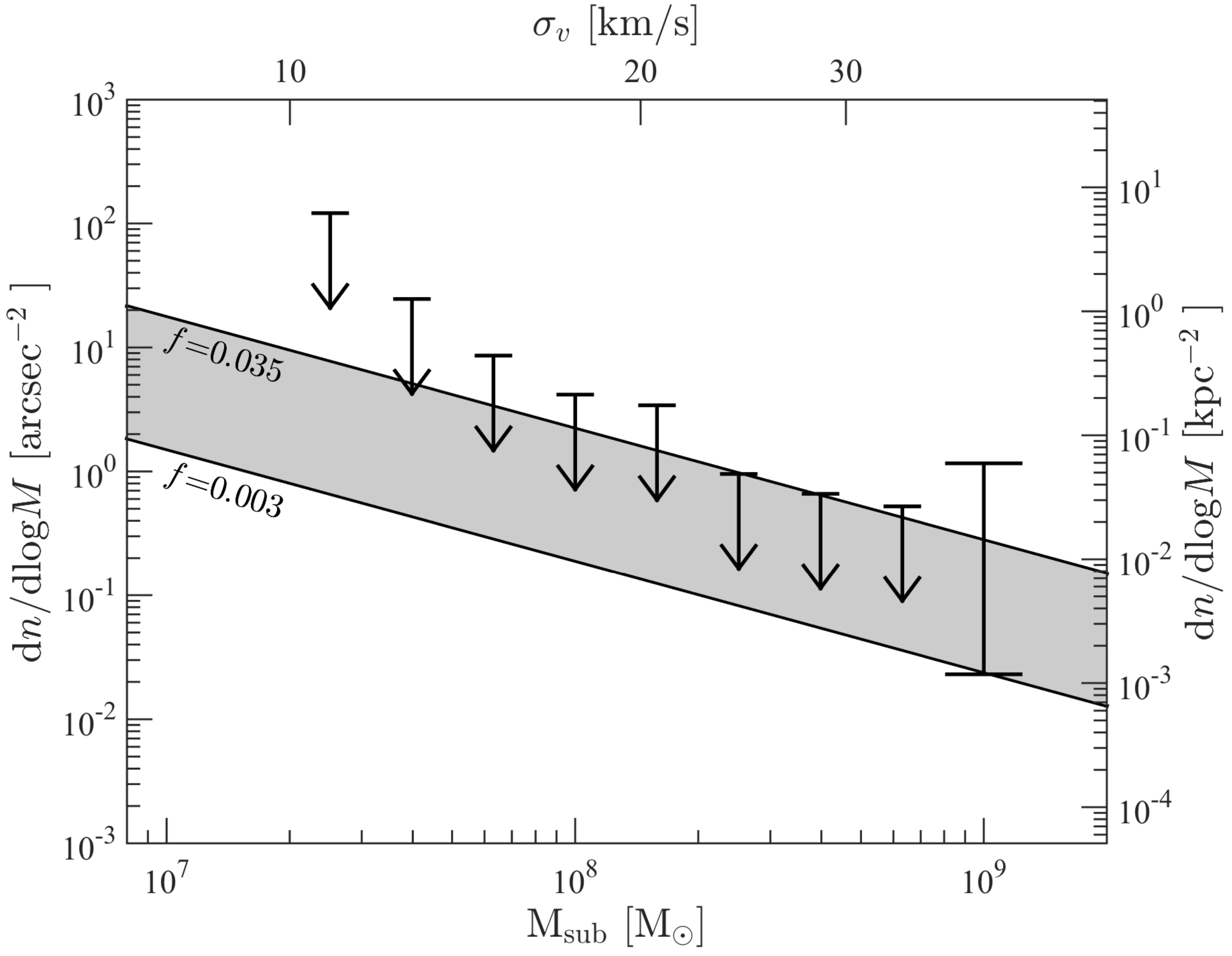


Effect is very subtle for realistic dark matter substructure

Strong lensing: conventional substructure searches

Constraints on subhalo mass function from detections of individual subhalos





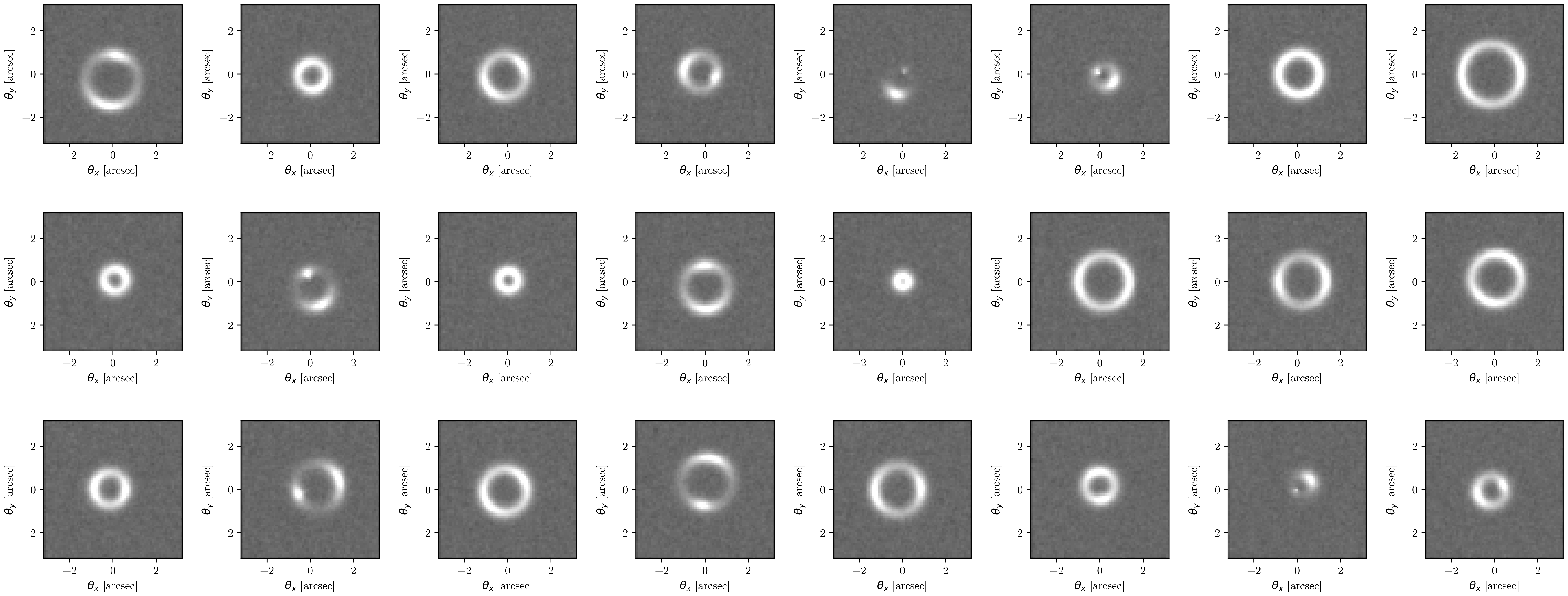
Sensitive to individual, massive subhalos

Hezaveh et al [1601.01388]

Goal

Future surveys like LSST, *Euclid* expected to deliver large samples of galaxy-galaxy strong lenses

$\mathcal{O}(10{,}000)$ Collett et al [1507.02657]

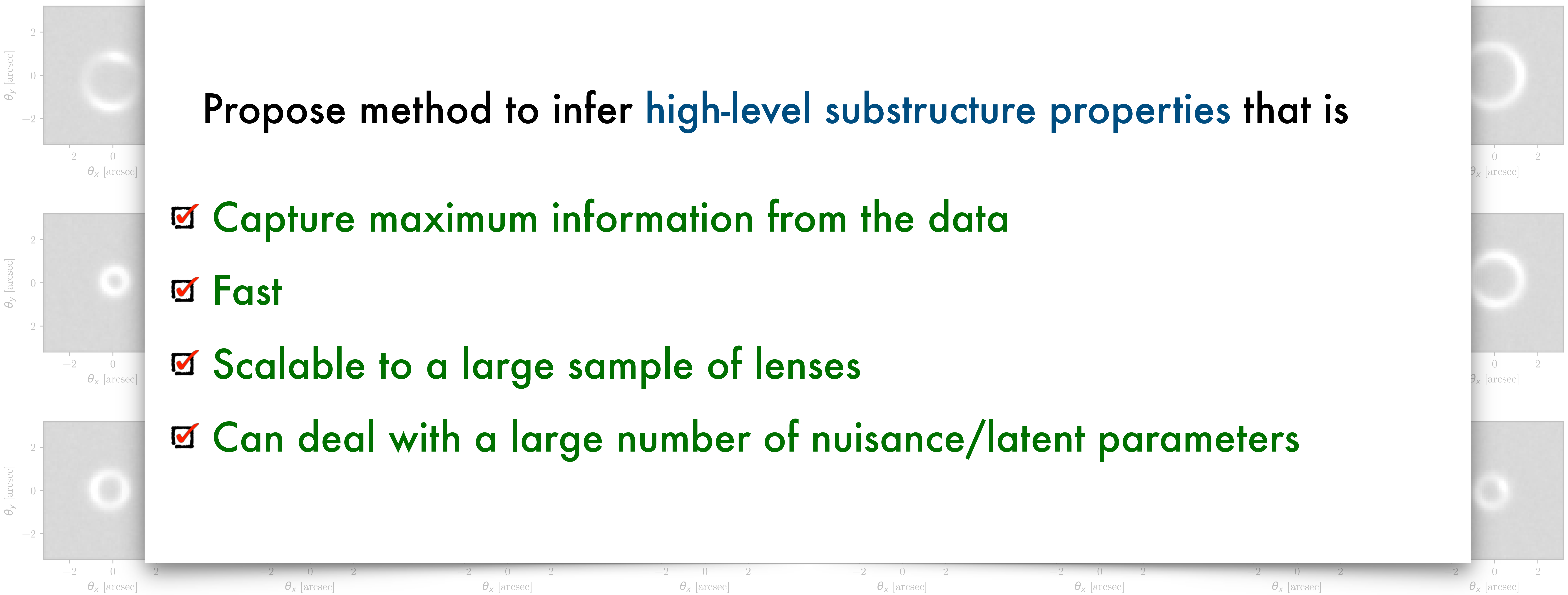


Goal

Future surveys like LSST, *Euclid* expected to deliver large samples of galaxy-galaxy strong lenses

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

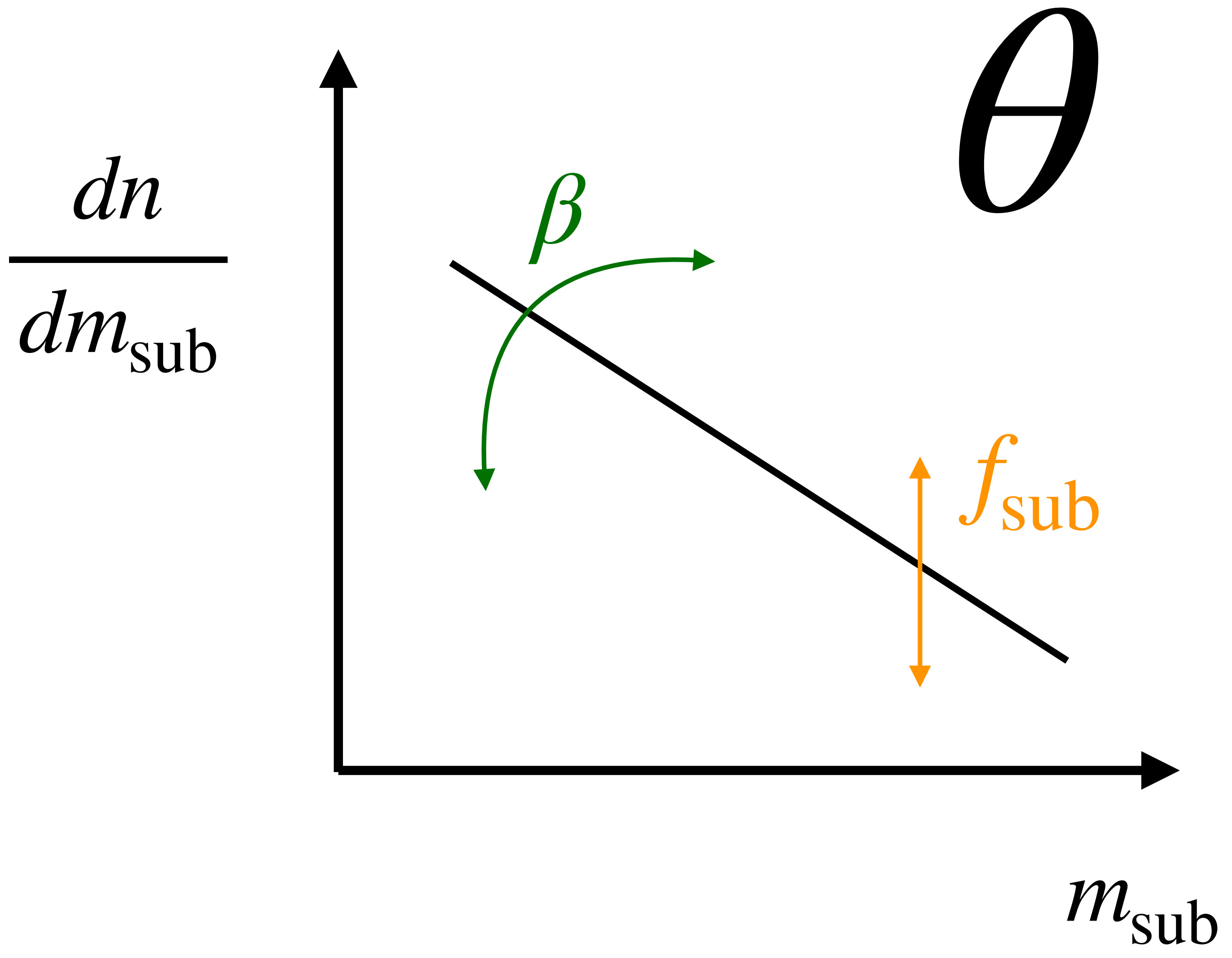
Collett et al [1507.02657]



Propose method to infer high-level substructure properties that is

- ☑ Capture maximum information from the data
- ☑ Fast
- ☑ Scalable to a large sample of lenses
- ☑ Can deal with a large number of nuisance/latent parameters

Substructure likelihood

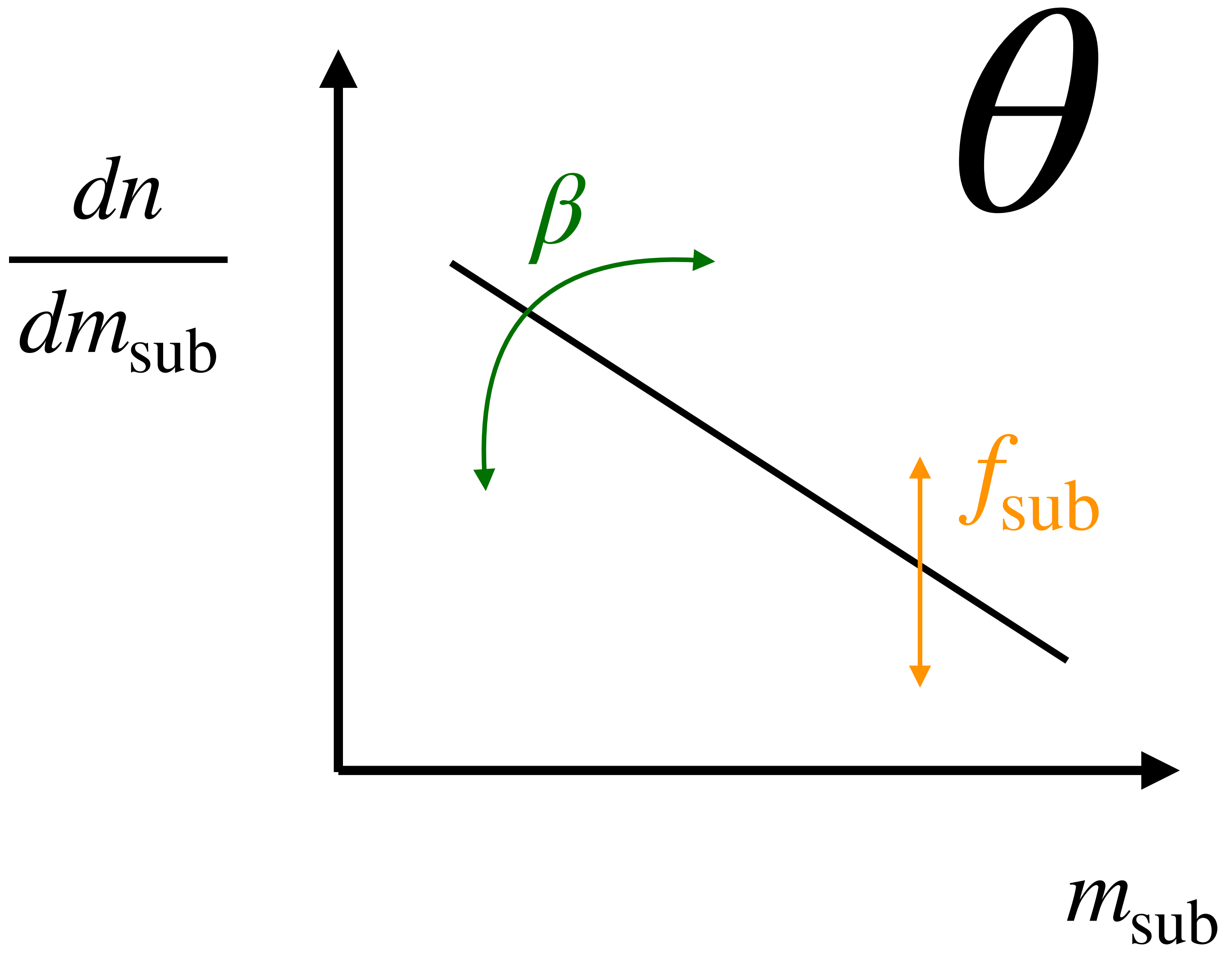


Parameters of interest

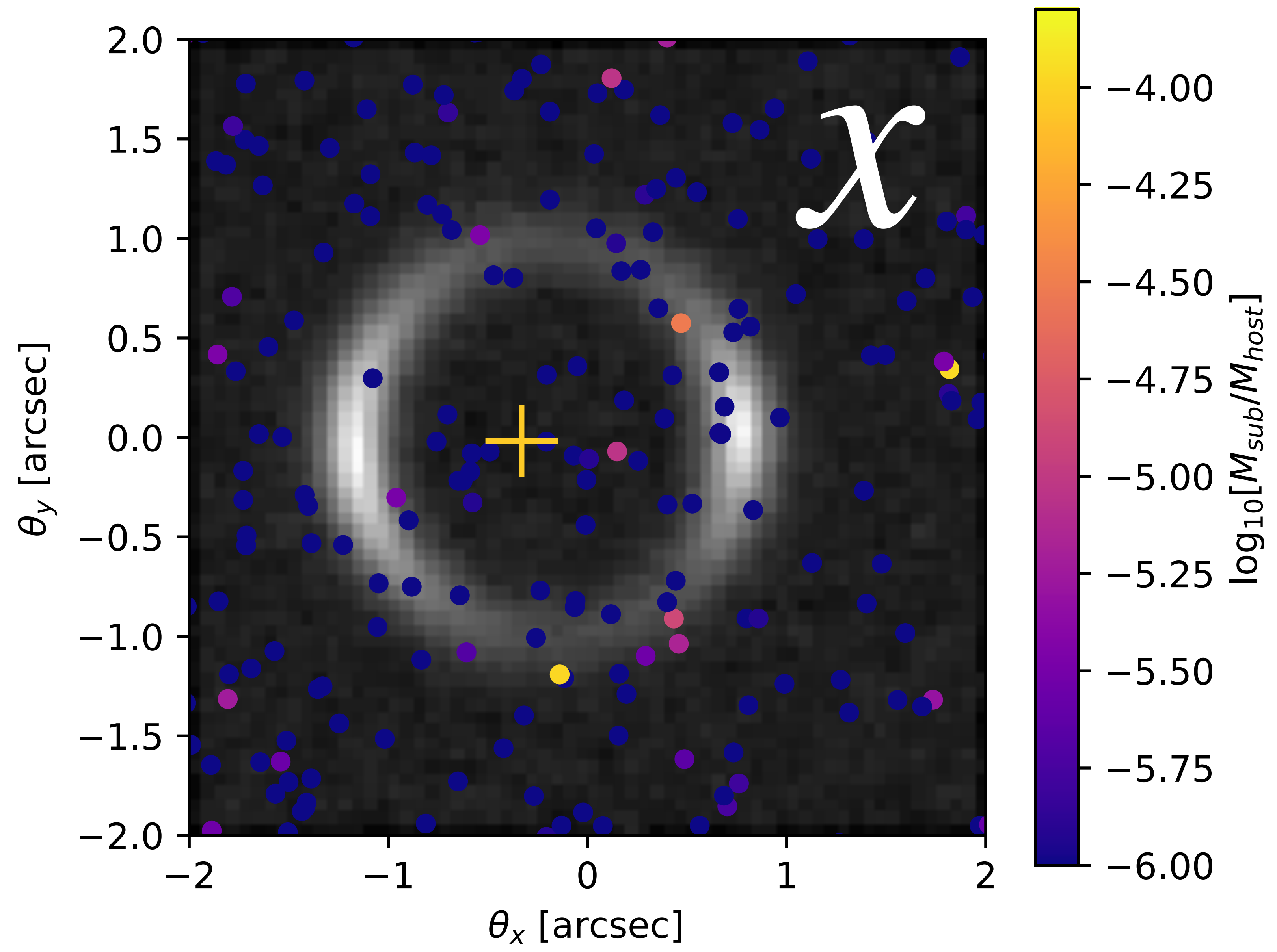
Subhalo population parameters

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

Substructure likelihood



Prediction

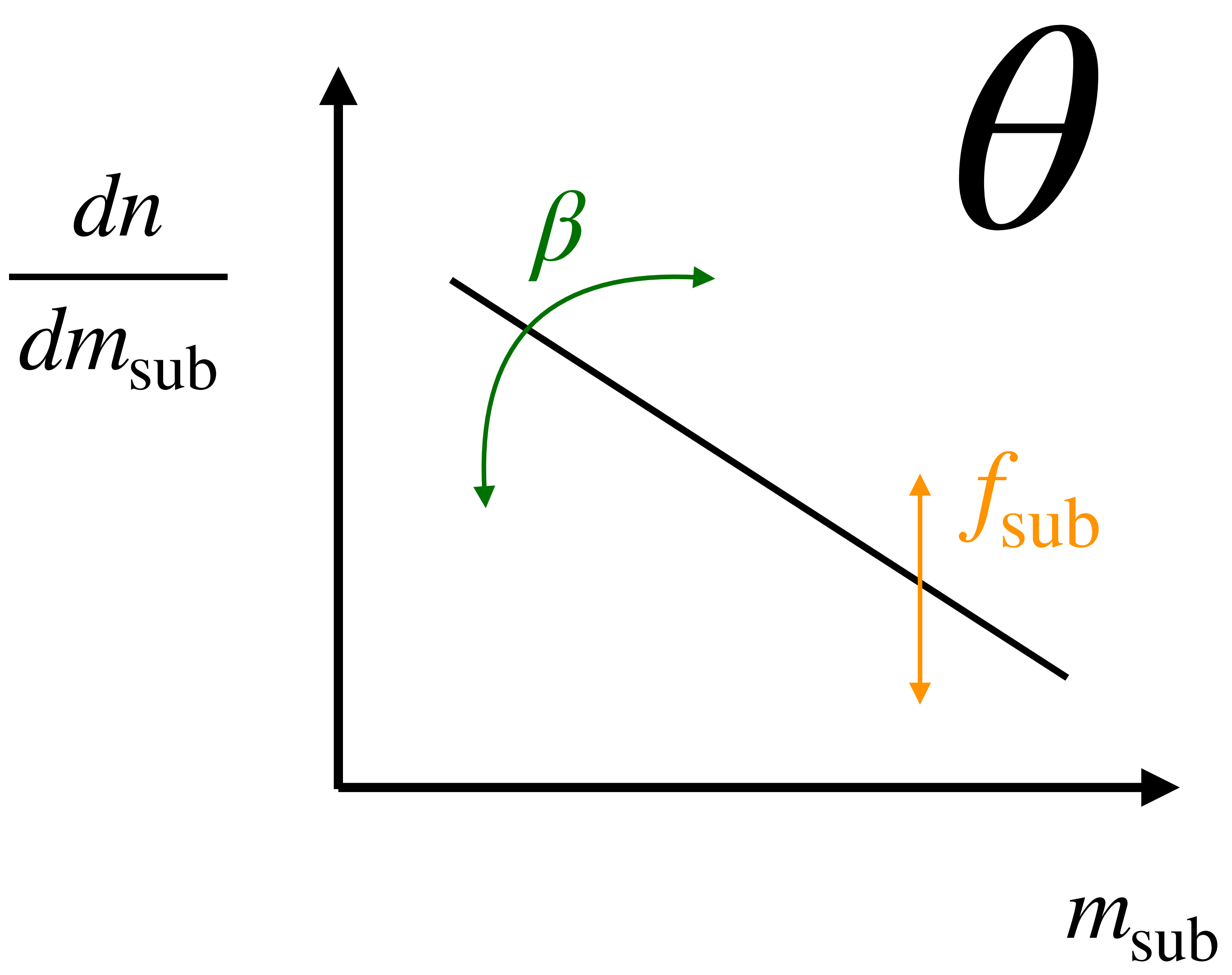


Parameters of interest

Subhalo population parameters

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

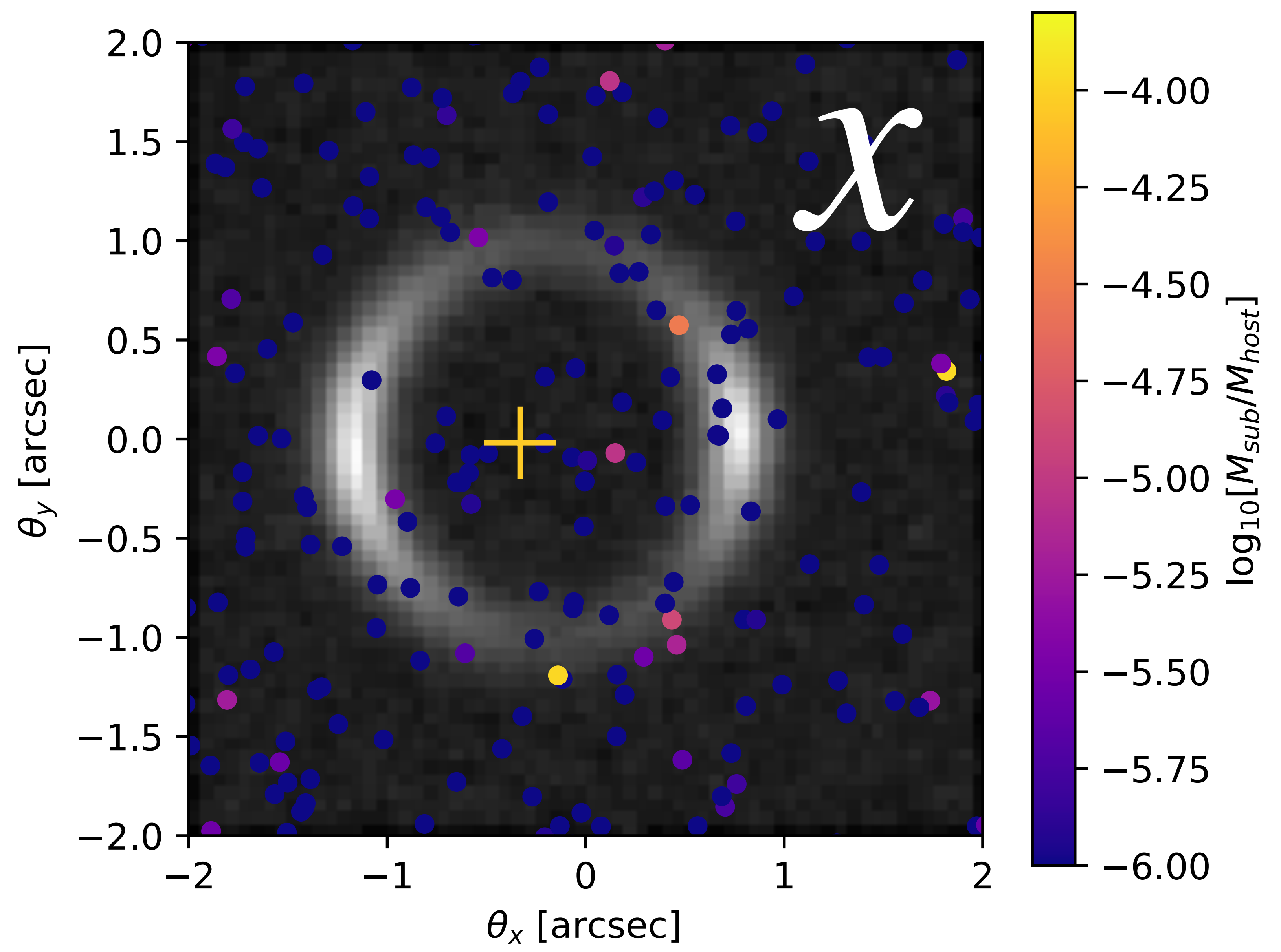
Substructure likelihood



Prediction

Inference

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



Parameters of interest

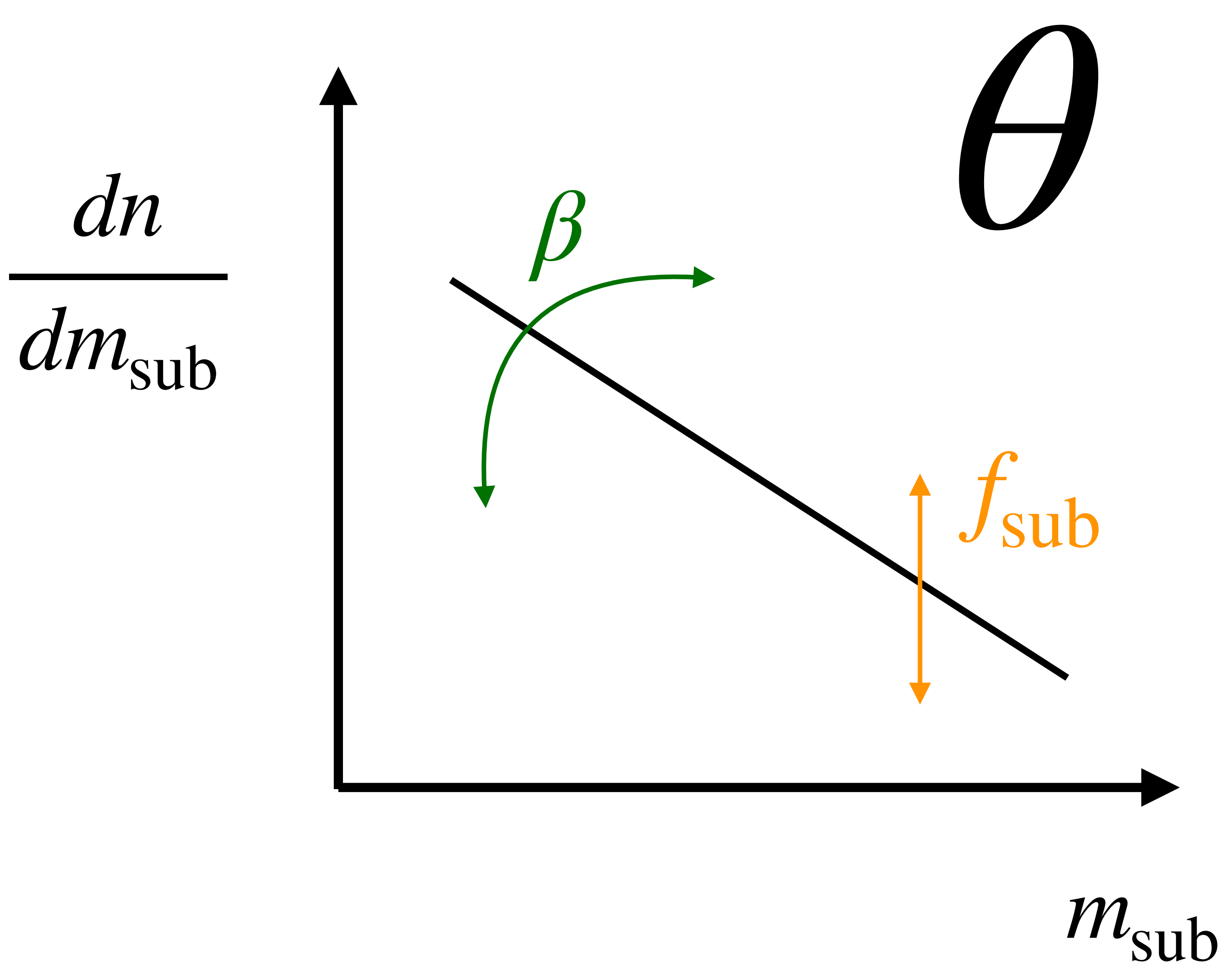
Subhalo population parameters

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

Latent variables

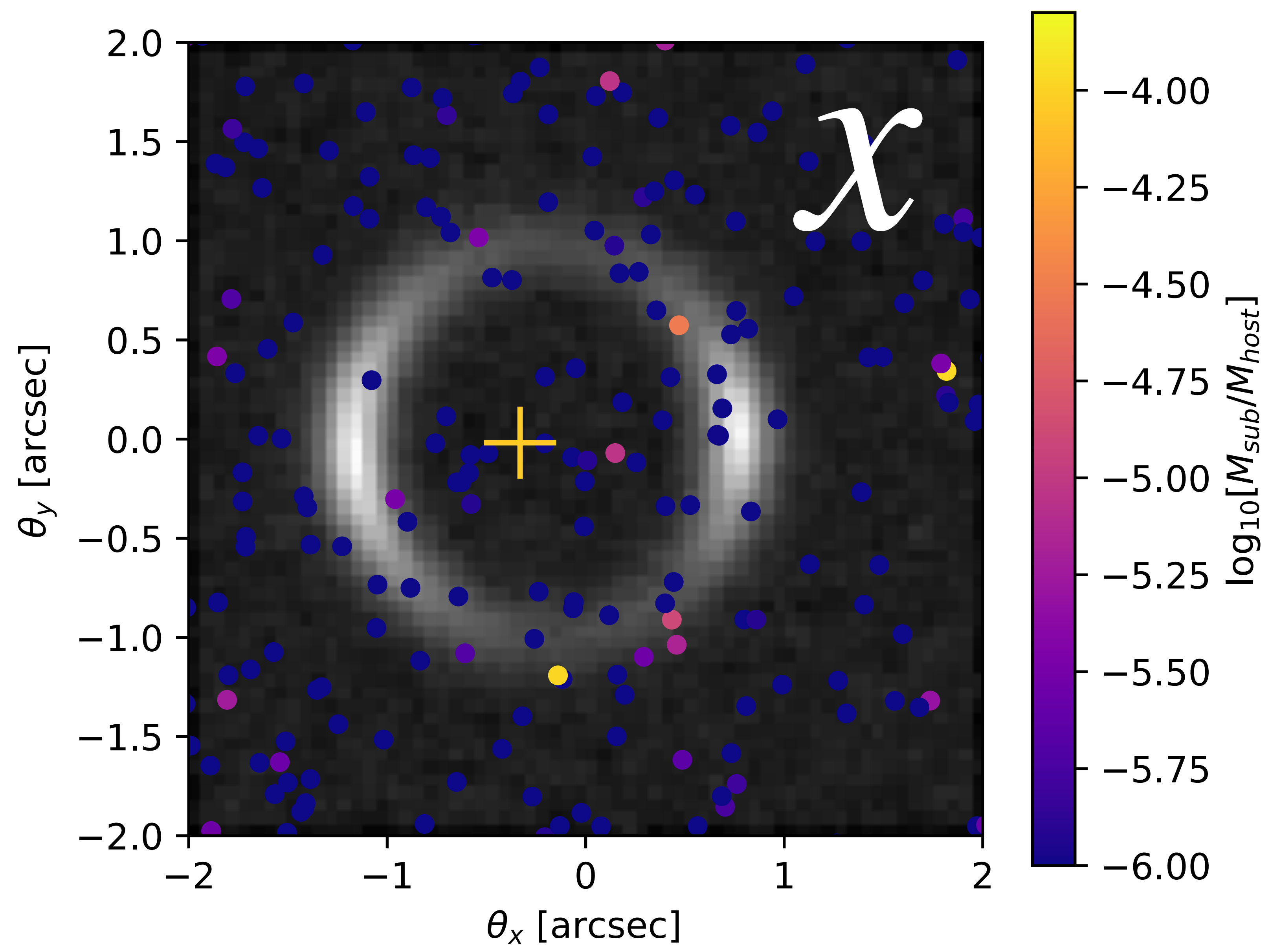
$$z = \{\vec{m}_{\mathrm{sub}}, \vec{r}_{\mathrm{sub}}\}$$

Substructure likelihood



Prediction →

← Inference



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

Parameters of interest

Subhalo population parameters

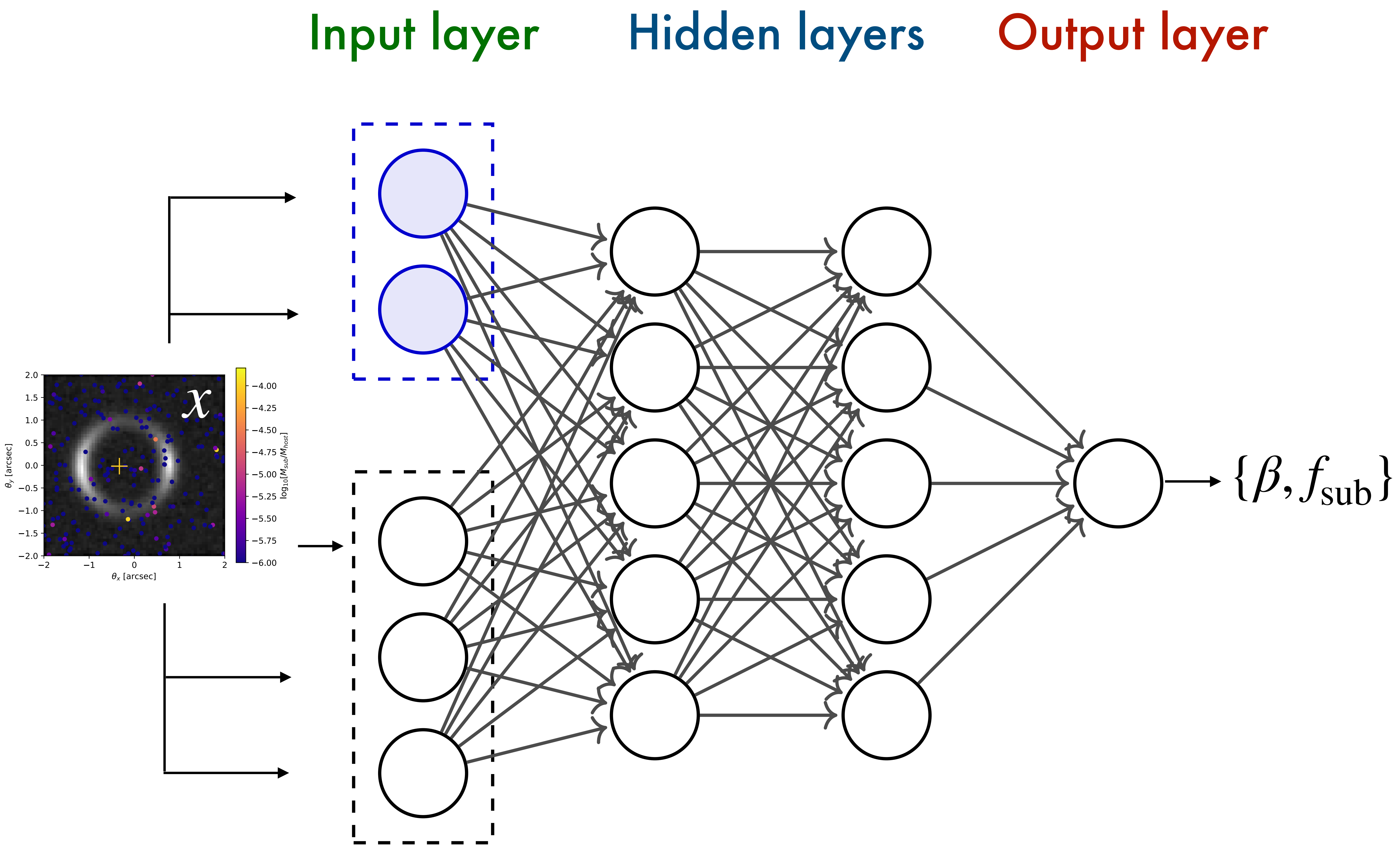
$$\theta = \{f_{\mathrm{sub}}, \beta\}$$

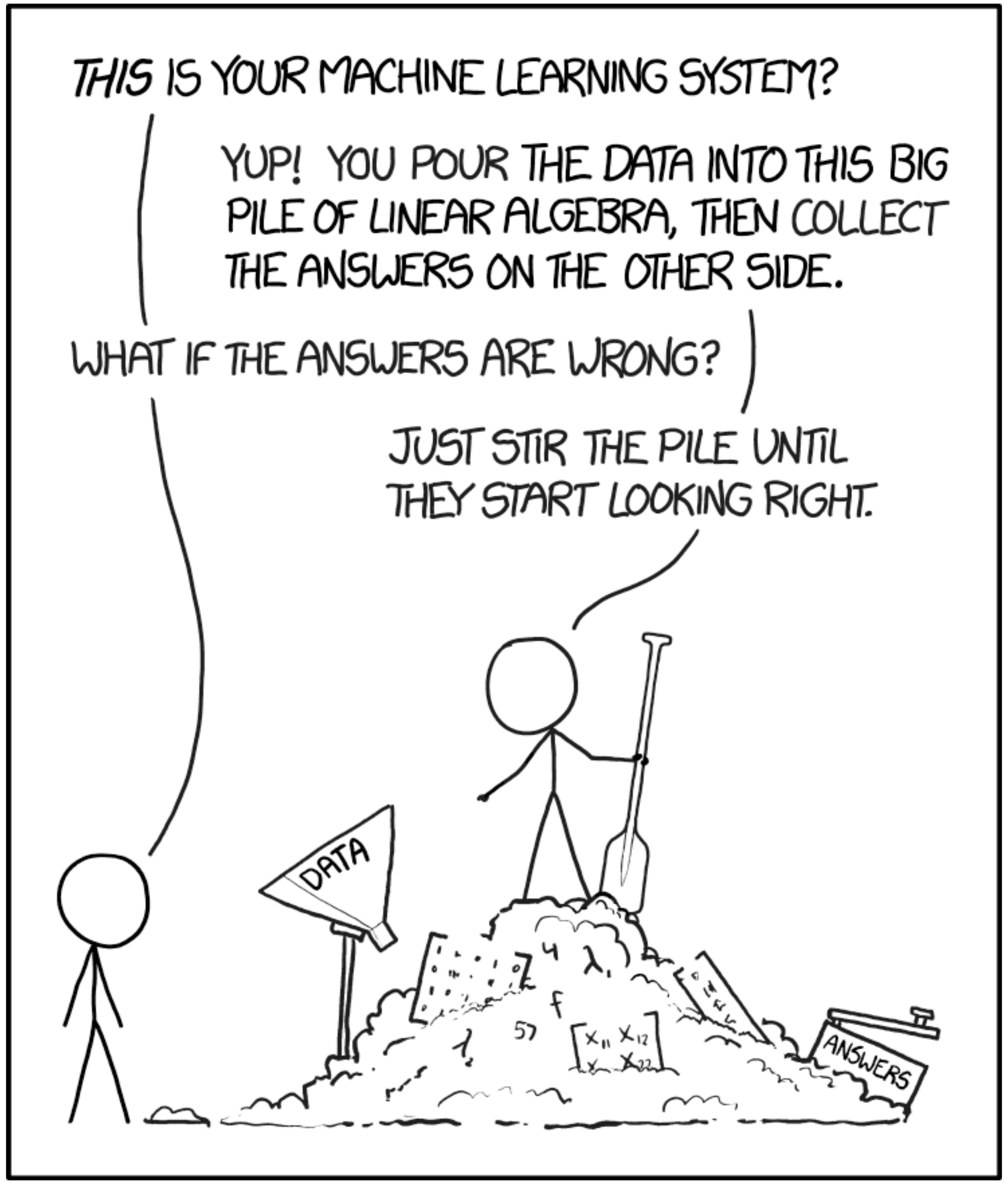
Latent variables

$$z = \{\vec{m}_{\mathrm{sub}}, \vec{r}_{\mathrm{sub}}\}$$

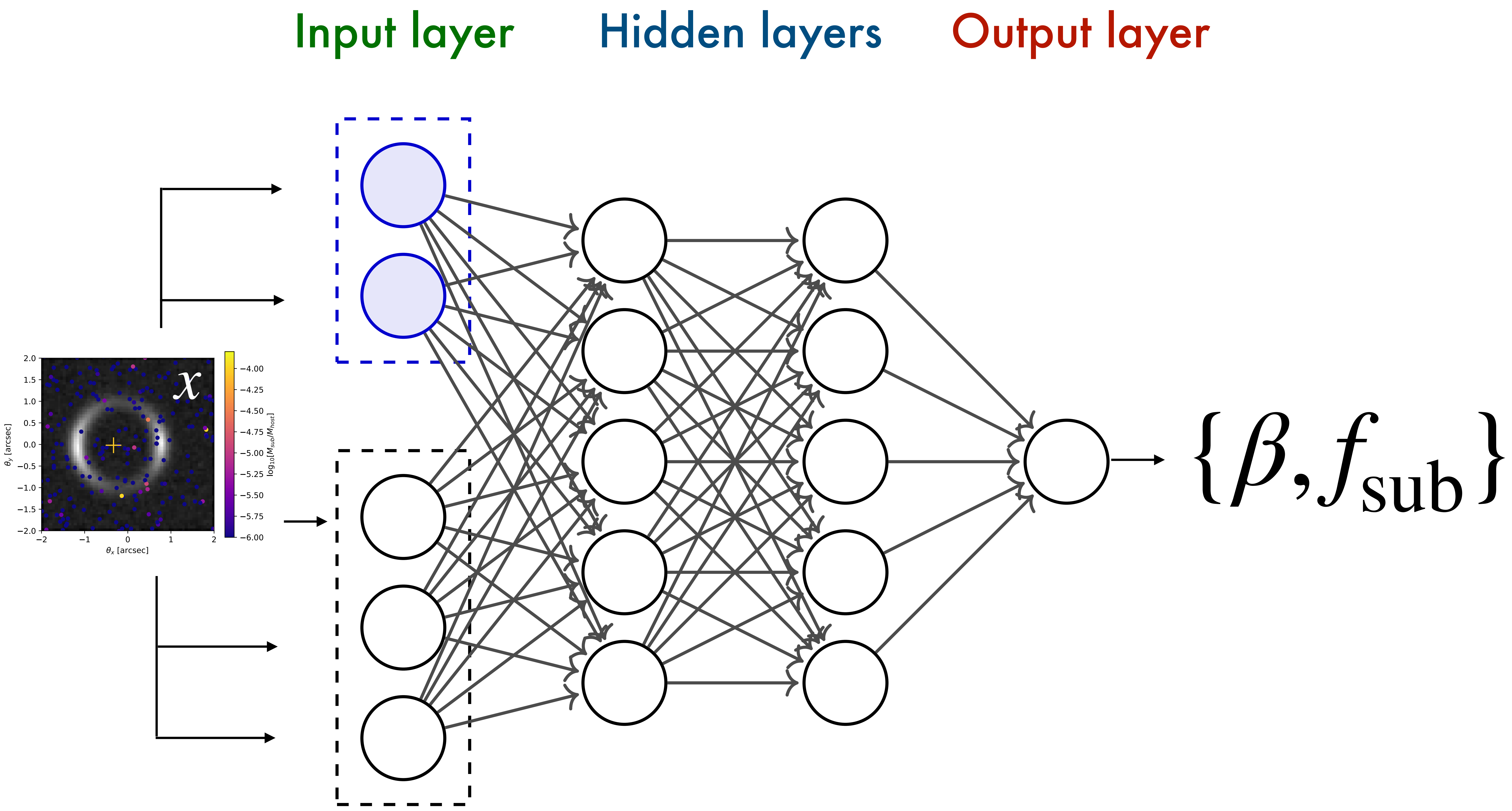
Huge latent space – full likelihood is intractable!

Neural networks





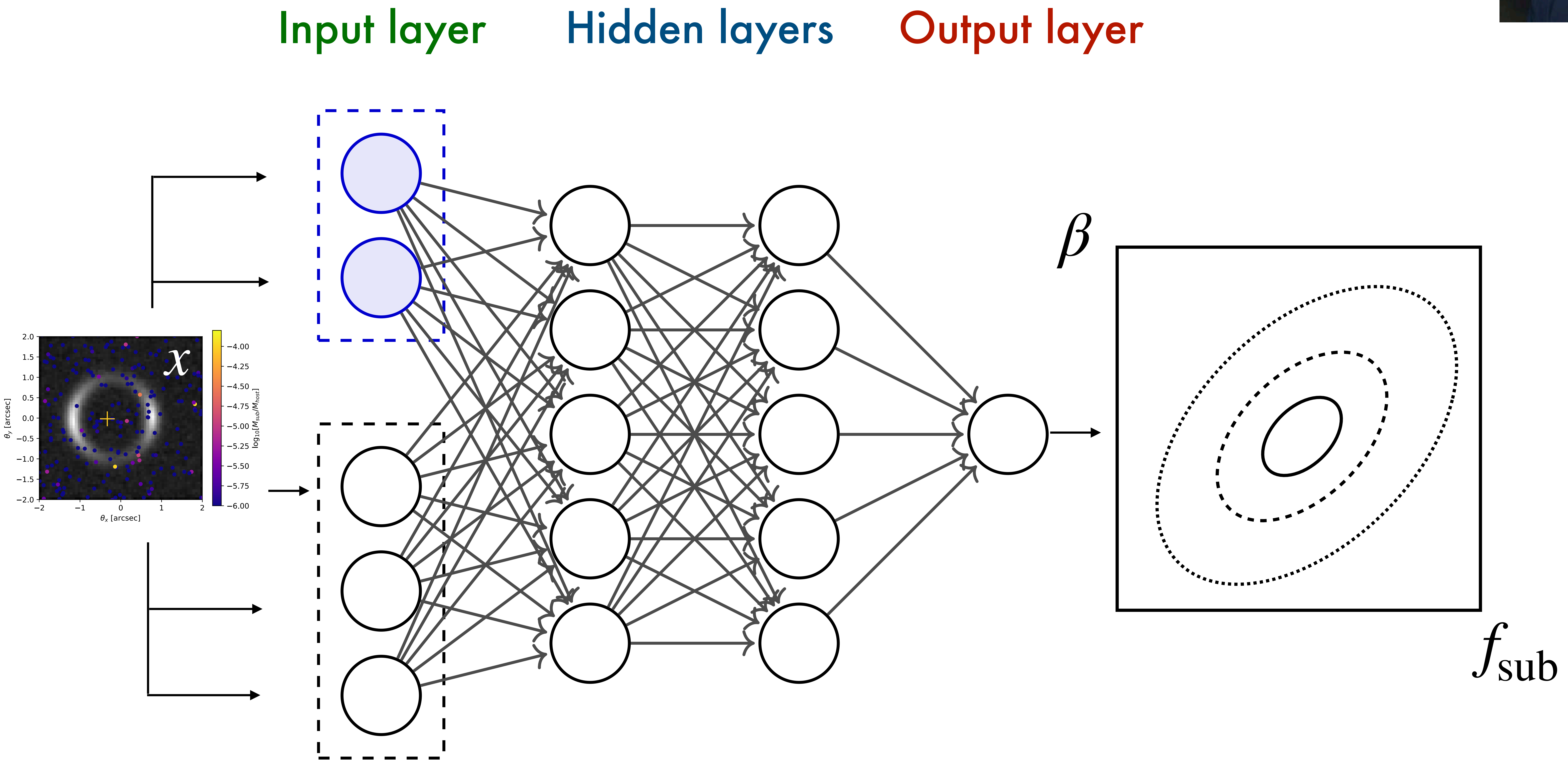
Neural networks



Neural networks

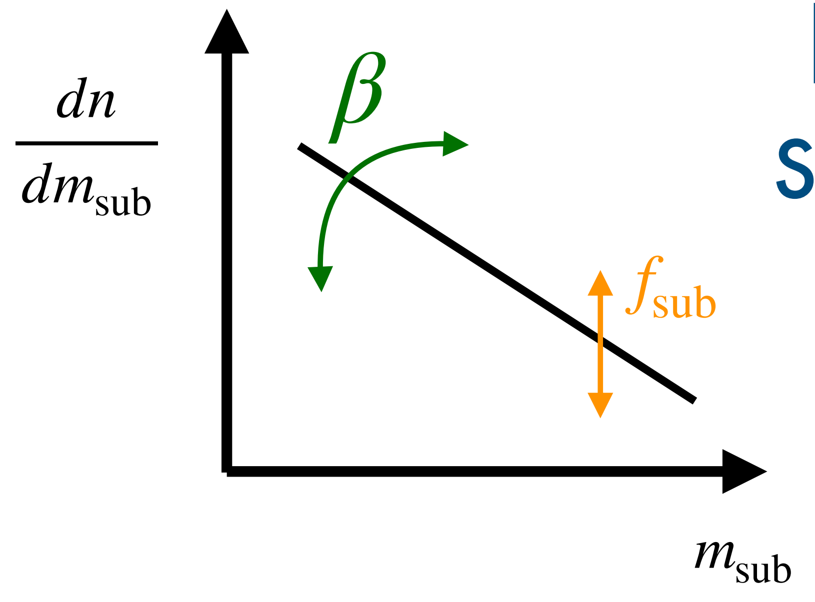
Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]

Slides courtesy of Johann Brehmer



Application to substructure in strong lenses

Application to substructure in strong lenses

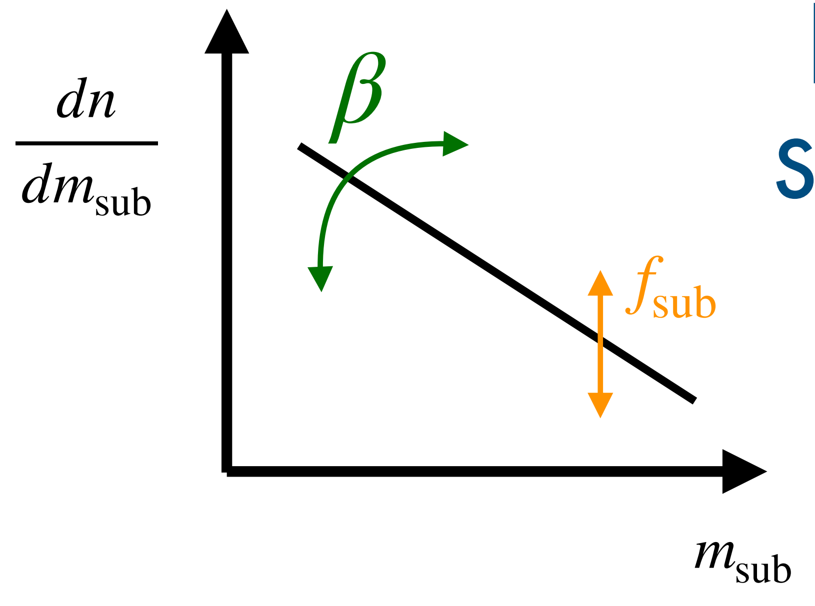


Parameters of interest

Subhalo population parameters

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

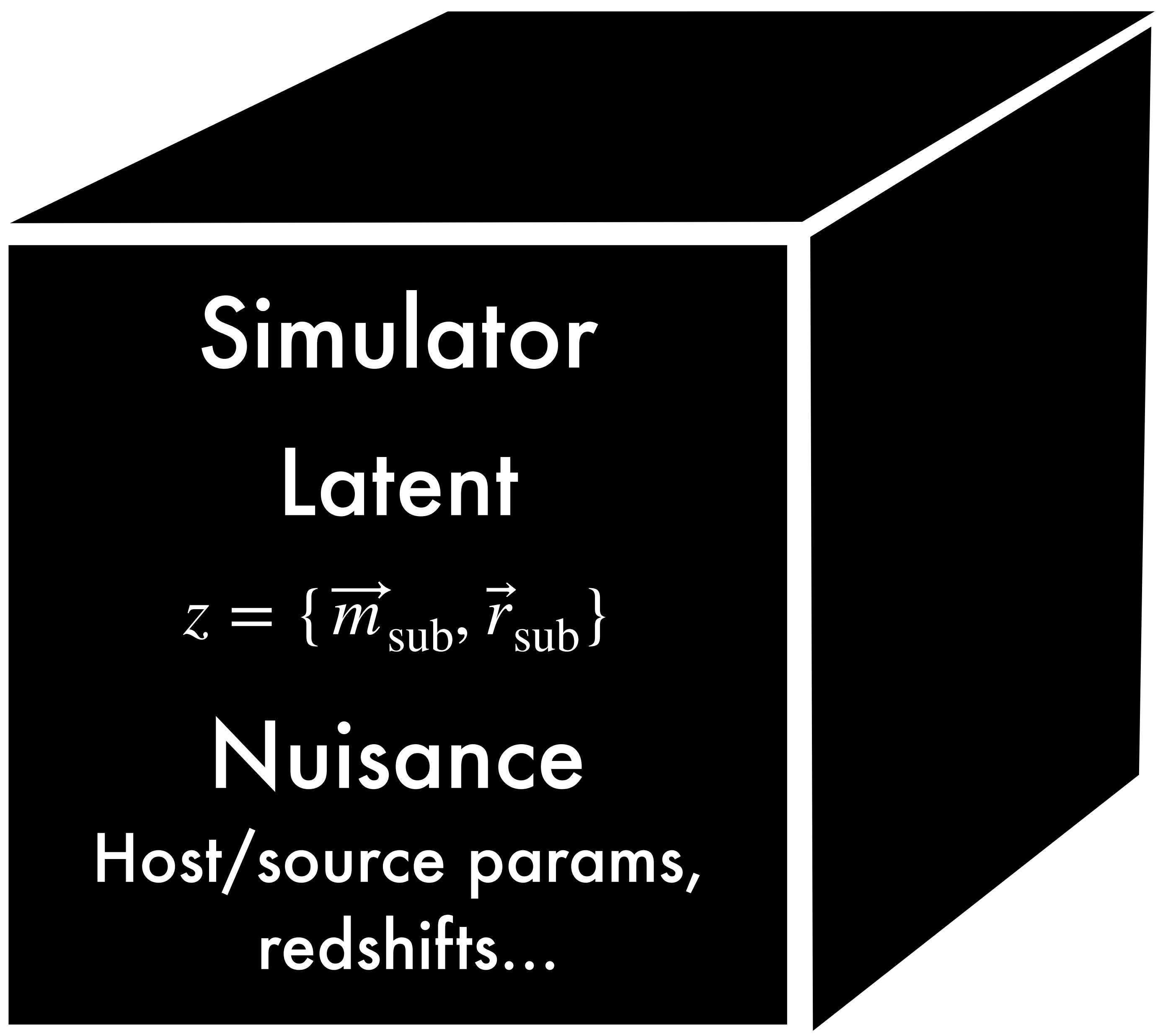
Application to substructure in strong lenses



Parameters of interest

Subhalo population parameters

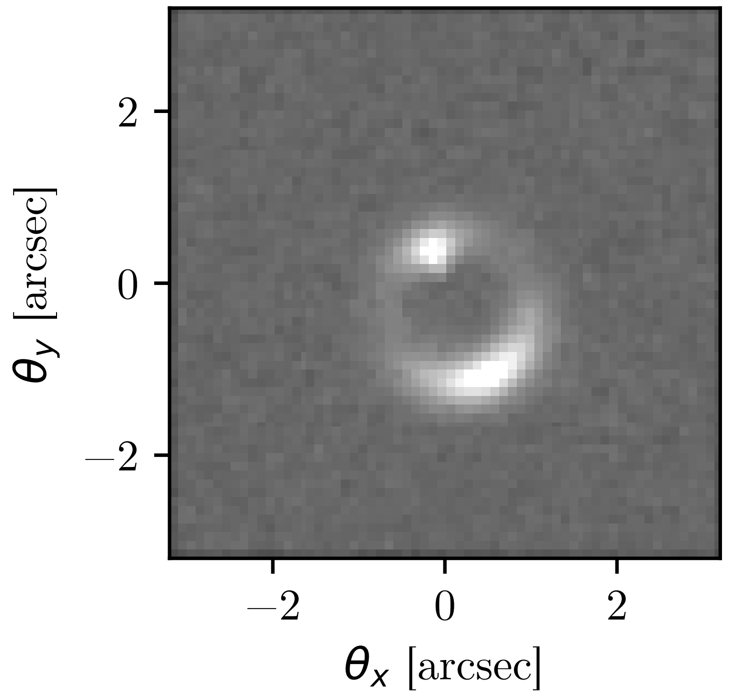
$\theta = \{f_{\rm sub}, \beta\}$



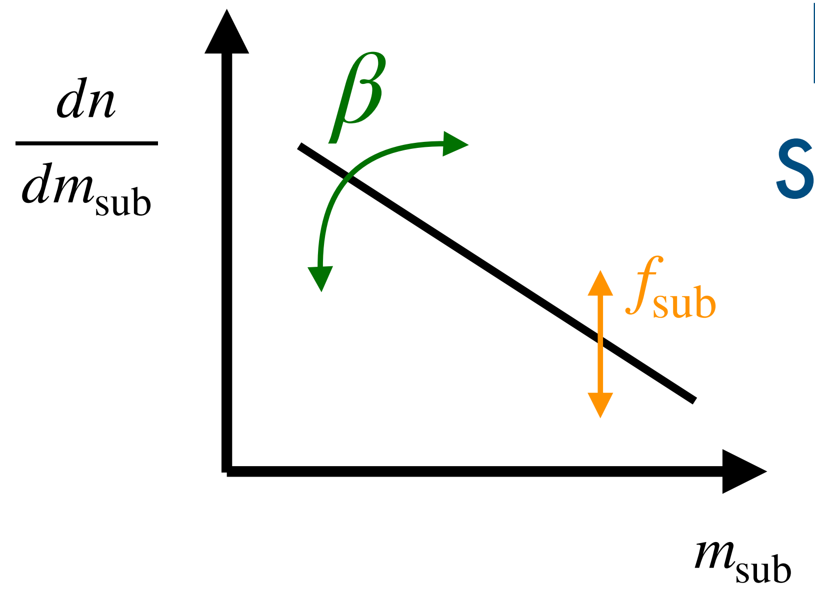
$x =$

Observables

Lensed image



Application to substructure in strong lenses



Parameters of interest

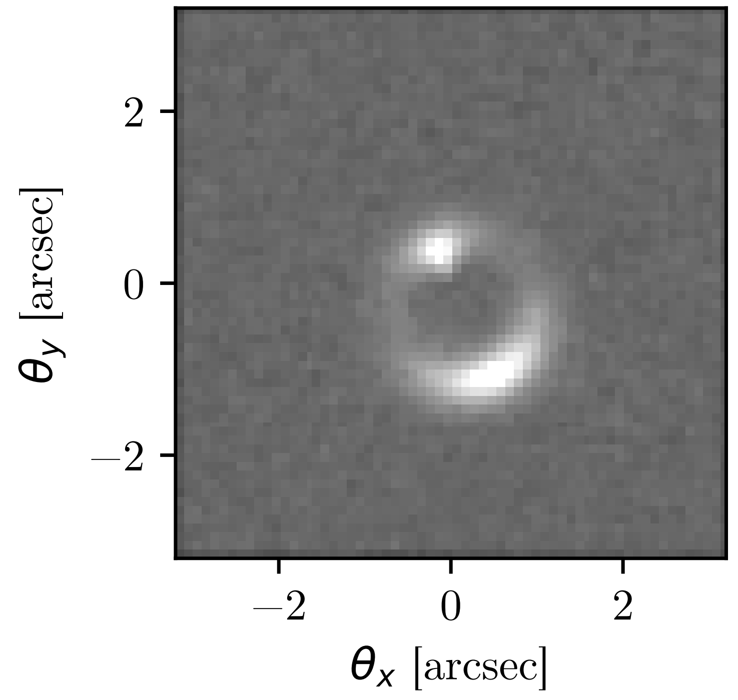
Subhalo population parameters

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

Observables

Lensed image

$x =$



Simulator

Latent

$z = \{\vec{m}_{\mathrm{sub}}, \vec{r}_{\mathrm{sub}}\}$

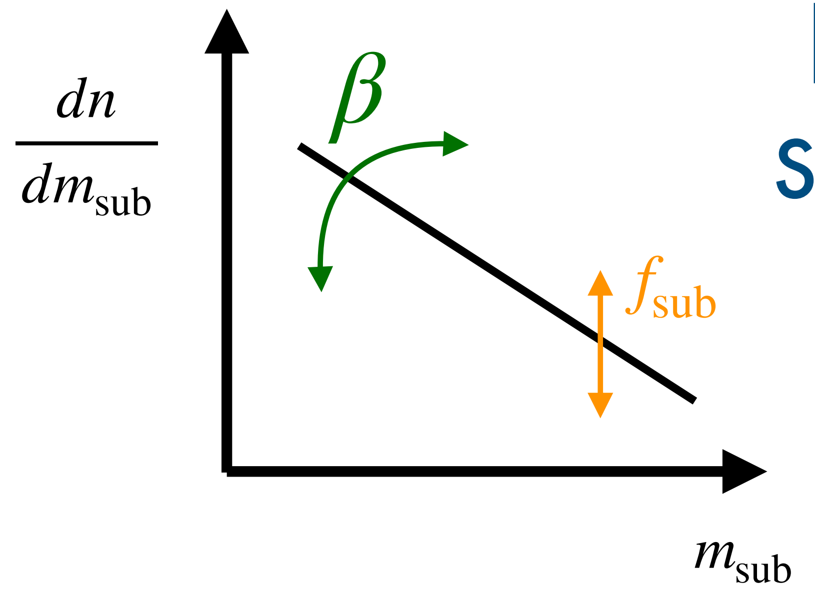
Nuisance

Host/source params, redshifts...

Joint likelihood ratio

$$r\left(x, z \,|\, \theta_0, \theta_1\right) = \frac{p(\vec{m}_{\mathrm{sub}}, \vec{r}_{\mathrm{sub}} \,|\, f_{\mathrm{sub},0}, \beta_0)}{p(\vec{m}_{\mathrm{sub}}, \vec{r}_{\mathrm{sub}} \,|\, f_{\mathrm{sub},1}, \beta_1)}$$

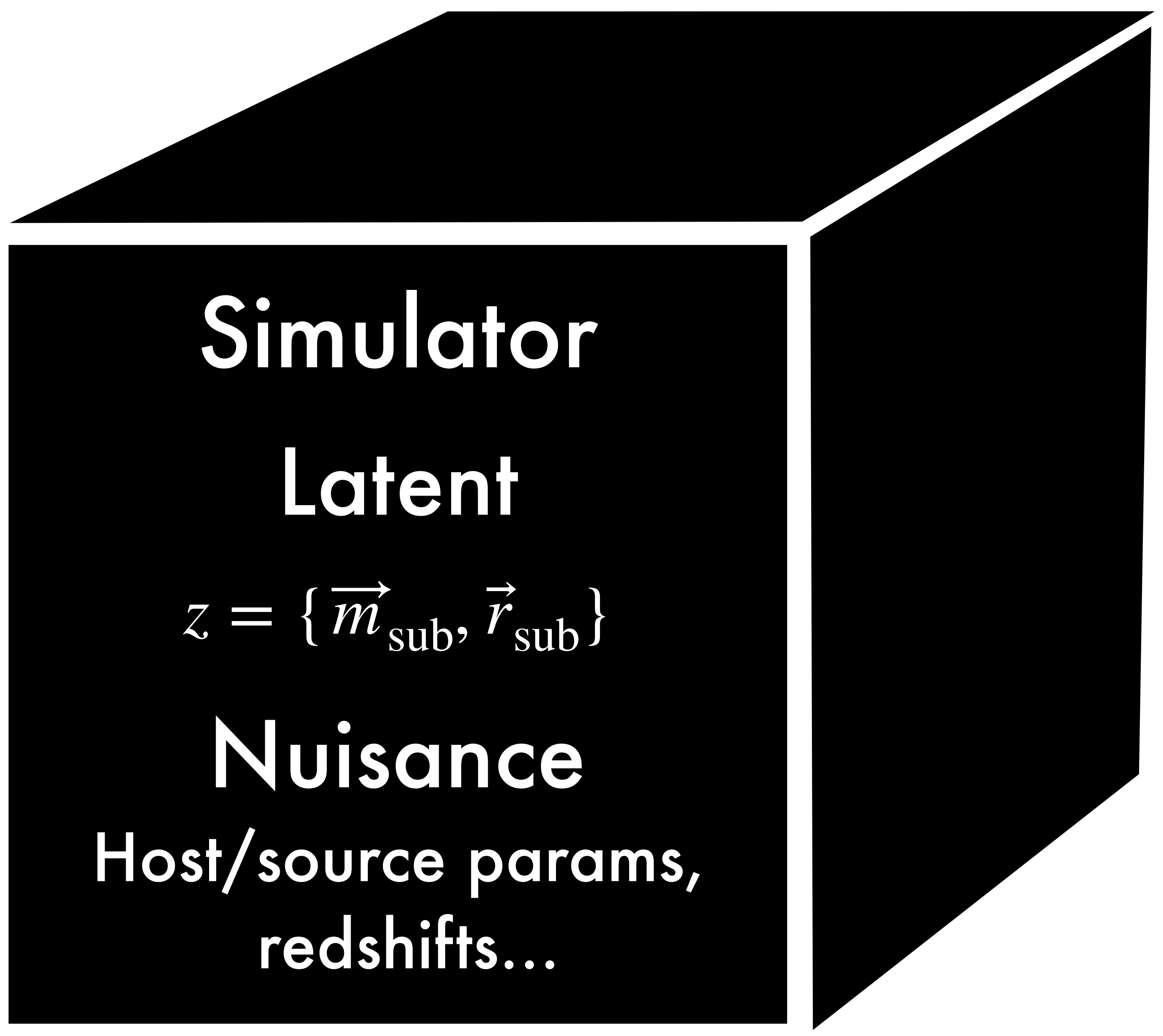
Application to substructure in strong lenses



Parameters of interest

Subhalo population parameters

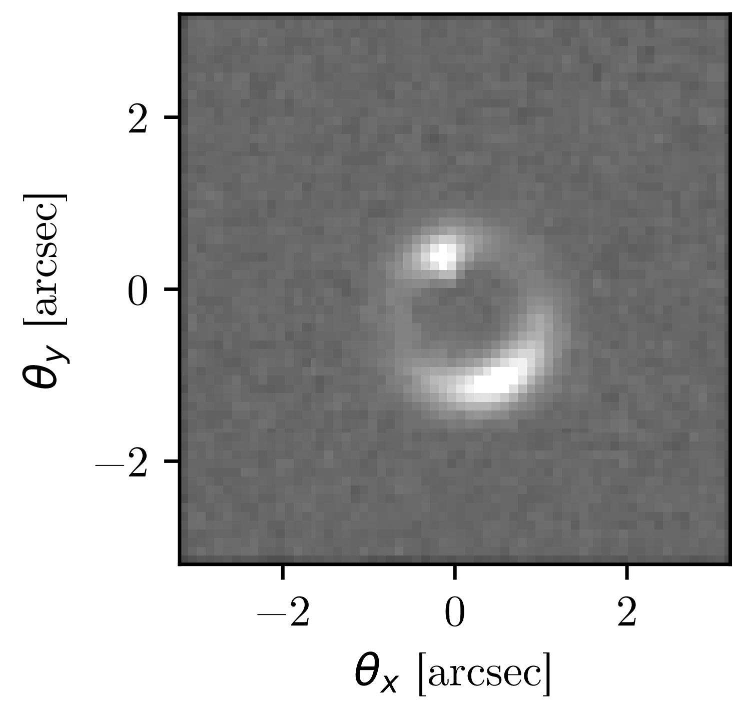
$\theta = \{f_{\rm sub}, \beta\}$

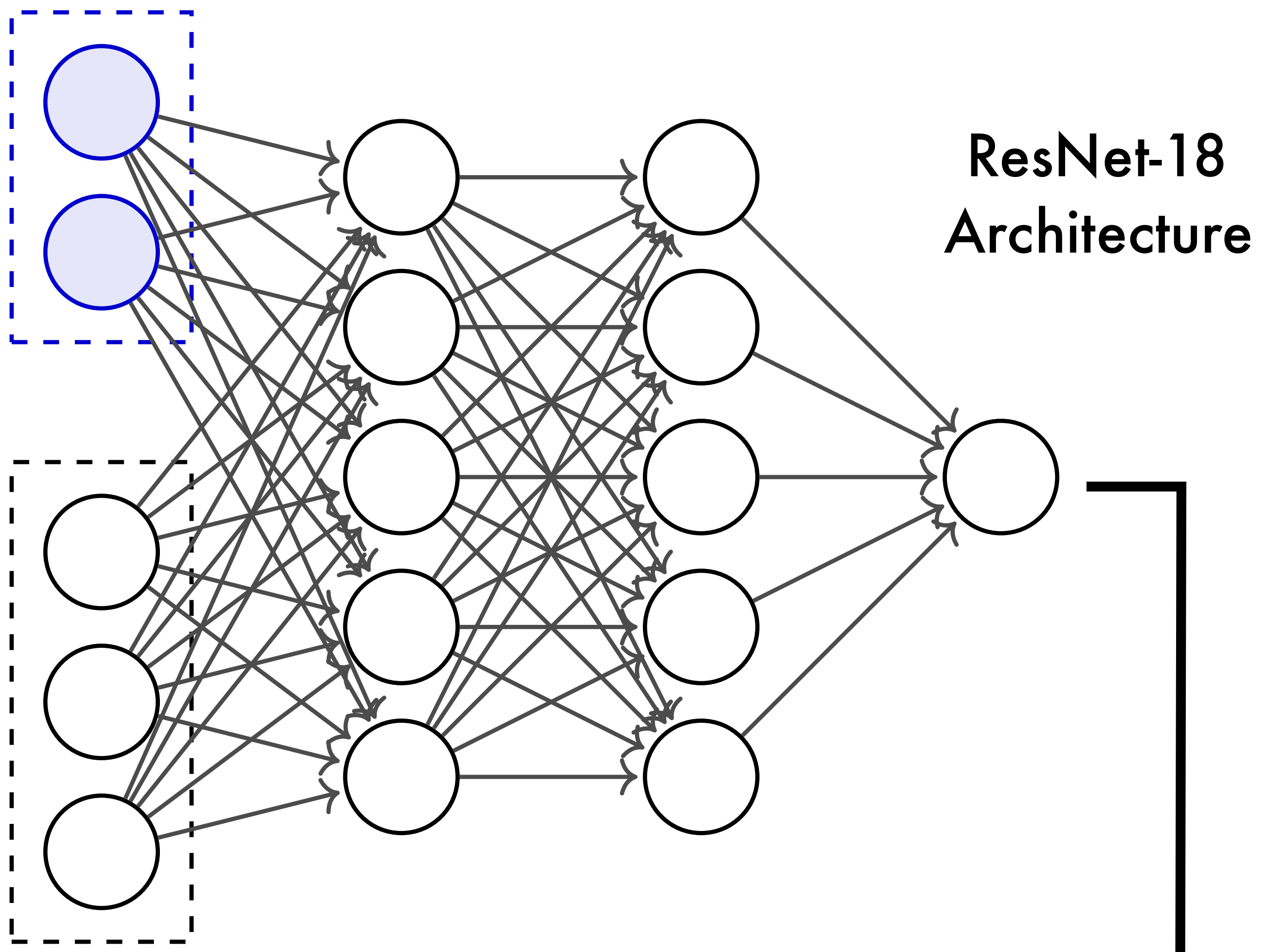


Observables

Lensed image

$x =$



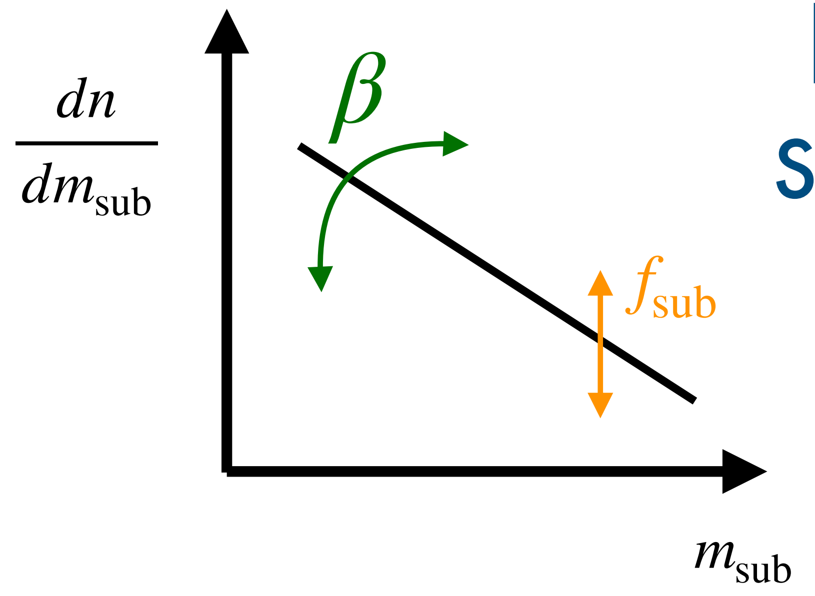


Joint likelihood ratio

$$r\left(x, z \,|\, \theta_0, \theta_1\right) = \frac{p(\vec{m}_{\rm sub}, \vec{r}_{\rm sub} \,|\, f_{\rm sub,0}, \beta_0)}{p(\vec{m}_{\rm sub}, \vec{r}_{\rm sub} \,|\, f_{\rm sub,1}, \beta_1)}$$

$$\arg\min_g L[g]$$

Application to substructure in strong lenses



Parameters of interest

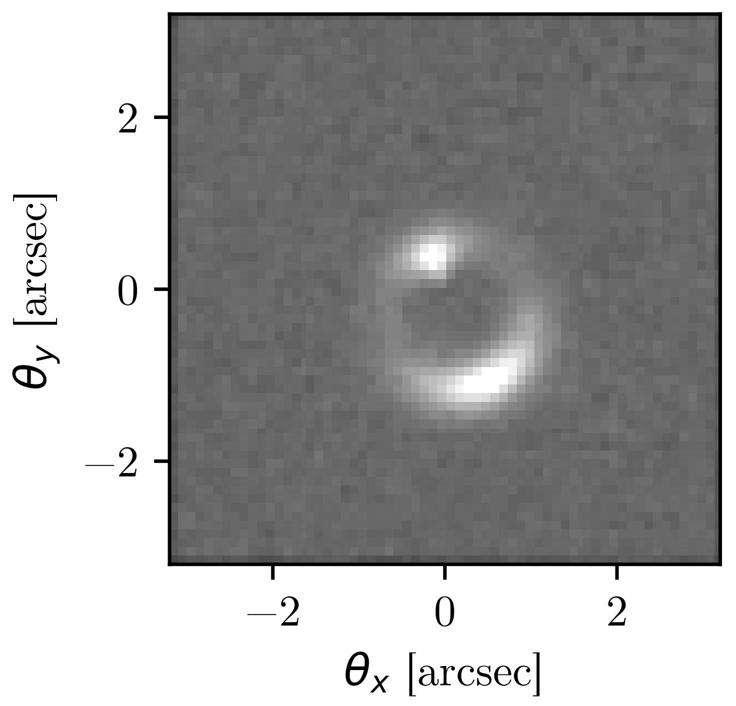
Subhalo population parameters

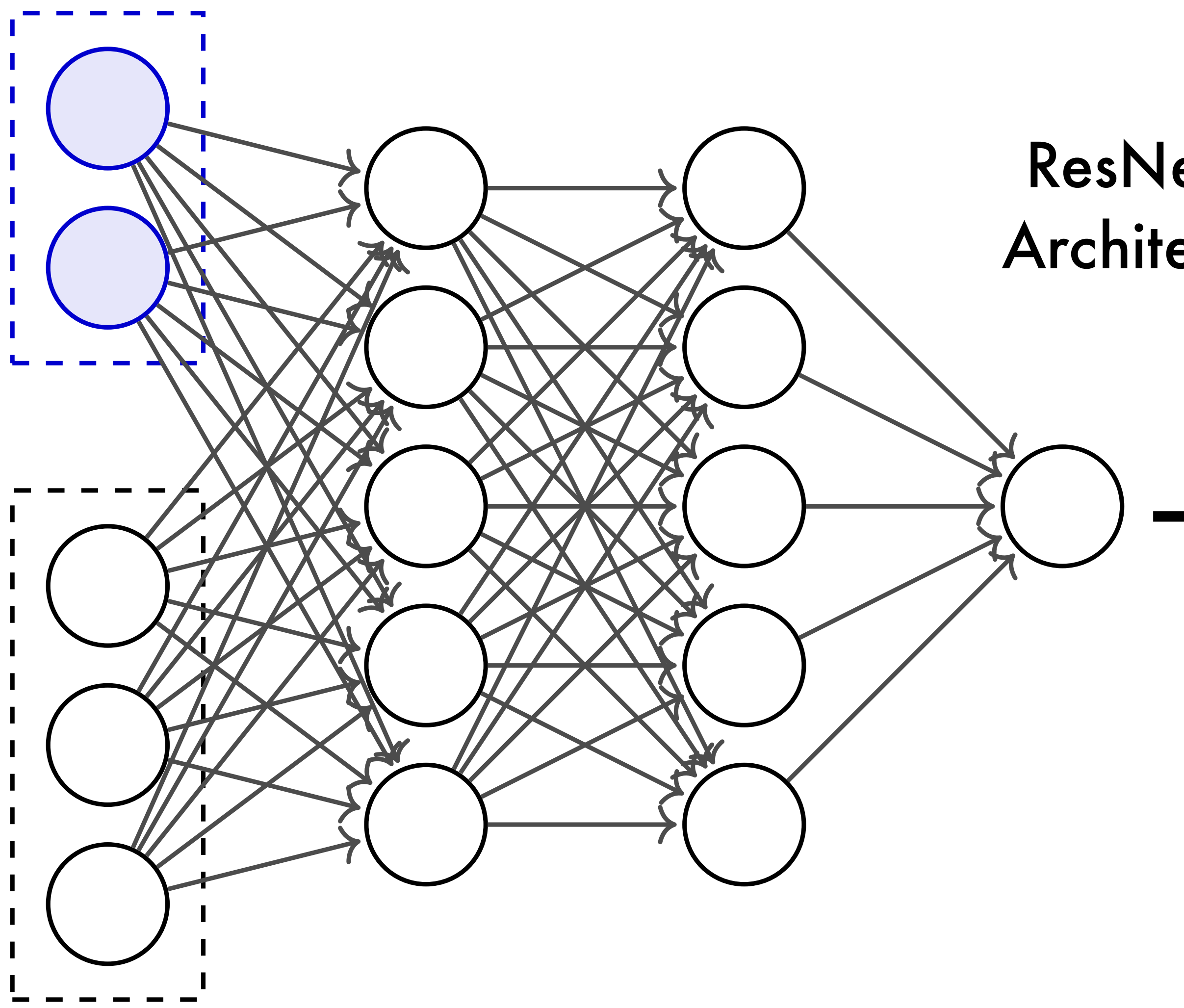
$\theta = \{f_{\rm sub}, \beta\}$

Observables

Lensed image

$x =$





Simulator

Latent

$z = \{\vec{m}_{\rm sub}, \vec{r}_{\rm sub}\}$

Nuisance

Host/source params, redshifts...

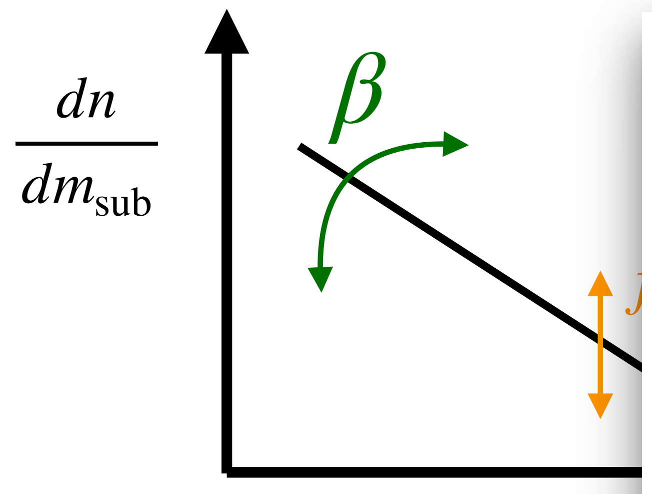
Joint likelihood ratio

$$r\,(x, z\,|\,\theta_0, \theta_1) = \frac{p(\vec{m}_{\rm sub}, \vec{r}_{\rm sub}\,|\,f_{\rm sub,0}, \beta_0)}{p(\vec{m}_{\rm sub}, \vec{r}_{\rm sub}\,|\,f_{\rm sub,1}, \beta_1)}$$

$$\arg\min_g L[g]$$

Likelihood ratio estimate $\hat{r}(x|\theta)$

Application to substructure in strong lenses



Simulat

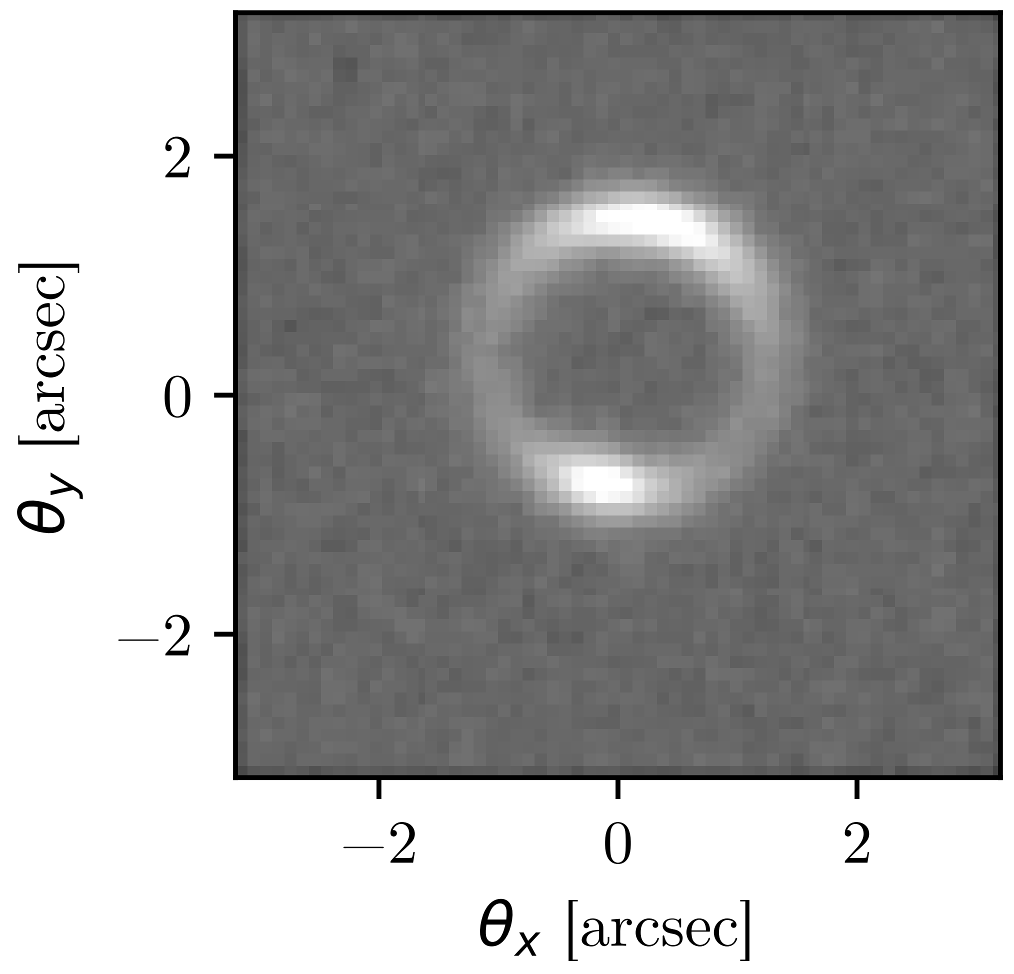
Latent

$z = \{\vec{m}_{\rm sub}, \vec{r}$

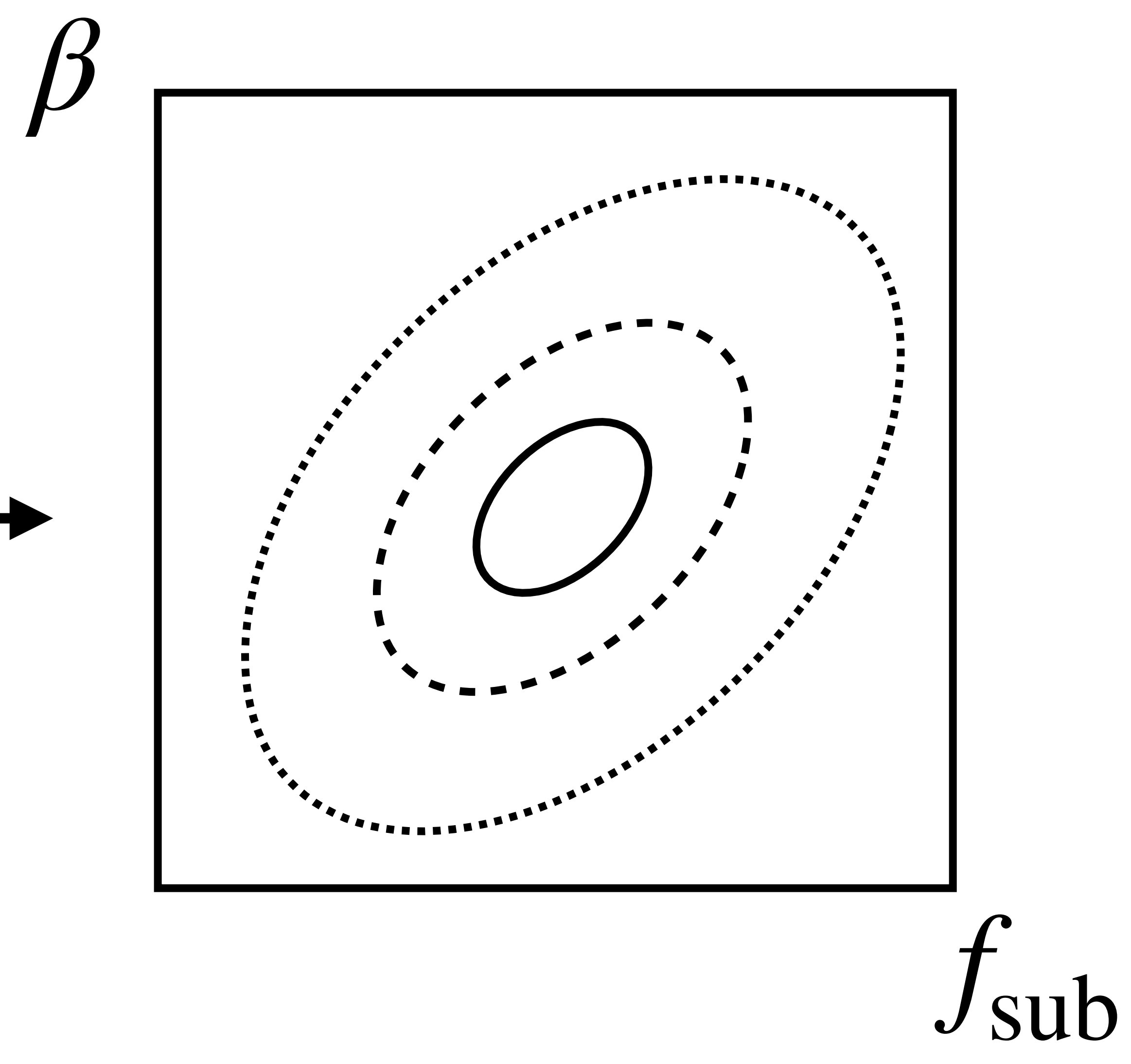
Nuisanc

Host/source p
redshifts...

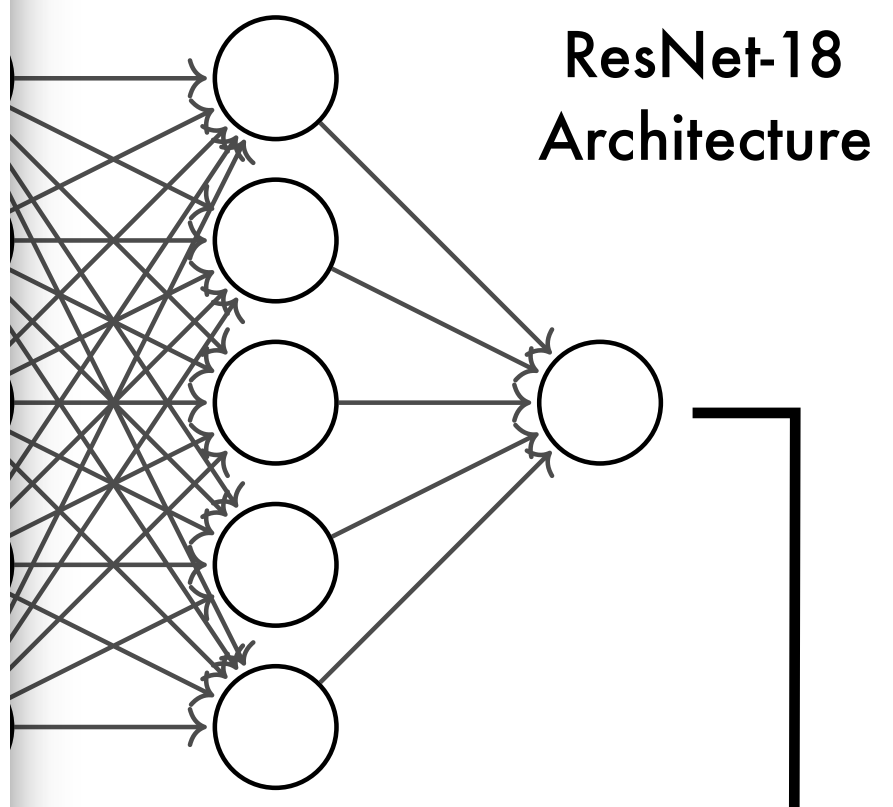
After training the parameterized LR estimator...



Inference



... can infer parameters of interest for given lensed image!



$\arg\min_g L[g]$

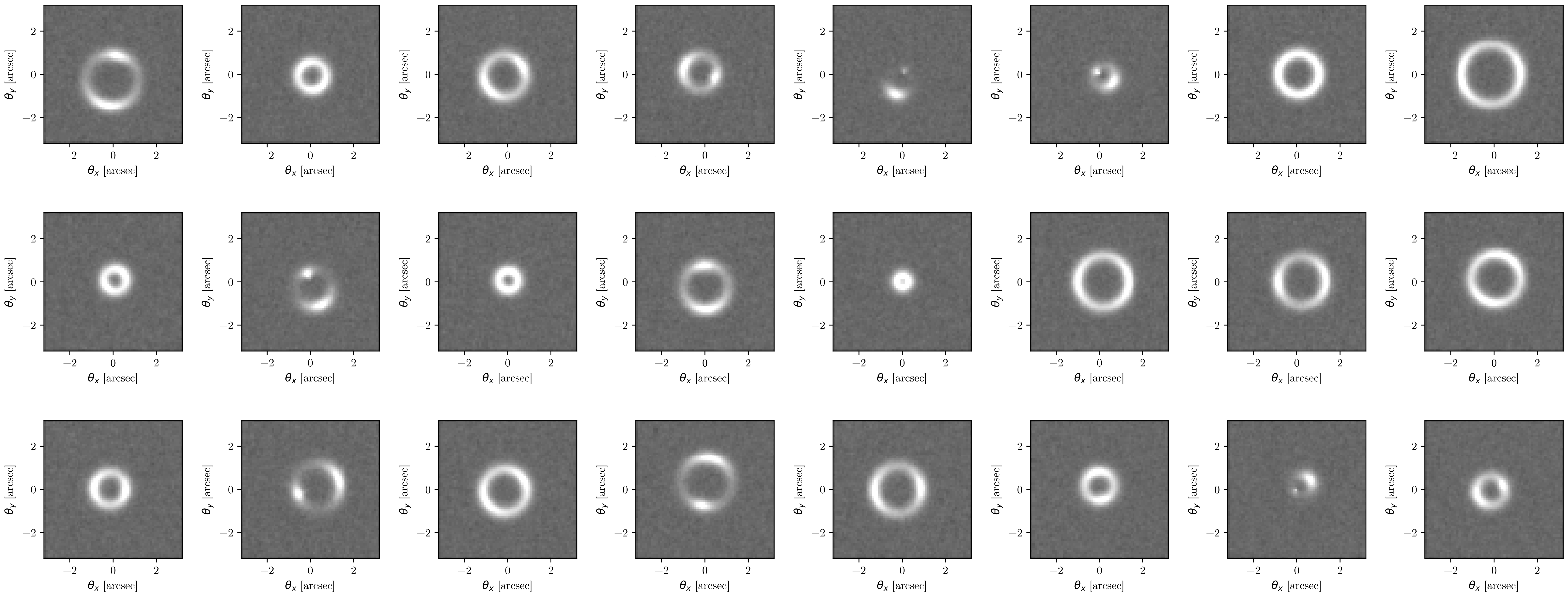
elihood ratio
estimate

$\hat{r}(x|\theta)$

Proof of principle

Use simulated ensemble of galaxy-galaxy lenses observable by *Euclid*

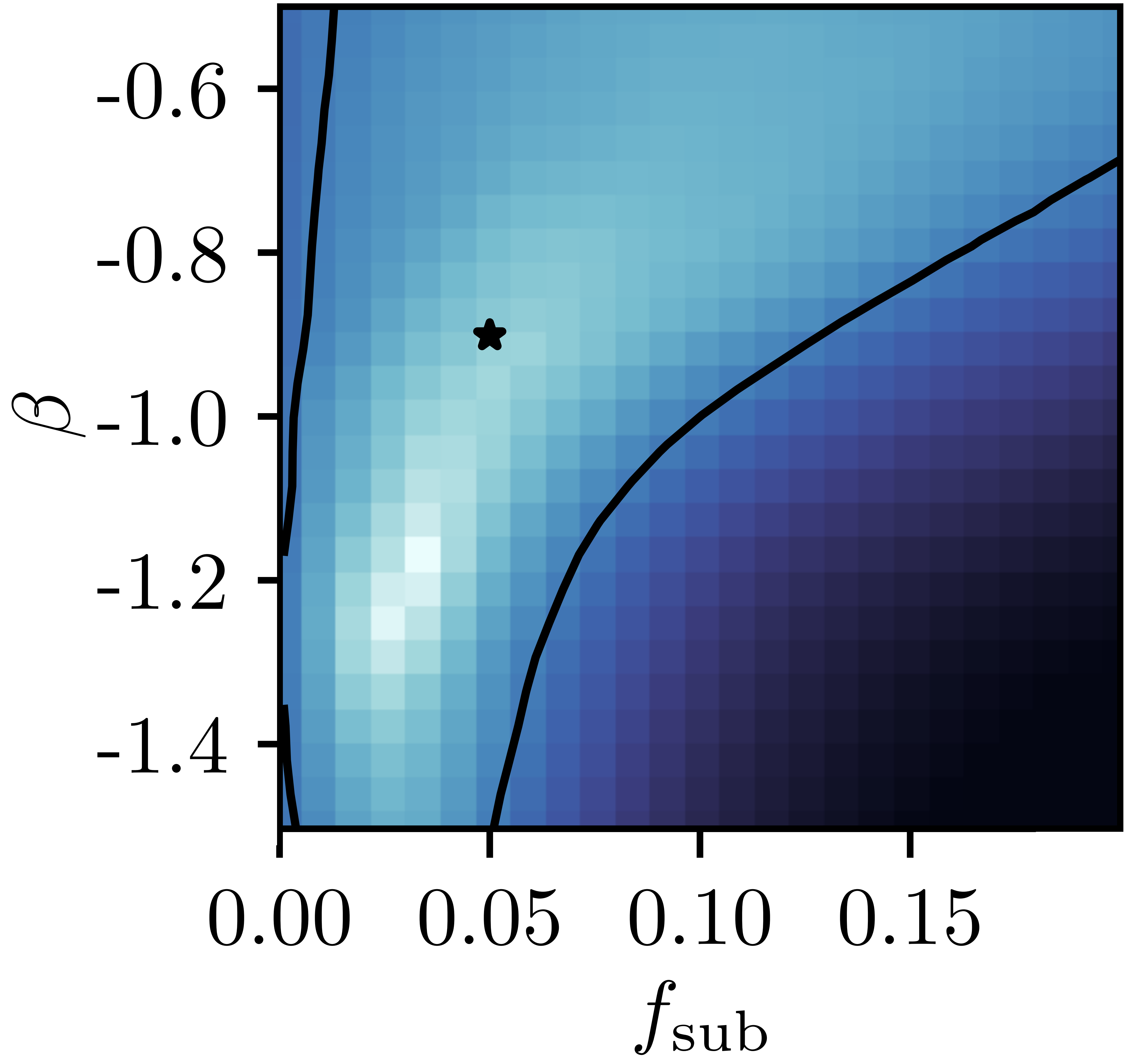
Collett et al [1507.02657]



1. Train likelihood ratio estimator with $f_{\mathrm{sub}} \sim [0, 0.2]$, $\beta \sim [-2.5, -1.5]$
2. Test on simulated data with $f_{\mathrm{sub}} = 0.05$, $\beta = -0.9$

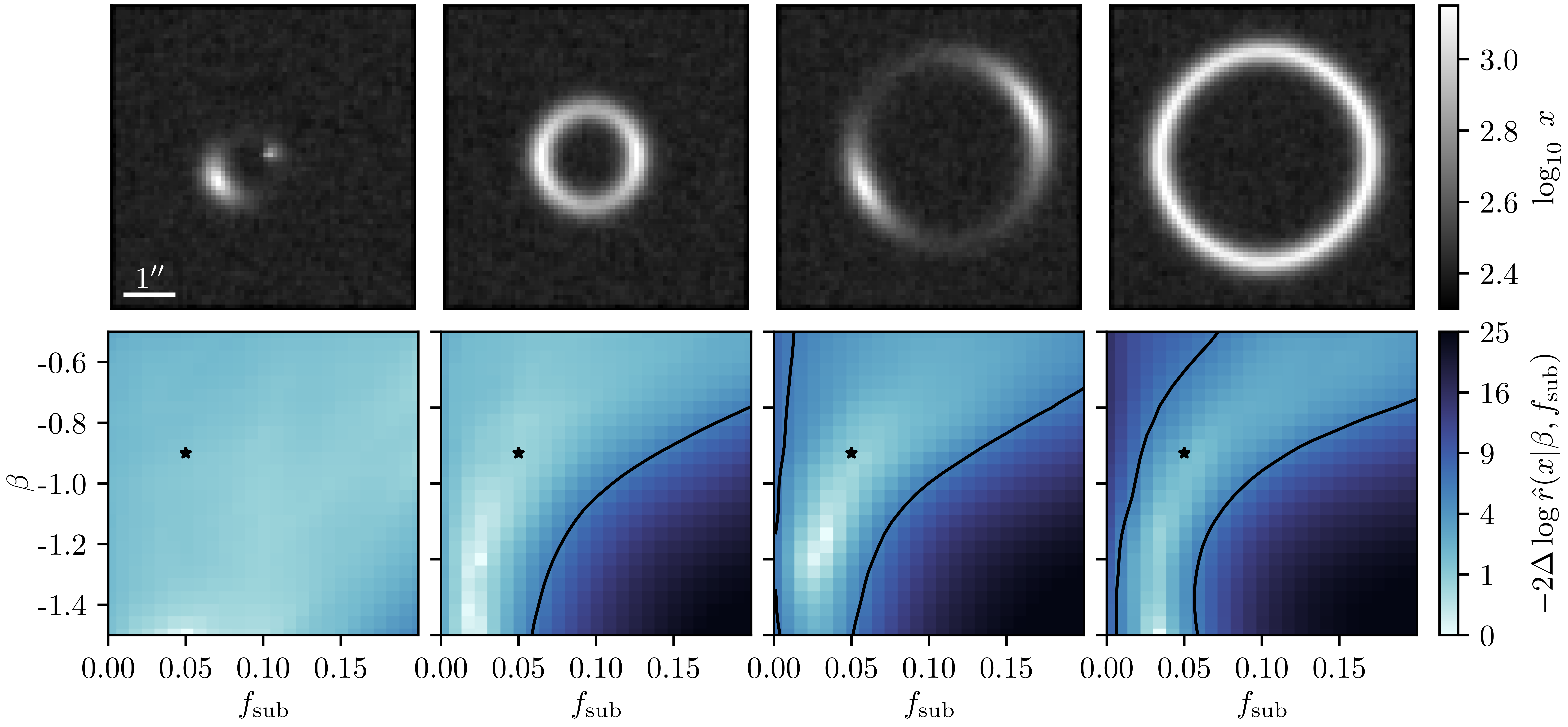
Inferred likelihood ratios

$f_{\mathrm{sub}} = 0.05, \beta = -0.9$



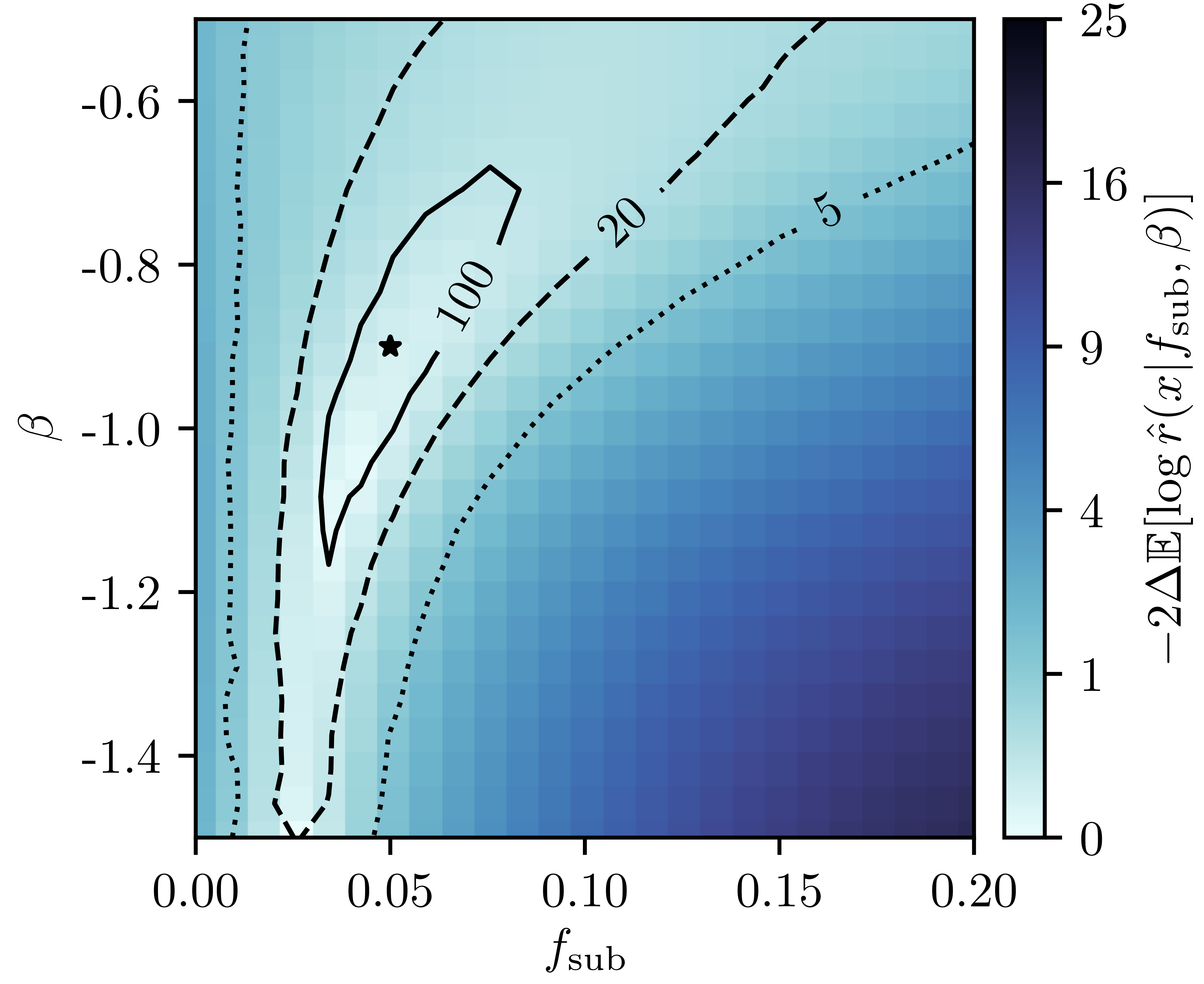
Inferred likelihood ratios

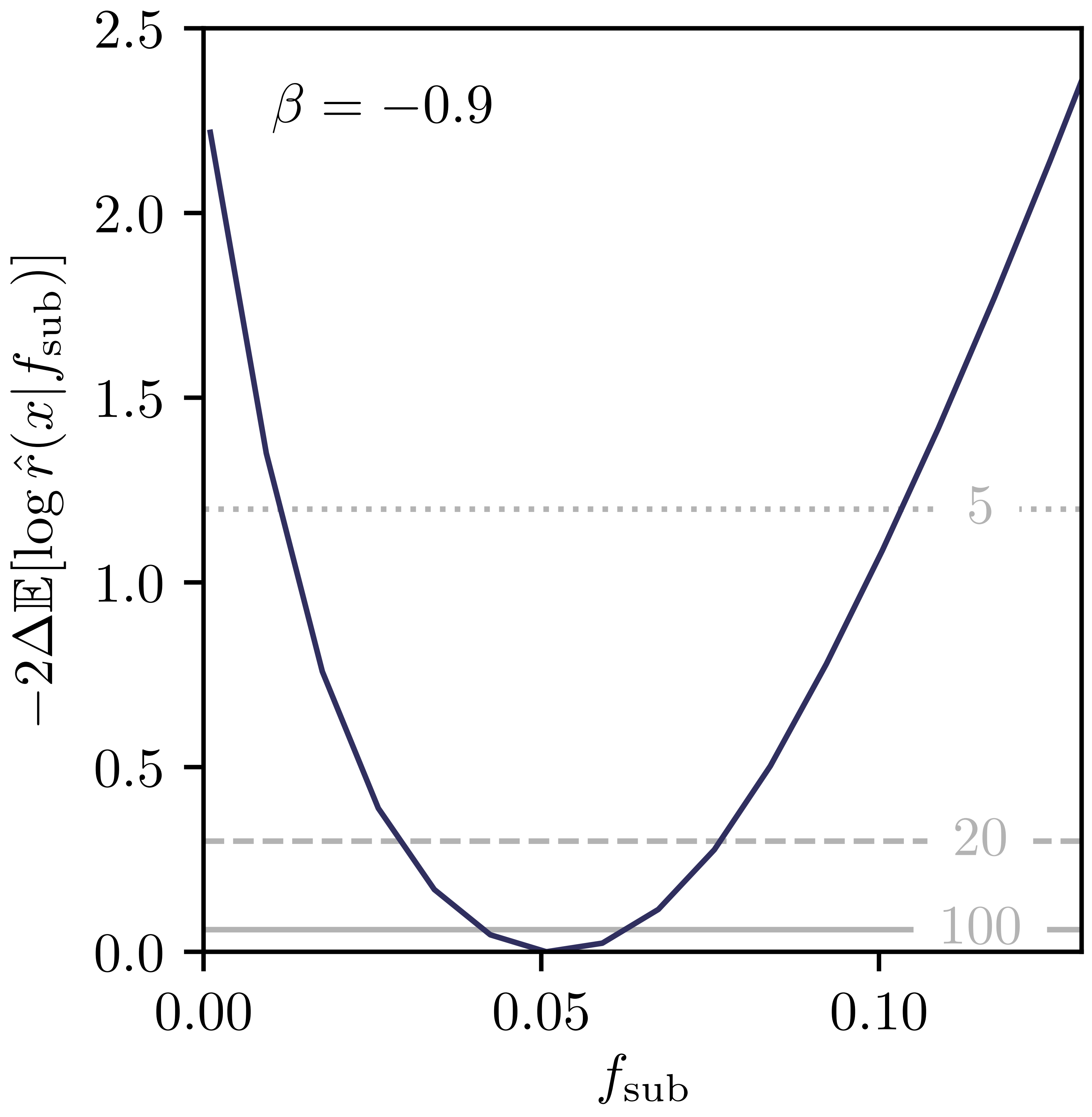
$f_{\rm sub} = 0.05,\ \beta = -0.9$



Inferred likelihood ratios

$f_{\mathrm{sub}} = 0.05,\ \beta = -0.9$

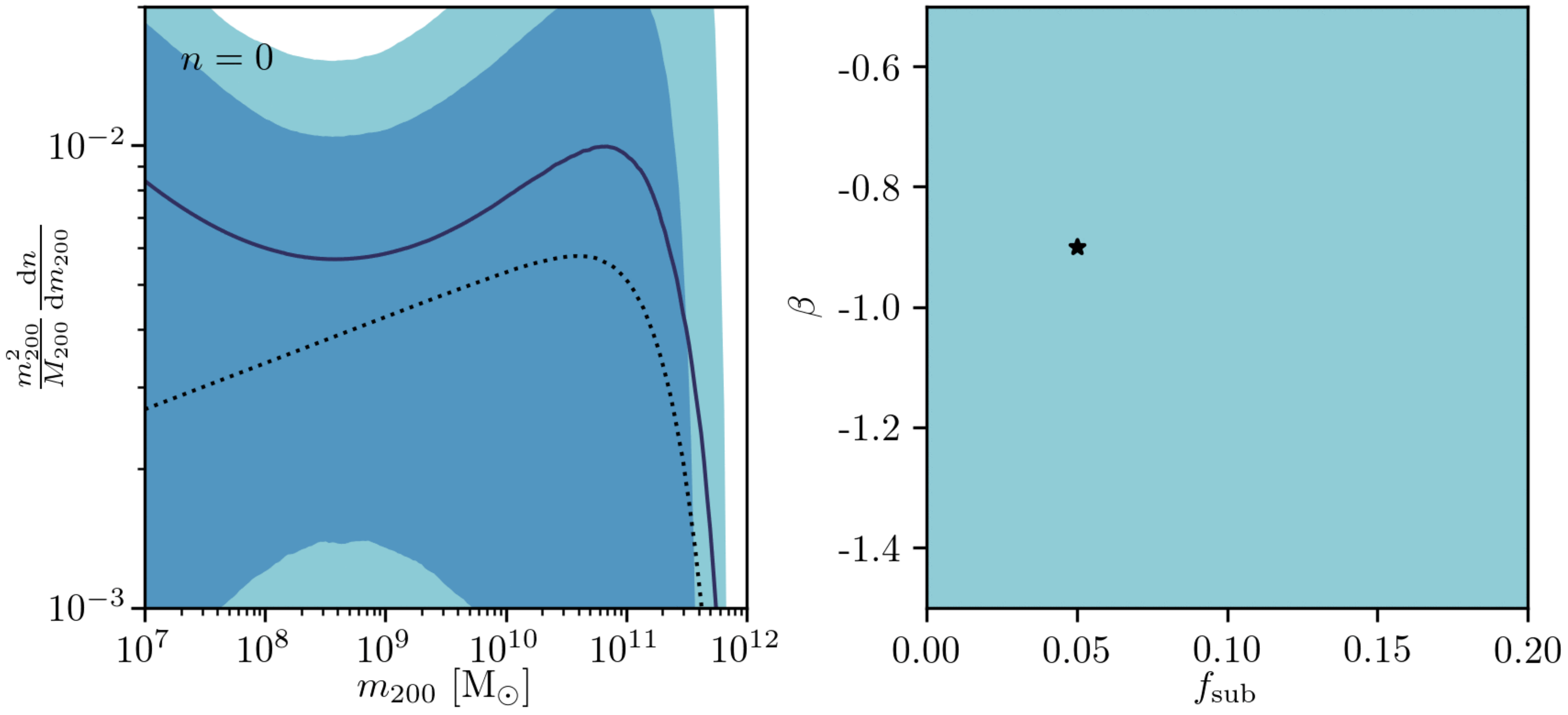




Bayesian interpretation

$f_{\mathrm{sub}} = 0.05,\ \beta = -0.9$

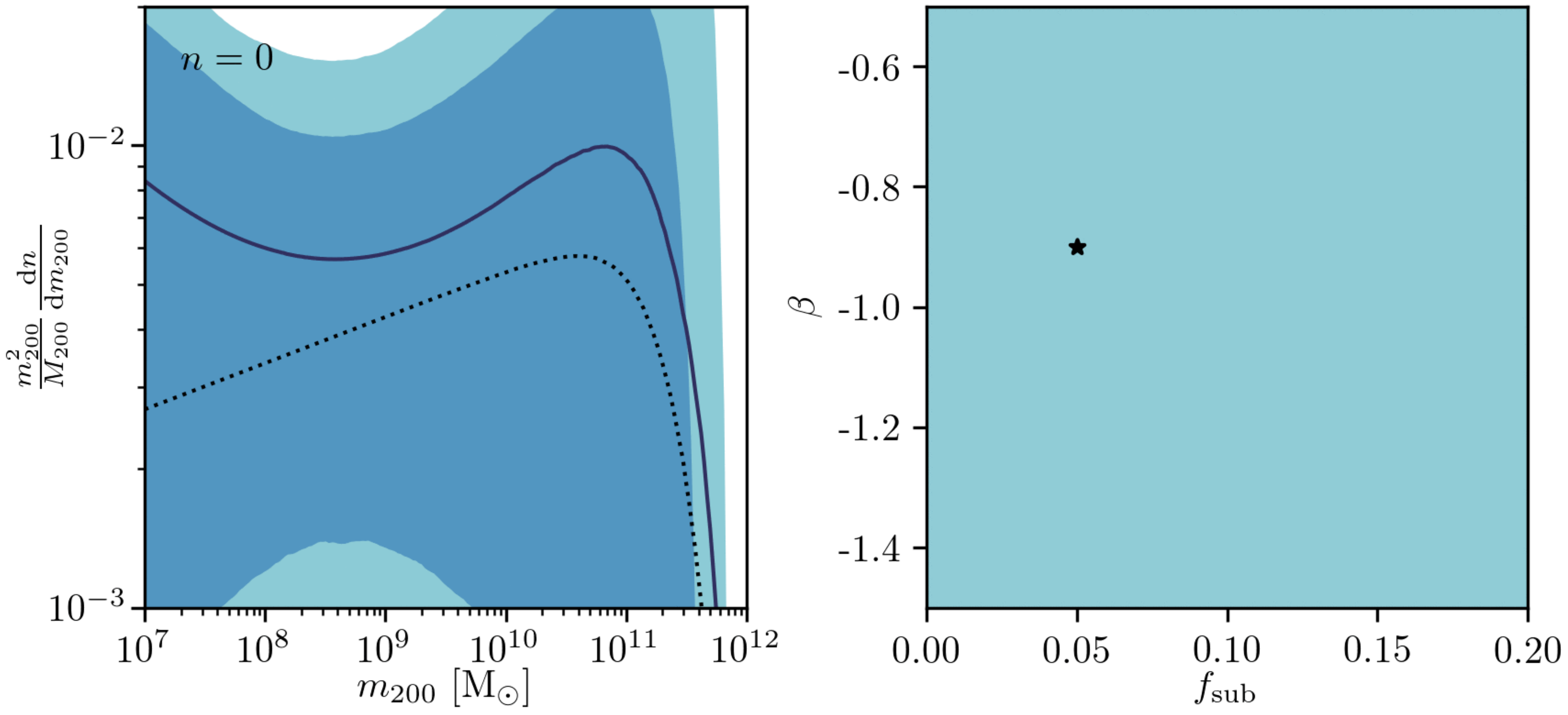
Gaussian prior for $\beta \sim \mathcal{N}(-0.9,\, 0.1)$



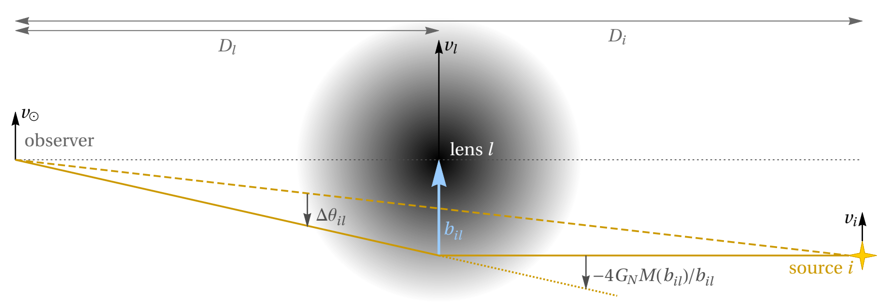
Bayesian interpretation

$f_{\mathrm{sub}} = 0.05, \beta = -0.9$

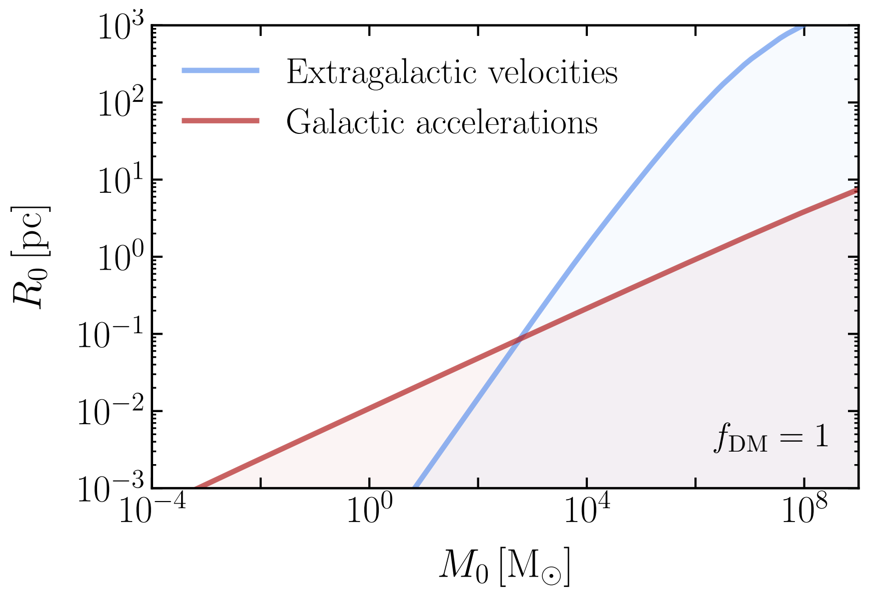
Gaussian prior for $\beta \sim \mathcal{N}(-0.9, 0.1)$



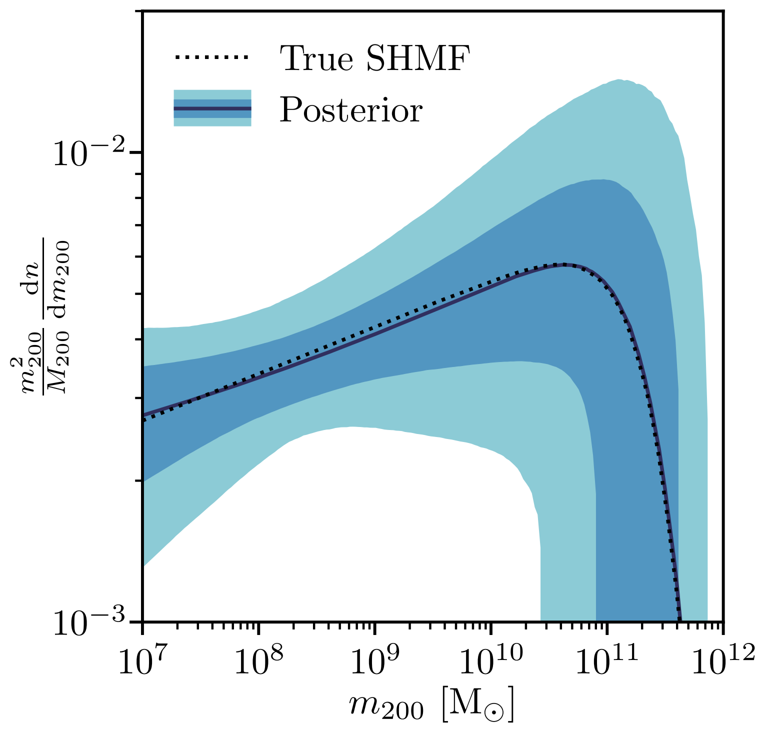
Conclusions



Gravitational Lensing
A probe of dark matter substructure



Inferring Galactic Substructure
With Astrometry & Weak Lensing



Inferring Extragalactic Substructure
With Likelihood-free Inference & Strong Lensing

Thanks!

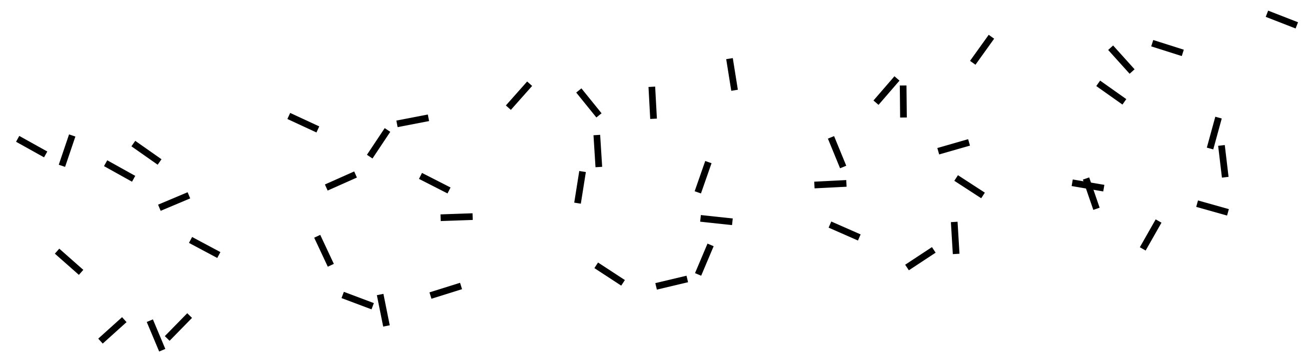

Thanks!

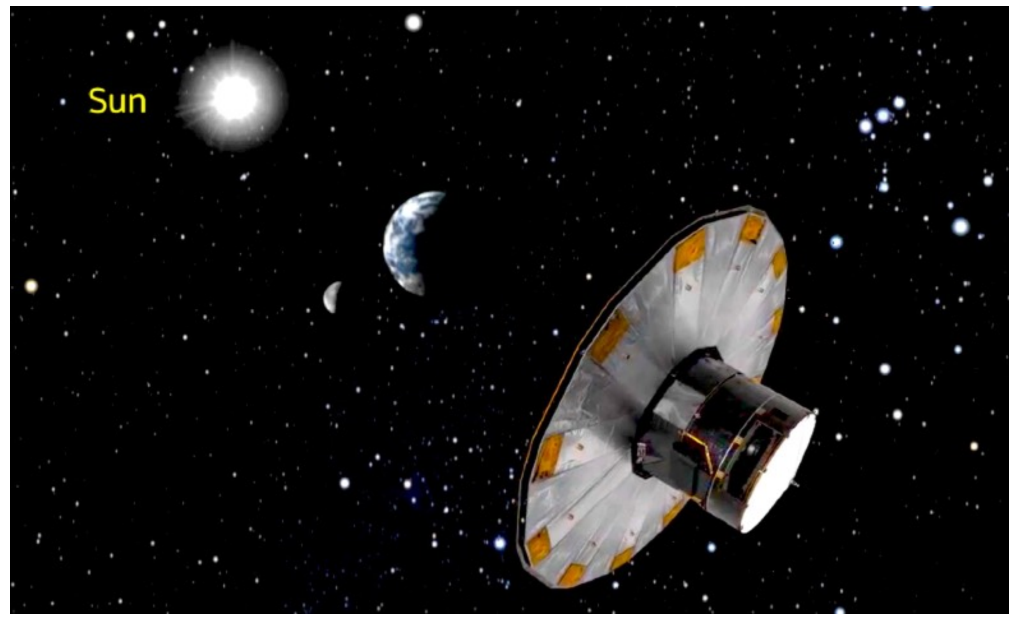
Backup Slides

Astrometry today

Repeatedly measure positions of celestial objects (stars, galaxies...) to get distances (through *parallax*) as well as time-domain information (velocities, accelerations)

$\boldsymbol{\theta}_i(t_n) \longrightarrow$



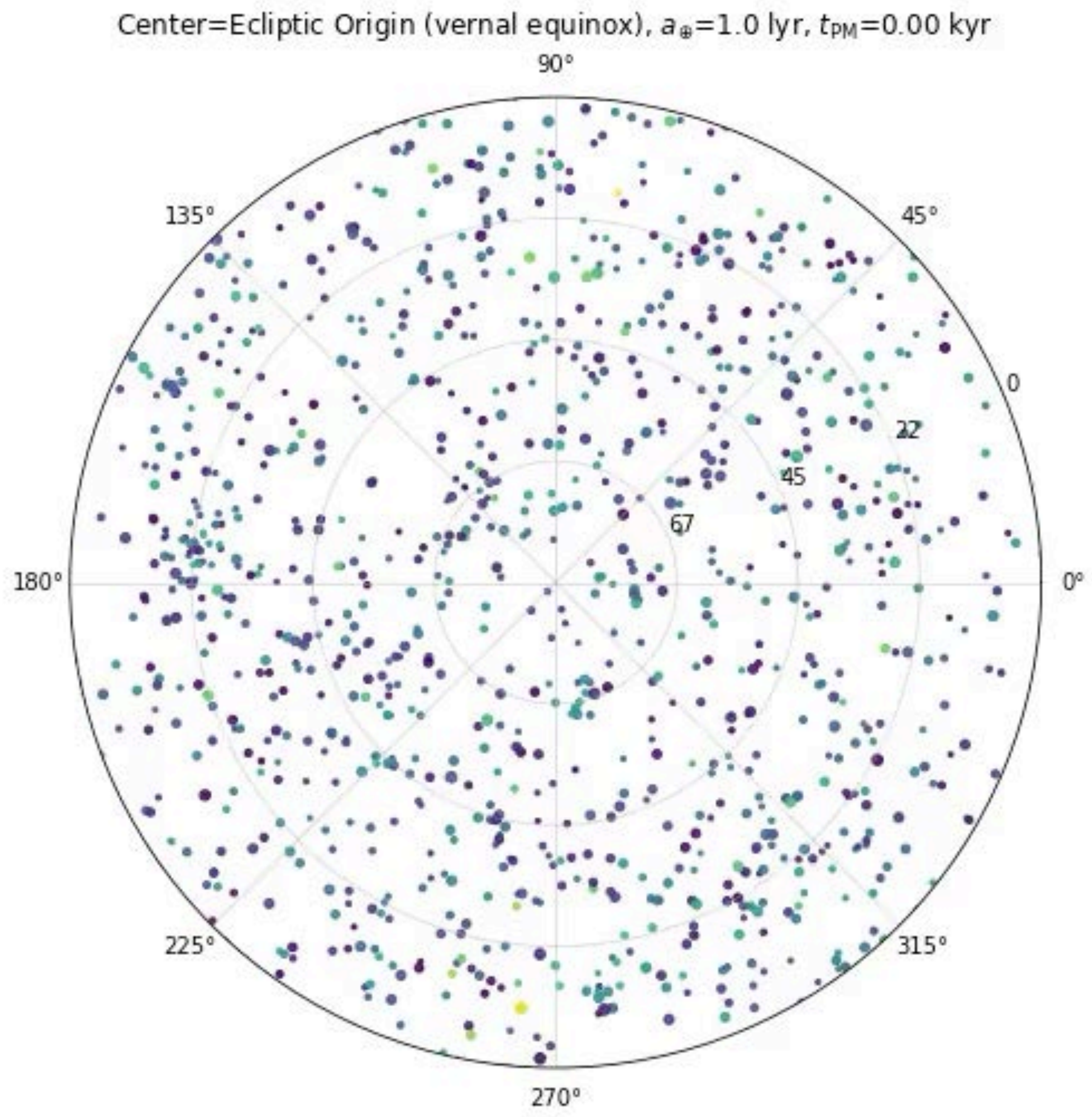


Noise configuration

$\sigma_\mu \sim 100\,\mu\mathrm{as\,yr}^{-1}$

$N_q \sim 10^6$

$f_\mathrm{sky} \sim 1$



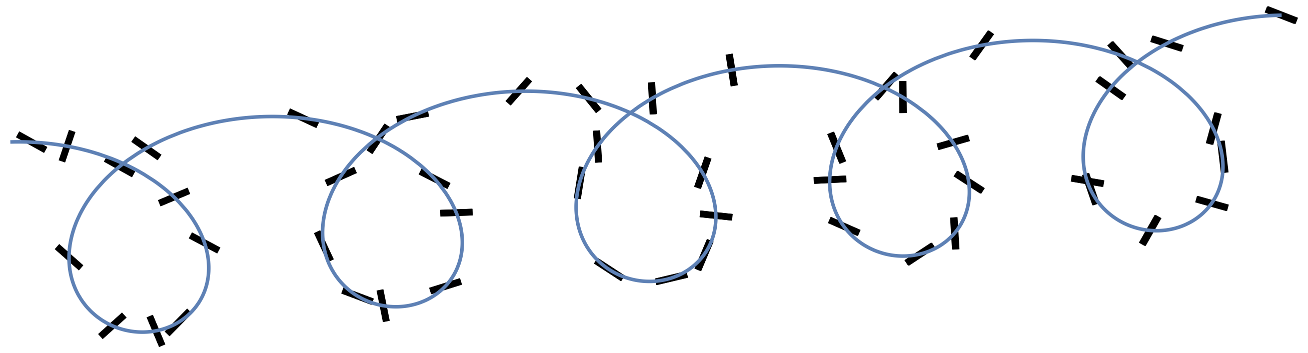
Credits: Erik Tollerud

https://gist.github.com/eteq/02a0065f15da3b3d8c2a9dea146a2a14

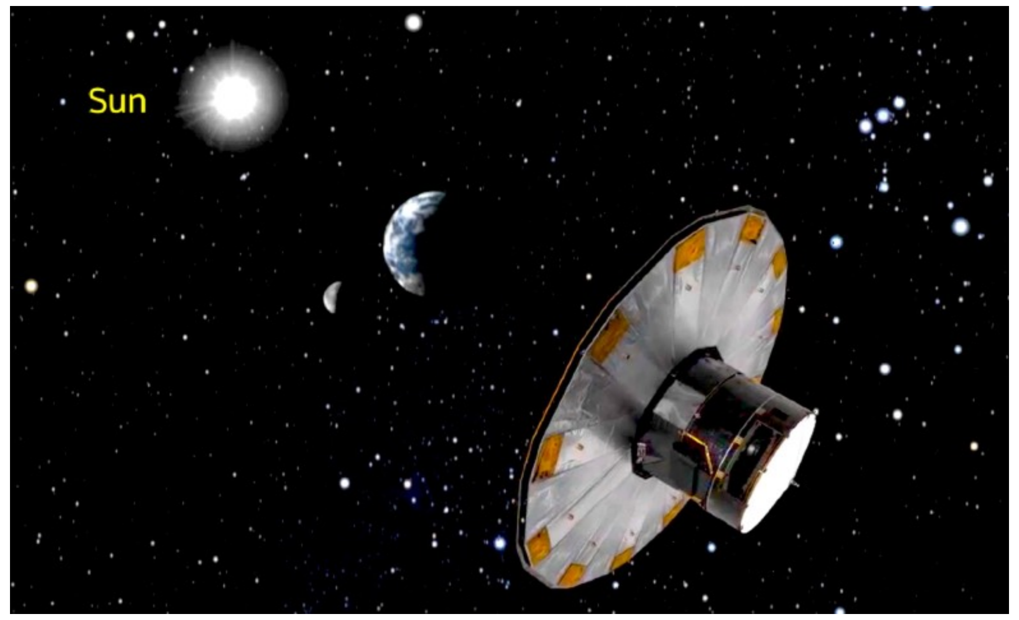
Astrometry today

Repeatedly measure positions of celestial objects (stars, galaxies...) to get distances (through *parallax*) as well as time-domain information (velocities, accelerations)

$\boldsymbol{\theta}_i(t_n) \longrightarrow$



$\longrightarrow \bar{\boldsymbol{\theta}}_i, \bar{\boldsymbol{\mu}}_i, \frac{\varpi_i}{\mathrm{mas}} = \frac{\mathrm{kpc}}{D_i}$

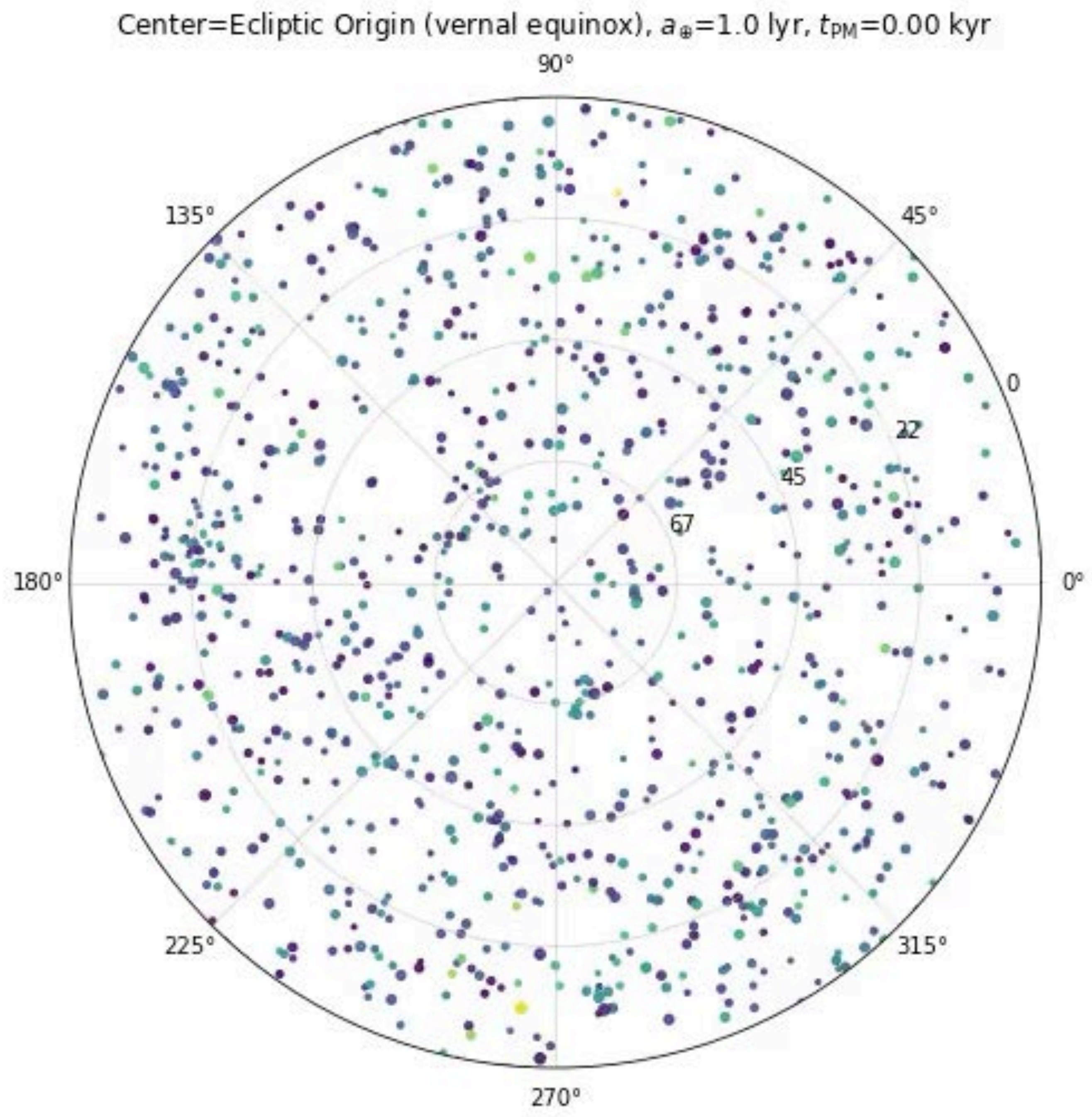


Noise configuration

$\sigma_\mu \sim 100\,\mu\mathrm{as\,yr}^{-1}$

$N_q \sim 10^6$

$f_\mathrm{sky} \sim 1$



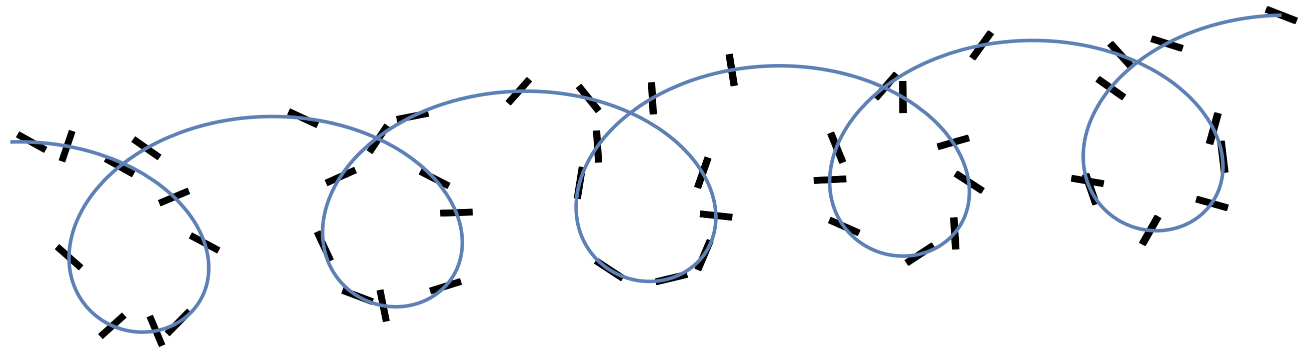
Credits: Erik Tollerud

https://gist.github.com/eteq/02a0065f15da3b3d8c2a9dea146a2a14

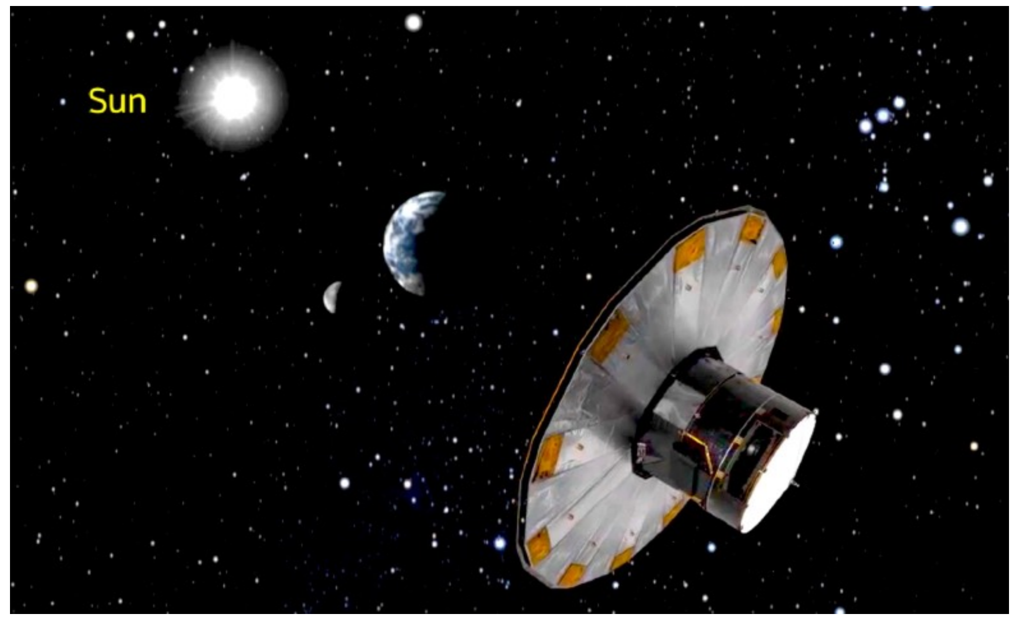
Astrometry today

Repeatedly measure positions of celestial objects (stars, galaxies…) to get distances (through *parallax*) as well as time-domain information (velocities, accelerations)

$\boldsymbol{\theta}_i(t_n) \longrightarrow$



$\longrightarrow \bar{\boldsymbol{\theta}}_i, \bar{\boldsymbol{\mu}}_i, \frac{\varpi_i}{\mathrm{mas}} = \frac{\mathrm{kpc}}{D_i}$

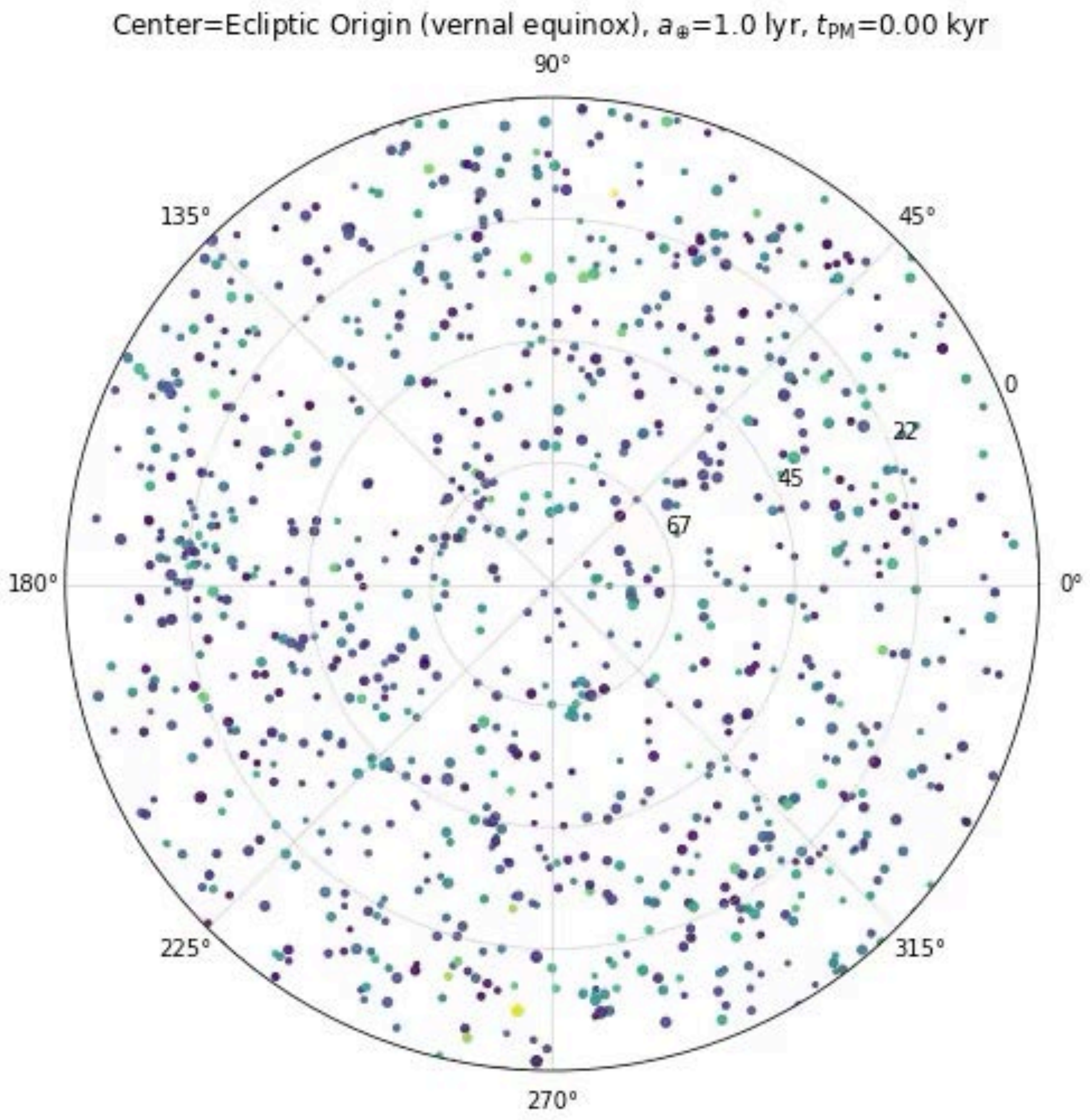


Noise configuration

$\sigma_\mu \sim 100\,\mu\mathrm{as\,yr}^{-1}$

$N_q \sim 10^6$

$f_{\mathrm{sky}} \sim 1$



Credits: Erik Tollerud

https://gist.github.com/eteq/02a0065f15da3b3d8c2a9dea146a2a14

Lens equation

$$\vec{\theta} D_{\mathsf{S}} = \vec{\beta} D_{\mathsf{S}} + \hat{\vec{\alpha}} D_{\mathsf{LS}}$$

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta})$$

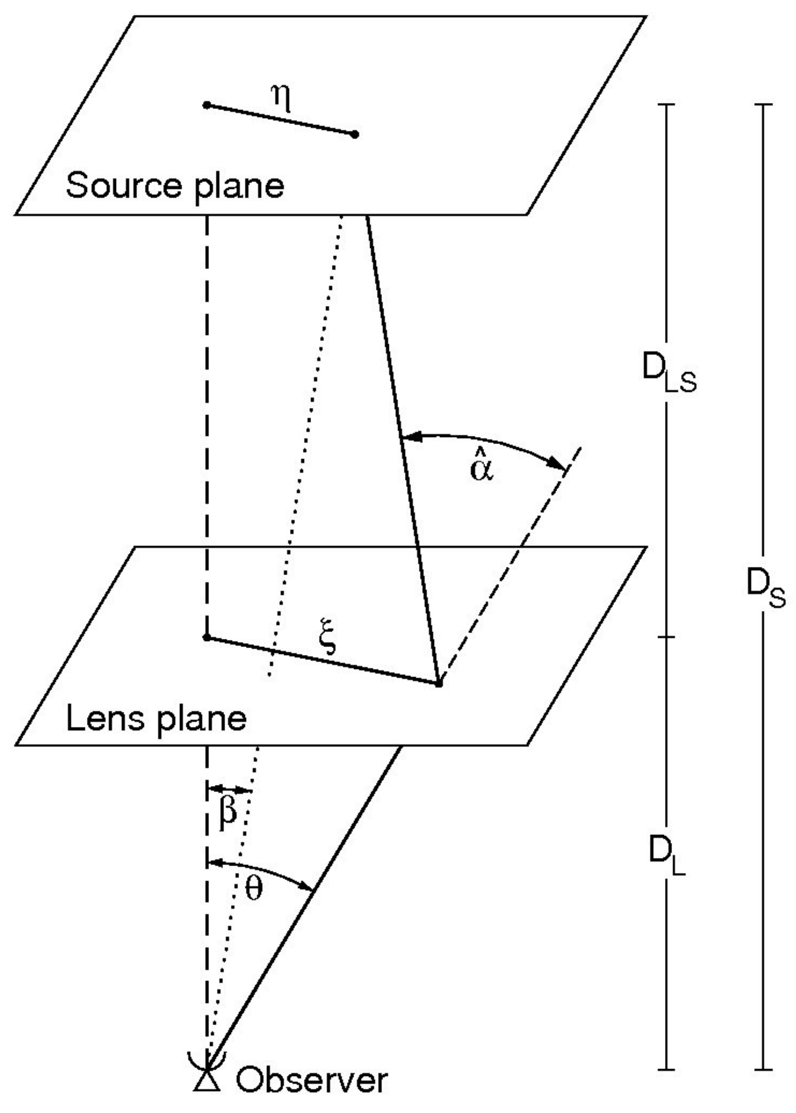
$$\vec{\alpha}(\vec{\theta}) \equiv \frac{D_{\mathsf{LS}}}{D_{\mathsf{S}}} \hat{\vec{\alpha}}(\vec{\theta})$$

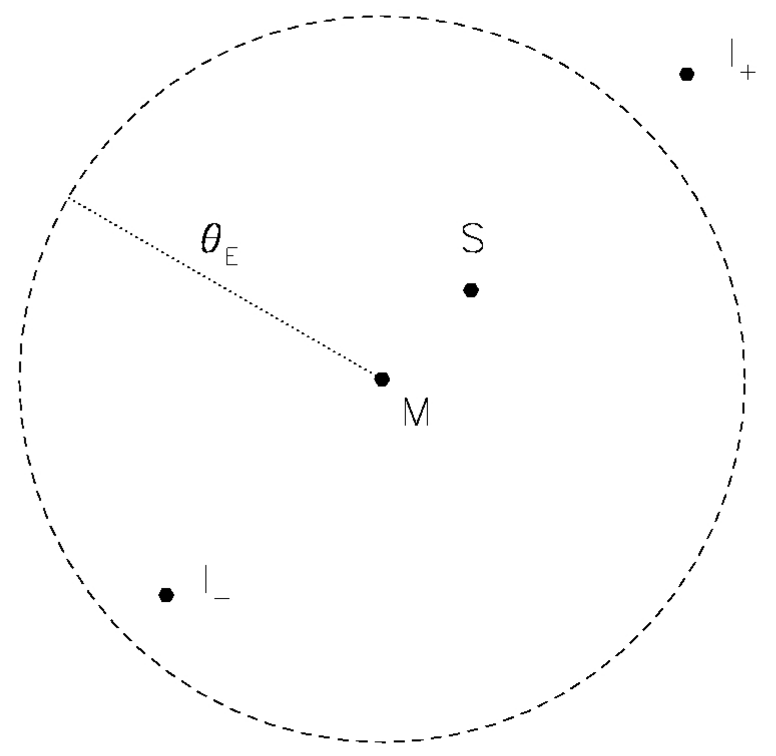
$$\hat{\vec{\alpha}} = -\frac{4GM}{c^2 b} \vec{e}_r$$

$$\beta = \theta - \frac{\theta_E^2}{\theta}$$

$$x_{\pm} = \frac{1}{2}\left[y \pm \sqrt{y^2 + 4}\right]$$

$y = \beta/\theta_E$ and $x = \theta/\theta_E$





Ultralight scalar field dark matter

An alternative to CDM is dark matter consisting of ultralight scalars $\sim 10^{-22}$ eV in mass

$$\phi(x) = \int \frac{đ^3 k}{\sqrt{2k^0}} \left(a_{\mathbf{k}} e^{-ik\cdot x} + a_{\mathbf{k}}^{\dagger} e^{+ik\cdot x} \right)$$

Large de Broglie wavelength generally suppresses amount of *bound* small-scale structure

$$\frac{\lambda}{2\pi} = \frac{\hbar}{mv} = 1.92\,\mathrm{kpc} \left(\frac{10^{-22}\mathrm{eV}}{m} \right) \left(\frac{10\mathrm{kms}^{-1}}{v} \right)$$

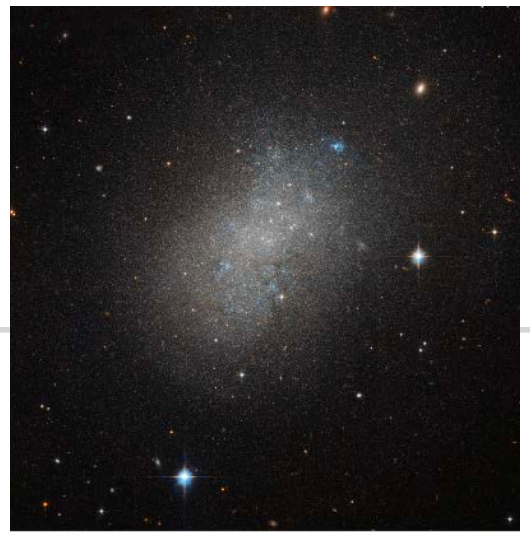
However, have *irreducible, unbound density fluctuations* at this scale

Typical mass and size of fluctuations

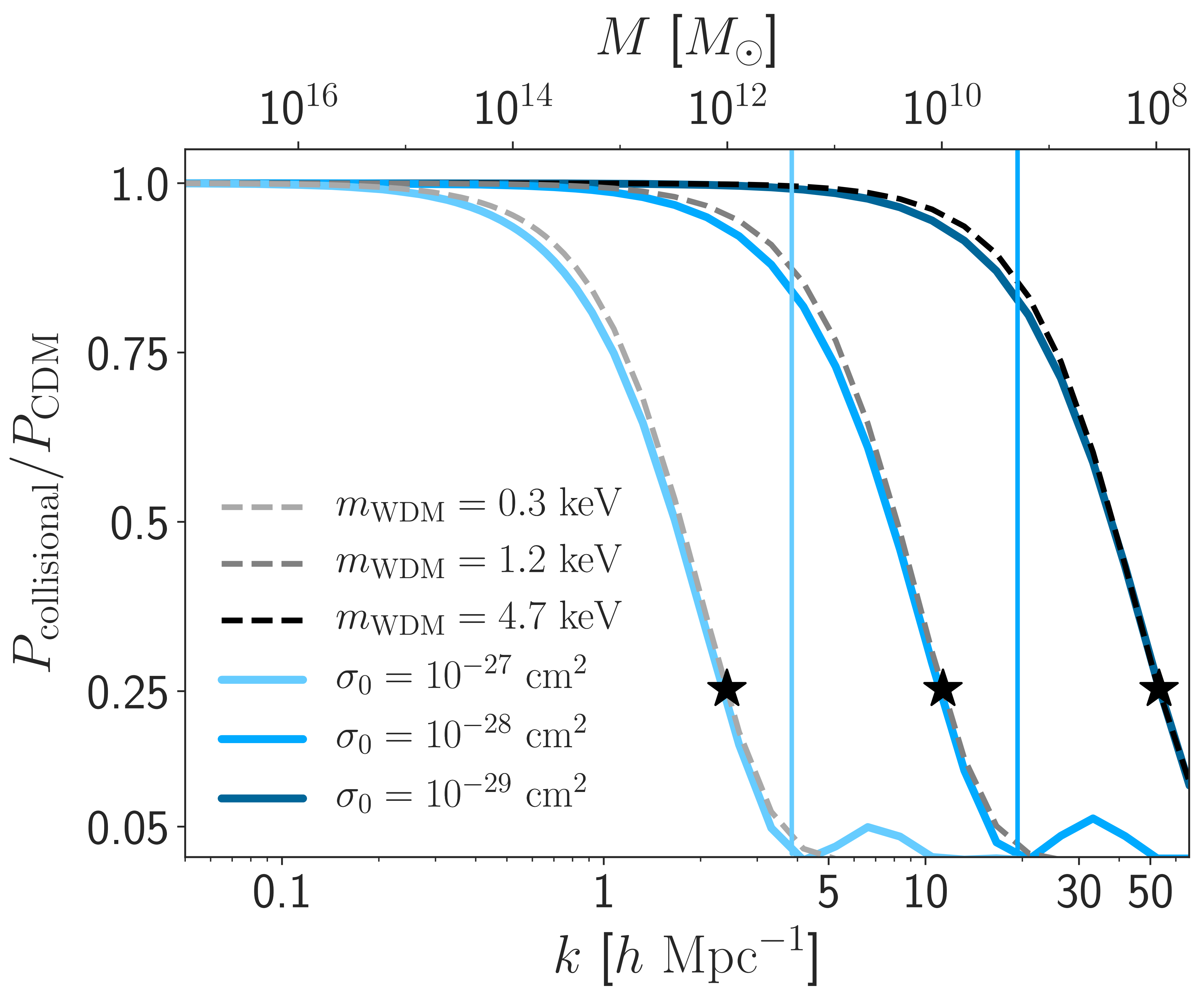
$$M_0 = C\rho_0 \left(\frac{\pi}{\sigma_k} \right)^3 \approx 5 \times 10^5 M_\odot\, C m_{22}^{-3}$$

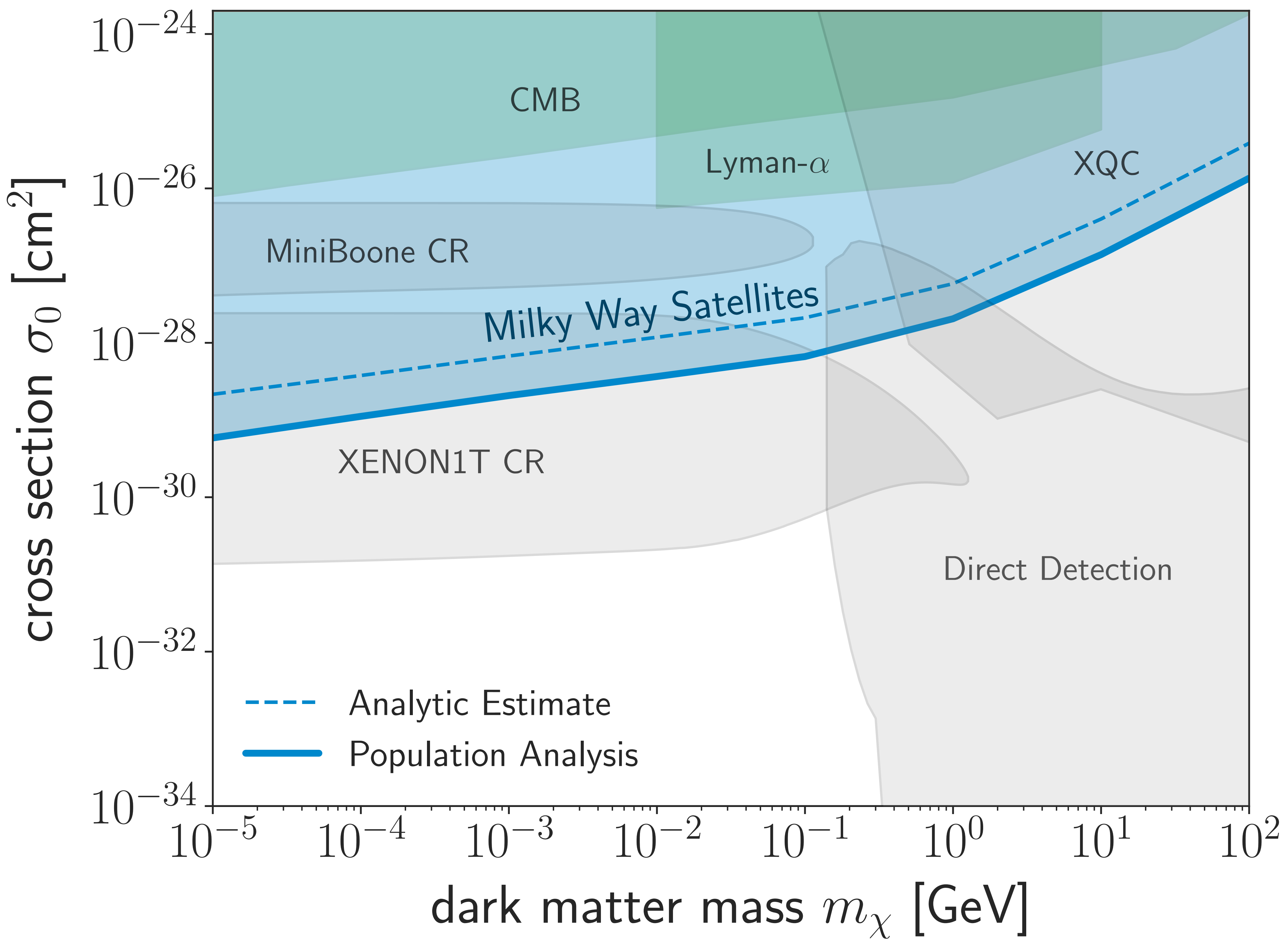
$$R_0 = \frac{1}{2\sigma_k} \approx 58\,\mathrm{pc}\, m_{22}^{-1}$$

Dwarf galaxies as probed of DM



Constraints on DM-baryon interactions from properties of Milky Way satellites

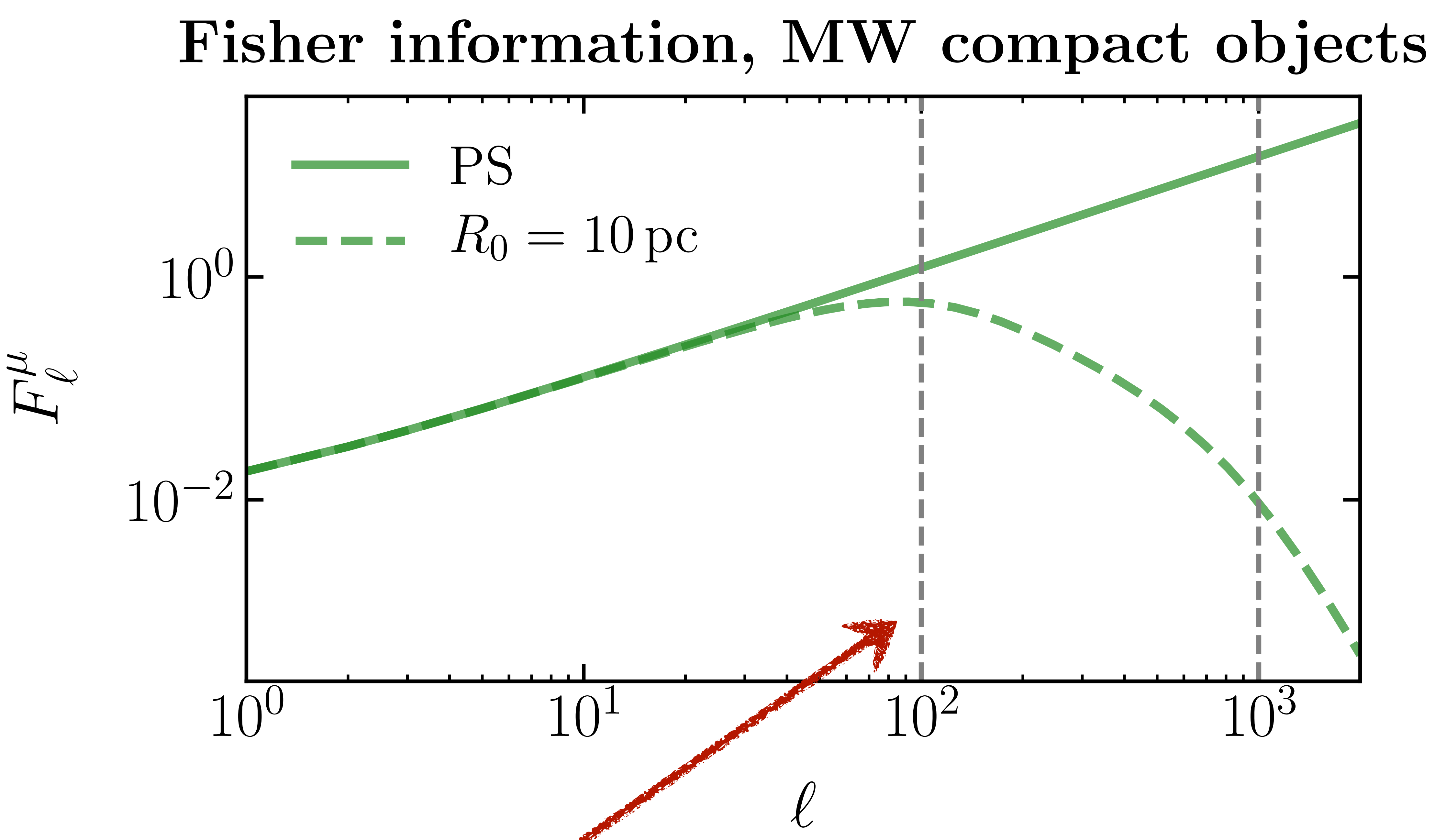


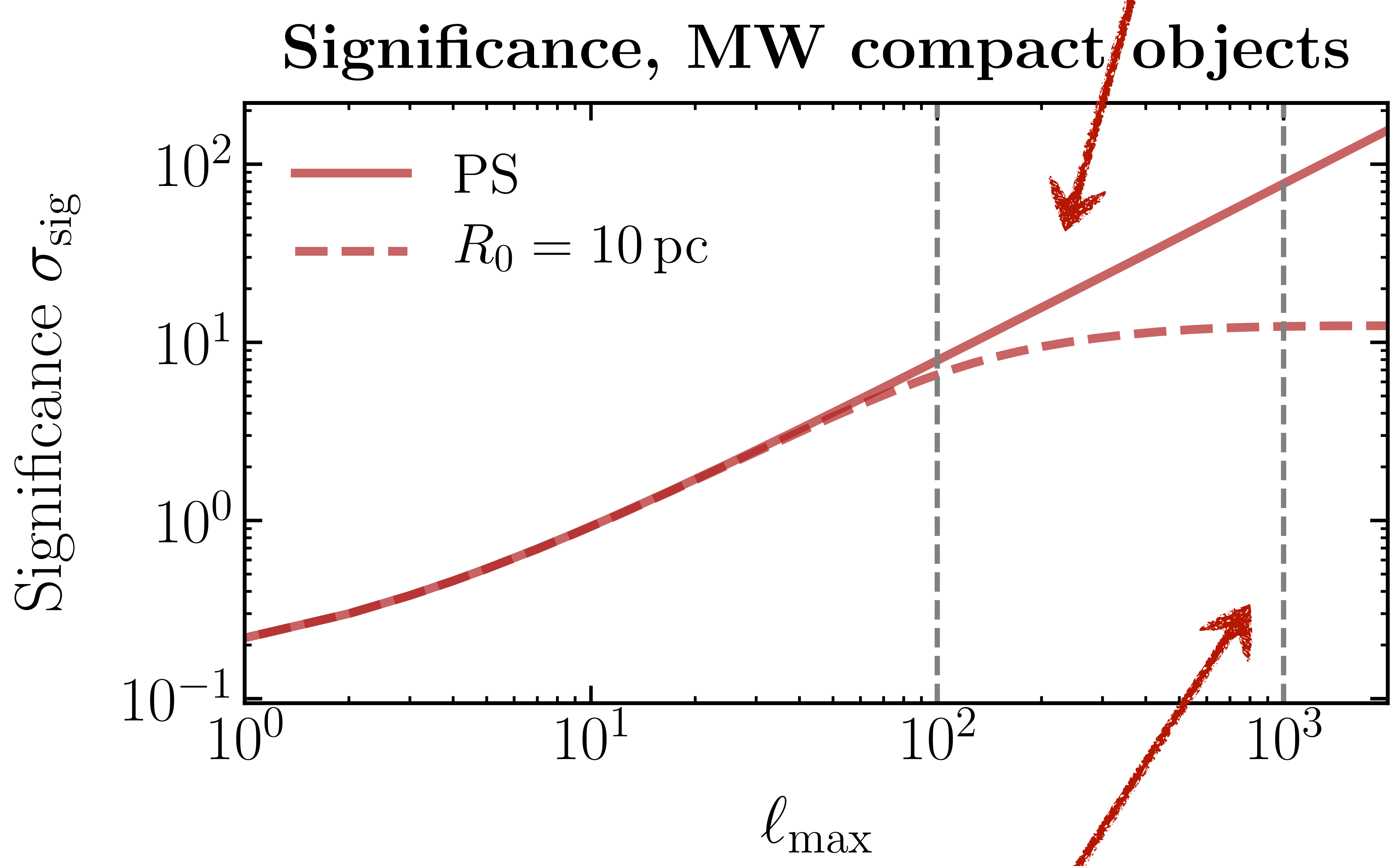


Nadler et al [1904.10000]

Compact objects in the Milky Way: intuition

For point-like objects sensitivity increases linearly with $\ell_{\rm max}$





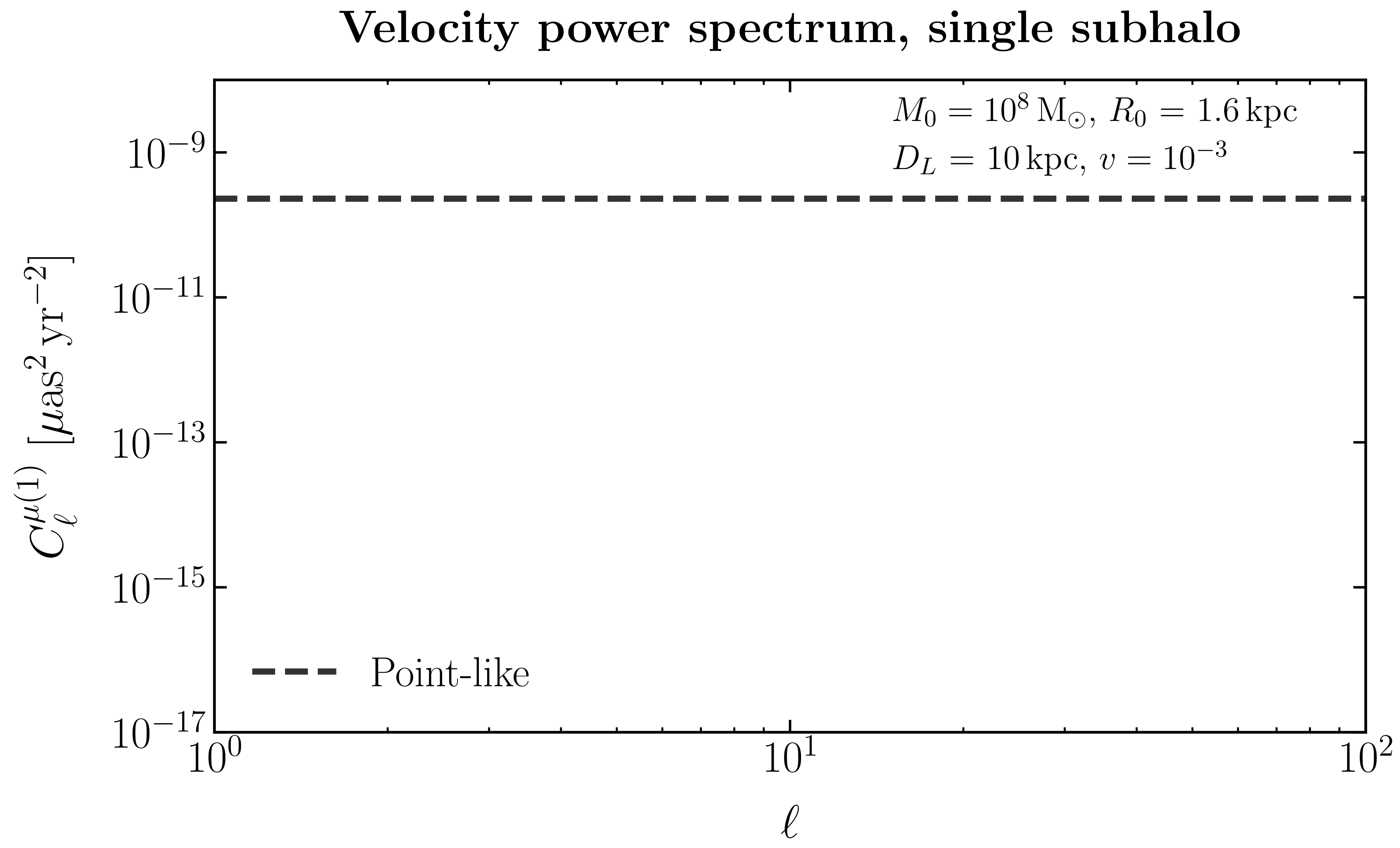
Maximum information comes from scales

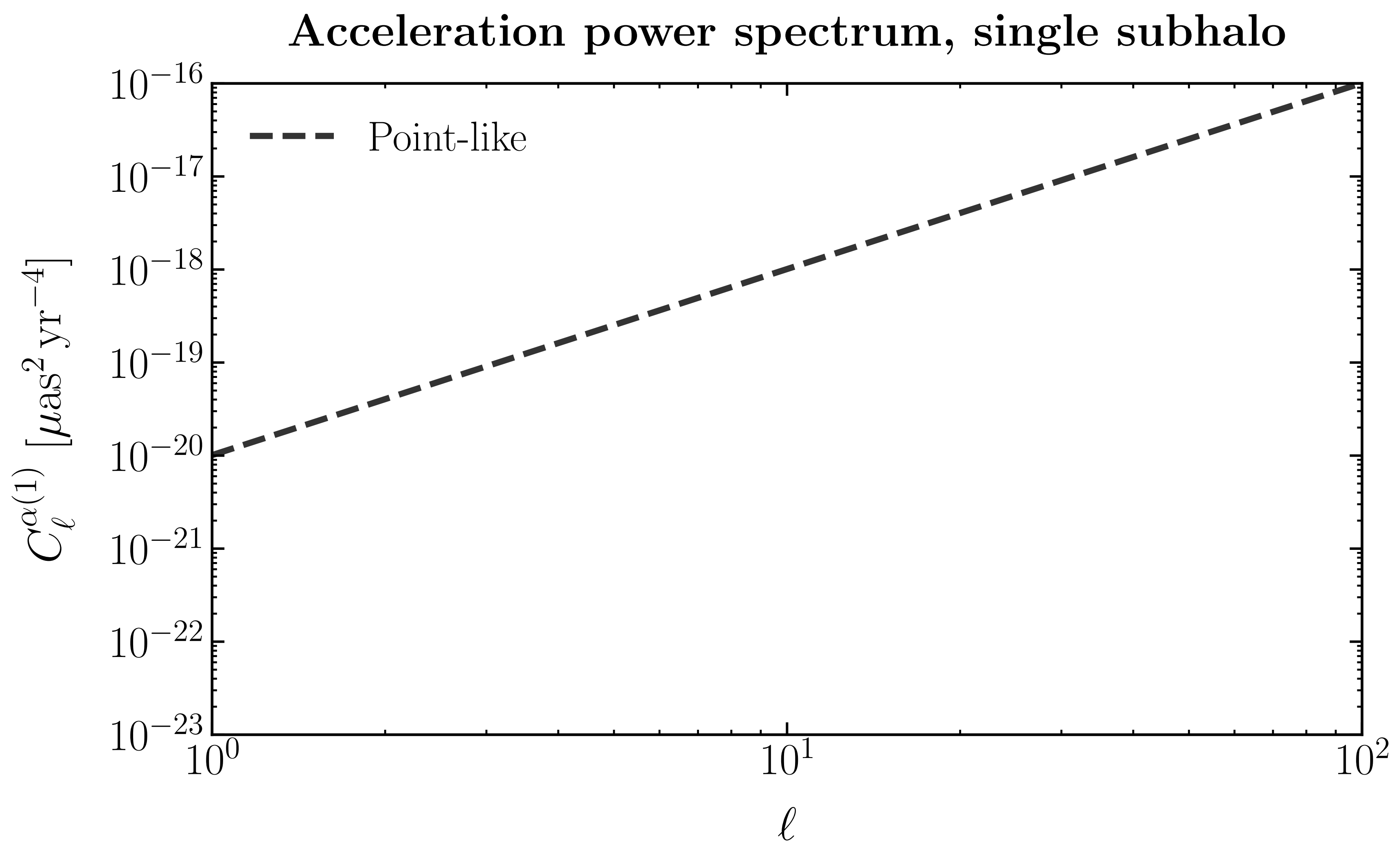
$\ell \sim 0.5\,\mathrm{kpc}/R_0$

Significance plateaus around

$\ell \sim 10\,\mathrm{kpc}/R_0$

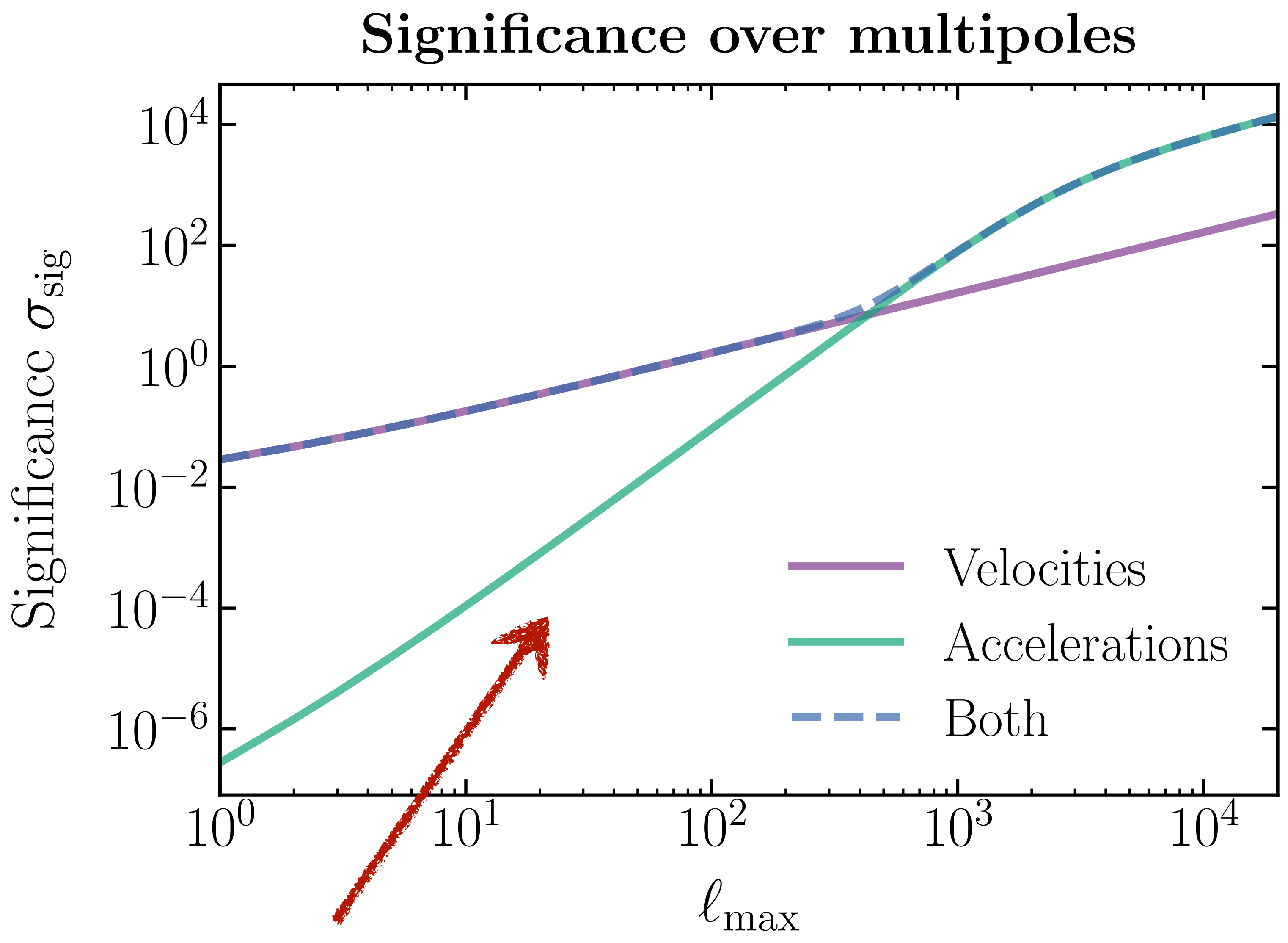
The lensing signal: point lenses





$$C_\ell^{\mu(1)} \simeq \left(\frac{4G_N M_0 v}{D_l^2}\right)^2 \frac{\pi}{2}$$

Acceleration preferentially probes smaller scales



Significance increases linearly with smaller scale

The lensing signal: realistic lenses

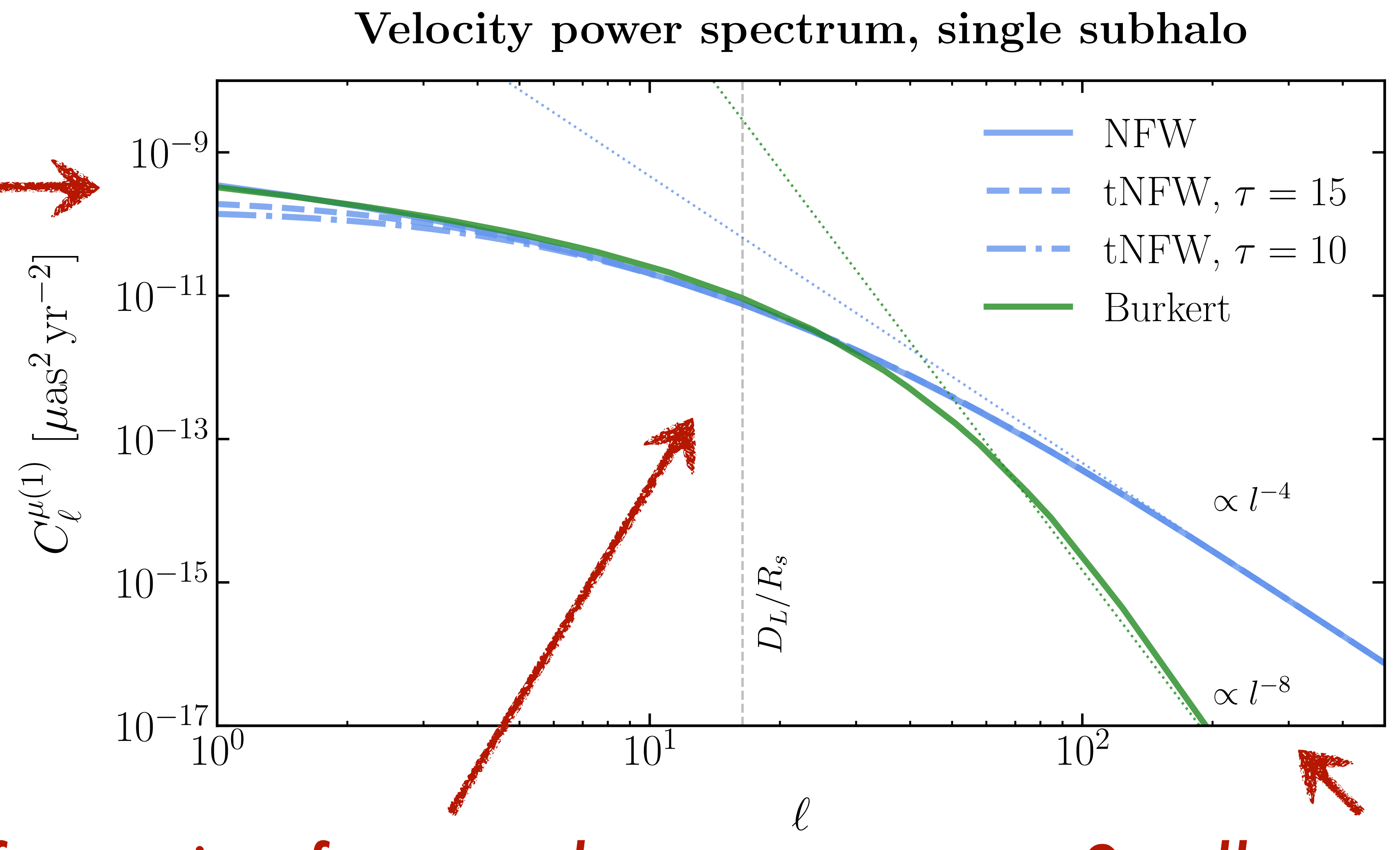
$$\rho_{\mathrm{tNFW}}(r) = \frac{M_s}{4\pi r (r + r_s)^2} \left(\frac{r_t^2}{r^2 + r_t^2} \right)$$

(truncated) NFW profile: cuspy, describes (tidally stripped) CDM subhalos

$$\rho_{\mathrm{Burkert}}(r) = \frac{M_b}{4\pi (r + r_b)(r^2 + r_b^2)}$$

Burkert profile: cored, describes subhalos e.g. in case of DM self-interactions

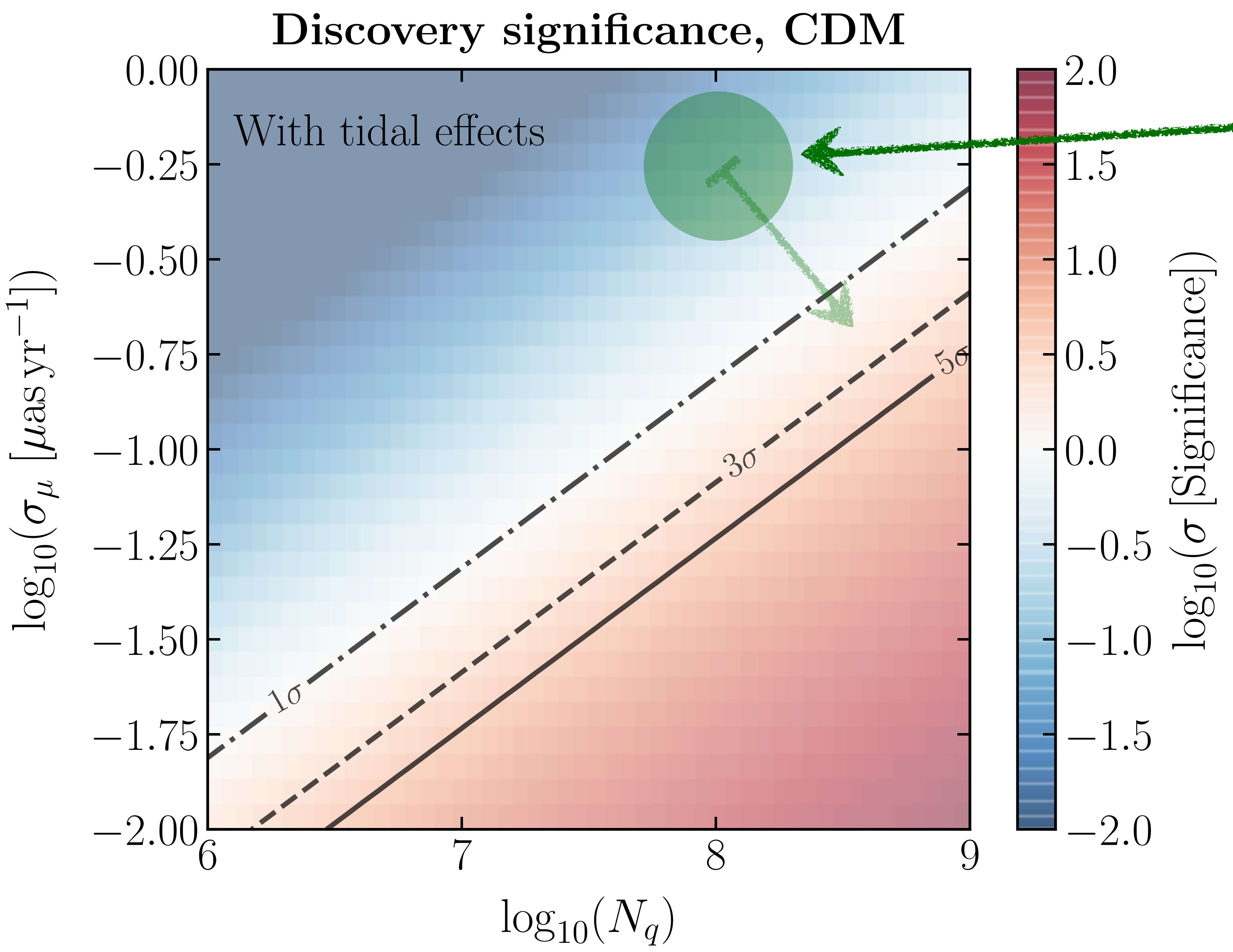
Truncation effects show up at larger scales



Peak information from scales corresponding to subhalo scale radius

Smaller scales probe subhalo core

Cold dark matter: discovery potential



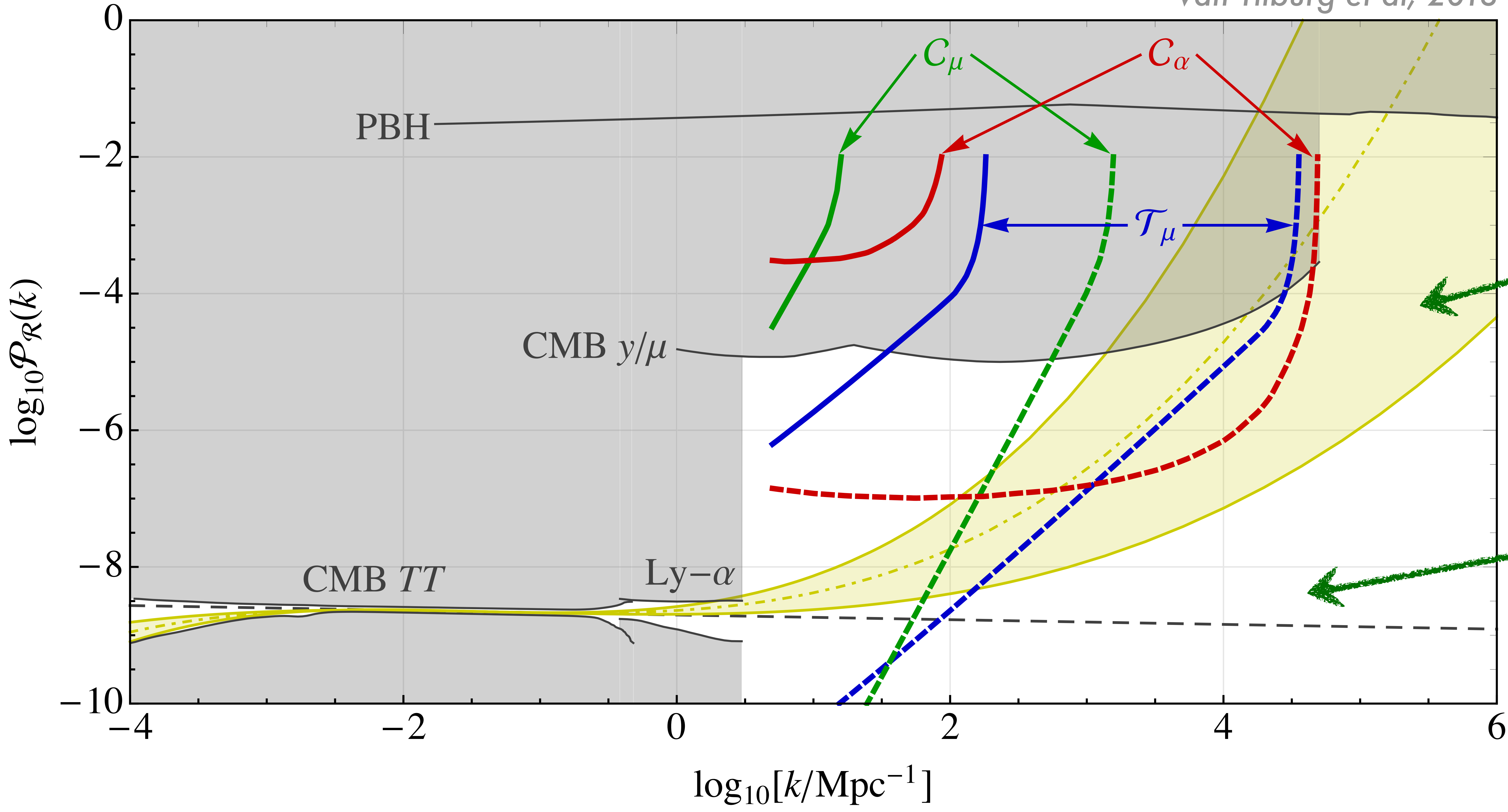
Anticipated noise levels with future surveys

A CDM subhalo population can be detected!

But more difficult for a highly depleted population...

Compact objects from primordial fluctuations

Van Tilburg et al, 2018



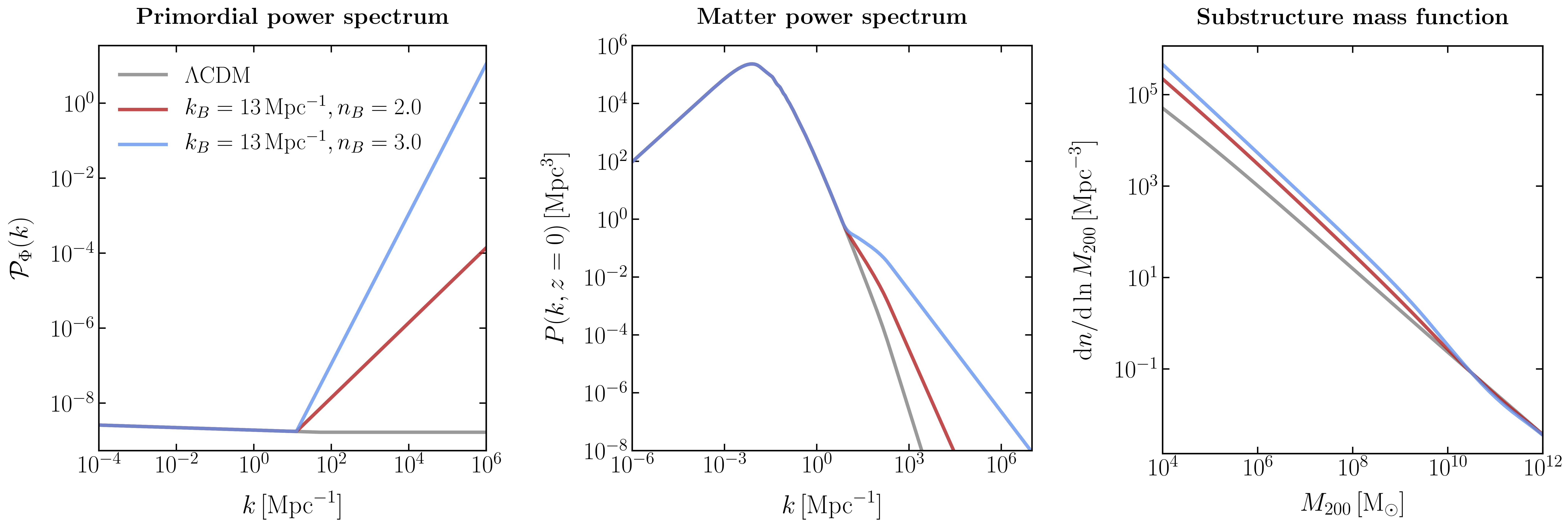
Running spectral index

Constant spectral index

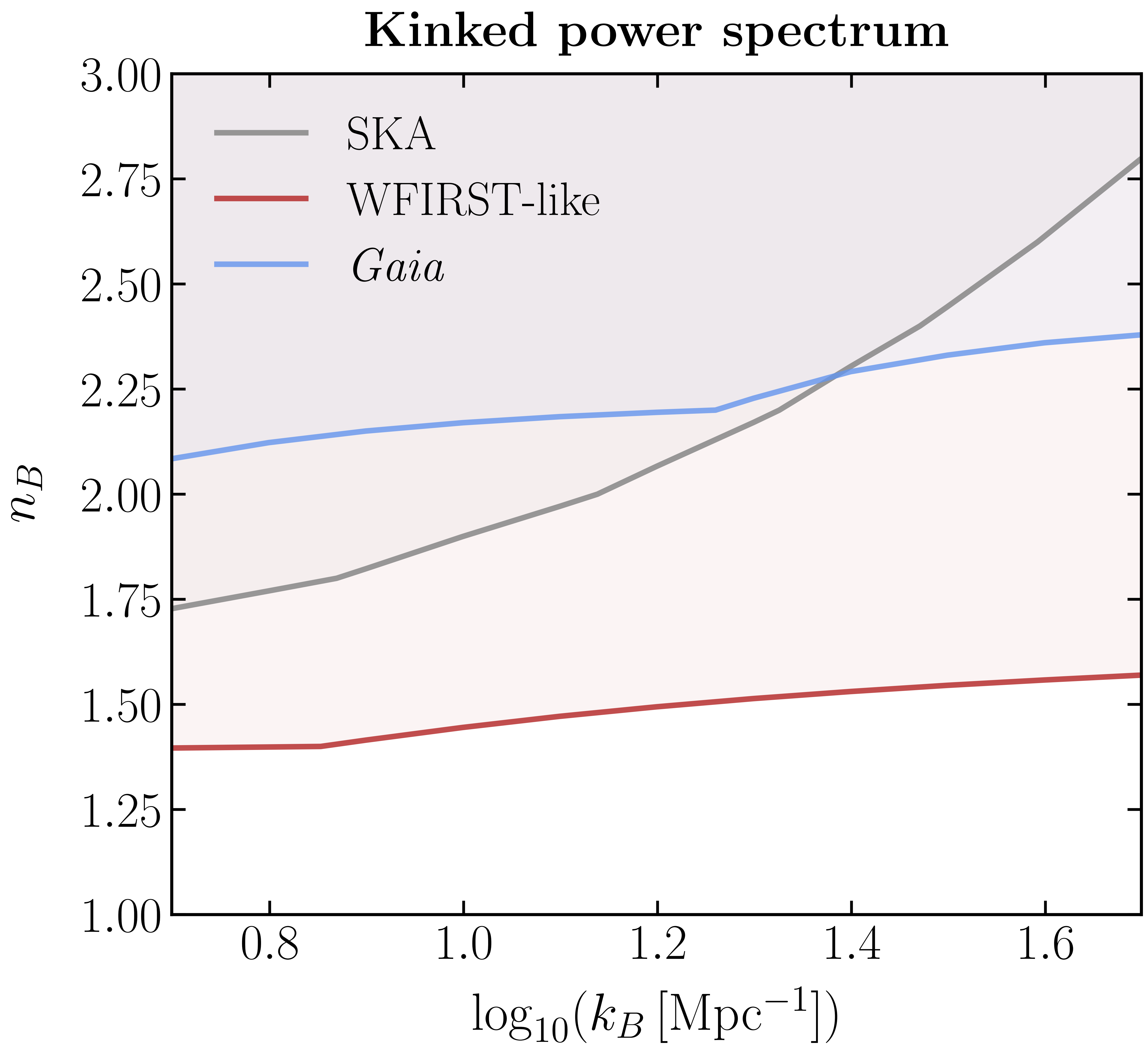
Enhancement in small-scale power is unconstrained and motivated —can have abundance of small clumps

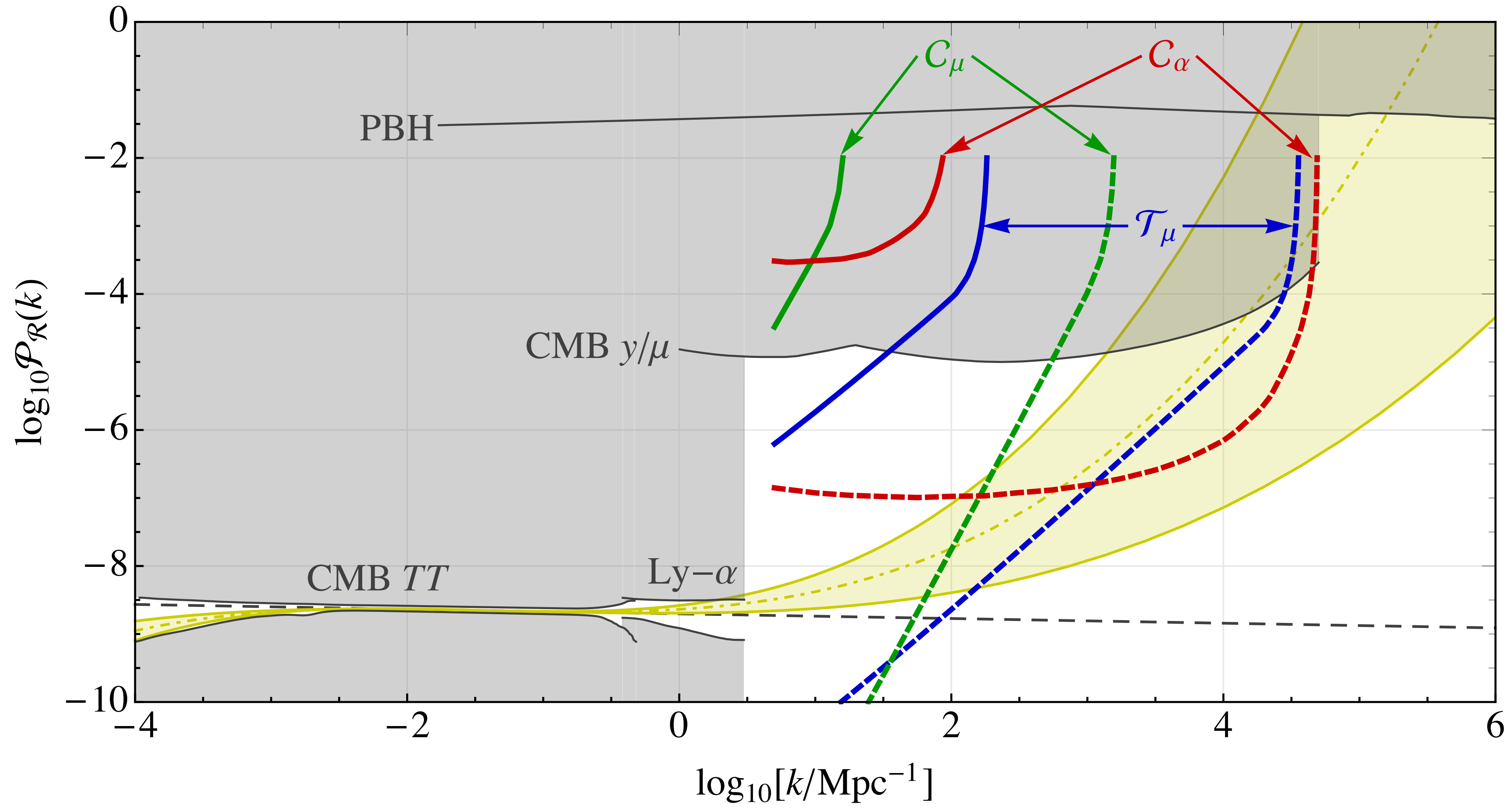
Kinked primordial power spectrum

$$\mathcal{P}_\Phi(k) = \begin{cases} A_s \left(\frac{k}{k_*}\right)^{n_s - 1} & k < k_B \\ A_s \left(\frac{k_B}{k_*}\right)^{n_s - 1} \left(\frac{k}{k_B}\right)^{n_B - 1} & k \geq k_B \end{cases}$$



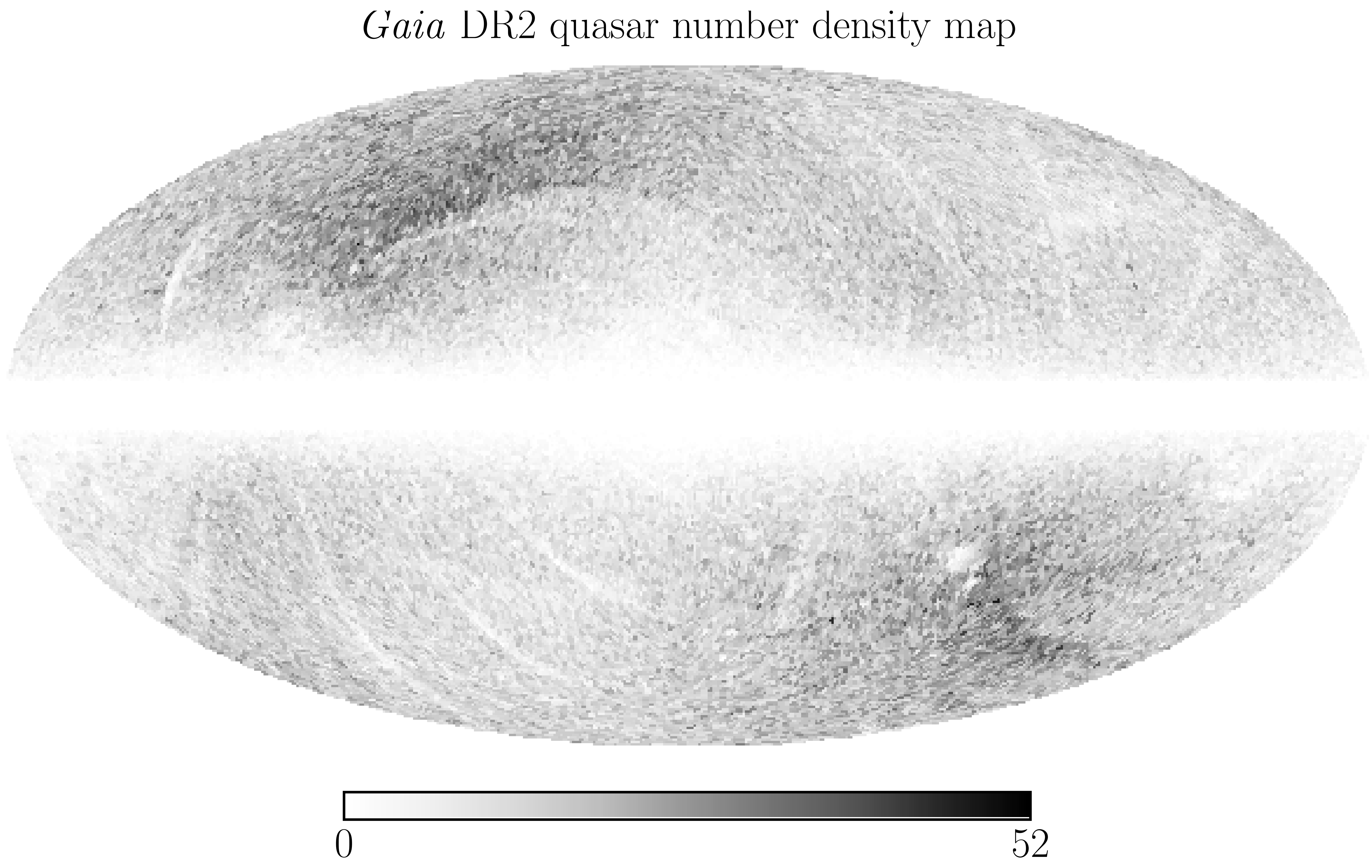
Kinked primordial power spectrum



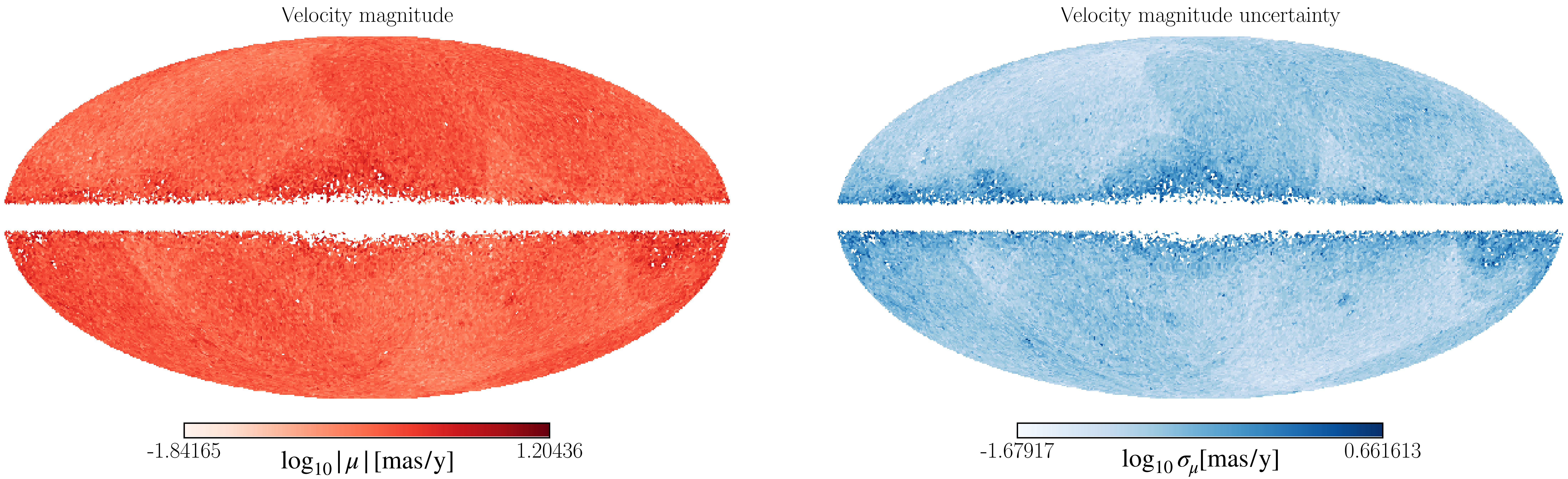


$$\mathcal{P}_\Phi(k) = \begin{cases} A_s\left(\frac{k}{k_*}\right)^{n_s-1} & k < k_B \\ A_s\left(\frac{k_B}{k_*}\right)^{n_s-1}\left(\frac{k}{k_B}\right)^{n_B-1} & k \geq k_B \end{cases}$$

Gaia DR2 quasars

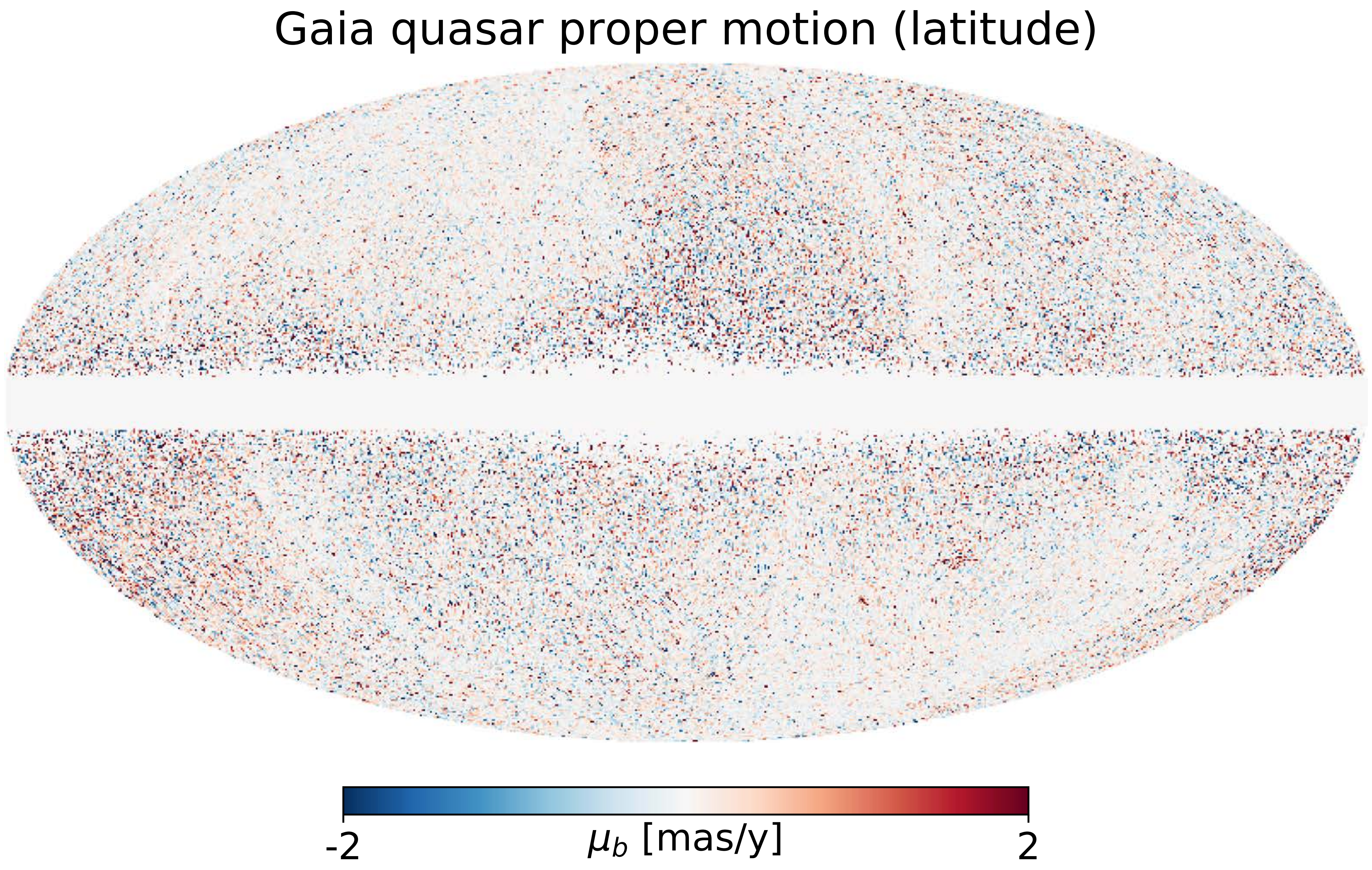


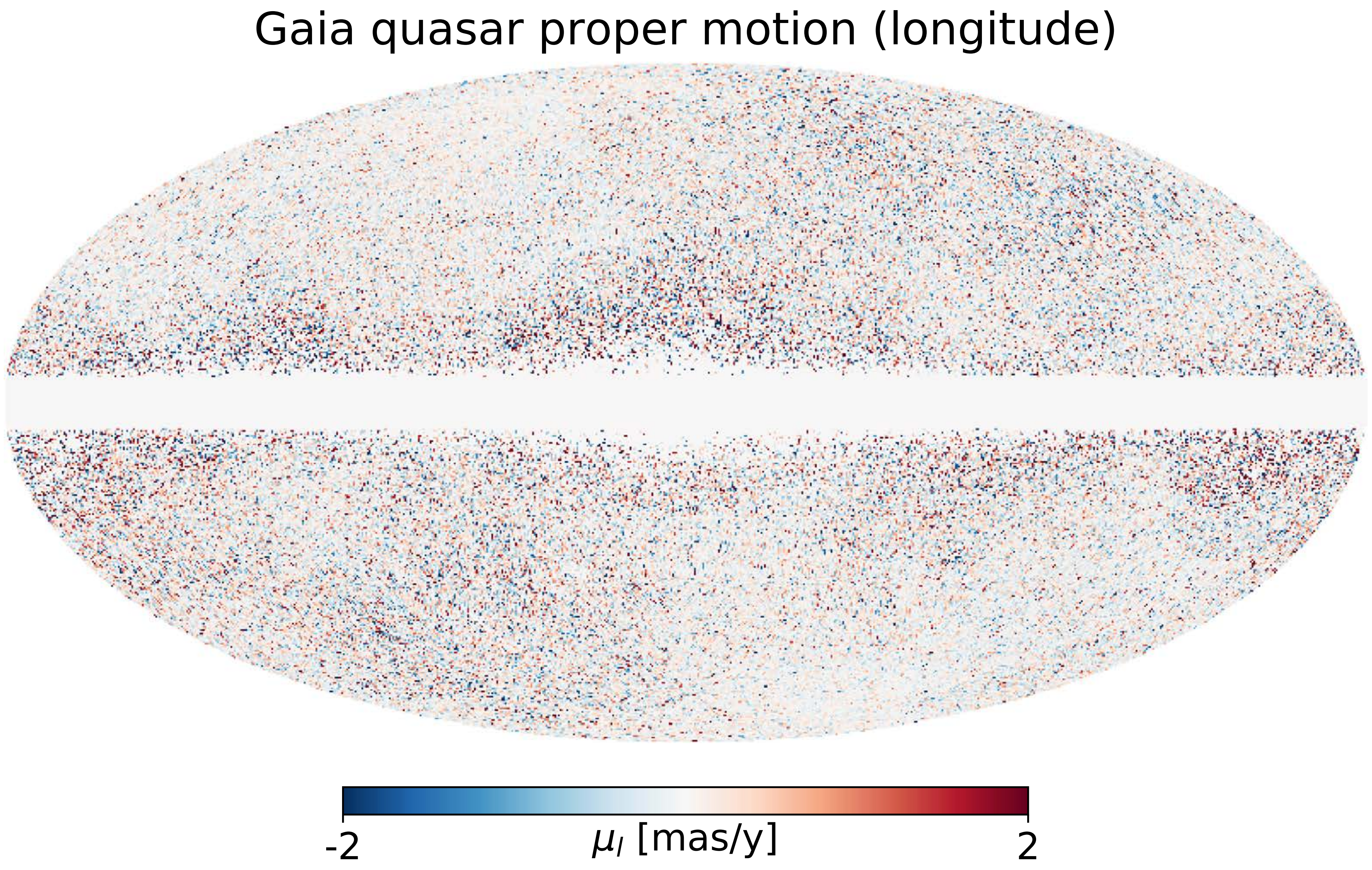
Proper motions



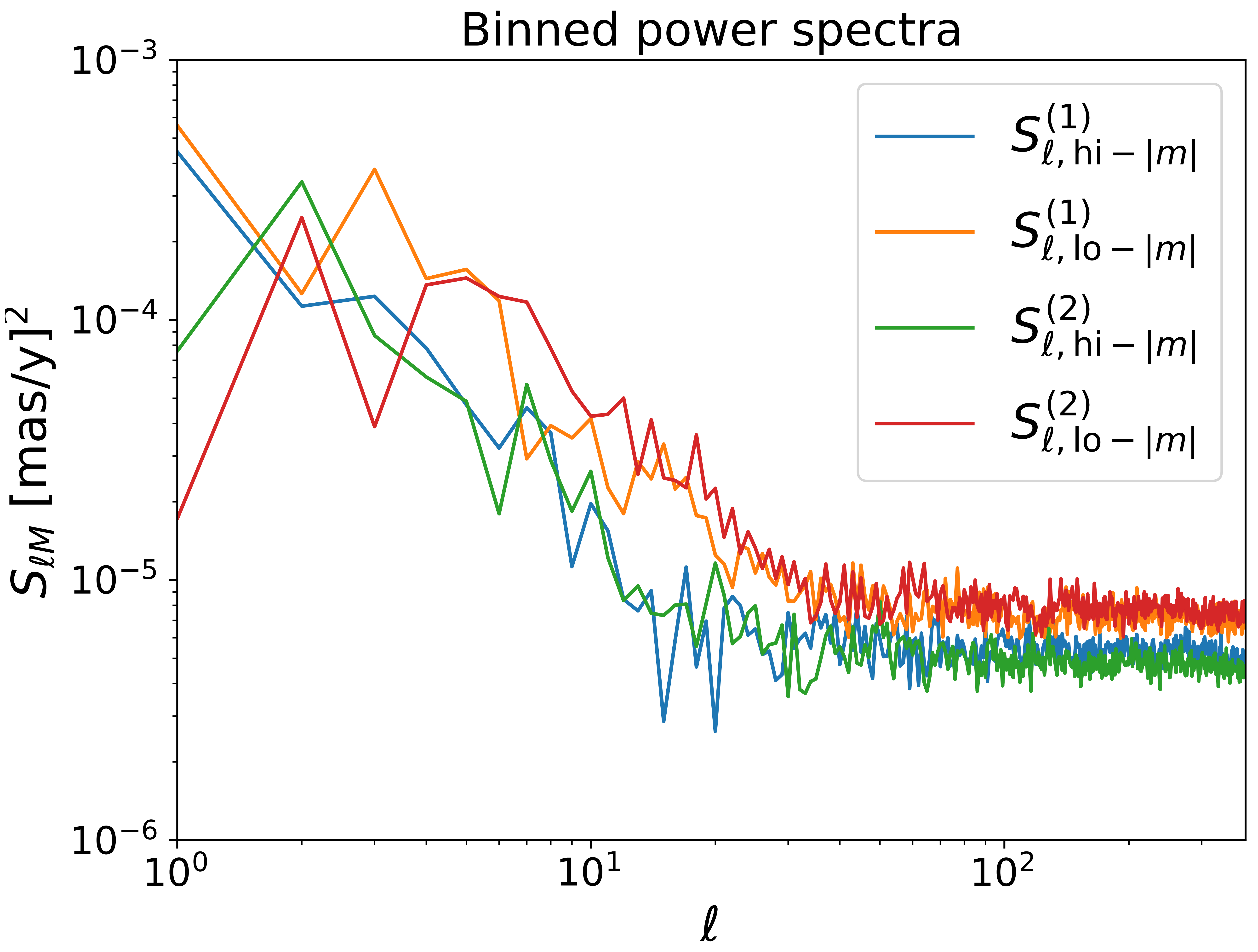
Practically, need *unbiased estimator* to account for non-uniform noise and sky sampling

Proper motions

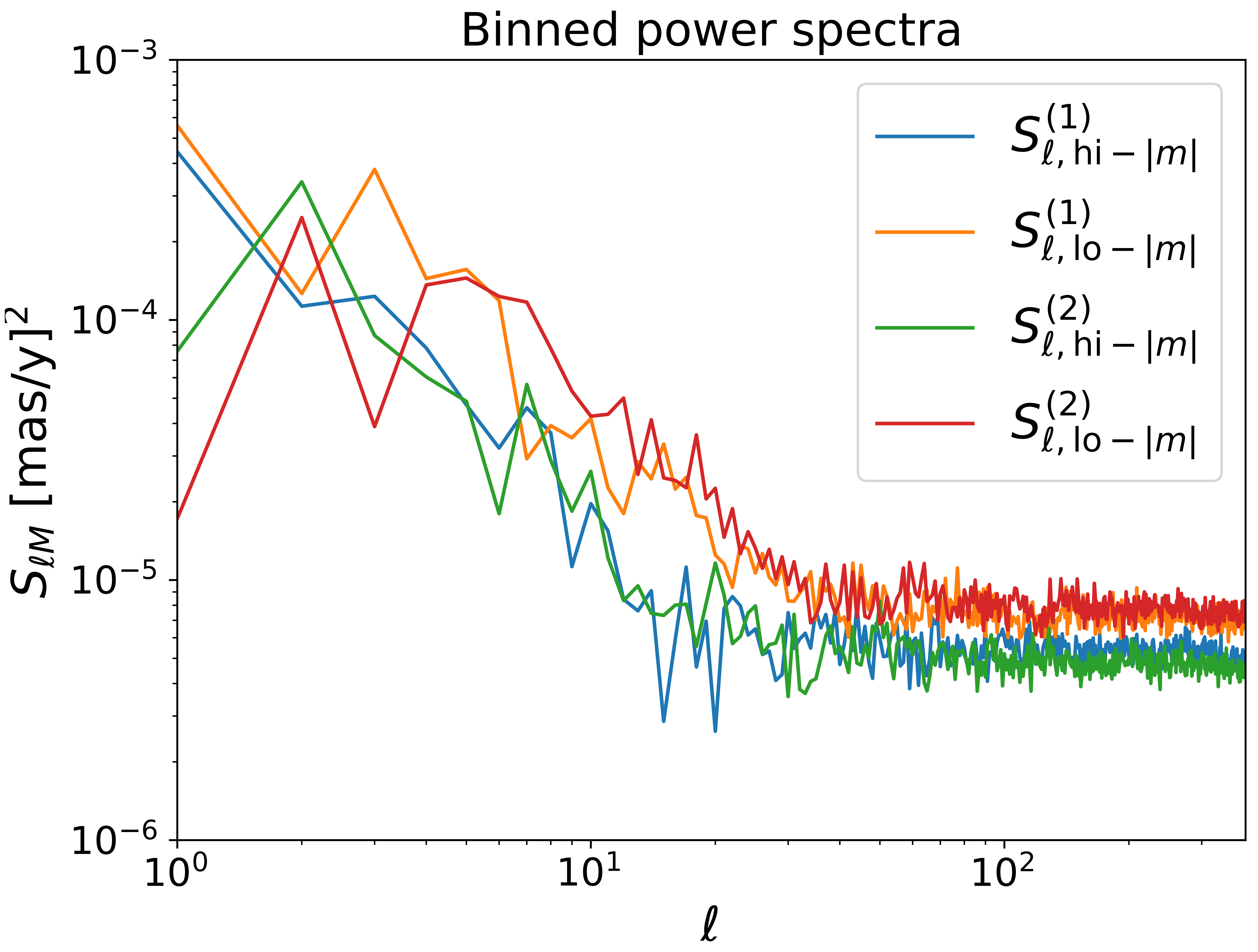


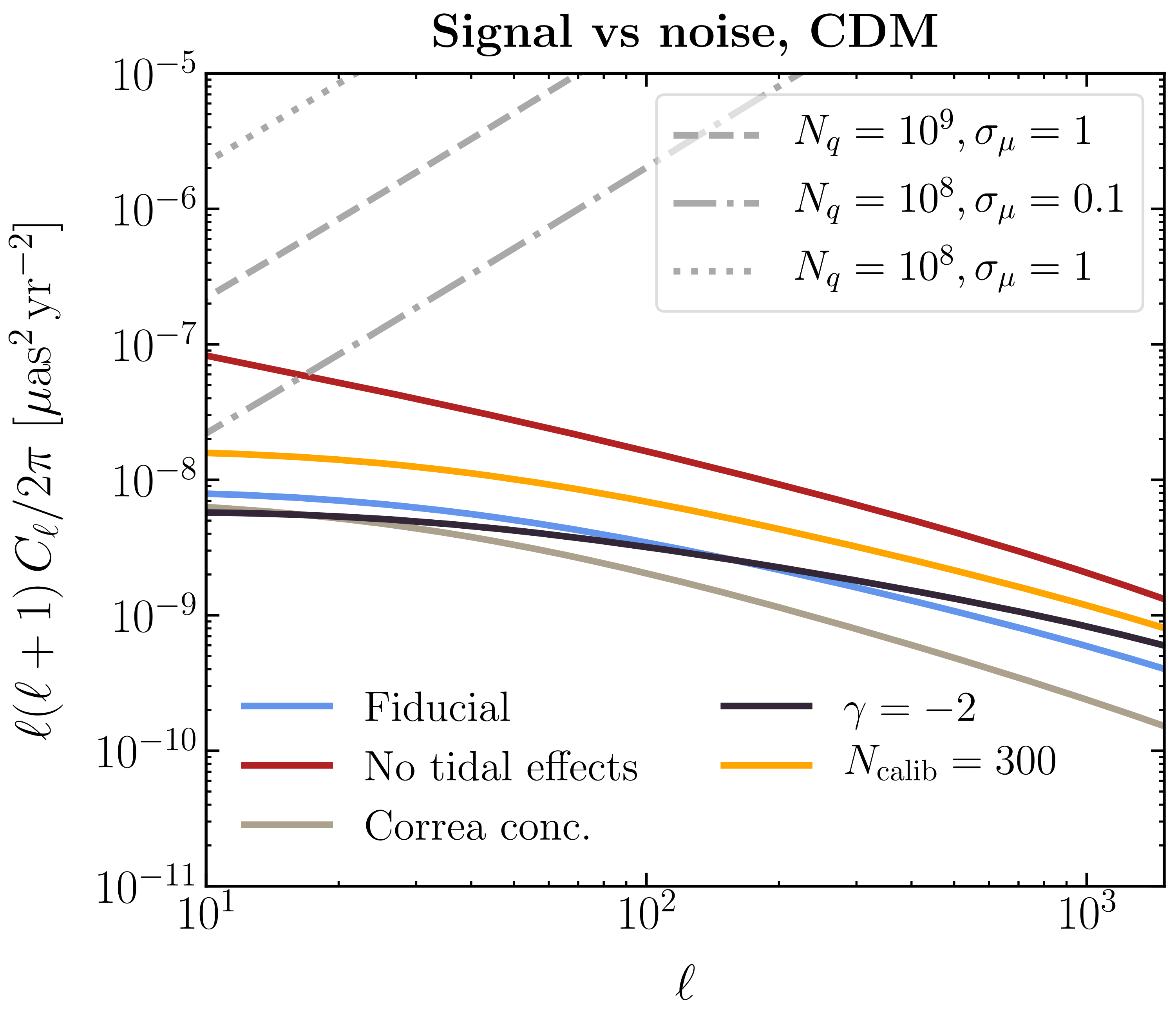


Quasar power spectrum estimator



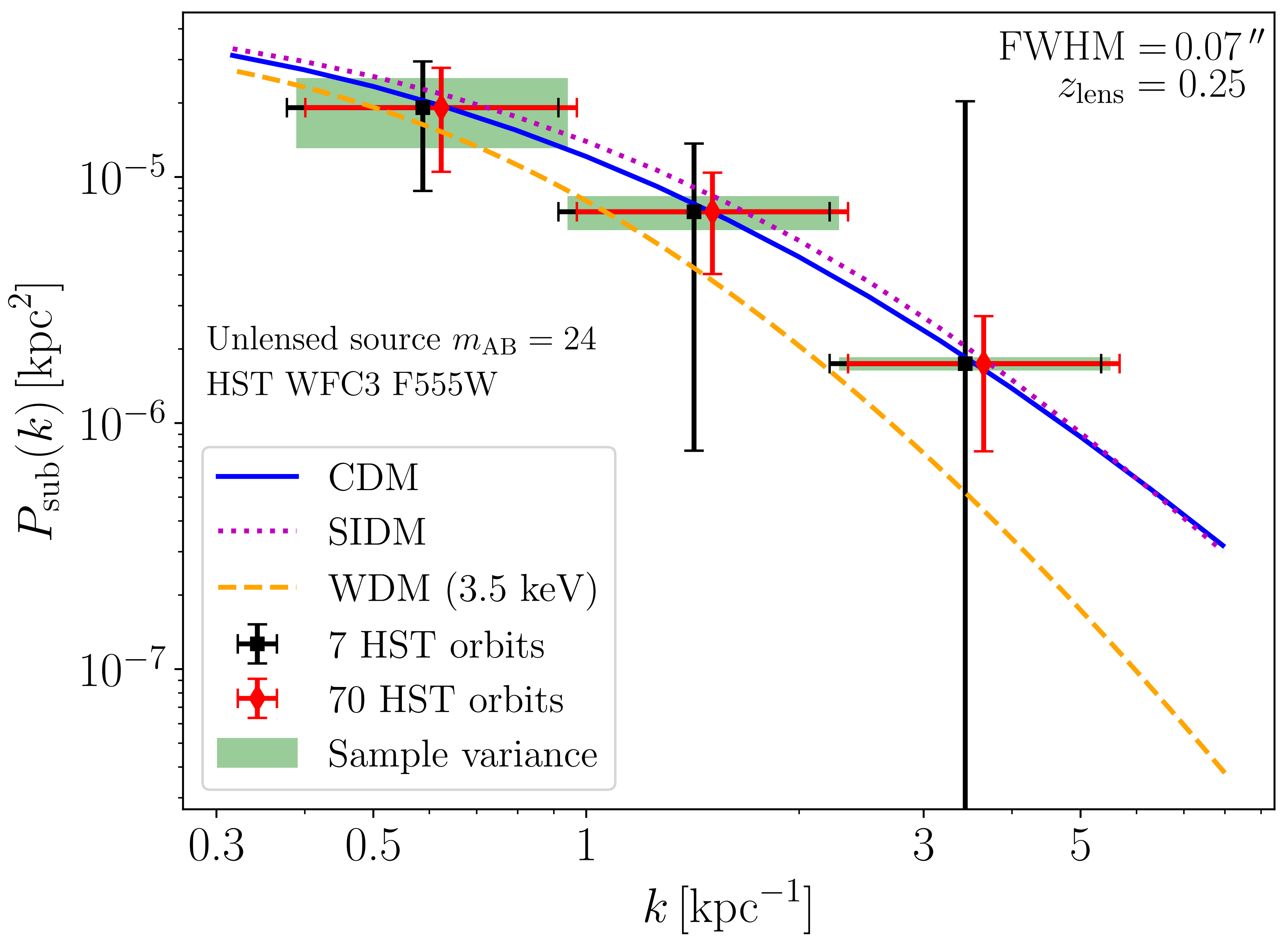
Quasar power spectrum estimator





Inferring substructure through collective effects

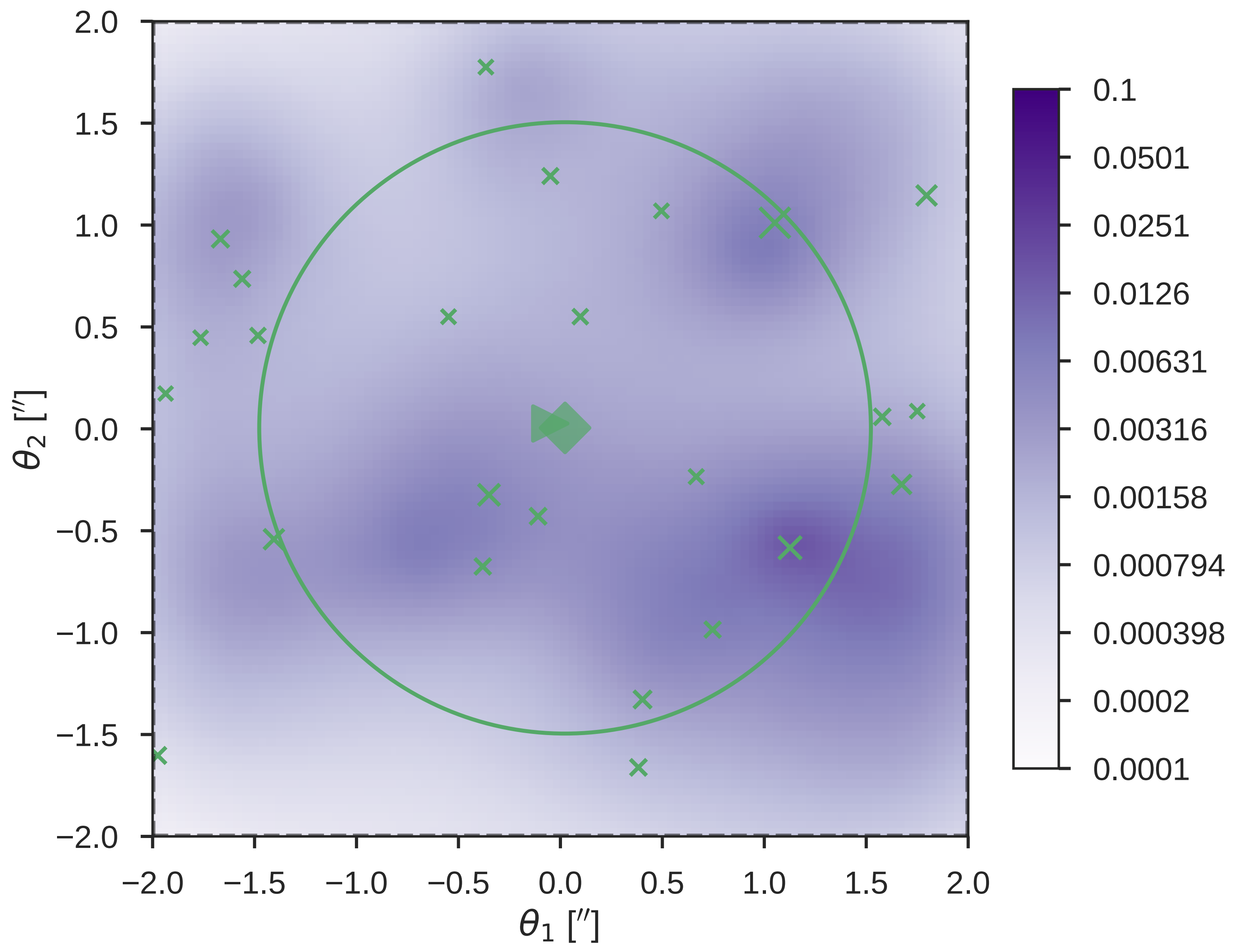
Power spectrum decomposition



Cyr-Racine et al [1806.07897]

See also
Rivero et al [1707.04590]
Rivero et al [1809.00004]
Brennan et al [1808.03501]
Hezaveh et al [1403.2720]

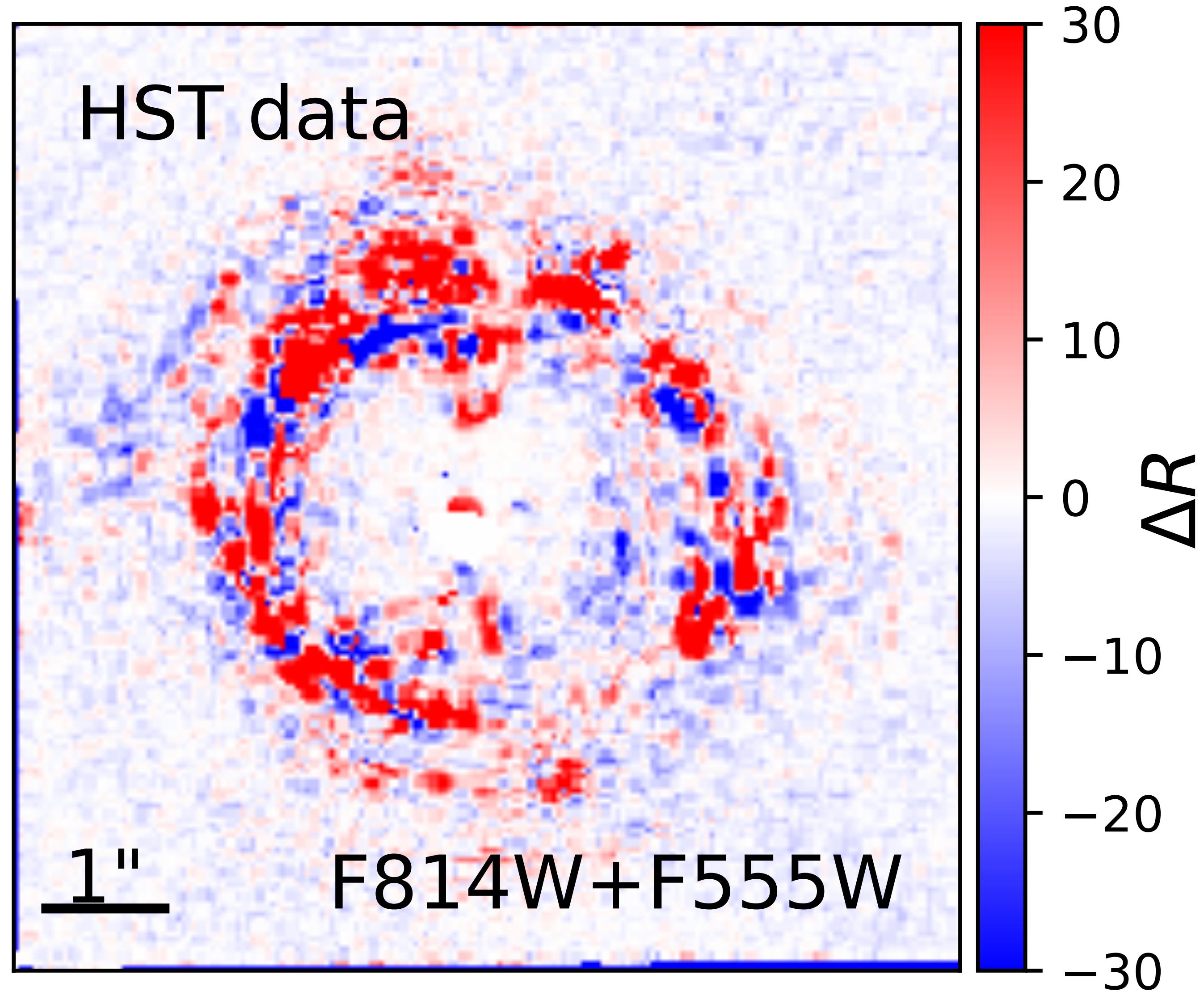
Trans-dimensional methods



Daylan et al [1706.06111]

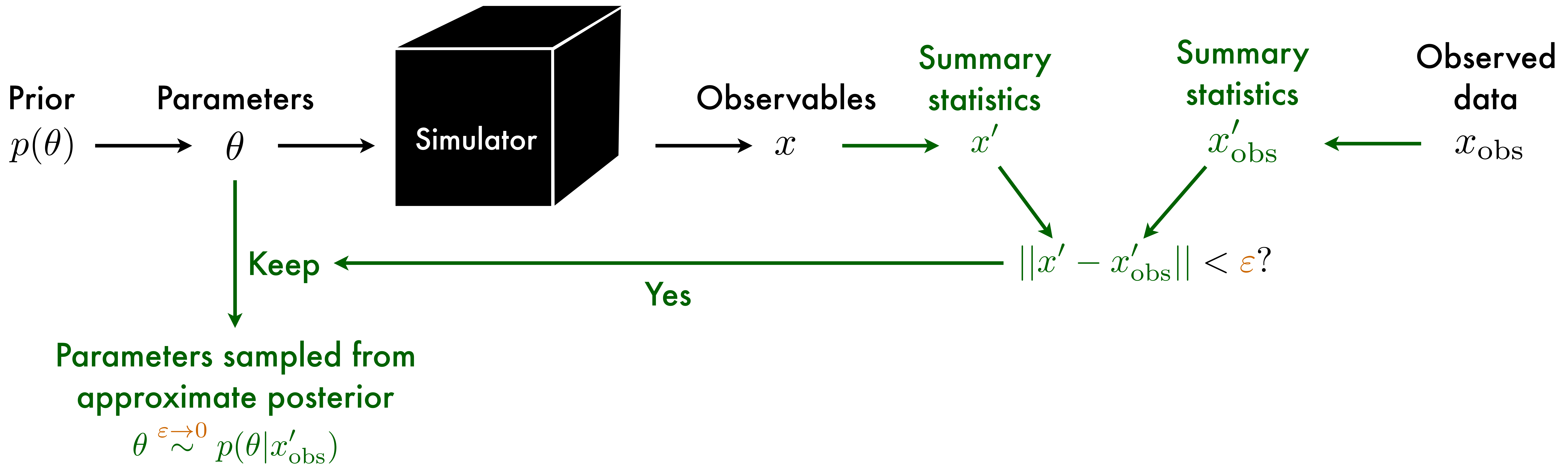
See also
Brewer et al [1508.00662]

Summary statistics



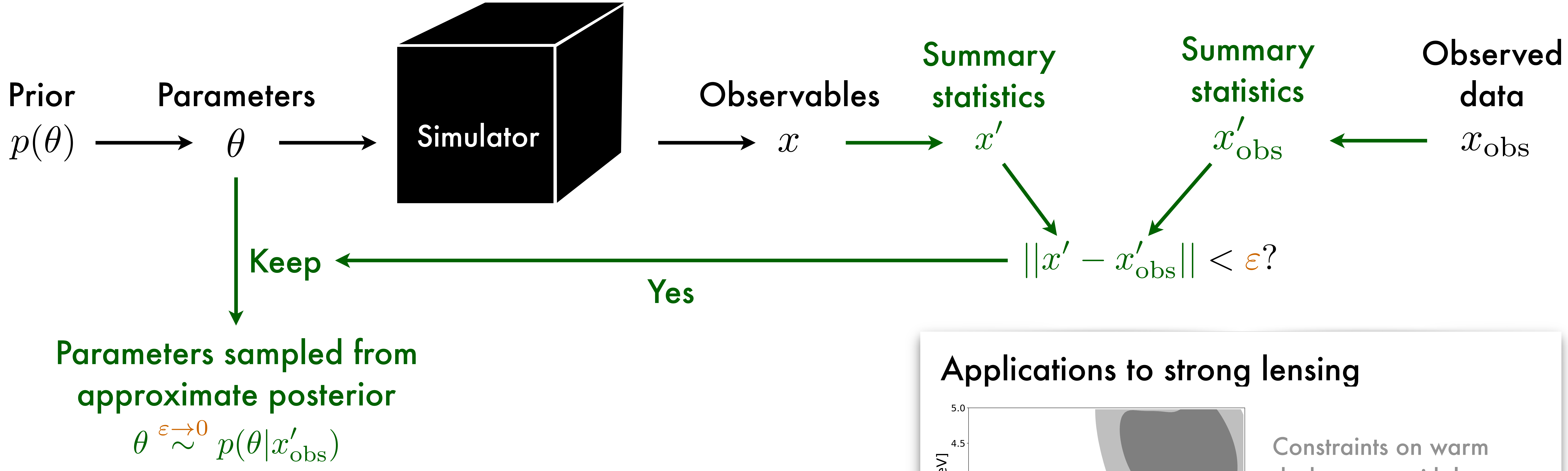
Birrer et al [1702.00009]

"Traditional" LFI: Approximate Bayesian Computation



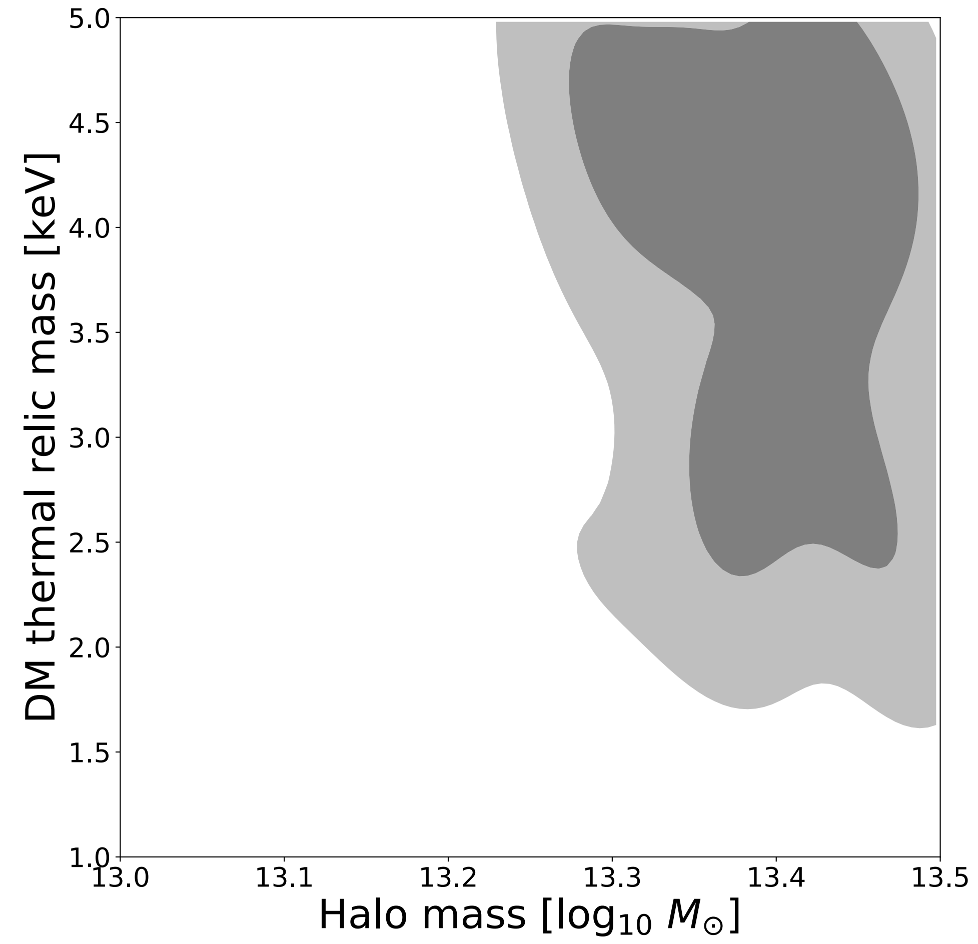
- How to choose x' ? "Curse of dimensionality"
- How to choose ε ? Precision vs efficiency tradeoff
- No tractable posterior
- Need to run new simulations for new data or new prior

"Traditional" LFI: Approximate Bayesian Computation



- How to choose x' ? "Curse of dimensionality"
- How to choose ε ? Precision vs efficiency tradeoff
- No tractable posterior
- Need to run new simulations for new data or new prior

Applications to strong lensing



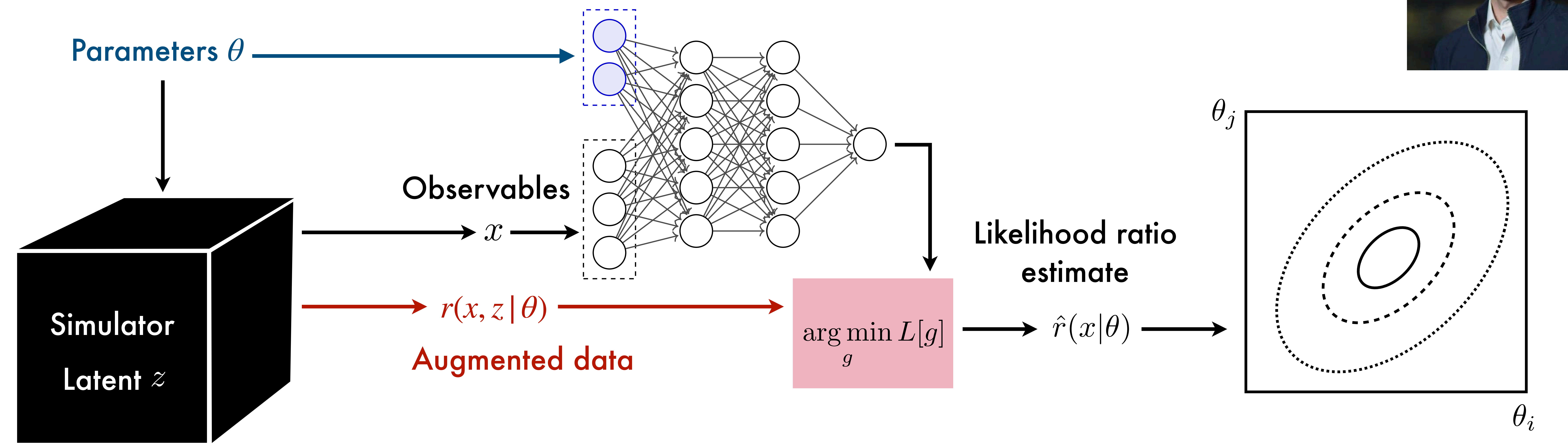
Constraints on warm dark matter with lens RXJ1131-1231

[Birrer et al 2017]

Overview

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]

Slides courtesy of Johann Brehmer



1. Simulation

2. Machine Learning

3. Inference

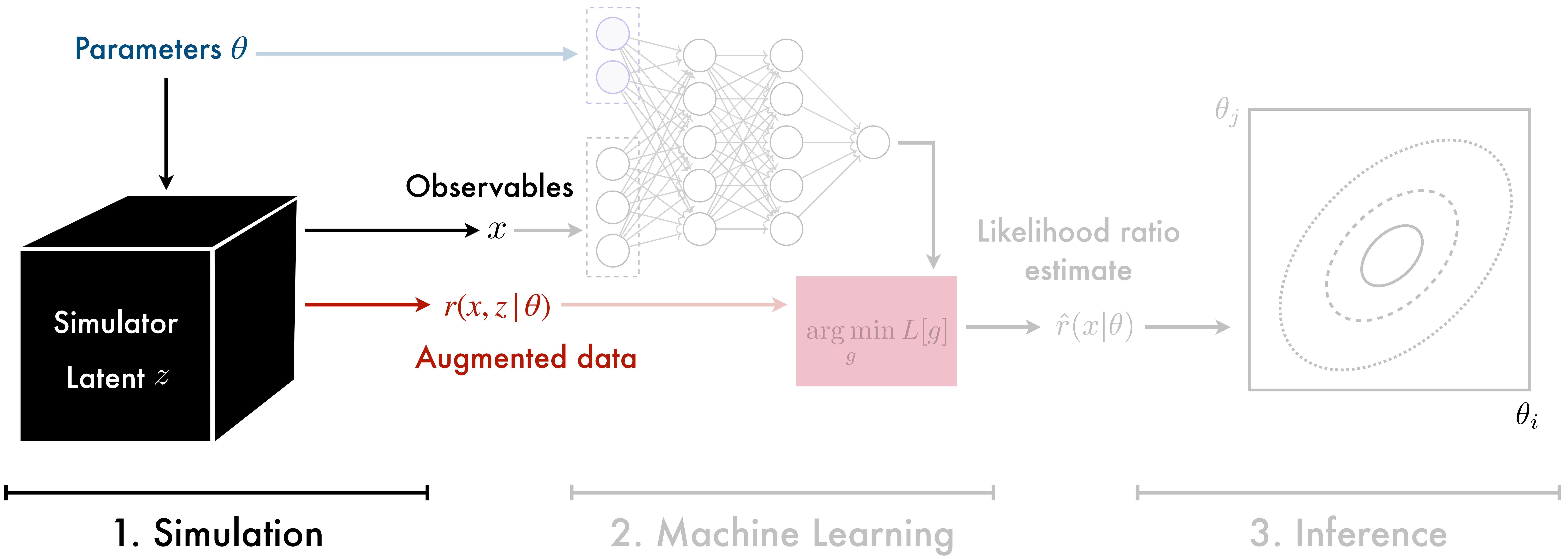
Extract additional information from simulator

Use this information to train estimator for likelihood ratio

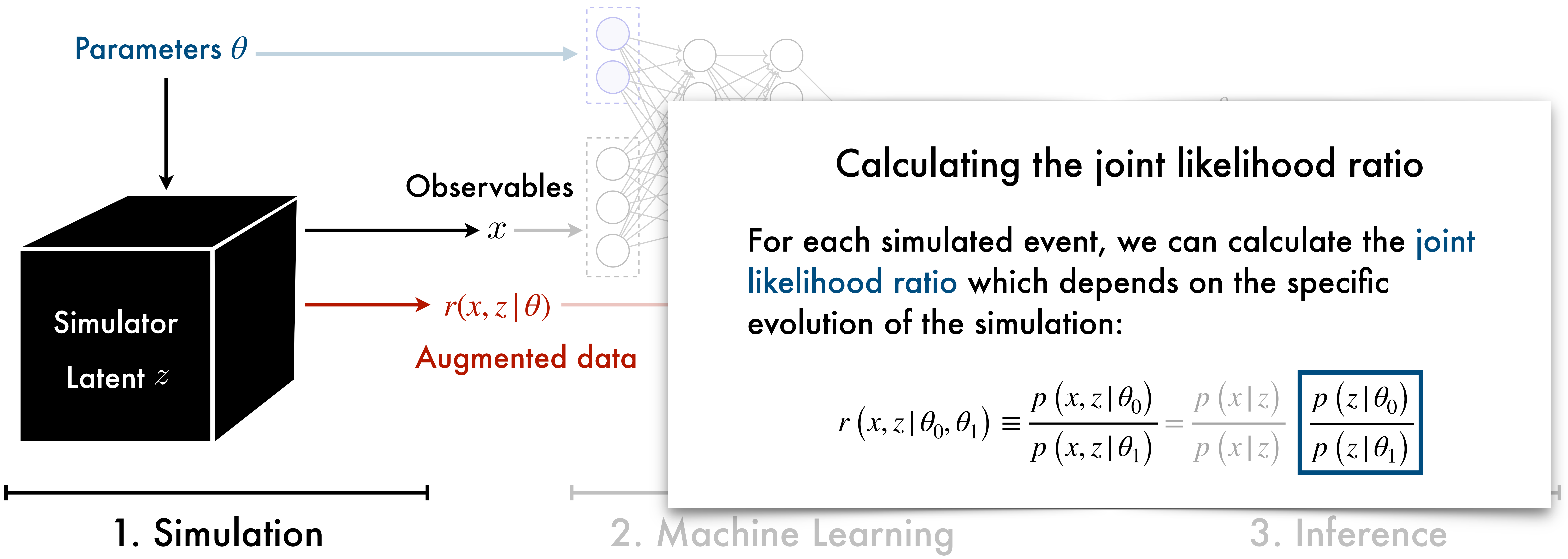
Limit setting with standard hypothesis tests

Simulation

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



Simulation



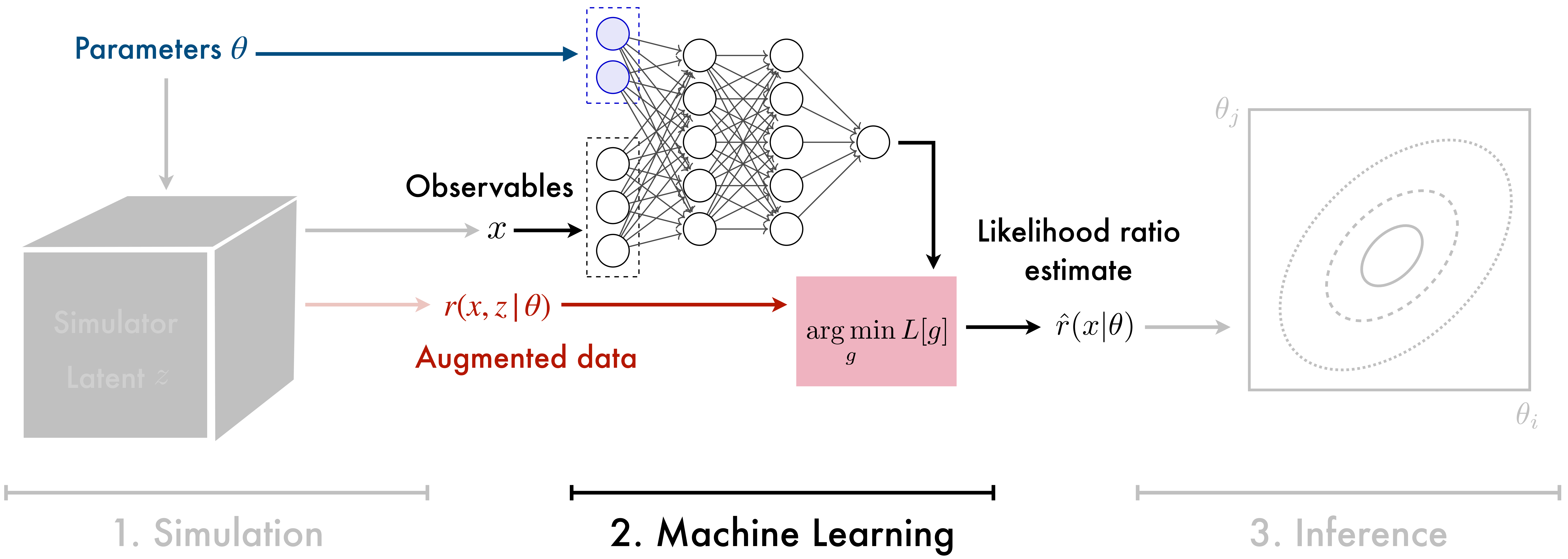
Calculating the joint likelihood ratio

For each simulated event, we can calculate the joint likelihood ratio which depends on the specific evolution of the simulation:

$$r(x, z|\theta_0, \theta_1) \equiv \frac{p(x, z|\theta_0)}{p(x, z|\theta_1)} = \frac{p(x|z)}{p(x|z)} \boxed{\frac{p(z|\theta_0)}{p(z|\theta_1)}}$$

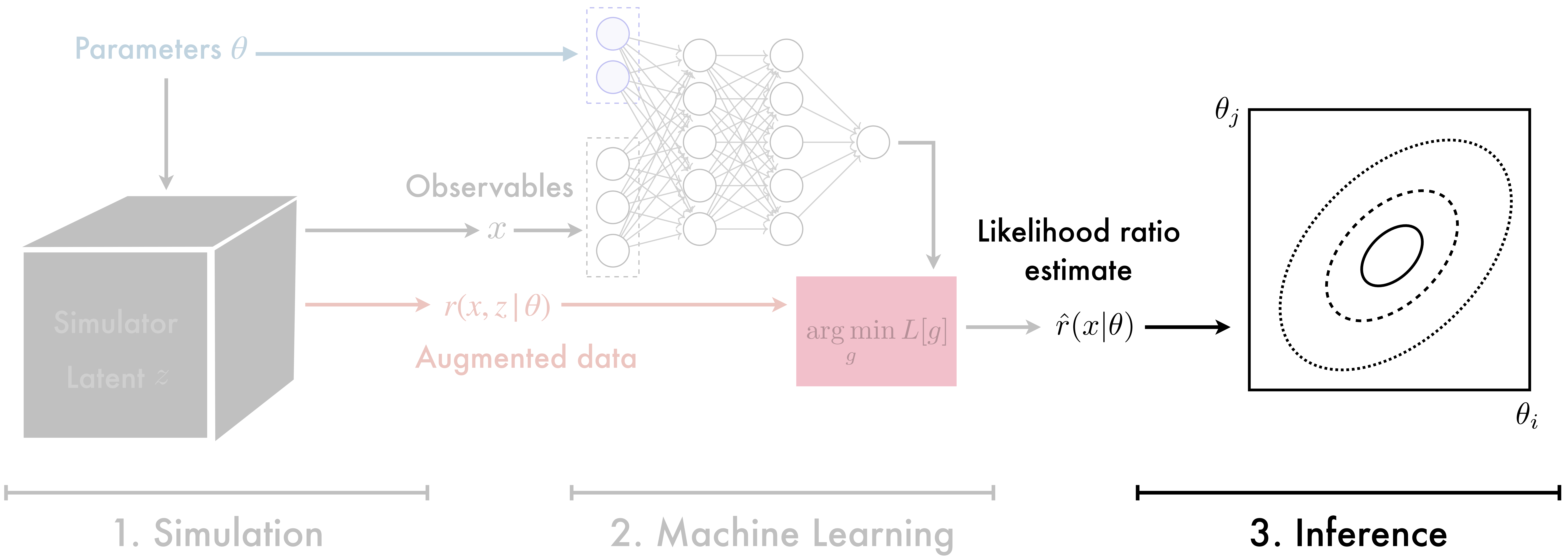
Machine learning

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



Inference

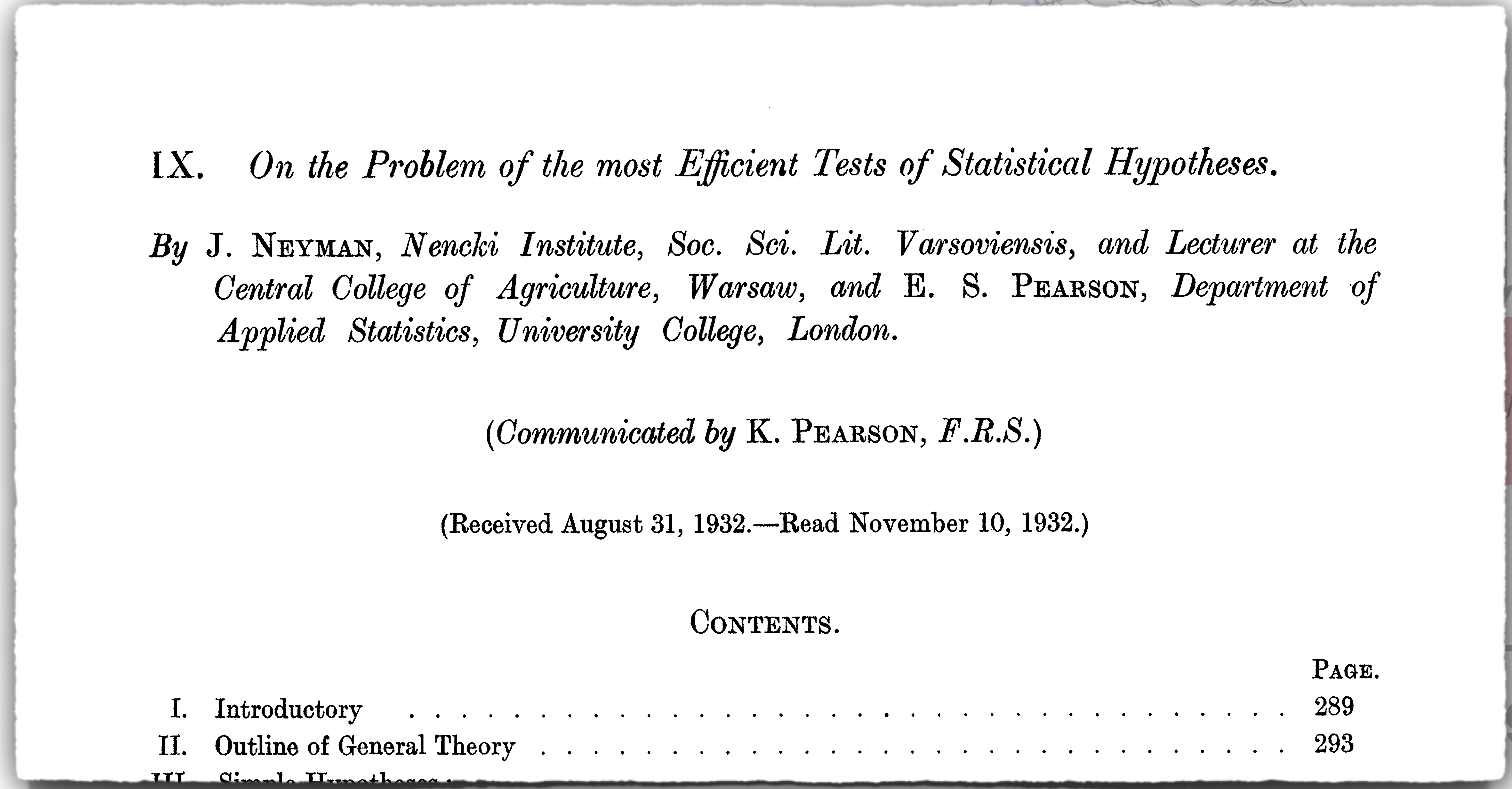
Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]



Inference

Brehmer et al [1805.00013]
Brehmer et al [1805.00020]
Stoye et al [1808.00973]

Parameters θ



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.*

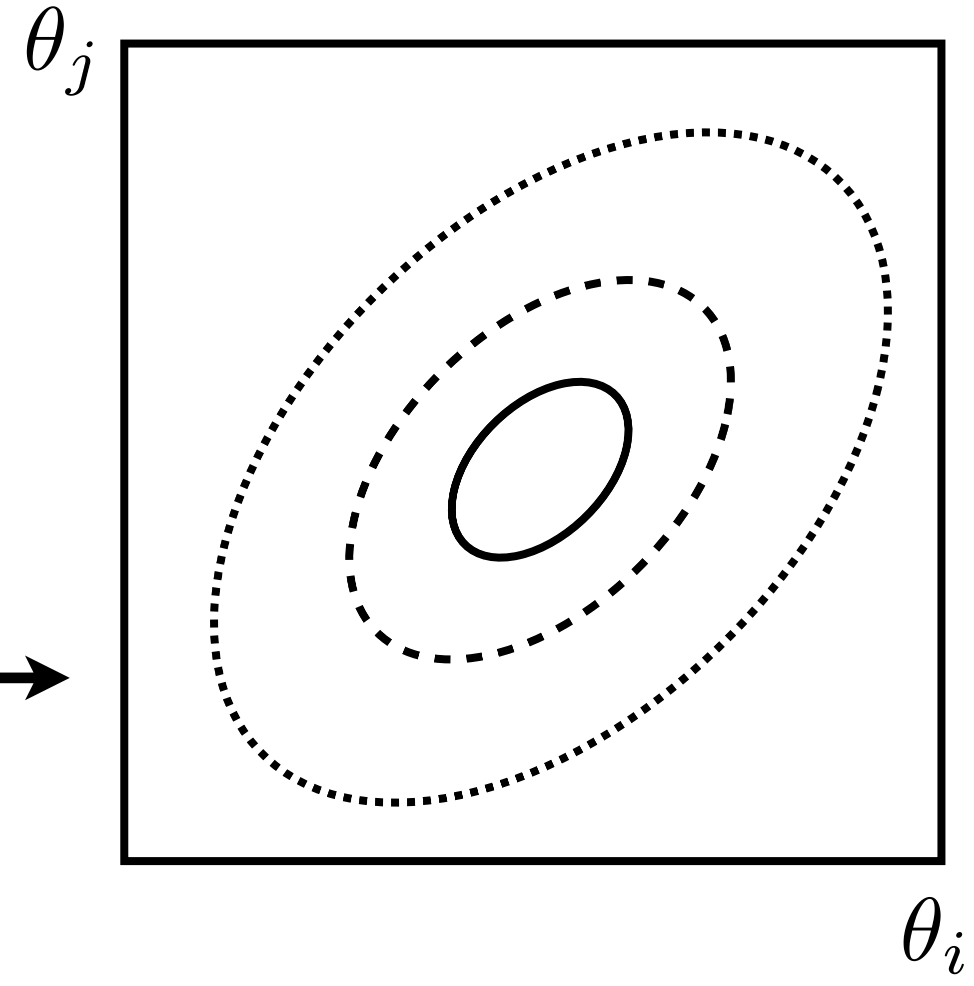
(*Communicated by* K. PEARSON, *F.R.S.*)

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

CONTENTS.

Likelihood ratio estimate

$\hat{r}(x|\theta)$



3. Inference

Machine learning

We can calculate the joint likelihood ratio

$$r(x, z \mid \theta_0, \theta_1) \equiv \frac{p(x, z \mid \theta_0)}{p(x, z \mid \theta_1)}$$

("How much more likely is this simulated event, including all intermediate states, for θ_0 compared to θ_1?")

We want the likelihood ratio function

$$r(x \mid \theta_0, \theta_1) \equiv \frac{p(x \mid \theta_0)}{p(x \mid \theta_1)}$$

("How much more likely is the observation for θ_0 compared to θ_1?")

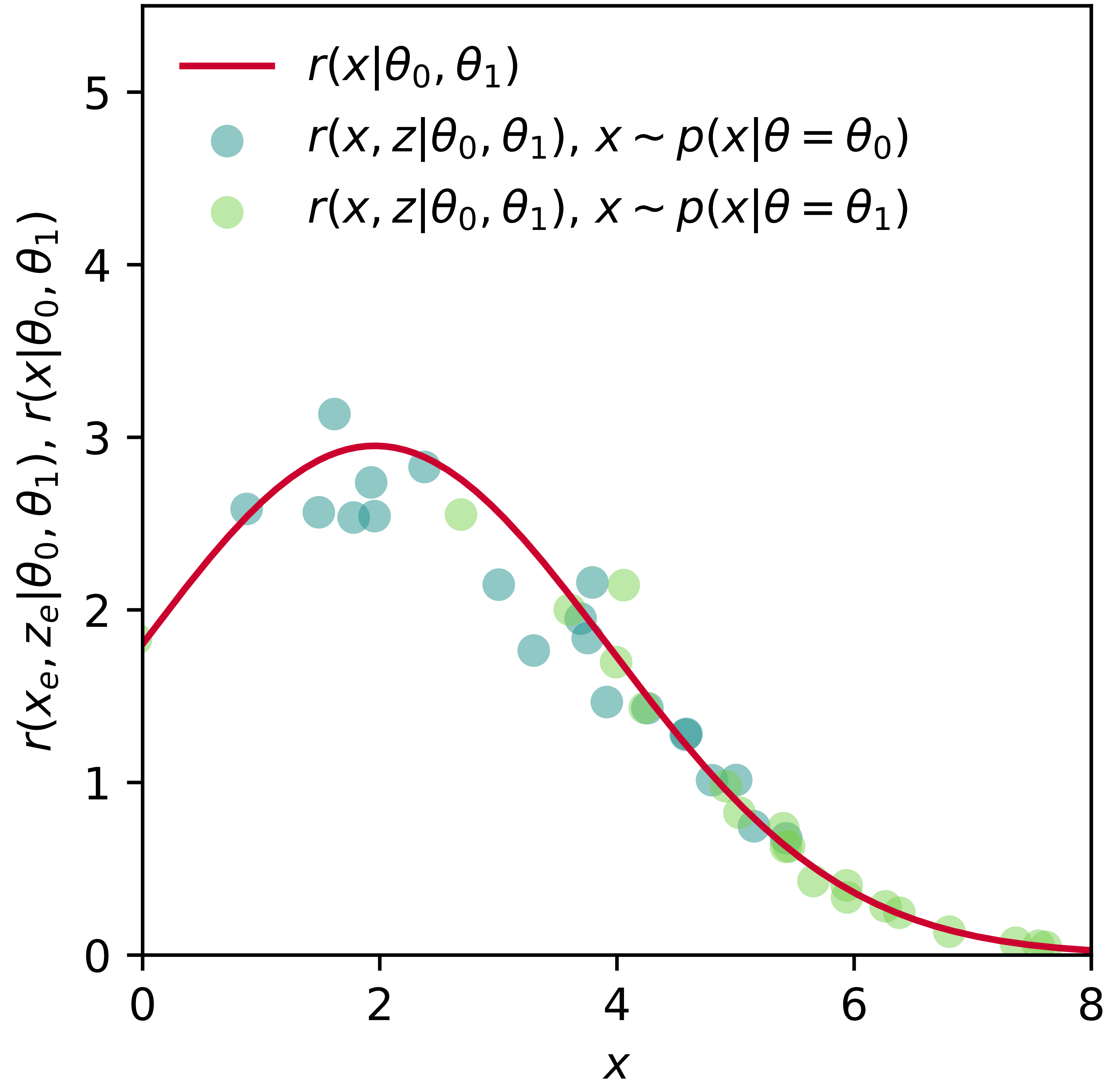
θ_i

Machine learning

We can calculate the joint likelihood ratio

$$r(x, z \mid \theta_0, \theta_1) \equiv \frac{p(x, z \mid \theta_0)}{p(x, z \mid \theta_1)}$$

$r(x, z|\theta_0, \theta_1)$
are scattered around
$r(x|\theta_0, \theta_1)$



We want the likelihood ratio function

$$r(x \mid \theta_0, \theta_1) \equiv \frac{p(x \mid \theta_0)}{p(x \mid \theta_1)}$$

θ_i

Machine learning

With $r(x, z|\theta_0, \theta_1)$, we define a functional like

$$L_r[\hat{r}(x|\theta_0, \theta_1)] = \int \mathrm{d}x \int \mathrm{d}z \; p(x, z|\theta_1) \left[\left(\hat{r}(x|\theta_0, \theta_1) - r(x, z|\theta_0, \theta_1) \right)^2 \right]$$

It is minimized by

$$r(x|\theta_0, \theta_1) = \underset{\hat{r}(x|\theta_0, \theta_1)}{\arg\min} \; L_r[\hat{r}(x|\theta_0, \theta_1)]$$

(And we can sample from $p(x, z|\theta)$ by running the simulator.)

θ_i

ResNet-18 architecture

Layer Name	Output Size	ResNet-18
conv1	$112 \times 112 \times 64$	7×7, 64, stride 2
conv2_x	$56 \times 56 \times 64$	3×3 max pool, stride 2
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	$1 \times 1 \times 512$	7×7 average pool
fully connected	1000	512×1000 fully connections
softmax	1000	