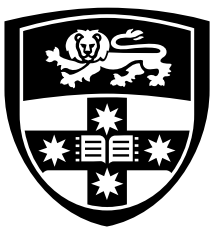
Wising up to CatWISE

Measuring the cosmic dipole with simulation-based inference

Oliver Oayda

PhD Candidate
Sydney Institute for Astronomy
The University of Sydney

Supervised by
Geraint Lewis
Tara Murphy



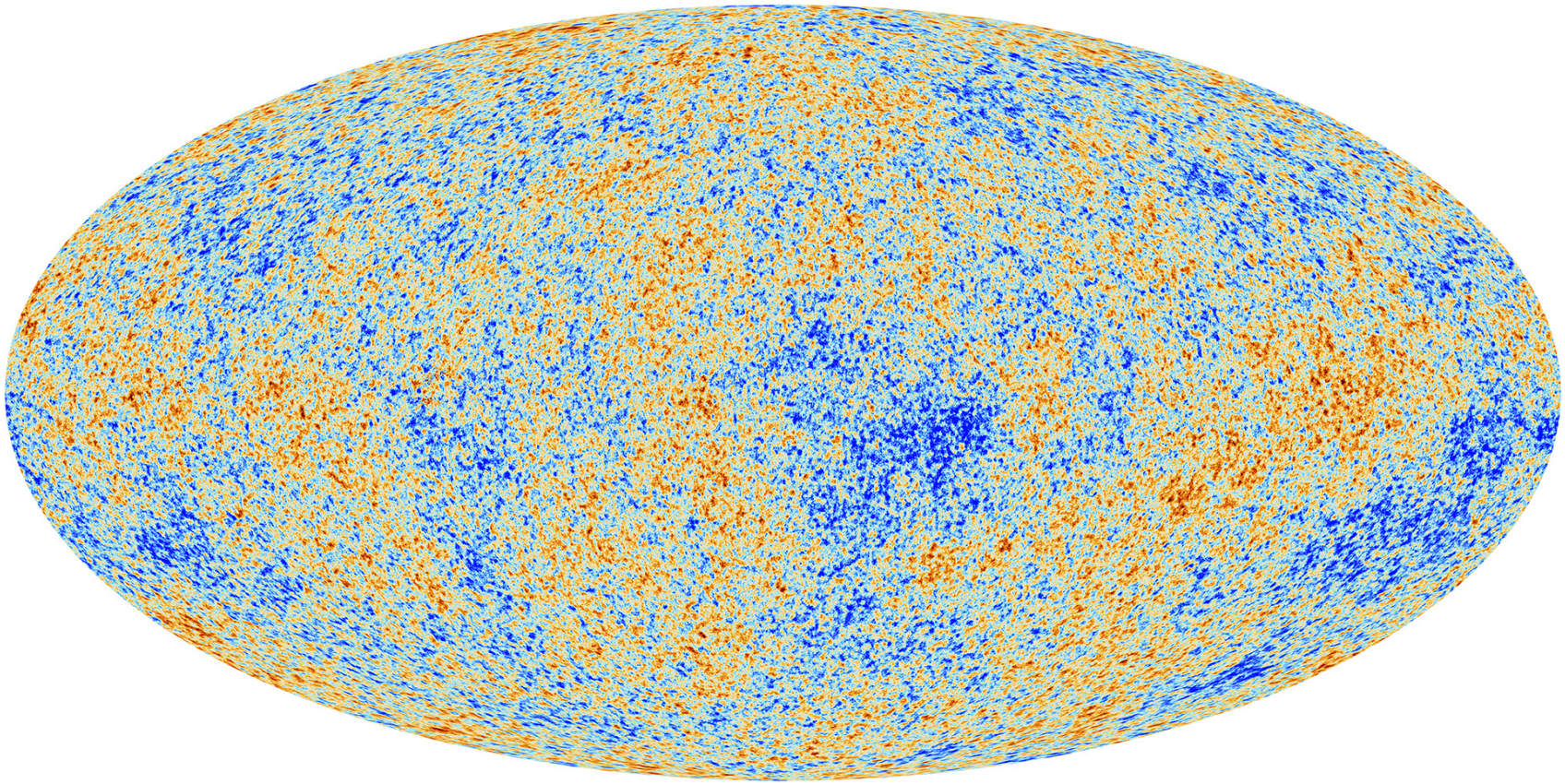
THE UNIVERSITY OF
SYDNEY

Euclid Dipole Seminar

June 3, 2026

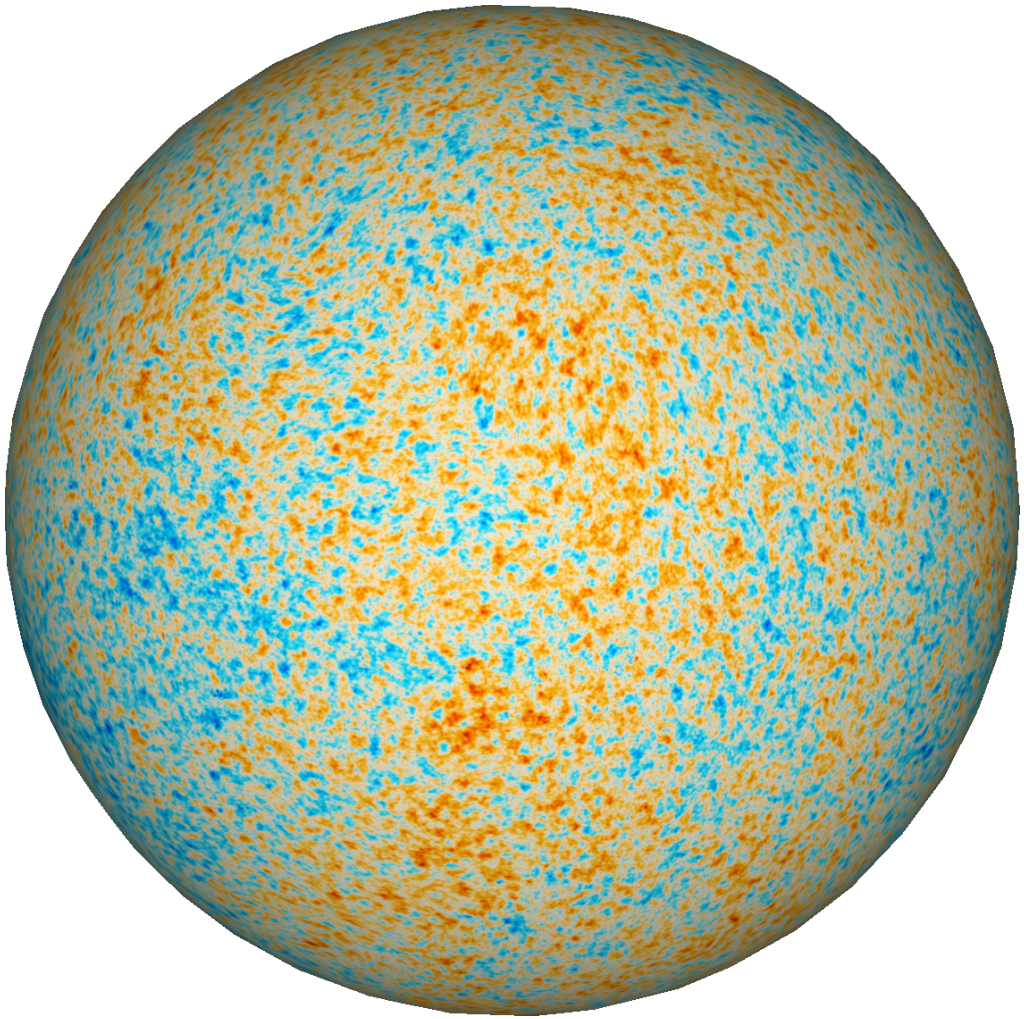
Simulation-based inference helps us understand the ecliptic bias in CatWISE.

The Kinematic Dipole



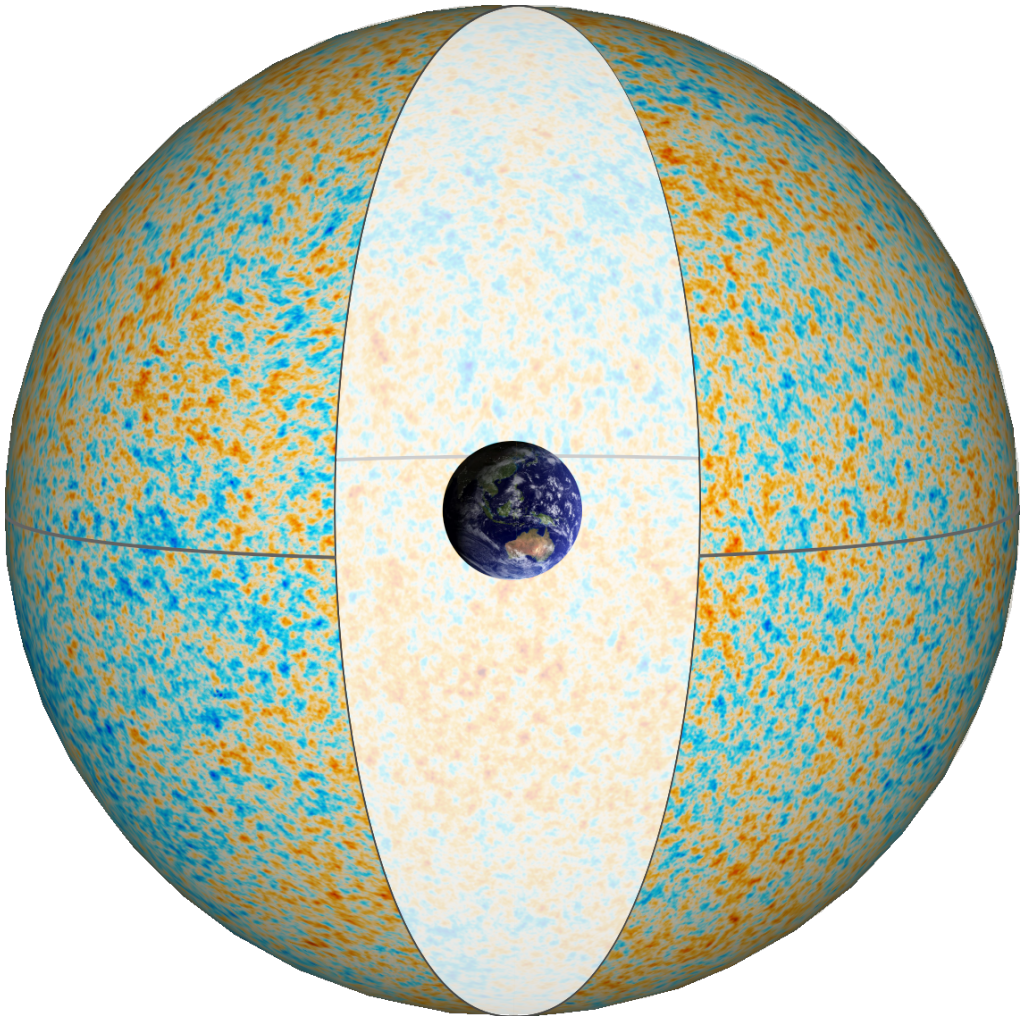
CMB temperature map (Planck satellite).

The Kinematic Dipole



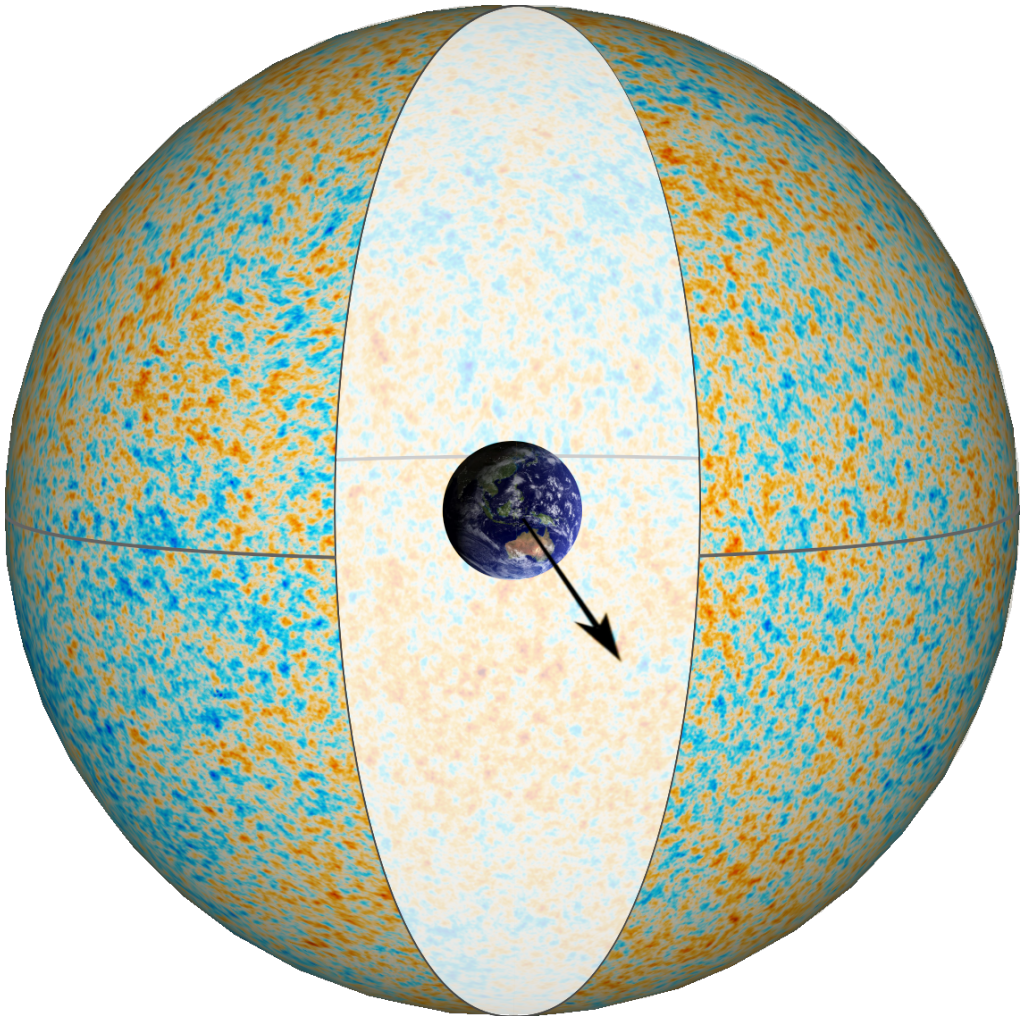
CMB as a sphere.

The Kinematic Dipole



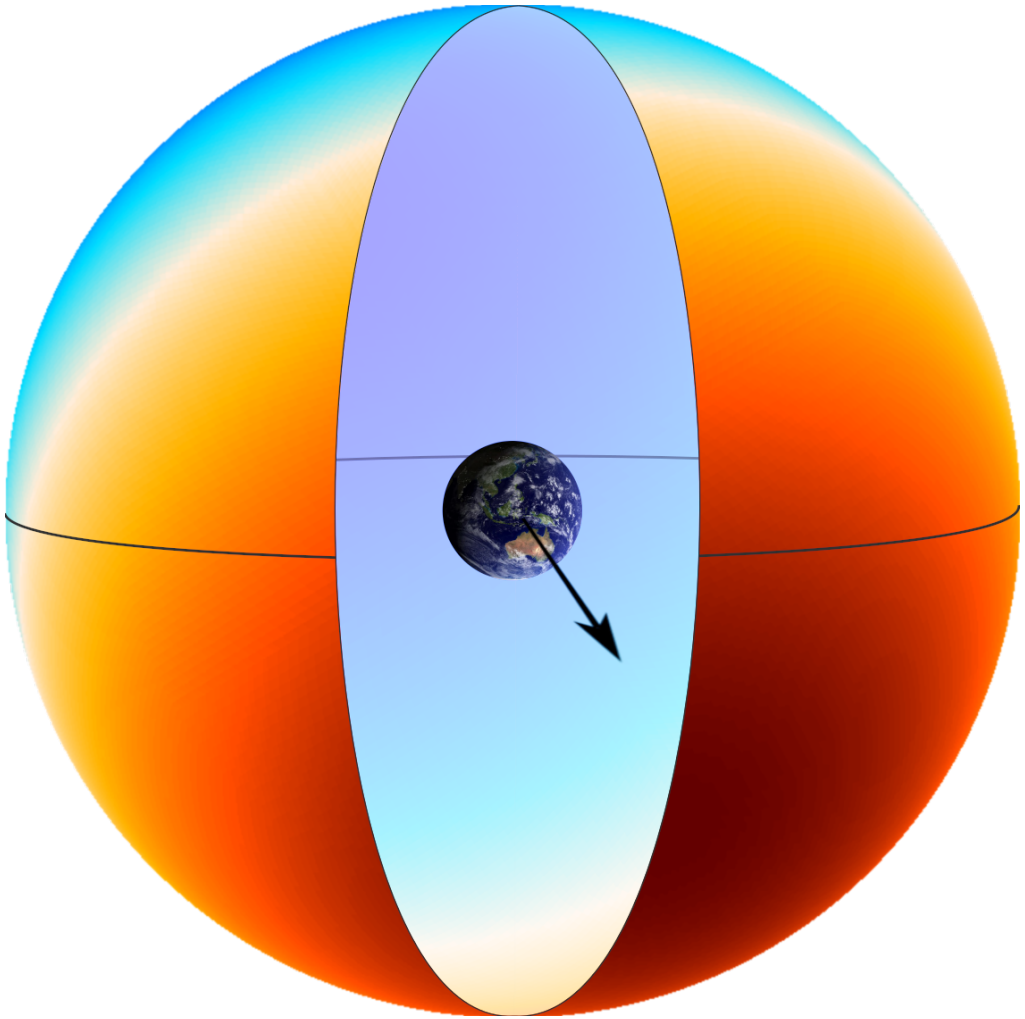
CMB as a sphere (Earth inside).

The Kinematic Dipole



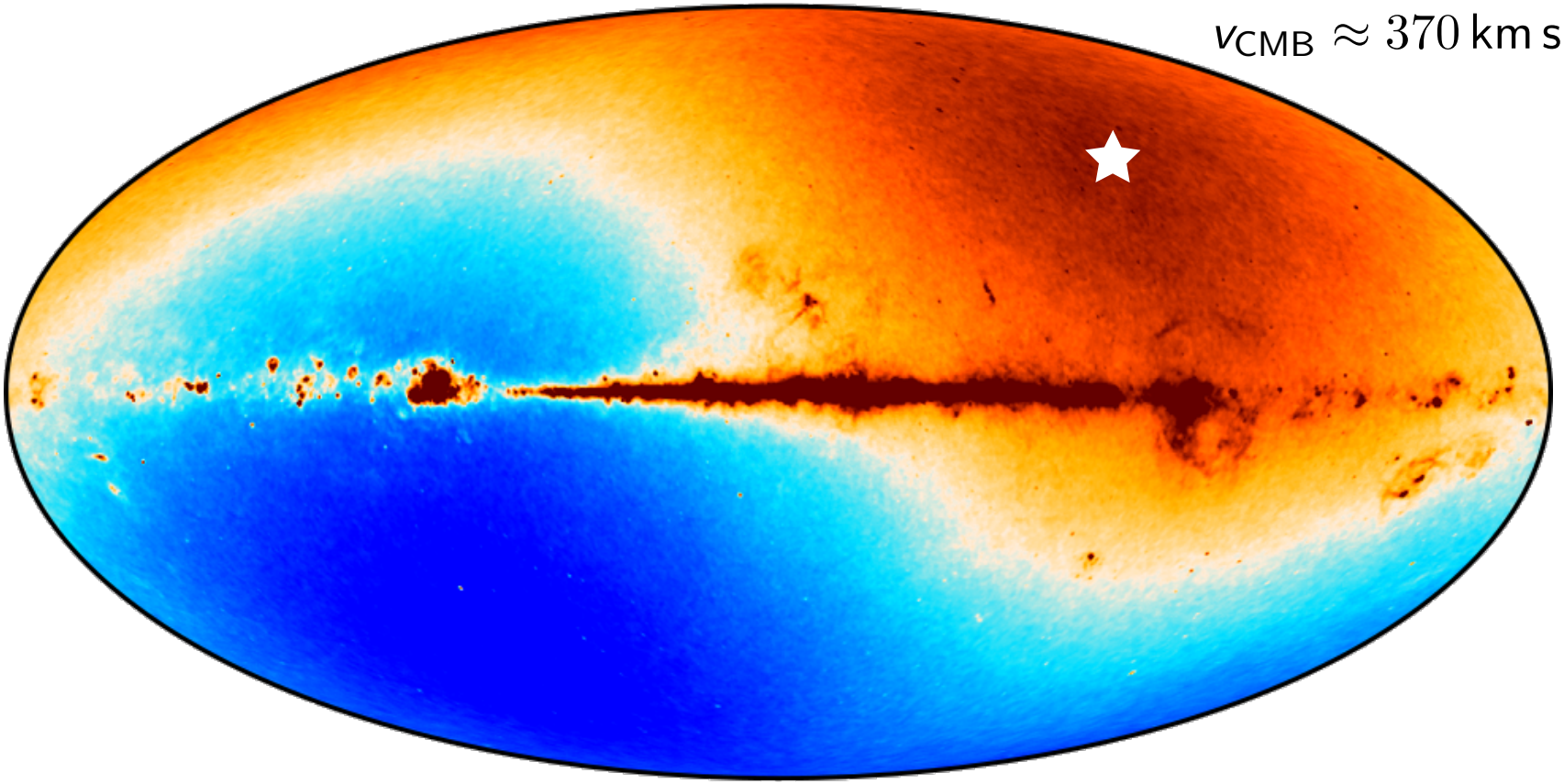
We're moving through the Universe!

The Kinematic Dipole



CMB dipole as sphere (Earth inside).

The Kinematic Dipole



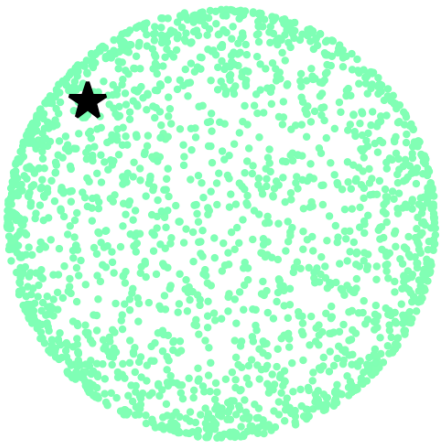
CMB temperature map (dipole included; BeyondPlanck).
★: dipole direction.

The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity: ➢ Aberration ➢ Doppler boosting

The Cosmic Dipole — Ellis & Baldwin (1984)

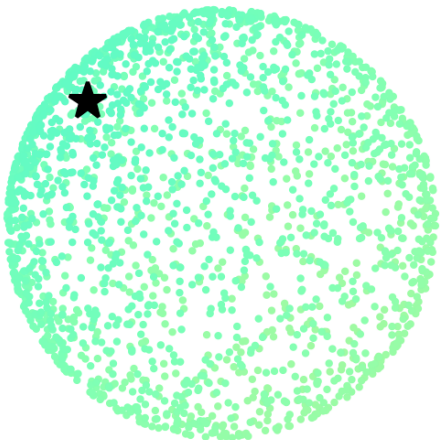
Special relativity: ➢ Aberration ➢ Doppler boosting



Speed (c) 0.025

The Cosmic Dipole — Ellis & Baldwin (1984)

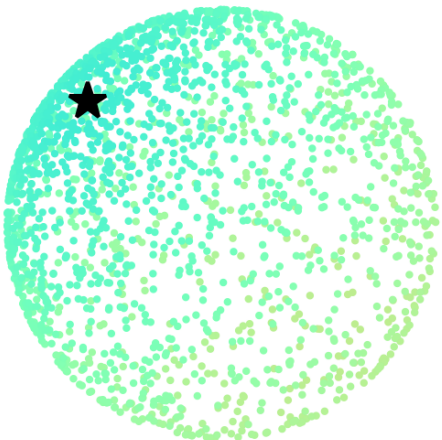
Special relativity: ➢ Aberration ➢ Doppler boosting



Speed (c) 0.244

The Cosmic Dipole — Ellis & Baldwin (1984)

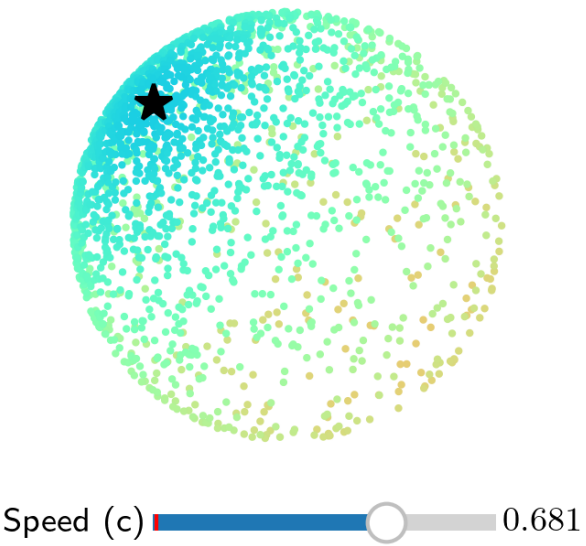
Special relativity: ➢ Aberration ➢ Doppler boosting



Speed (c) 0.462

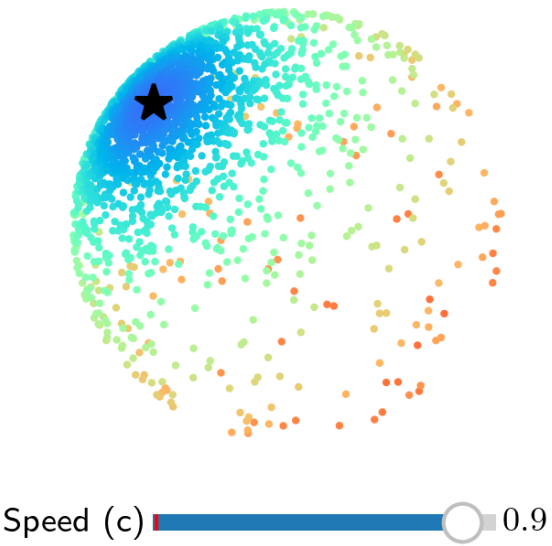
The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity: ➢ Aberration ➢ Doppler boosting



The Cosmic Dipole — Ellis & Baldwin (1984)

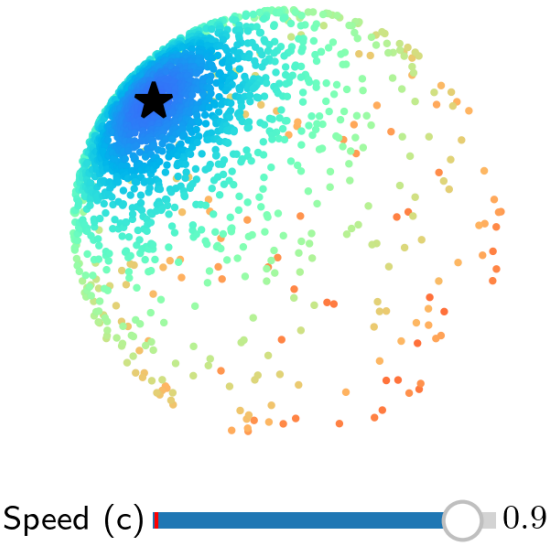
Special relativity: ➢ Aberration ➢ Doppler boosting

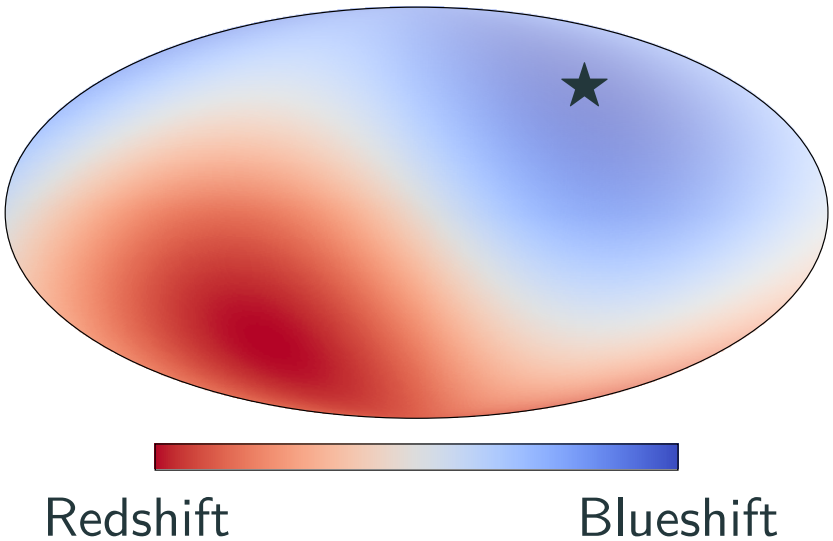


The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity:

- Aberration
- Doppler boosting

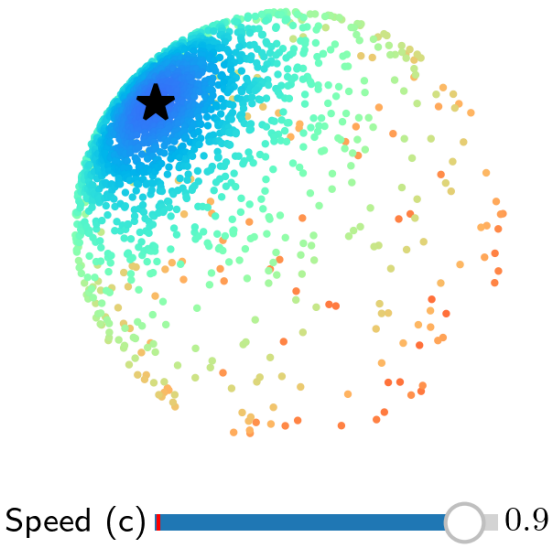


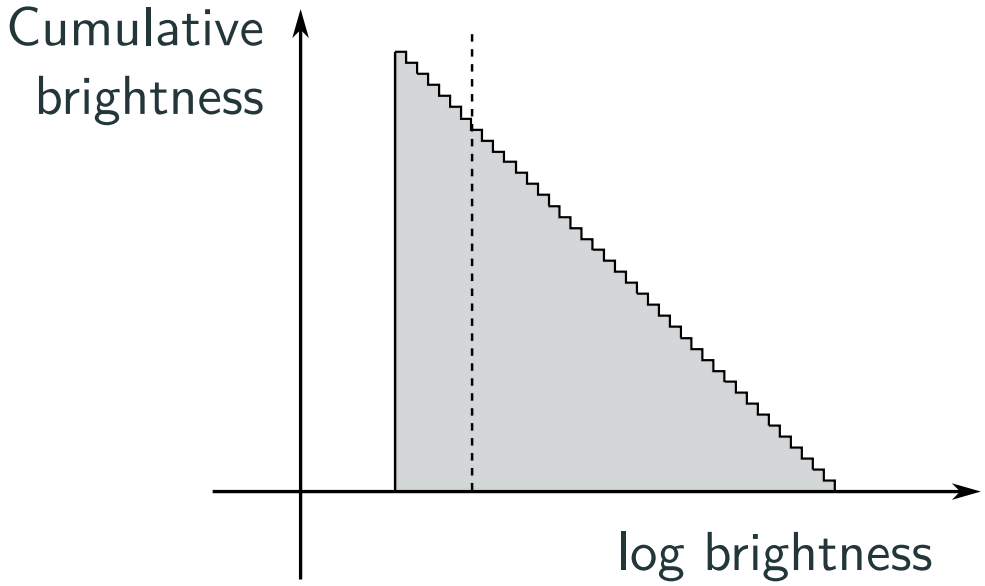


The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity:

- Aberration
- Doppler boosting

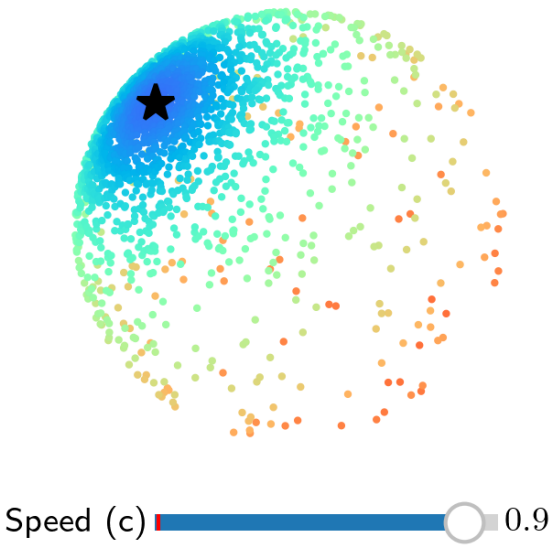


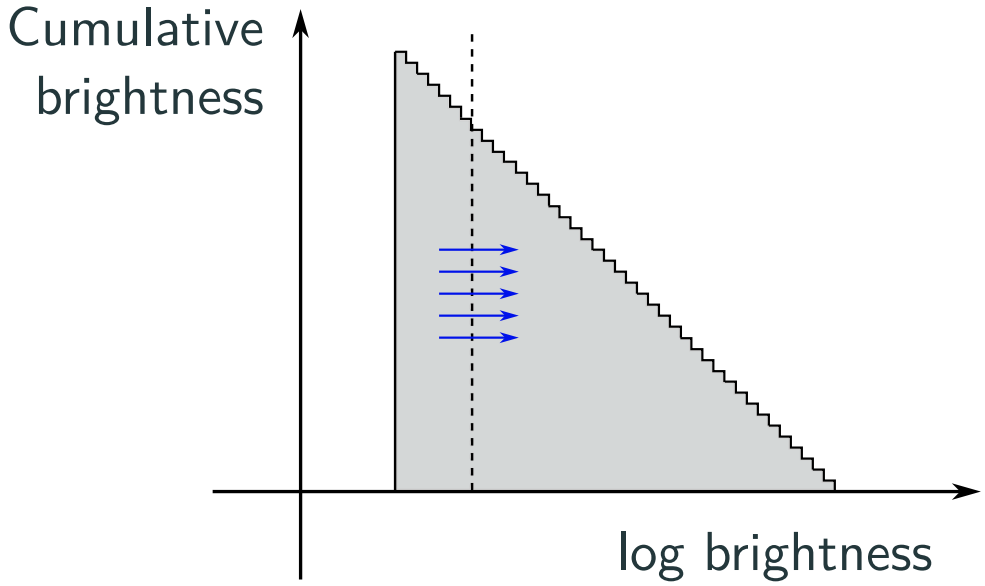


The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity:

- Aberration
- Doppler boosting



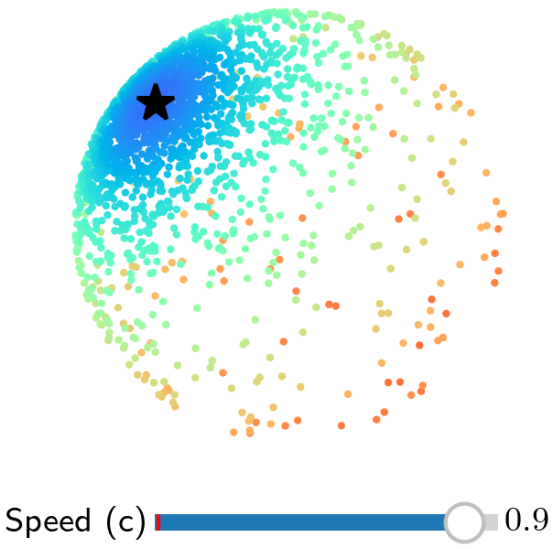


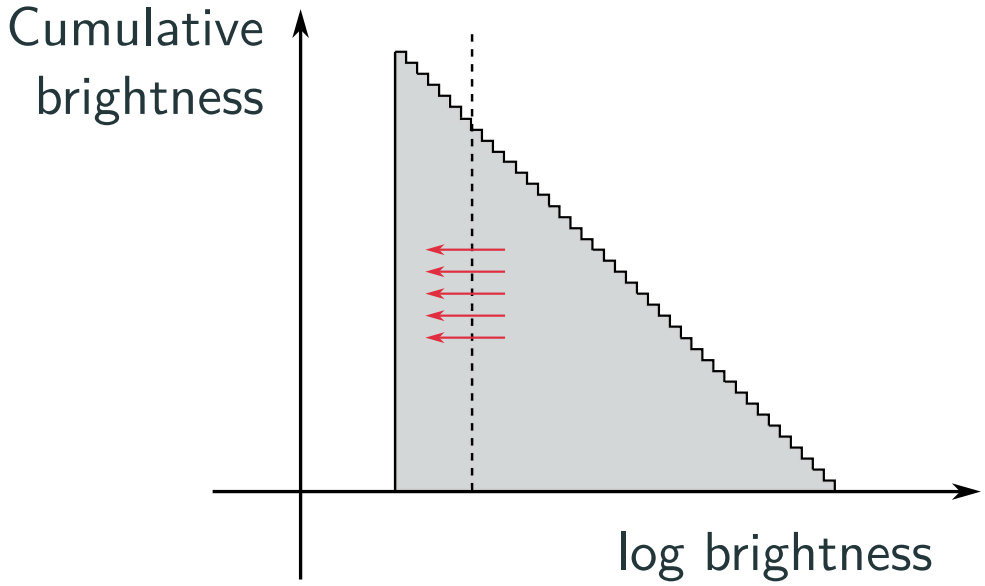
The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity:

➢ Aberration

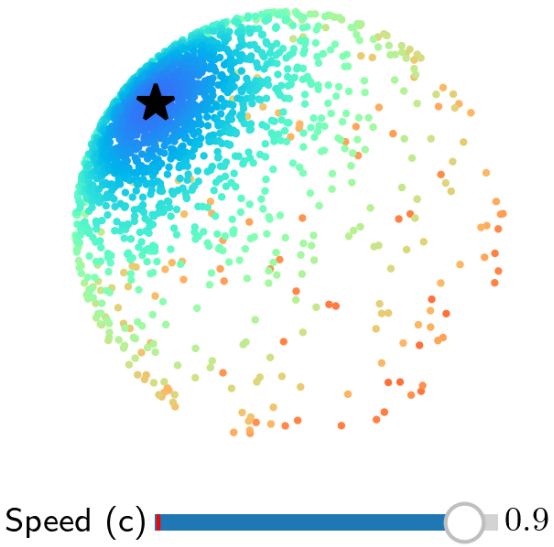
➢ Doppler boosting

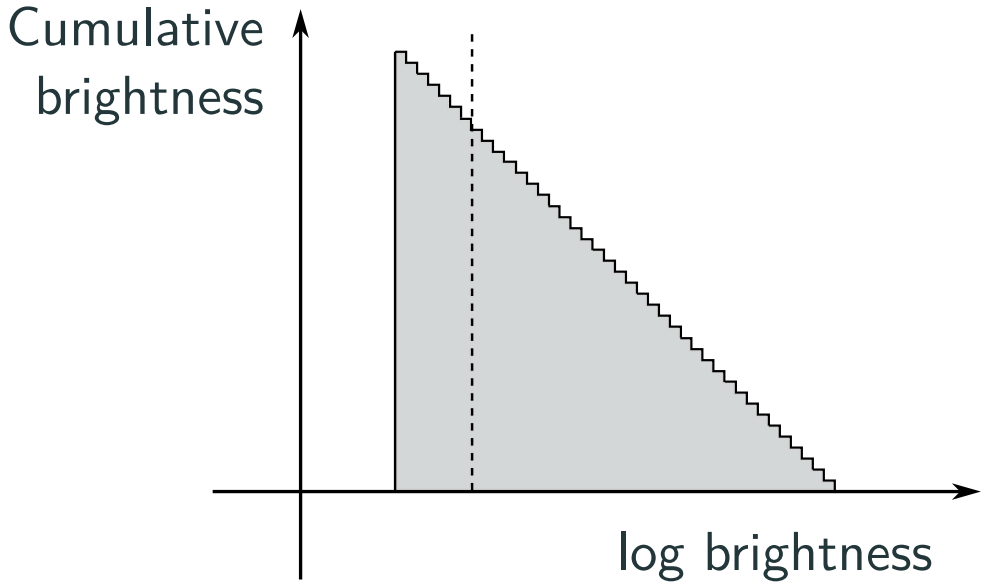




The Cosmic Dipole — Ellis & Baldwin (1984)

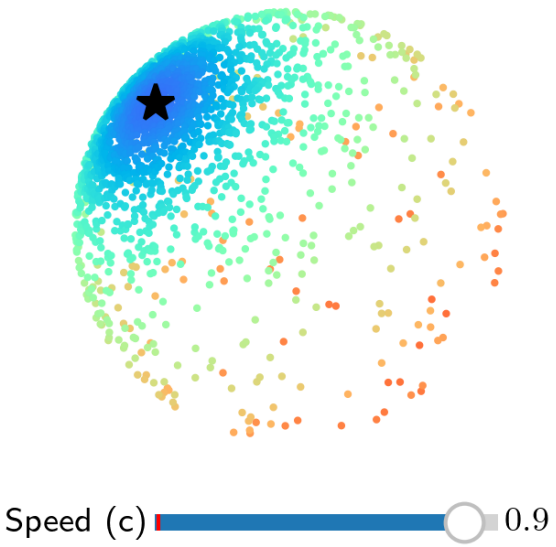
Special relativity: ➢ Aberration ➢ Doppler boosting

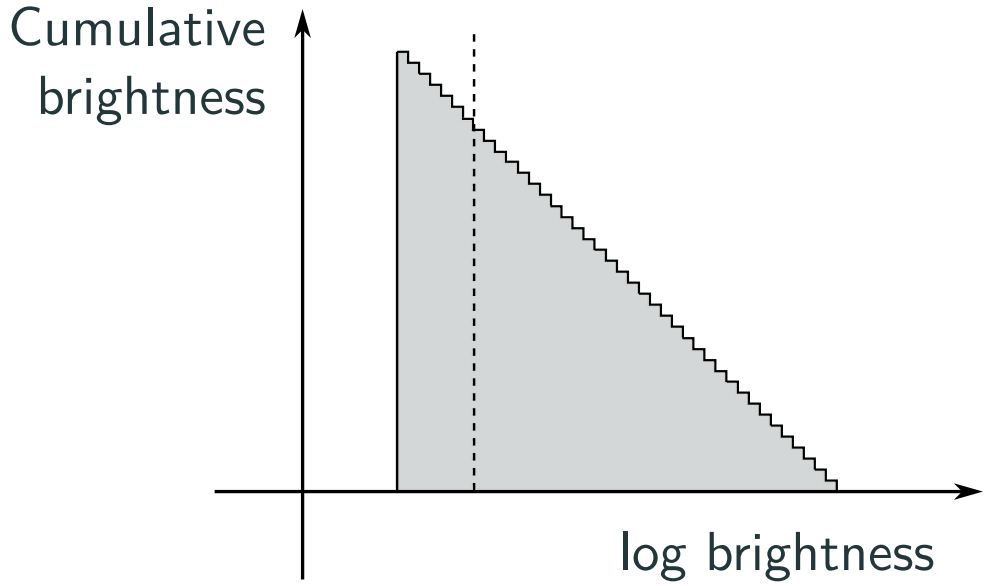




The Cosmic Dipole — Ellis & Baldwin (1984)

Special relativity: ➢ Aberration ➢ Doppler boosting



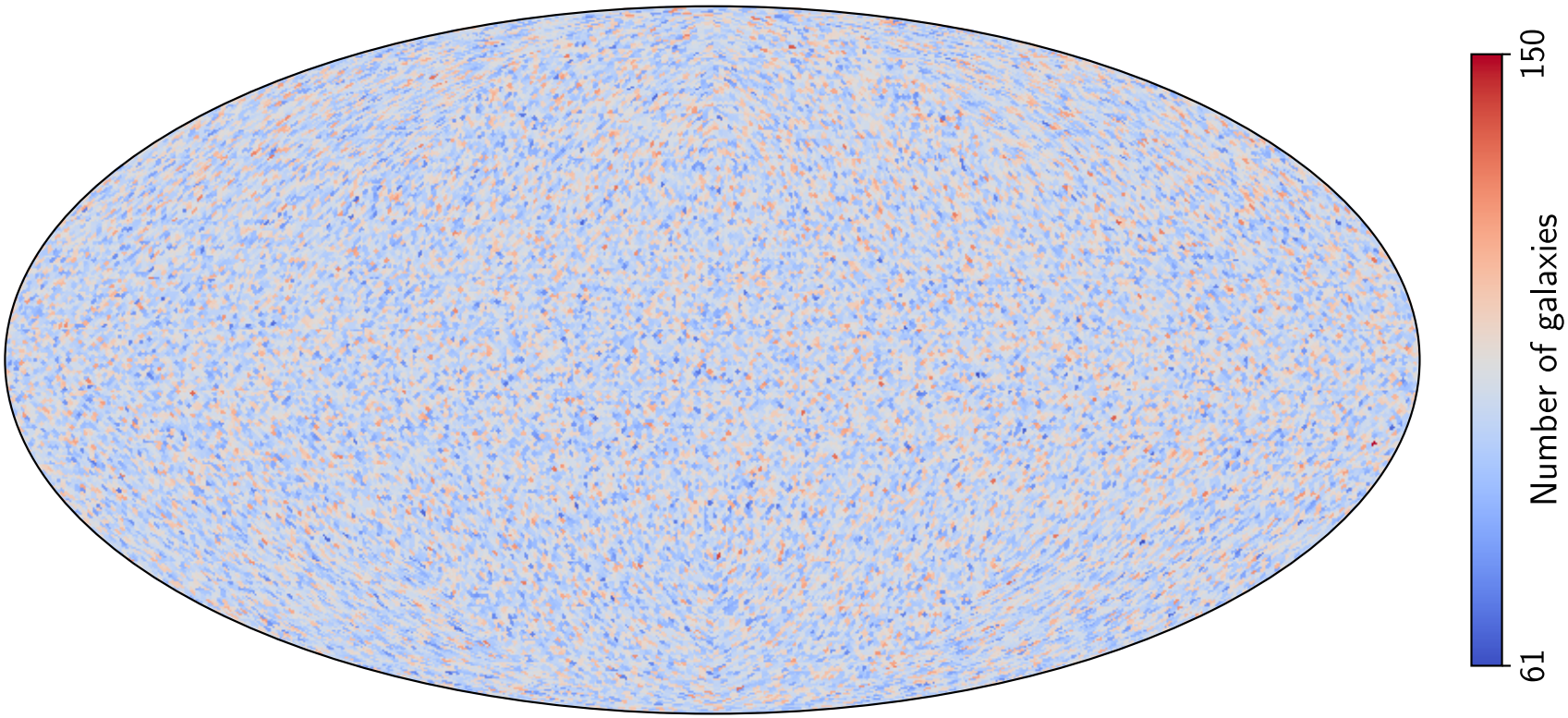


$$\mathcal{D}_{\mathrm{CMB}} = [2 + x\,(1 + \alpha)]\,\frac{v_{\mathrm{CMB}}}{c}.$$

Typical values: $\approx 0.5\%$

Counting Galaxies

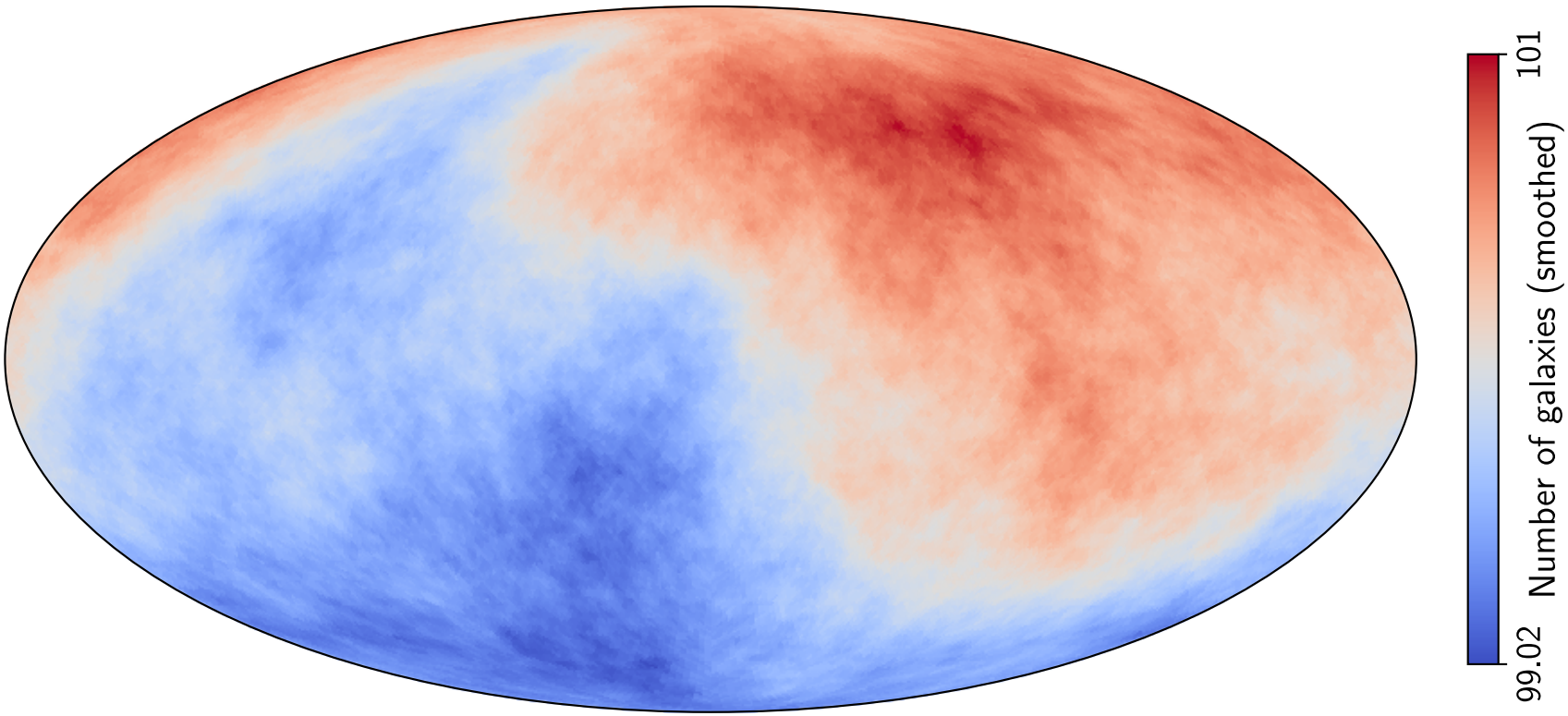
$$N_i = \overline{N}(1 + \mathcal{D}\cos\theta_i)$$



Simulated isotropic galaxy map (+ dipole).

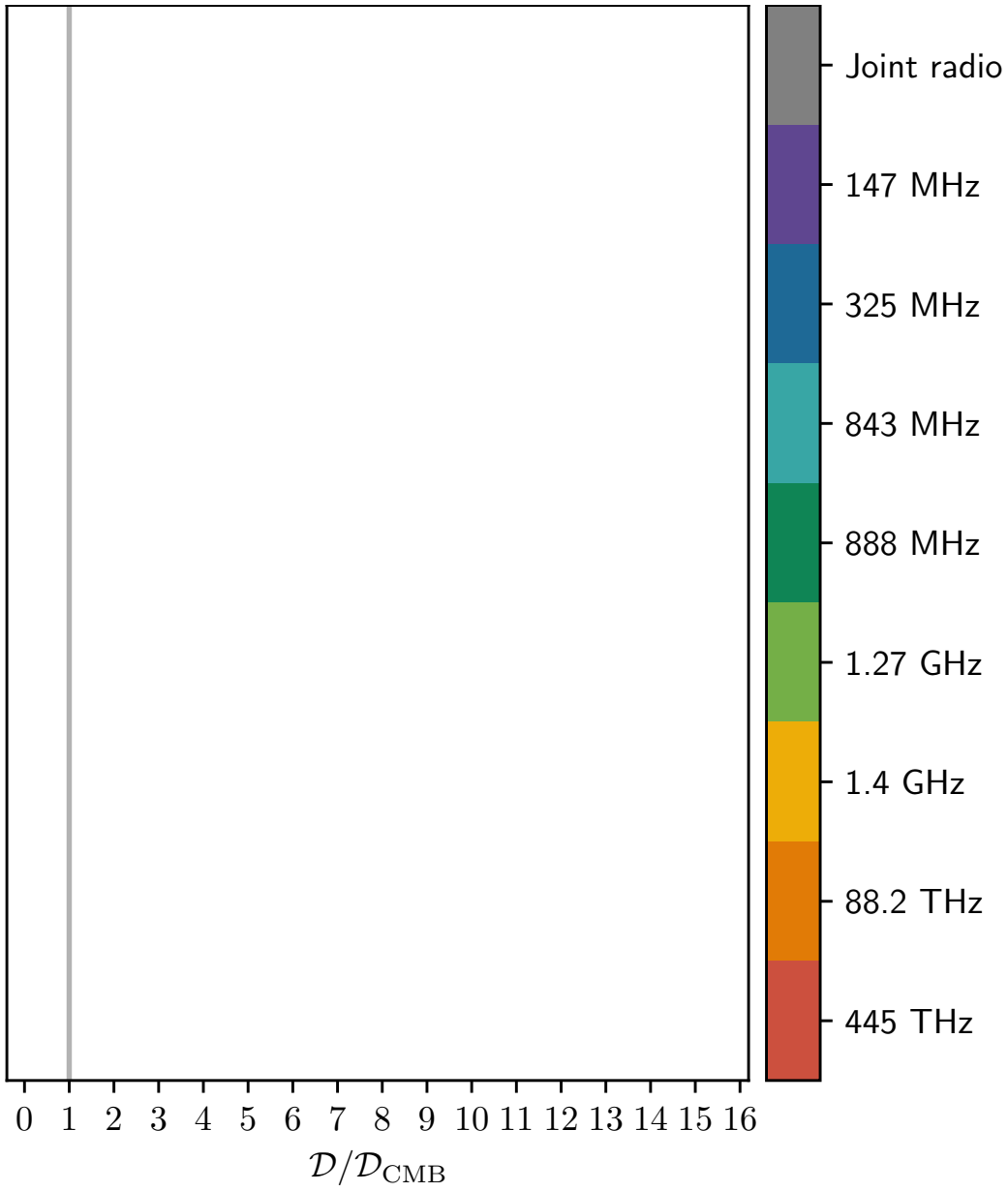
Counting Galaxies

$$N_i = \overline{N}(1 + \mathcal{D}\cos\theta_i)$$



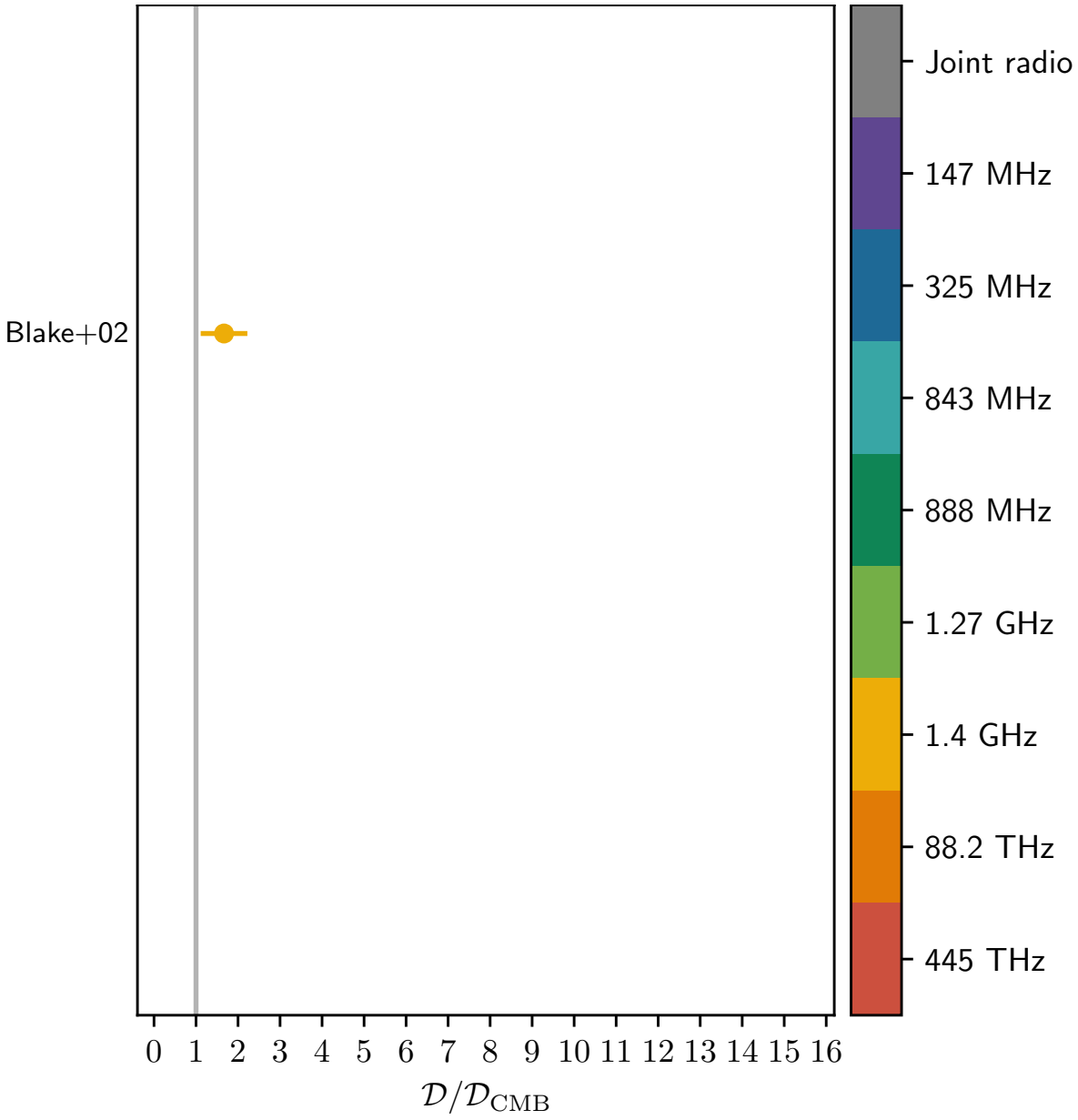
Simulated isotropic galaxy map (+ dipole).

The Amplitude Excess



The cosmic dipole should be consistent with CMB dipole.

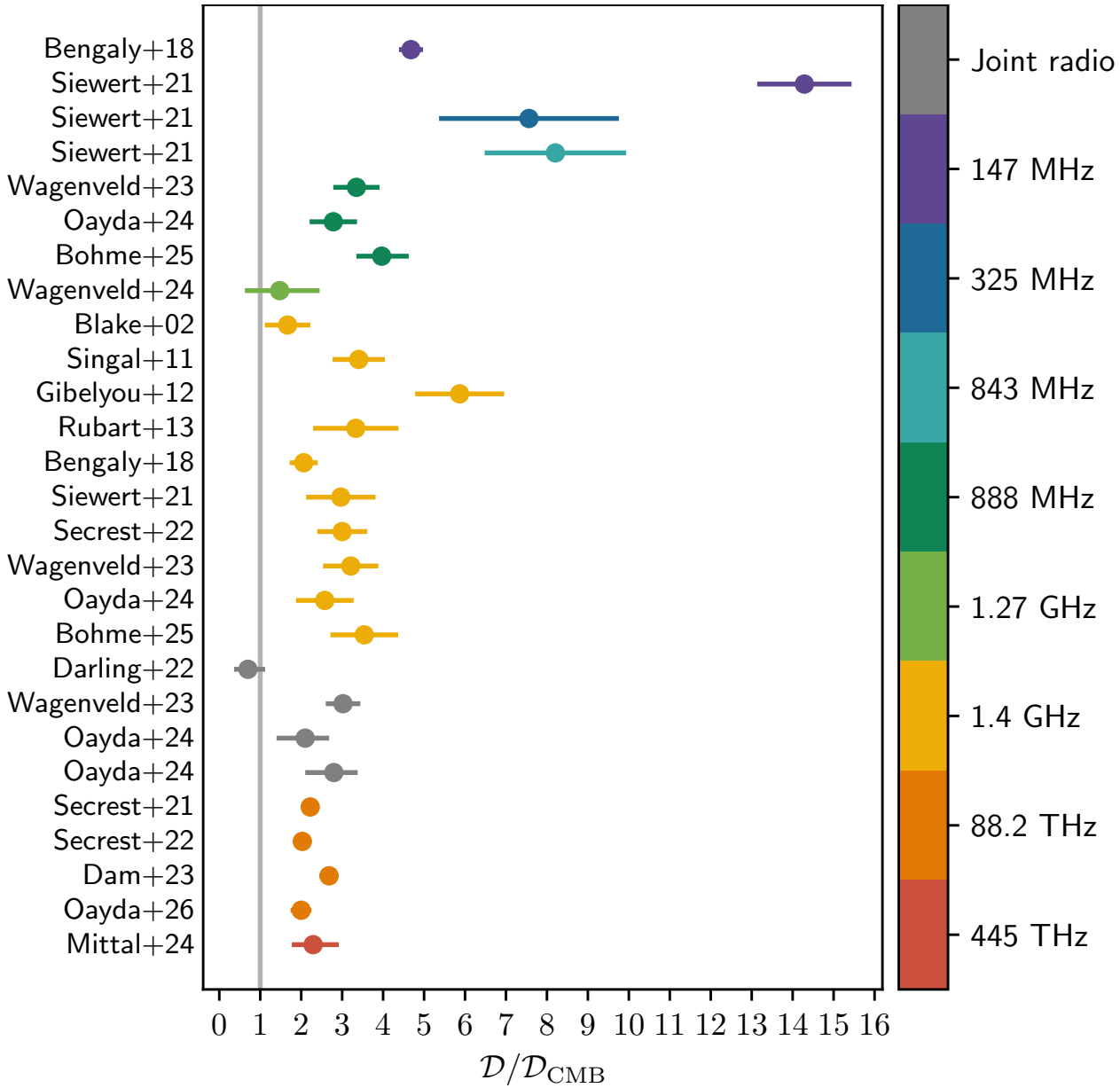
The Amplitude Excess



The cosmic dipole should be consistent with CMB dipole.

All is well!

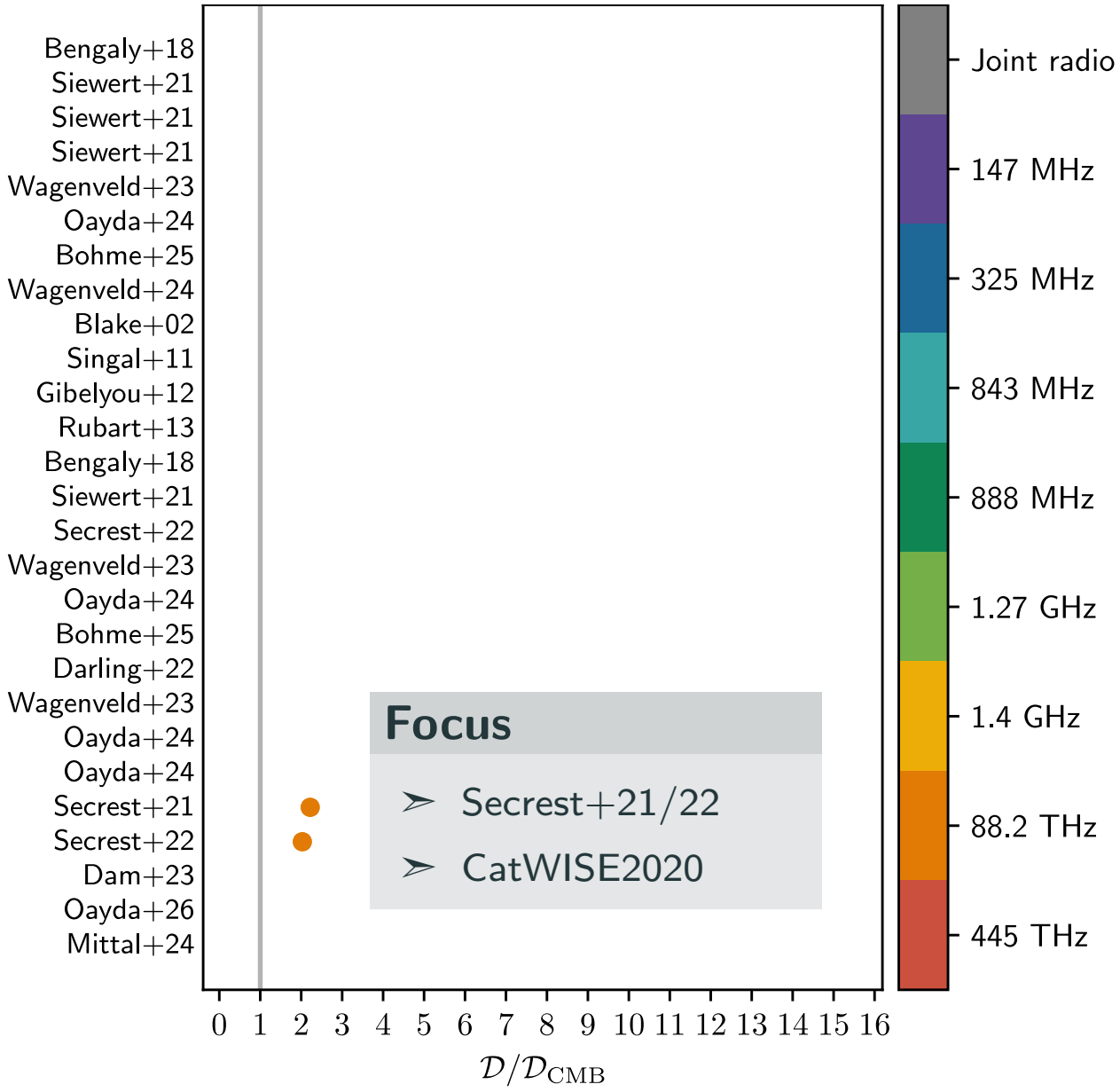
The Amplitude Excess



The cosmic dipole should be consistent with CMB dipole.

All is well! Wait...

The Amplitude Excess



The cosmic dipole should be consistent with CMB dipole.

All is well! Wait...

The CatWISE Sample (Secrest+21)

Two cuts ➢ $\mathrm{W1} < 16.4$ ➢ $\mathrm{W1} - \mathrm{W2} \geq 0.8$

The CatWISE Sample (Secrest+21)

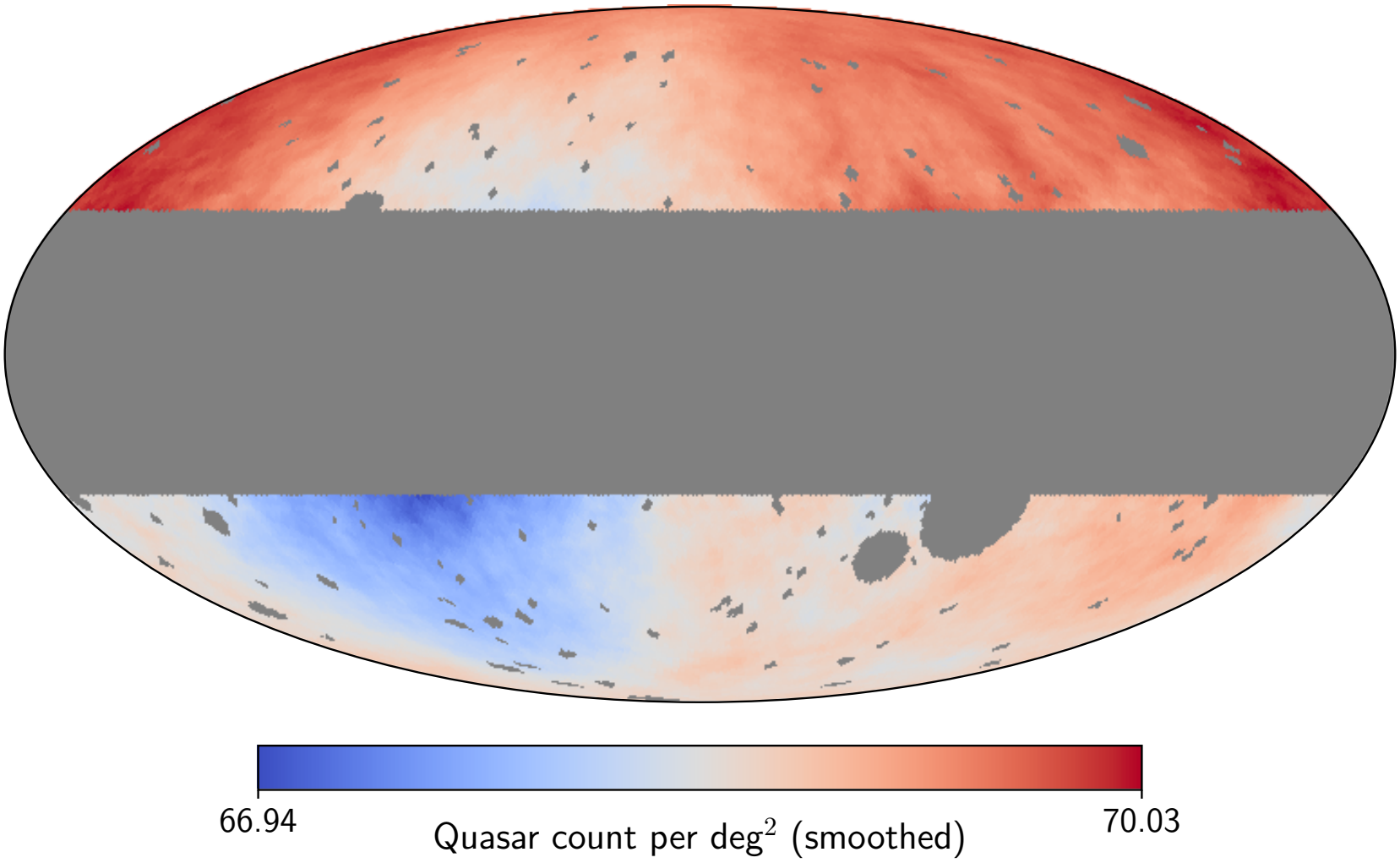
Two cuts ➢ W1 < 16.4 ➢ W1 $-$ W2 ≥ 0.8

The CatWISE Sample (Secrest+21)

Two cuts ➢ W1 < 16.4 ➢ W1 $-$ W2 ≥ 0.8

The CatWISE Sample (Secrest+21)

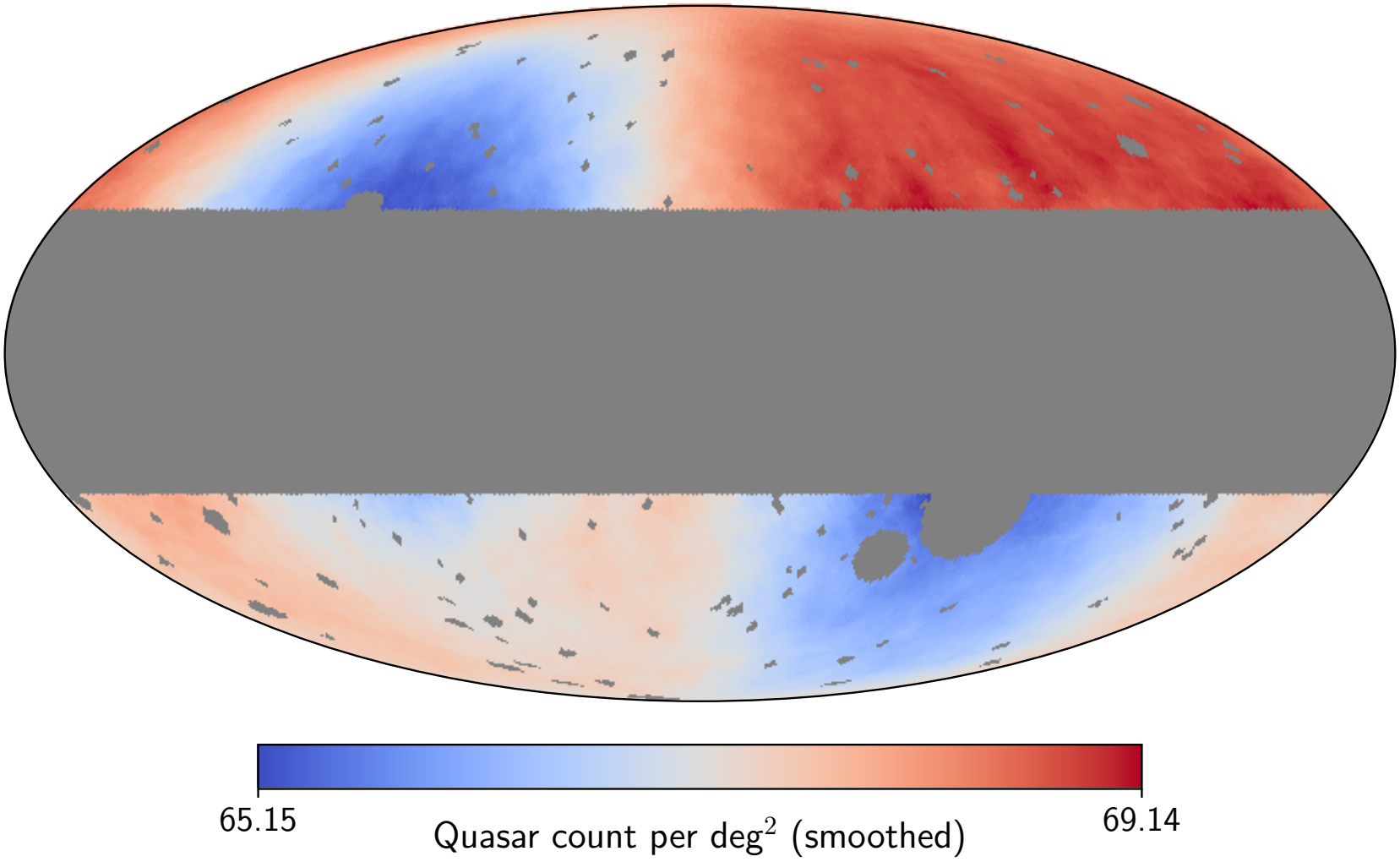
Two cuts ➢ $\mathrm{W1} < 16.4$ ➢ $\mathrm{W1} - \mathrm{W2} \geq 0.8$



CatWISE quasar map from Secrest+21.

The CatWISE Sample (Secrest+21)

Two cuts ➢ $\mathrm{W1} < 16.4$ ➢ $\mathrm{W1} - \mathrm{W2} \geq 0.8$



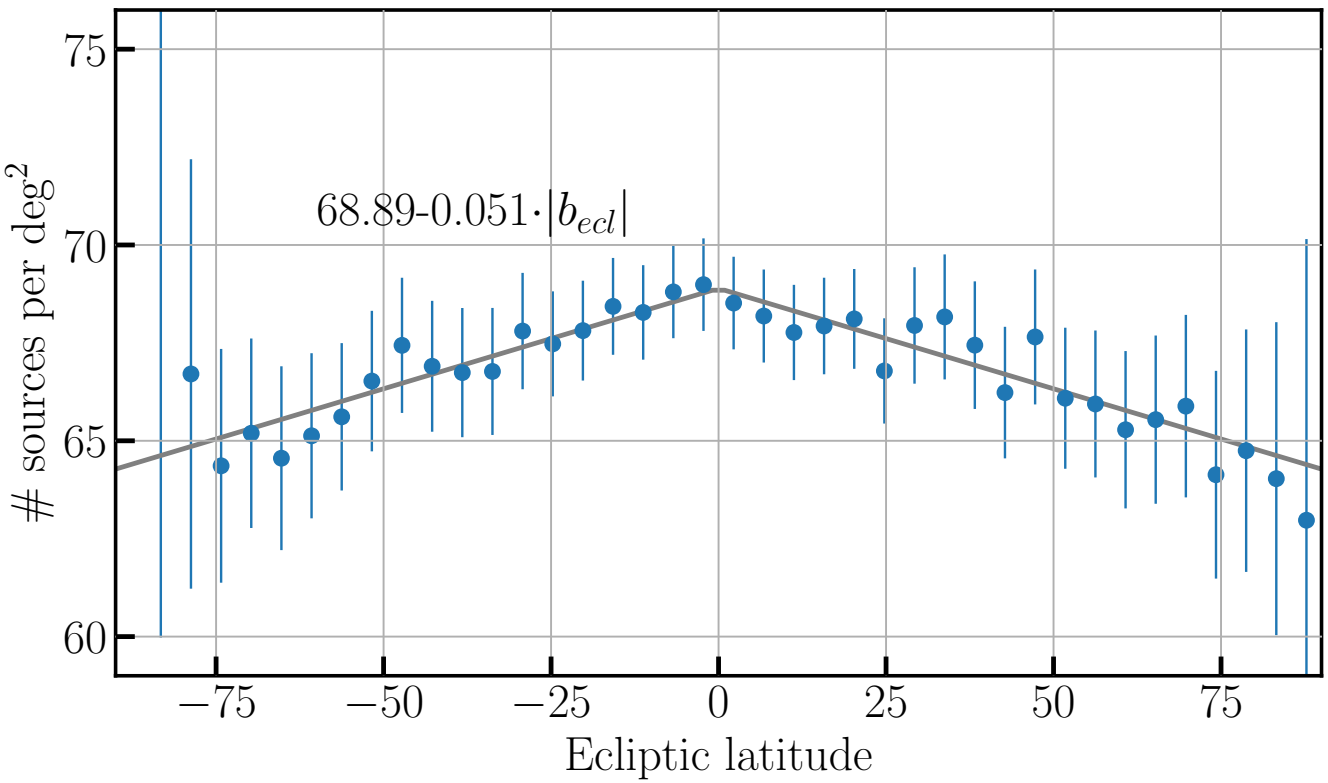
CatWISE quasar map, no linear weighting.

The CatWISE Sample (Secrest+21)

The Astrophysical Journal Letters, 908:L51 (6pp), 2021 February 20 https://doi.org/10.3847/2041-8213/abdd40

A Test of the Cosmological Principle with Quasars

Nathan J. Secrest[1], Sebastian von Hausegger[2,3,4], Mohamed Rameez[5], Roya Mohayaee[3], Subir Sarkar[4], and Jacques Colin[3]



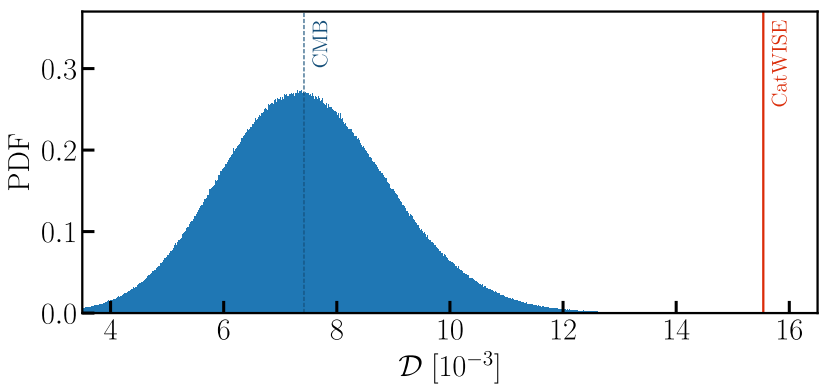
Linear fit to density vs. declination (Secrest+21).

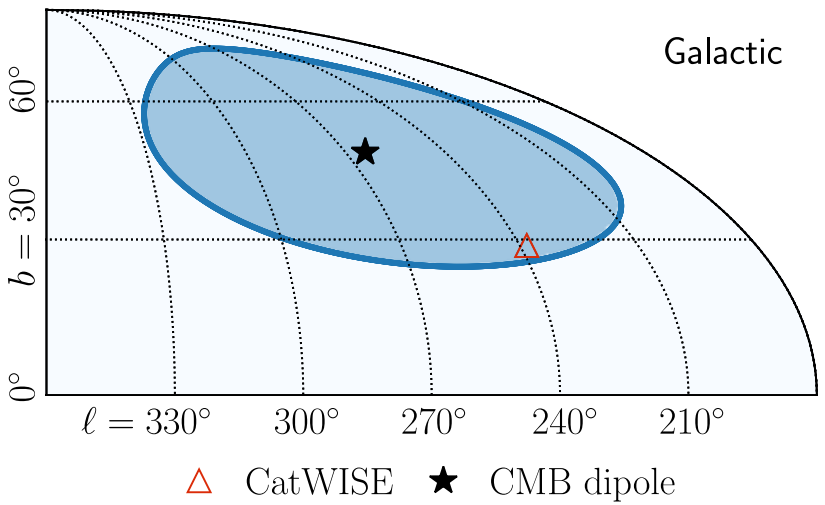
The CatWISE Sample (Secrest+21)

The Astrophysical Journal Letters, 908:L51 (6pp), 2021 February 20 https://doi.org/10.3847/2041-8213/abdd40

A Test of the Cosmological Principle with Quasars

Nathan J. Secrest[1], Sebastian von Hausegger[2,3,4], Mohamed Rameez[5], Roya Mohayaee[3], Subir Sarkar[4], and Jacques Colin[3]





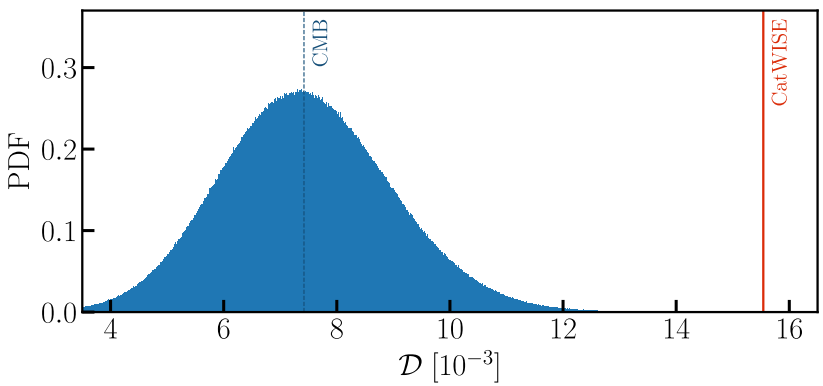
Null significance test (dipole amplitude + direction).

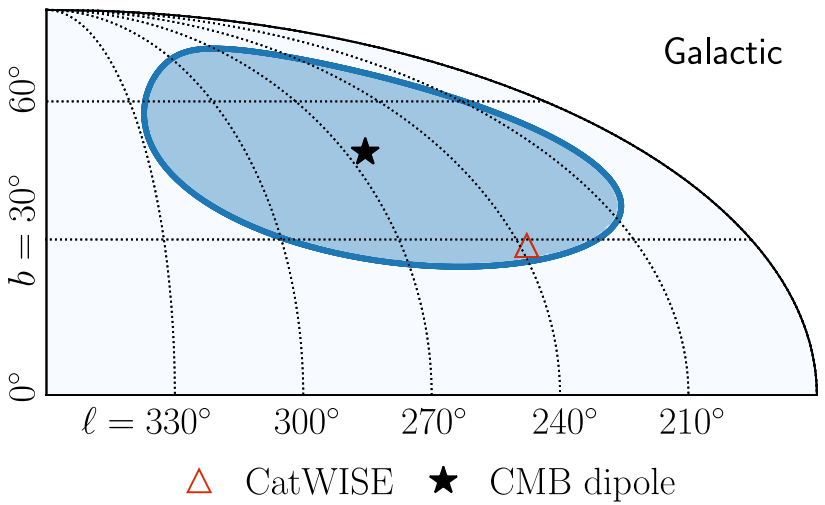
The CatWISE Sample (Secrest+21)

The Astrophysical Journal Letters, 908:L51 (6pp), 2021 February 20
https://doi.org/10.3847/2041-8213/abdd40

A Test of the Cosmological Principle with Quasars

Nathan J. Secrest[1], Sebastian von Hausegger[2,3,4], Mohamed Rameez[5], Roya Mohayaee[3], Subir Sarkar[4], and Jacques Colin[3]





Null significance test (dipole amplitude + direction).

$\approx 5\sigma$ disagreement with the expected amplitude!

The CatWISE Sample (Secrest+22)

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1 https://doi.org/10.3847/2041-8213/ac88c0

A Challenge to the Standard Cosmological Model

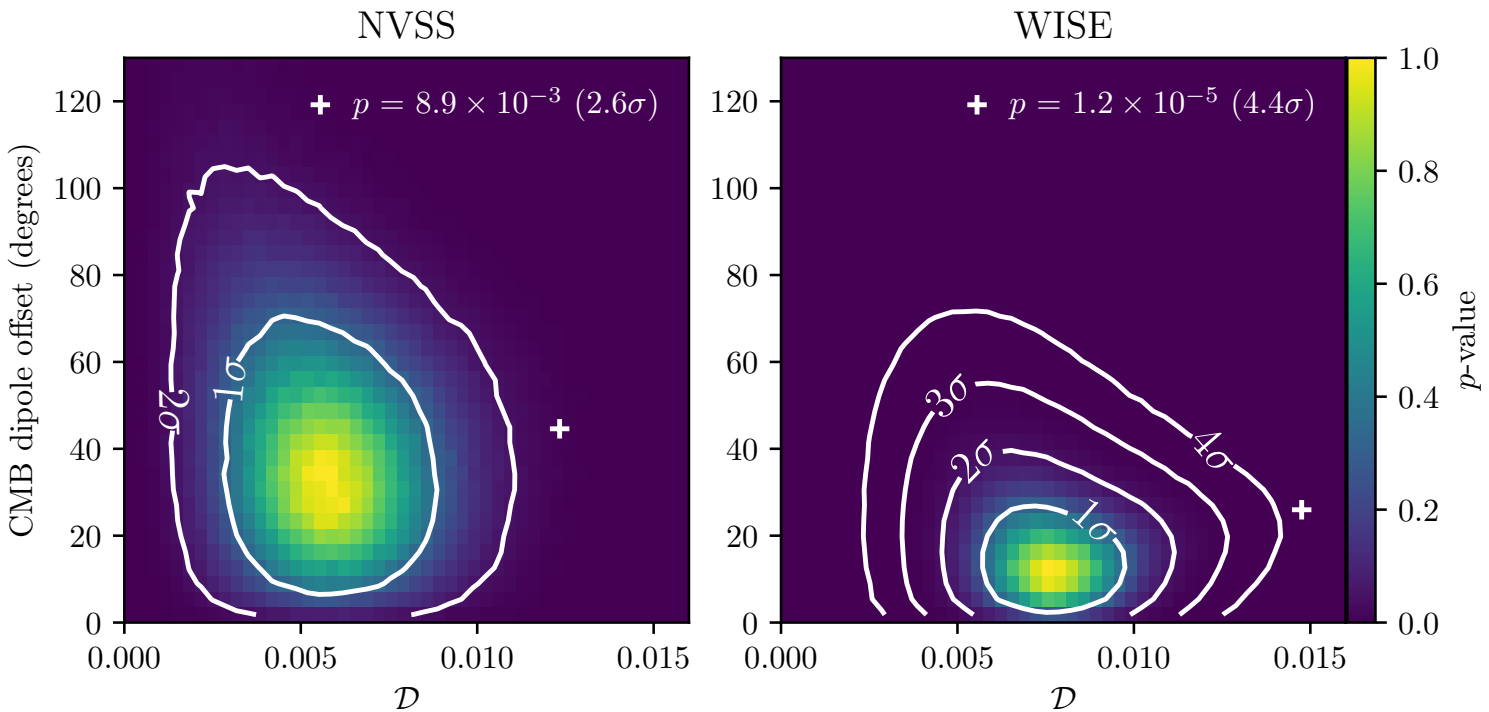
Nathan J. Secrest[1], Sebastian von Hausegger[2], Mohamed Rameez[3], Roya Mohayaee[2,4], and Subir Sarkar[2]

The CatWISE Sample (Secrest+22)

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1 https://doi.org/10.3847/2041-8213/ac88c0

A Challenge to the Standard Cosmological Model

Nathan J. Secrest[1], Sebastian von Hausegger[2], Mohamed Rameez[3], Roya Mohayaee[2,4], and Subir Sarkar[2]



Null significance test (dipole amplitude + offset).

Bayesian Treatment of CatWISE

Testing the cosmological principle with CatWISE quasars: a bayesian analysis of the number-count dipole

Lawrence Dam [1,2★] Geraint F. Lewis [1★] and Brendon J. Brewer[3]

- Frequentist: ecliptic bias is a correction/weighting to $\mathbf{D}$.

Bayesian Treatment of CatWISE

Testing the cosmological principle with CatWISE quasars: a bayesian analysis of the number-count dipole

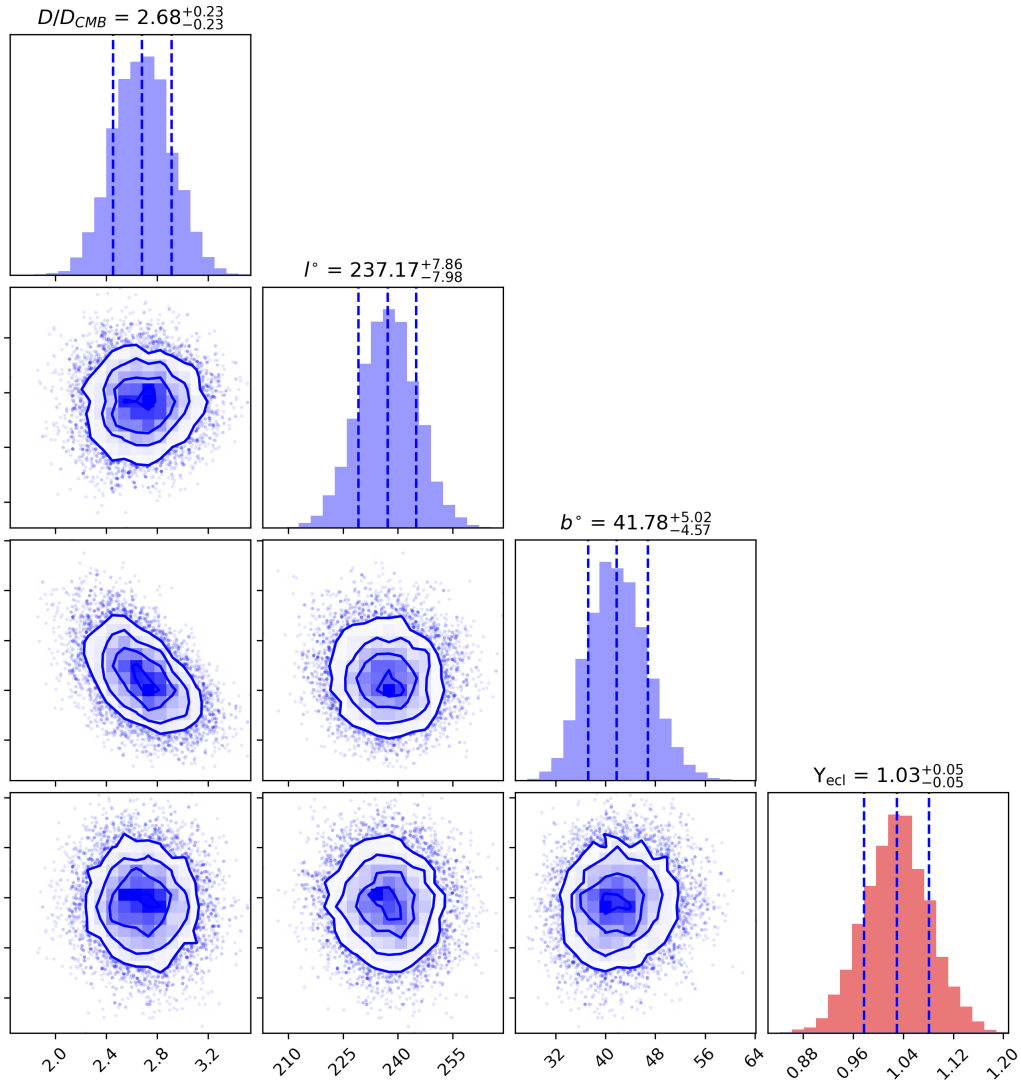
Lawrence Dam [1,2★] Geraint F. Lewis [1★] and Brendon J. Brewer[3]

- Frequentist: ecliptic bias is a correction/weighting to $\mathbf{D}$.
- Bayesian: ecliptic bias is a parameter for our model $P(\mathbf{D}|\Theta, M)$.

Bayesian Treatment of CatWISE

Testing the cosmological principle with CatWISE quasars: a bayesian analysis of the number-count dipole

Lawrence Dam [1,2★] Geraint F. Lewis [1★] and Brendon J. Brewer[3]

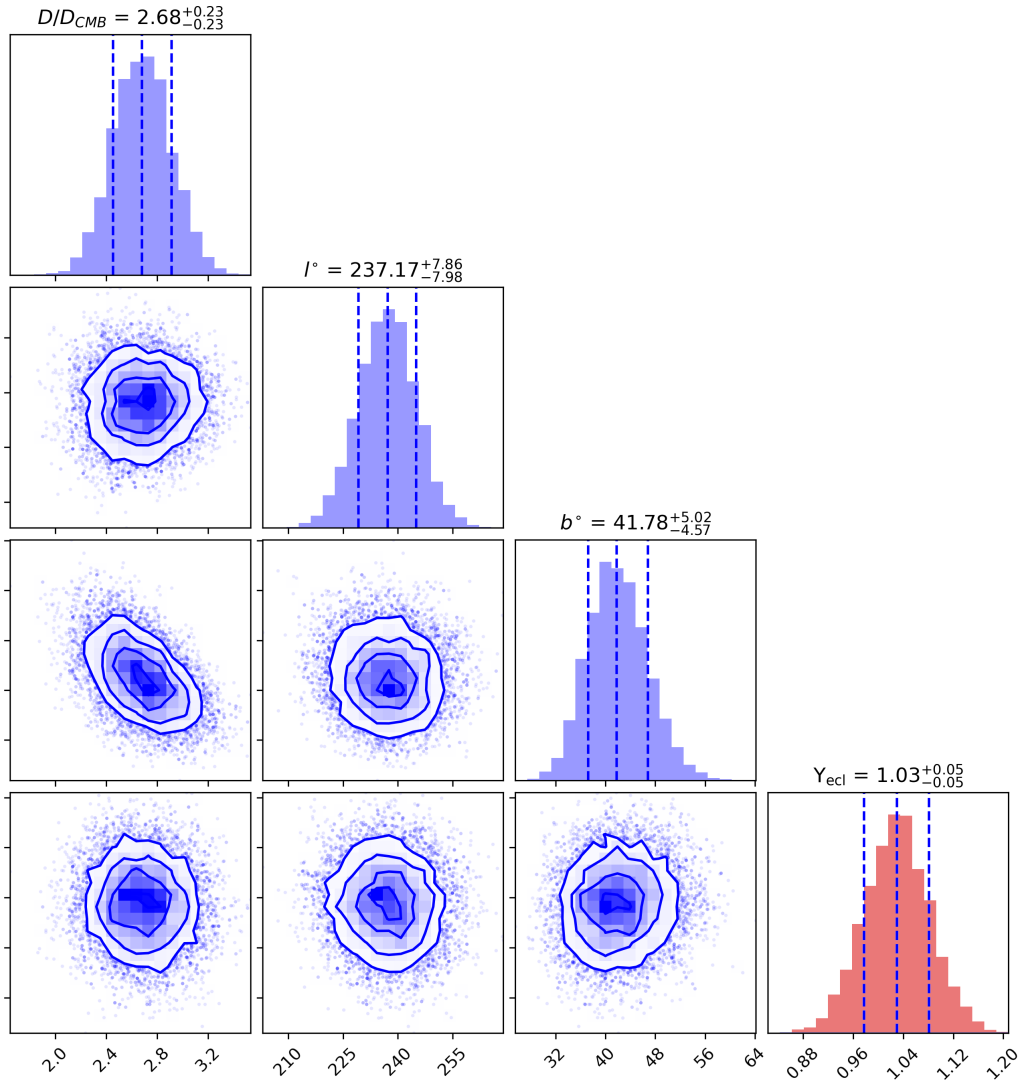


$$N(\hat{\mathrm{p}}_i) = \overline{N} f_{\mathrm{ecl.}} (1 + \mathcal{D} \cos\theta_i)$$

Bayesian Treatment of CatWISE

Testing the cosmological principle with CatWISE quasars: a bayesian analysis of the number-count dipole

Lawrence Dam [1,2★] Geraint F. Lewis [1★] and Brendon J. Brewer[3]



$$N(\hat{p}_i) = \overline{N} f_{ecl.} (1 + \mathcal{D} \cos\theta_i)$$

Model	ln B
Null (no dipole)	0.0
Free dipole, no bias	83.5
Free dipole	263.5

Can we produce and do inference with forward simulations that explain the ecliptic bias?

Cause of the Ecliptic Bias

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1

objects, and a flux cut of $W1 < 16.4$ mag for uniform sensitivity across the sky. Objects below an absolute Galactic latitude of 30° were excluded because of the drop in sky pixel density due to source confusion near the Galactic plane. A slight inverse linear trend was also observed between ecliptic latitude and sky density, which is potentially attributable to two effects. First, deeper coverage near the ecliptic poles increases sensitivity to faint sources that, while excluded by our flux cut, may cause deblending issues with brighter sources and lead to a loss of completeness. Second, shallower coverage near the ecliptic equator may lead to AGNs slightly bluer than $W1-W2 \geqslant 0.8$ scattering redward due to photometric error, increasing apparent source density if bluer AGNs are more common, as is implied in Figure 2 of Secrest et al. (2021). A detailed characterization of these effects is beyond the scope of this work; for our purposes it suffices that the ecliptic latitude trend is easy to correct for.

- Source confusion?
- Eddington bias?
- Abghari+24: stellar contamination?

But clearly it must be related to the scanning law.

Cause of the Ecliptic Bias

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1

objects, and a flux cut of $W1 < 16.4$ mag for uniform sensitivity across the sky. Objects below an absolute Galactic latitude of 30° were excluded because of the drop in sky pixel density due to source confusion near the Galactic plane. A slight inverse linear trend was also observed between ecliptic latitude and sky density, which is potentially attributable to two effects. First, deeper coverage near the ecliptic poles increases sensitivity to faint sources that, while excluded by our flux cut, may cause deblending issues with brighter sources and lead to a loss of completeness. Second, shallower coverage near the ecliptic equator may lead to AGNs slightly bluer than $W1-W2 \geqslant 0.8$ scattering redward due to photometric error, increasing apparent source density if bluer AGNs are more common, as is implied in Figure 2 of Secrest et al. (2021). A detailed characterization of these effects is beyond the scope of this work; for our purposes it suffices that the ecliptic latitude trend is easy to correct for.

- Source confusion?
- Eddington bias?
- Abghari+24: stellar contamination?

But clearly it must be related to the scanning law.

Cause of the Ecliptic Bias

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1

objects, and a flux cut of $W1 < 16.4$ mag for uniform sensitivity across the sky. Objects below an absolute Galactic latitude of 30° were excluded because of the drop in sky pixel density due to source confusion near the Galactic plane. A slight inverse linear trend was also observed between ecliptic latitude and sky density, which is potentially attributable to two effects. First, deeper coverage near the ecliptic poles increases sensitivity to faint sources that, while excluded by our flux cut, may cause deblending issues with brighter sources and lead to a loss of completeness. Second, shallower coverage near the ecliptic equator may lead to AGNs slightly bluer than $W1-W2 \geqslant 0.8$ scattering redward due to photometric error, increasing apparent source density if bluer AGNs are more common, as is implied in Figure 2 of Secrest et al. (2021). A detailed characterization of these effects is beyond the scope of this work; for our purposes it suffices that the ecliptic latitude trend is easy to correct for.

- Source confusion?
- Eddington bias?
- Abghari+24: stellar contamination?

But clearly it must be related to the scanning law.

Cause of the Ecliptic Bias

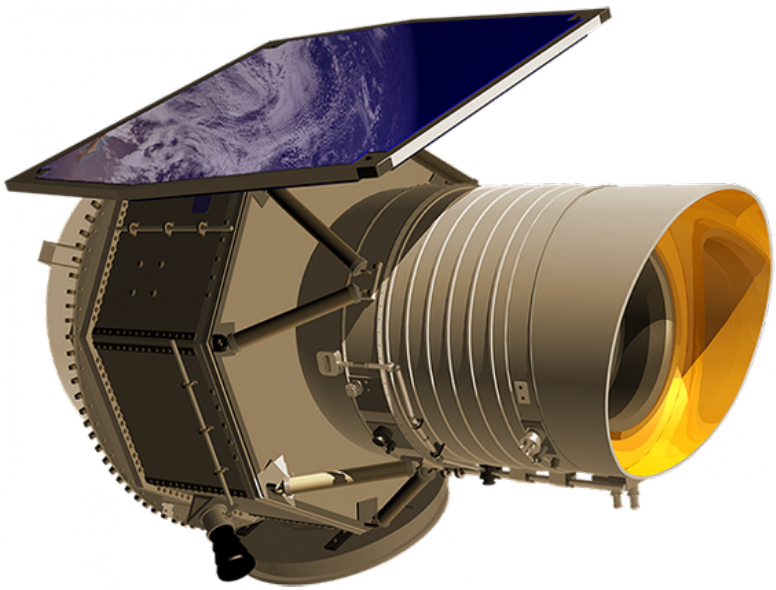
The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1

objects, and a flux cut of $W1 < 16.4$ mag for uniform sensitivity across the sky. Objects below an absolute Galactic latitude of 30° were excluded because of the drop in sky pixel density due to source confusion near the Galactic plane. A slight inverse linear trend was also observed between ecliptic latitude and sky density, which is potentially attributable to two effects. First, deeper coverage near the ecliptic poles increases sensitivity to faint sources that, while excluded by our flux cut, may cause deblending issues with brighter sources and lead to a loss of completeness. Second, shallower coverage near the ecliptic equator may lead to AGNs slightly bluer than $W1-W2 \geqslant 0.8$ scattering redward due to photometric error, increasing apparent source density if bluer AGNs are more common, as is implied in Figure 2 of Secrest et al. (2021). A detailed characterization of these effects is beyond the scope of this work; for our purposes it suffices that the ecliptic latitude trend is easy to correct for.

- Source confusion?
- Eddington bias?
- Abghari+24: stellar contamination?

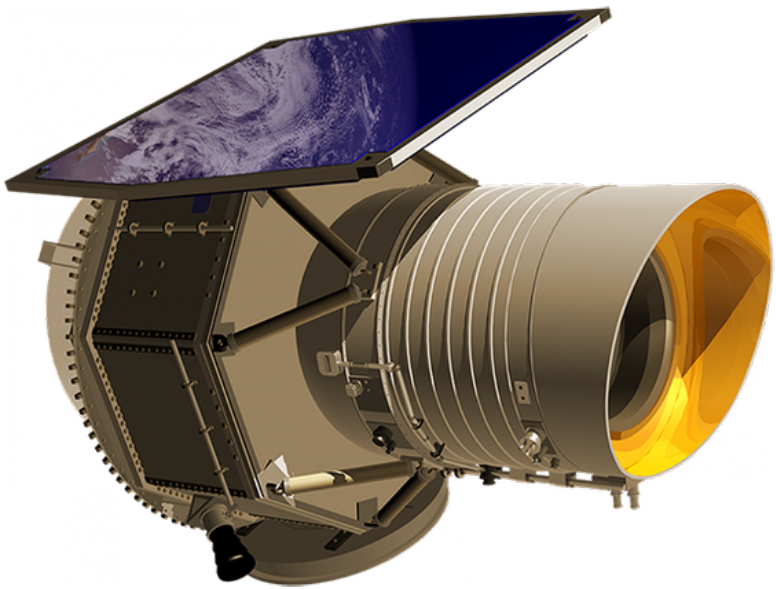
But clearly it must be related to the scanning law.

WISE's Scanning Law

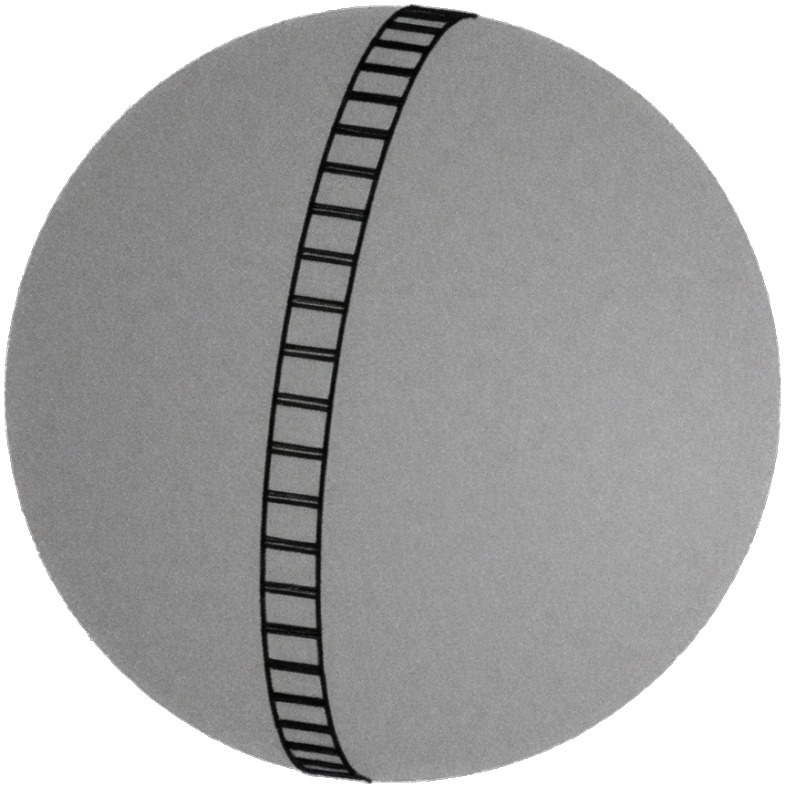


Render of the WISE satellite (NASA JPL).

WISE's Scanning Law

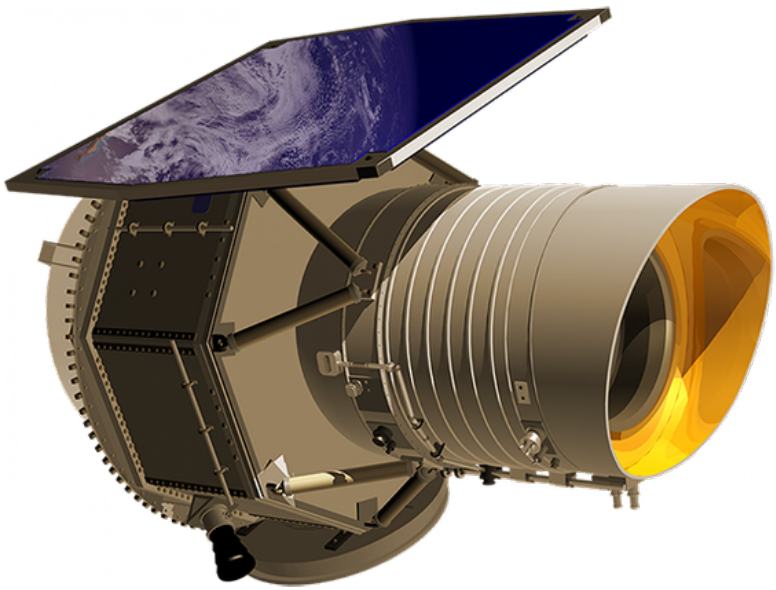


Render of the WISE satellite (NASA JPL).

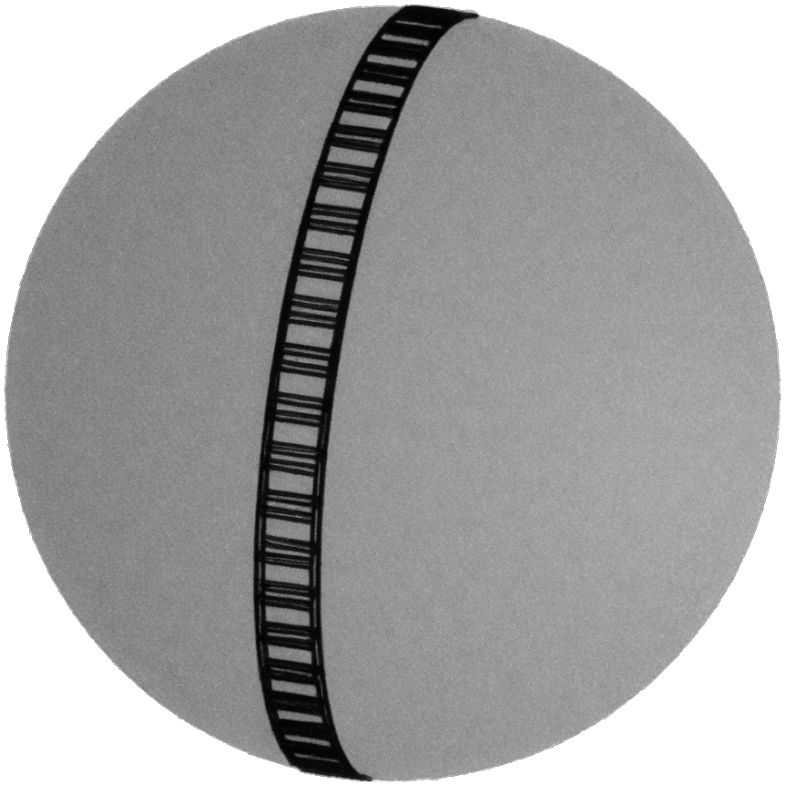


Frames seen over 1 orbit.

WISE's Scanning Law

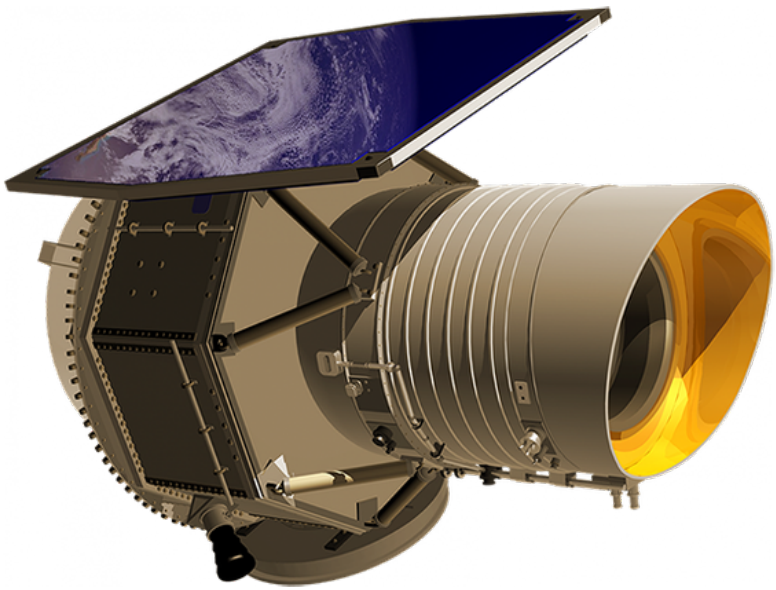


Render of the WISE satellite (NASA JPL).

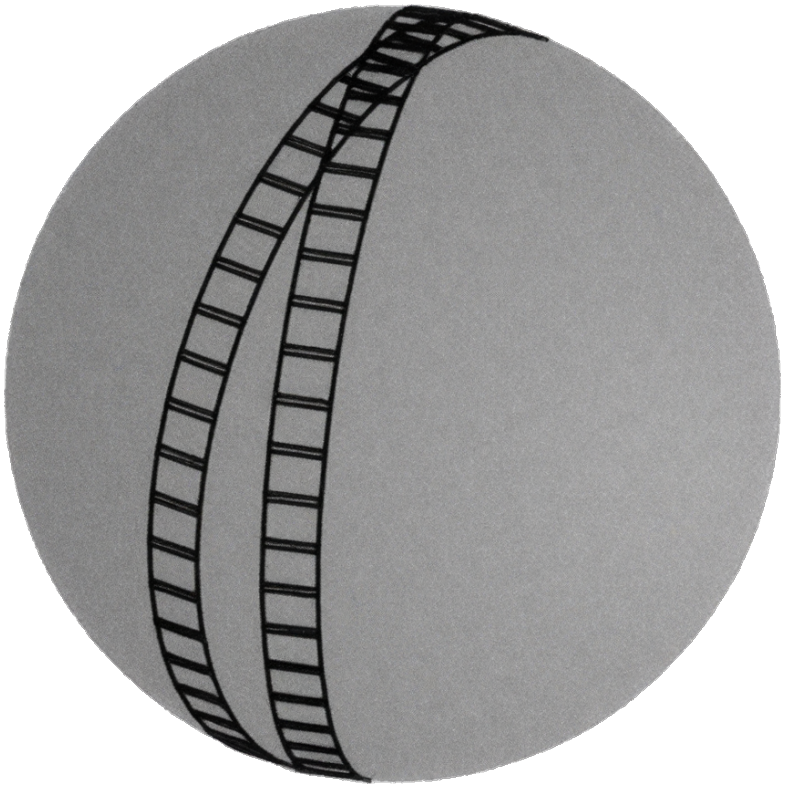


Frames seen over 2 orbits.

WISE's Scanning Law

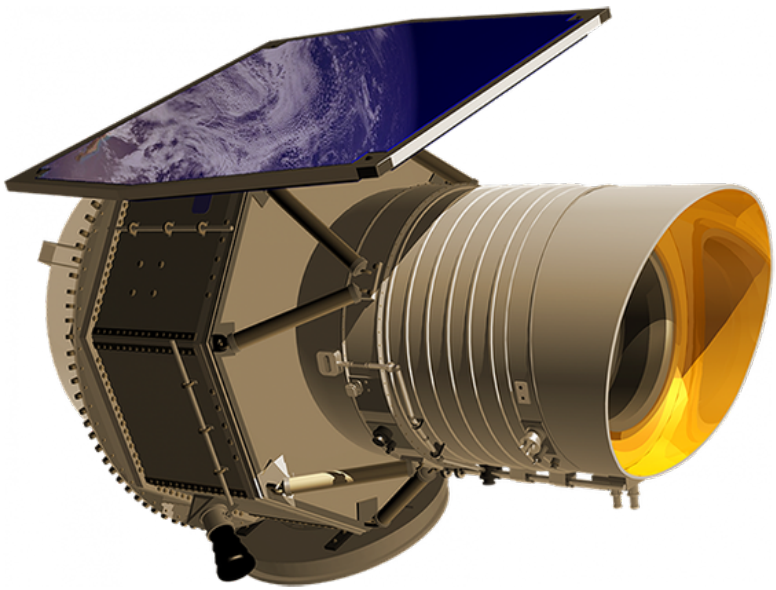


Render of the WISE satellite (NASA JPL).

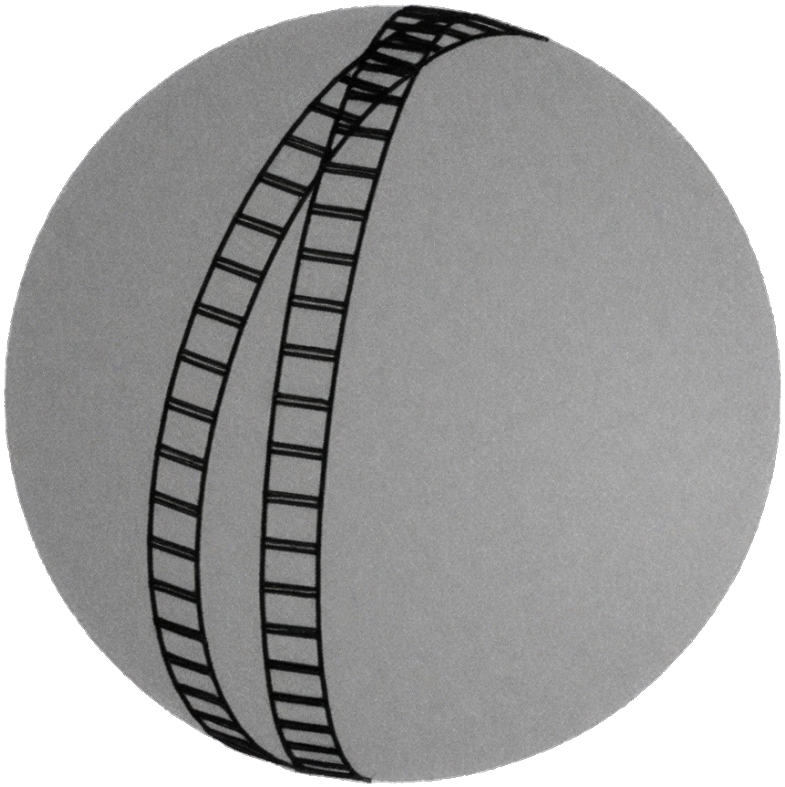


Frames seen over 2 orbits 20 days apart.

WISE's Scanning Law



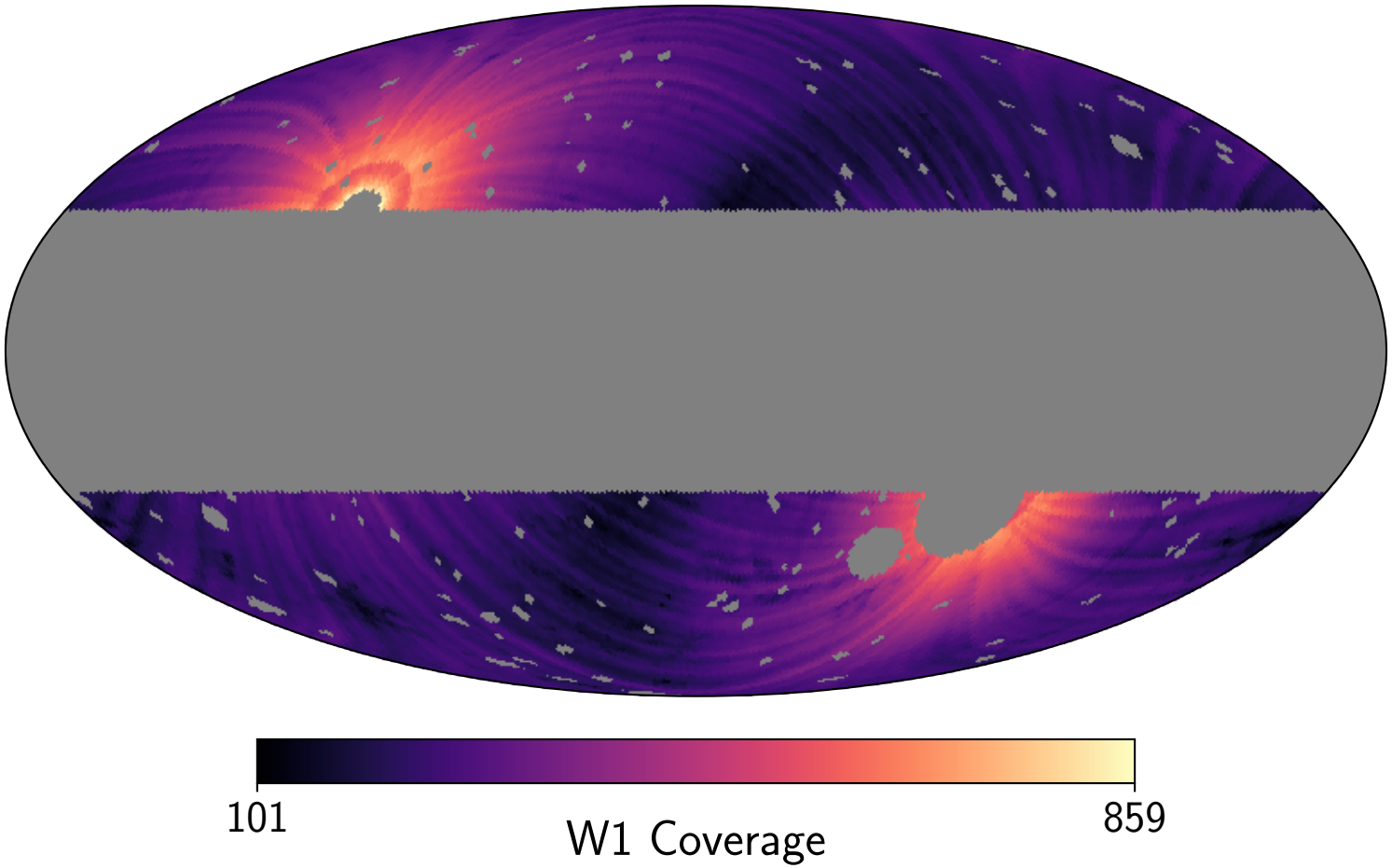
Render of the WISE satellite (NASA JPL).



Frames seen over 2 orbits 20 days apart.

Obeys a scanning law over the survey's lifetime.

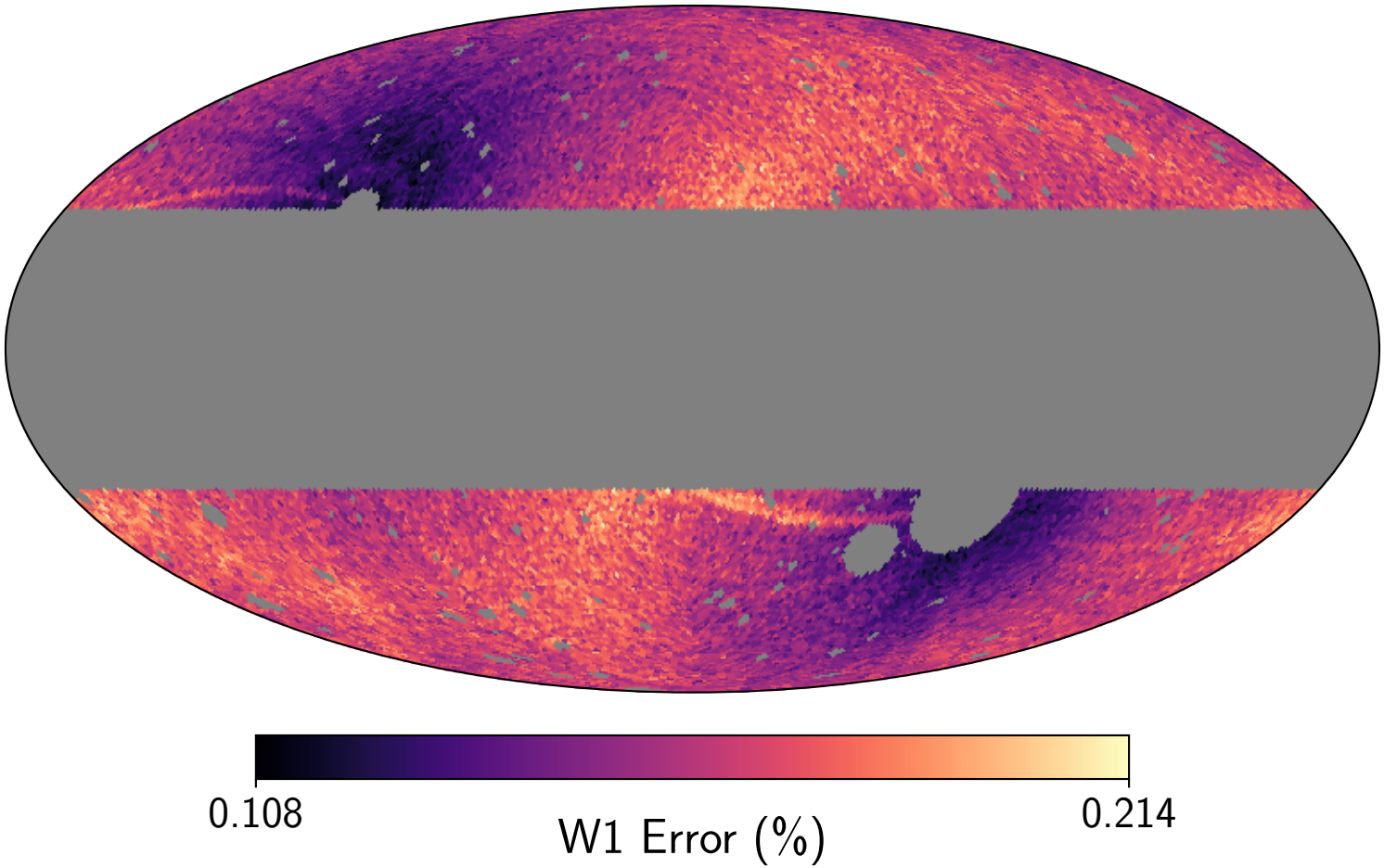
Photometric Errors



WISE coverage in W1 band for the Secrest+21 sample.

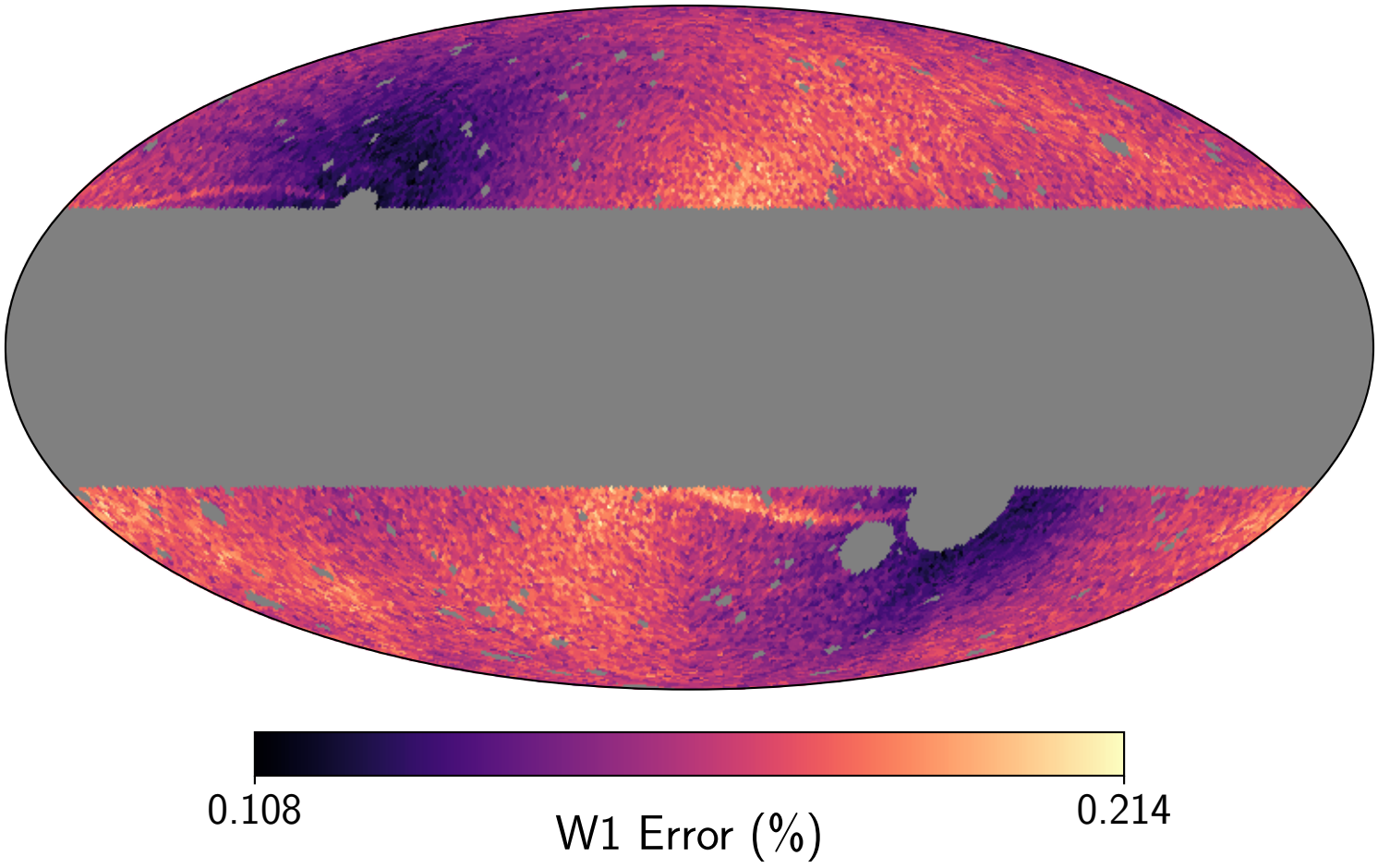
Oayda & Lewis (2026) MNRAS 546(4)

Photometric Errors



Source photometric error (%) in W1 band for the Secrest+21 sample.

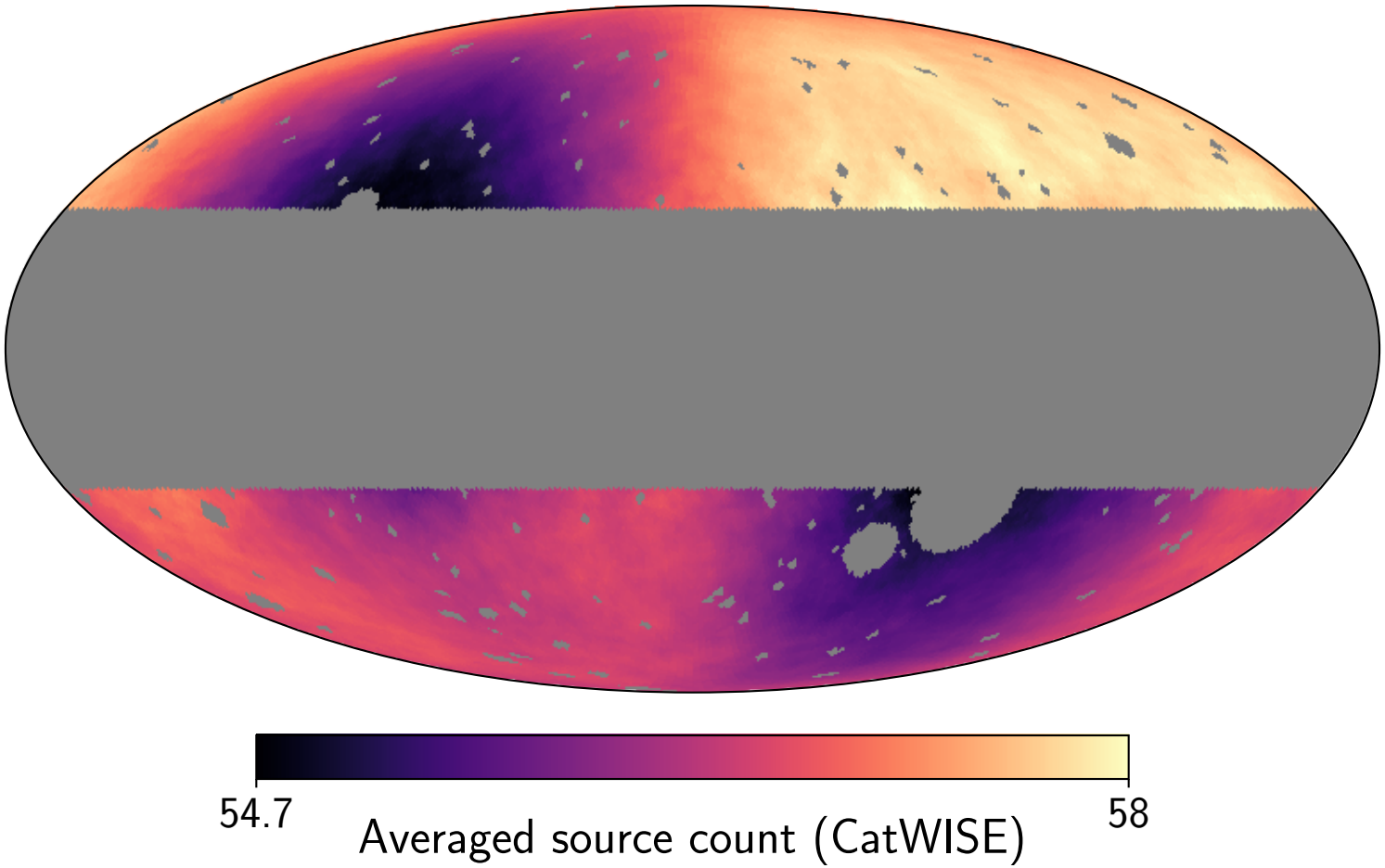
Photometric Errors



Source photometric error (%) in W1 band for the Secrest+21 sample.

We should hope $\sigma \propto \frac{1}{\sqrt{\text{Cov}}}$ — photon counting.

Photometric Errors



CatWISE quasar map, no linear weighting.

We should hope $\sigma \propto \frac{1}{\sqrt{\text{Cov}}}$ — photon counting.

Oayda & Lewis (2026) MNRAS 546(4)

Cause of the Ecliptic Bias

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1

objects, and a flux cut of $W1 < 16.4$ mag for uniform sensitivity across the sky. Objects below an absolute Galactic latitude of 30° were excluded because of the drop in sky pixel density due to source confusion near the Galactic plane. A slight inverse linear trend was also observed between ecliptic latitude and sky density, which is potentially attributable to two effects. First, deeper coverage near the ecliptic poles increases sensitivity to faint sources that, while excluded by our flux cut, may cause deblending issues with brighter sources and lead to a loss of completeness. Second, shallower coverage near the ecliptic equator may lead to AGNs slightly bluer than $W1-W2 \geqslant 0.8$ scattering redward due to photometric error, increasing apparent source density if bluer AGNs are more common, as is implied in Figure 2 of Secrest et al. (2021). A detailed characterization of these effects is beyond the scope of this work; for our purposes it suffices that the ecliptic latitude trend is easy to correct for.

- Source confusion?
- Eddington bias?
- Abghari+24: stellar contamination?

But clearly it must be related to the scanning law.

Cause of the Ecliptic Bias

The Astrophysical Journal Letters, 937:L31 (9pp), 2022 October 1

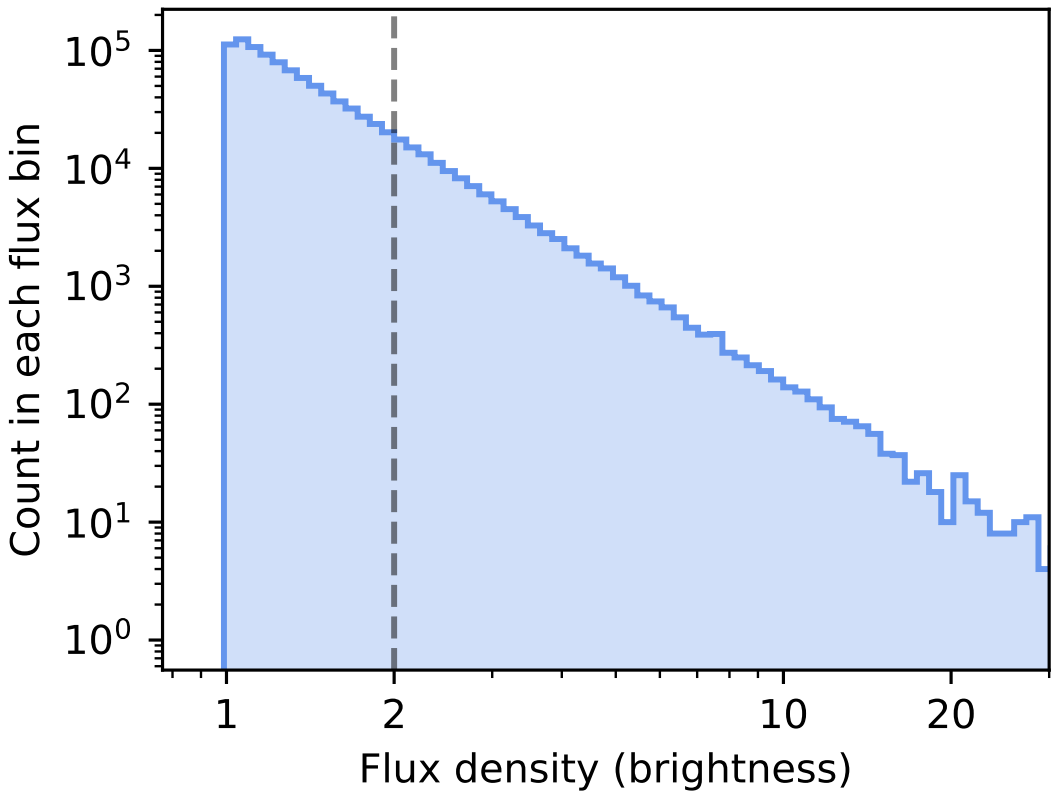
objects, and a flux cut of $W1 < 16.4$ mag for uniform sensitivity across the sky. Objects below an absolute Galactic latitude of 30° were excluded because of the drop in sky pixel density due to source confusion near the Galactic plane. A slight inverse linear trend was also observed between ecliptic latitude and sky density, which is potentially attributable to two effects. First, deeper coverage near the ecliptic poles increases sensitivity to faint sources that, while excluded by our flux cut, may cause deblending issues with brighter sources and lead to a loss of completeness. Second, shallower coverage near the ecliptic equator may lead to AGNs slightly bluer than $W1-W2 \geqslant 0.8$ scattering redward due to photometric error, increasing apparent source density if bluer AGNs are more common, as is implied in Figure 2 of Secrest et al. (2021). A detailed characterization of these effects is beyond the scope of this work; for our purposes it suffices that the ecliptic latitude trend is easy to correct for.

- Source confusion?
- Eddington bias?
- Abghari+24: stellar contamination?

But clearly it must be related to the scanning law.

Eddington Bias

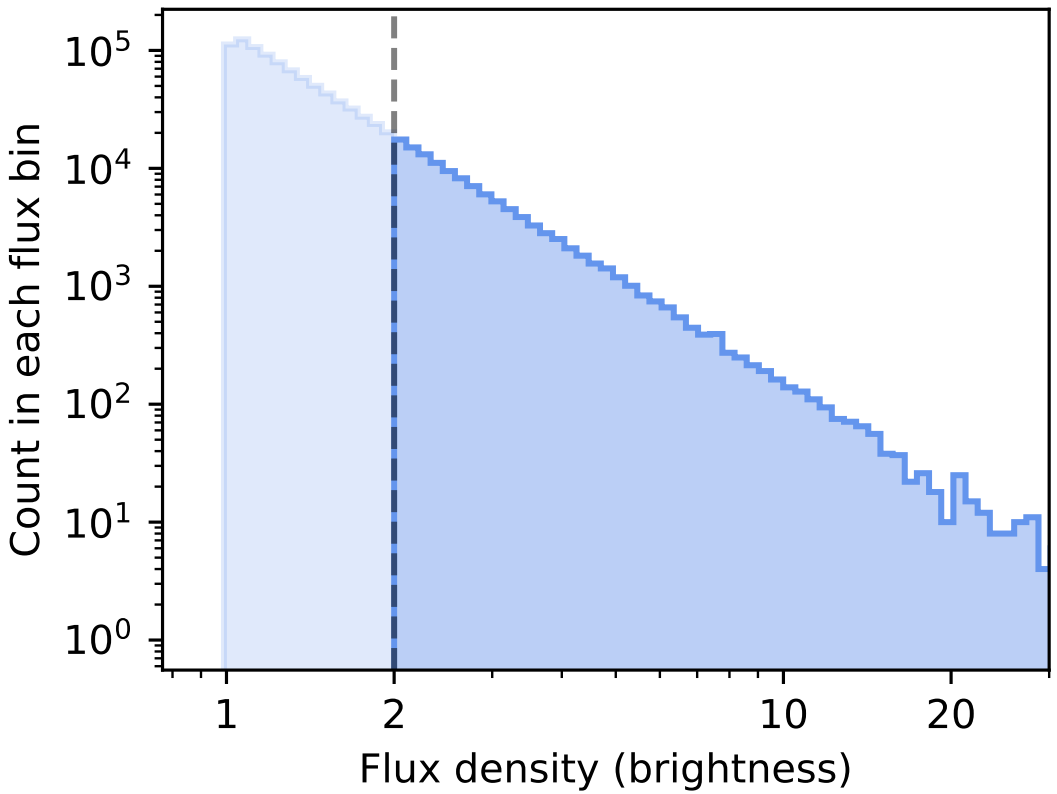
Sampled flux = True flux



True distribution of flux densities.

Eddington Bias

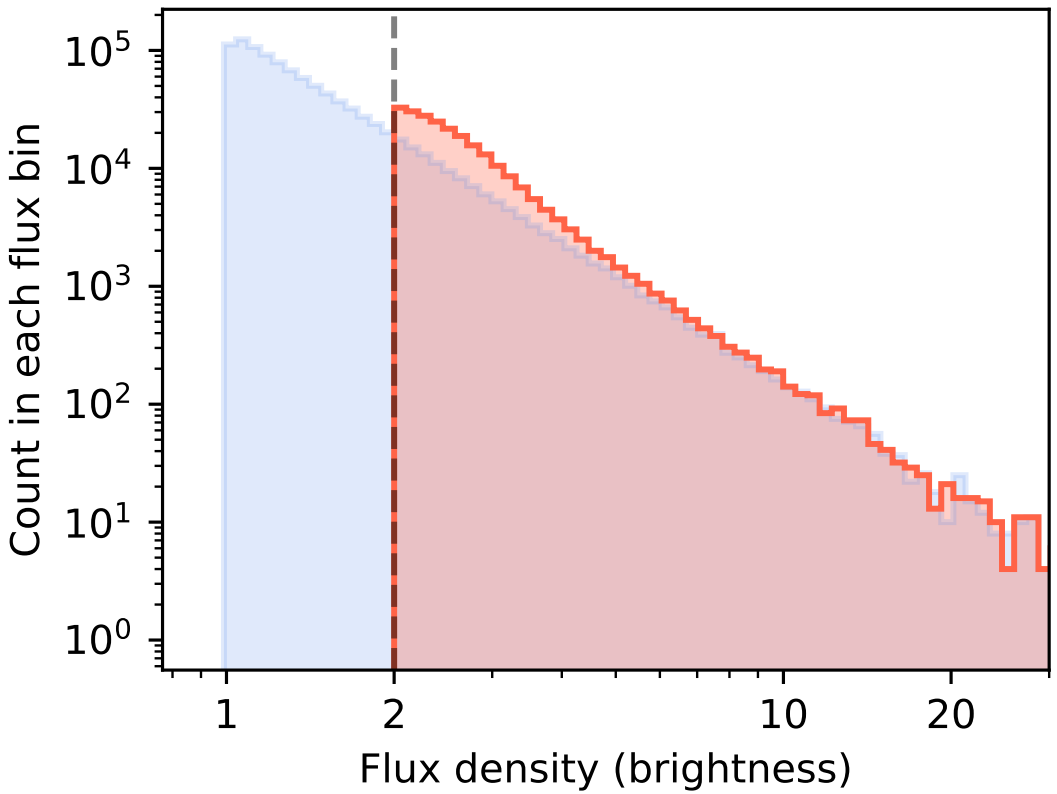
Sampled flux (> 2) = True flux (> 2)



True distribution of flux densities with **flux cut**.

Eddington Bias

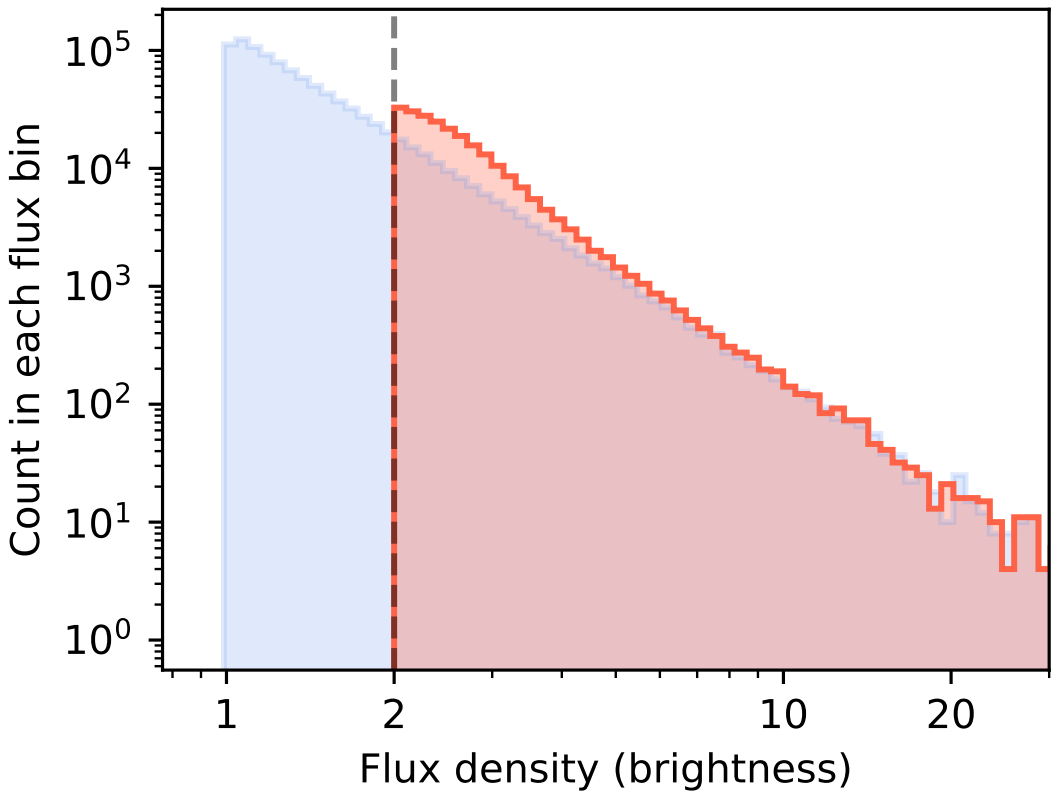
Sampled flux (> 2) = True flux (> 2) + Noise



Noisy distribution of flux densities with **flux cut**.

Eddington Bias

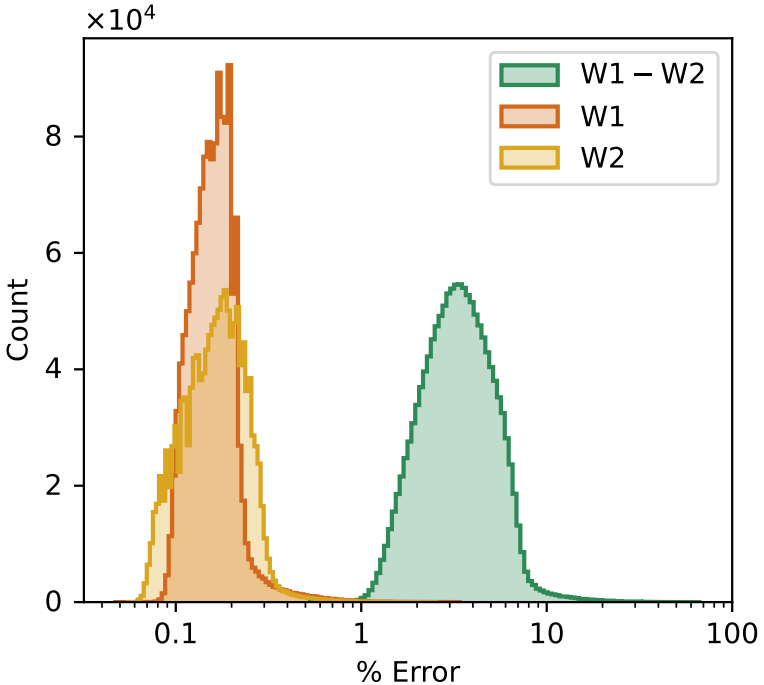
$$\text{Sampled flux } (> 2) = \text{True flux } (> 2) + \text{Noise}$$

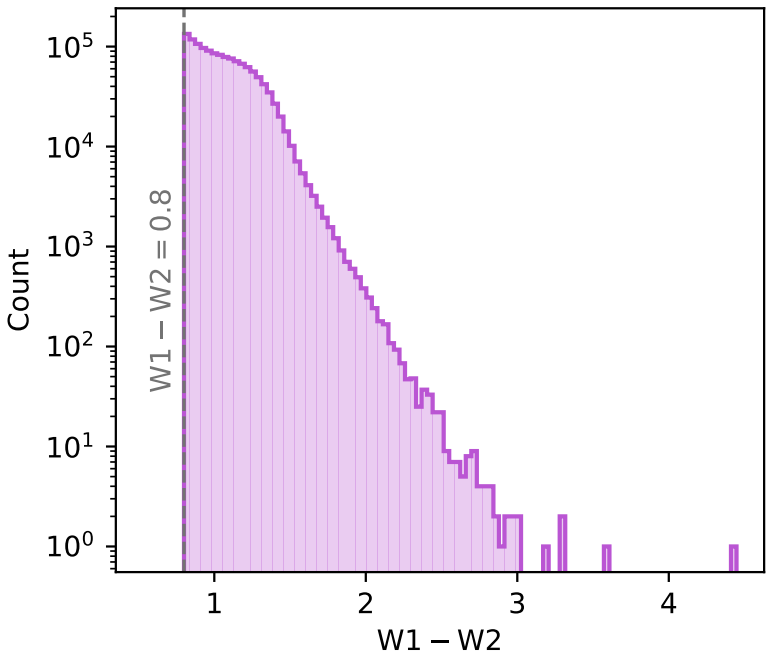


Noisy distribution of flux densities with **flux cut**.

What if the noise varies over the sky? 🤔

CatWISE Photometric Errors





- Colour % errors orders of magnitude larger
- The likely culprit
- Bluer AGNs scatter past the limit

Turning to Simulations

➢ CatSIM — our CatWISE simulation pipeline

Oayda & Lewis (2026) MNRAS 546(4)

Turning to Simulations

➢ CatSIM — our CatWISE simulation pipeline

Draw sources on isotropic sky
sample magnitudes

↓

Do special relativity

↓

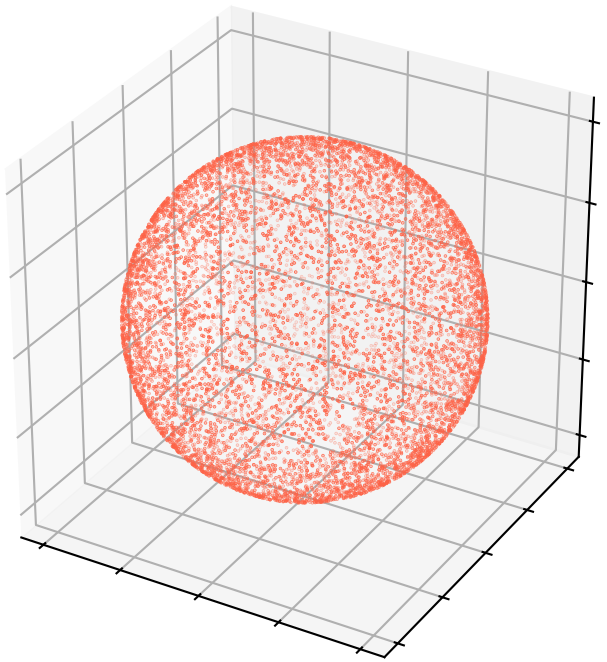
Add photometric errors
$m'_\nu \to m'_\nu + \Delta m'_\nu$

↓

Make magnitude cuts & mask

↓

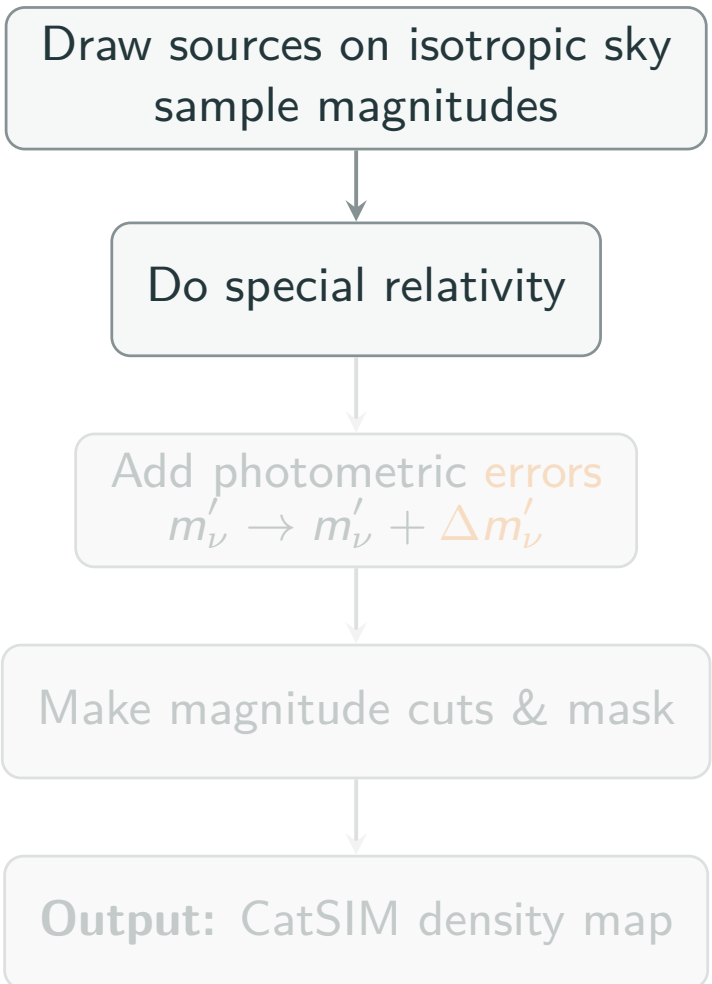
Output: CatSIM density map



Oayda & Lewis (2026) MNRAS 546(4)

Turning to Simulations

➢ CatSIM — our CatWISE simulation pipeline



$$(l^\circ, b^\circ) \to (l'^\circ, b'^\circ)$$
$$m_\nu \to m'_\nu$$
$$d\Omega \to d\Omega'$$

Oayda & Lewis (2026) MNRAS 546(4)

Turning to Simulations

➢ CatSIM — our CatWISE simulation pipeline

Draw sources on isotropic sky
sample magnitudes

↓

Do special relativity

↓

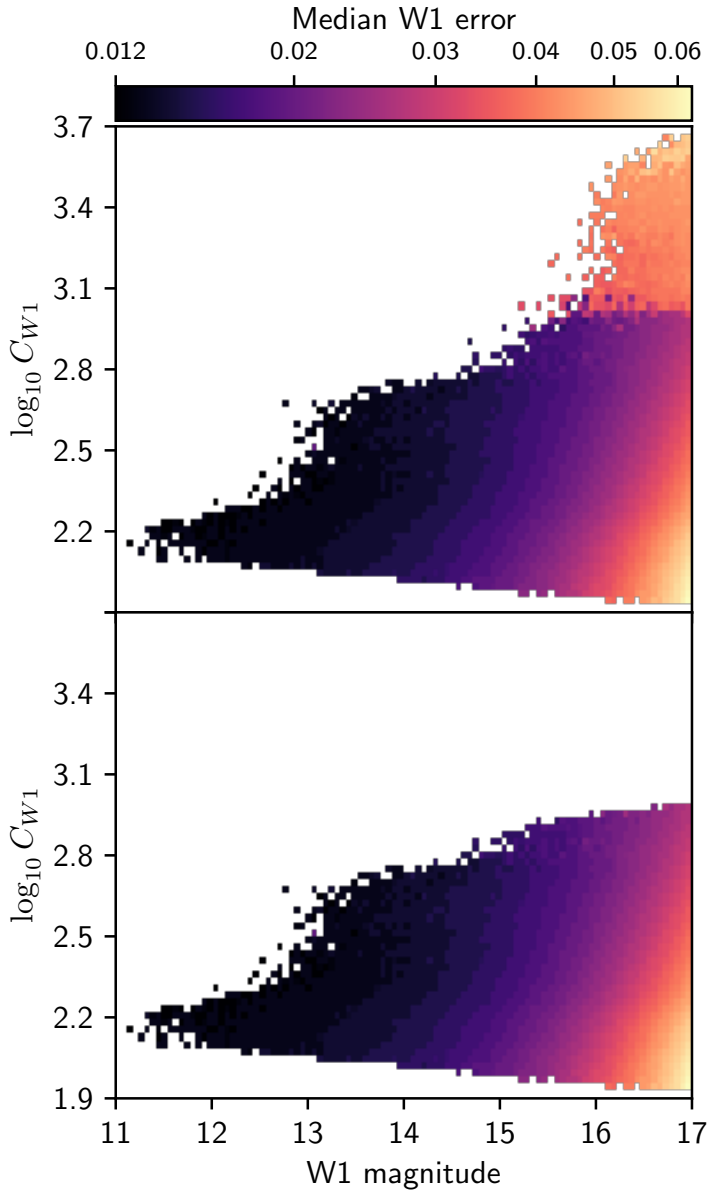
Add photometric errors
$m'_\nu \rightarrow m'_\nu + \Delta m'_\nu$

↓

Make magnitude cuts & mask

↓

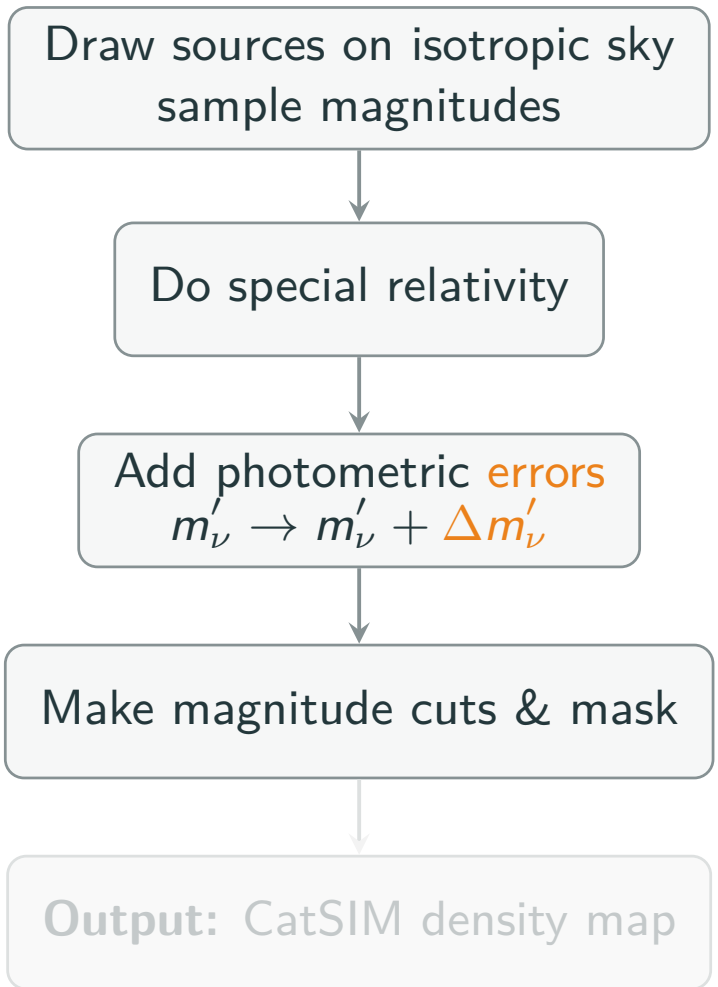
Output: CatSIM density map



Oayda & Lewis (2026) MNRAS 546(4)

Turning to Simulations

➢ CatSIM — our CatWISE simulation pipeline

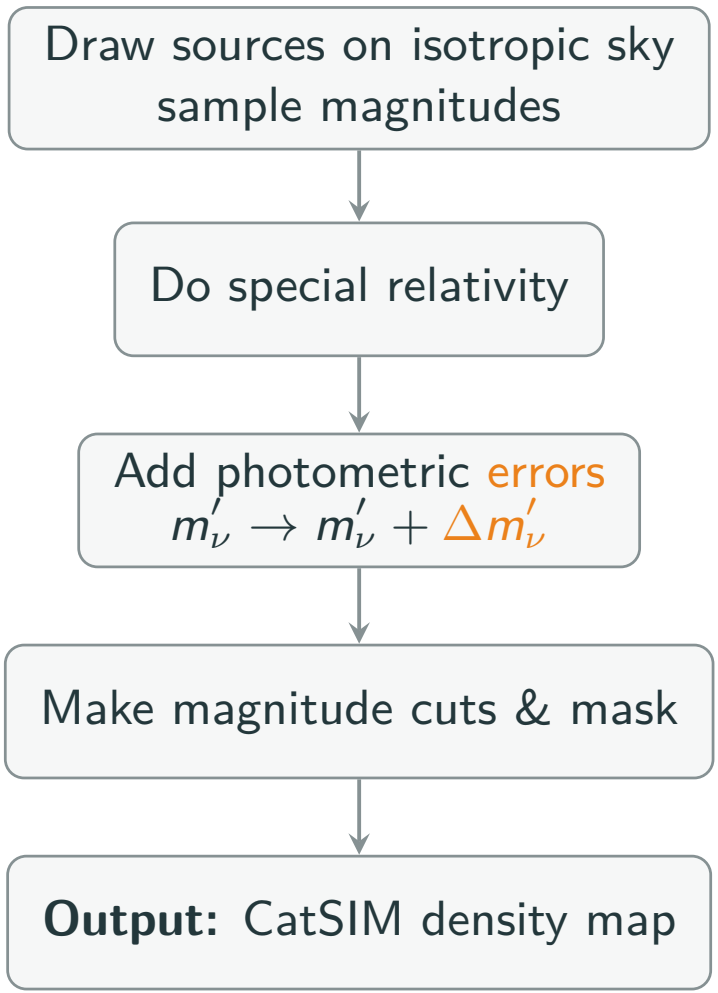


$$9 < m'_{\mathrm{W1}} < 16.4$$
$$m'_{\mathrm{W1}} - m'_{\mathrm{W2}} > 0.8$$

Oayda & Lewis (2026) MNRAS 546(4)

Turning to Simulations

➢ CatSIM — our CatWISE simulation pipeline



Oayda & Lewis (2026) MNRAS 546(4)

Simulation-based Inference

What Now...?

What Now...?

The classic approach:

$$P(\boldsymbol{\theta}|\mathbf{x}, M) = \frac{\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}, M)\,\pi(\boldsymbol{\theta}|M)}{\mathcal{Z}(\mathbf{x}|M)}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}.$$

What Now...?

The classic approach:

$$P(\boldsymbol{\theta}|\mathbf{x}, M) = \frac{\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}, M)\,\pi(\boldsymbol{\theta}|M)}{\mathcal{Z}(\mathbf{x}|M)}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}.$$

$$\mathcal{L} = \prod^{N_{\text{pix.}}} P(N_i|\lambda_i, M_{\text{dipole}})$$

where $\lambda_i = \bar{N}(1 + \mathcal{D}\cos\theta)$ and $P \to$ Pois, GenPo

What Now...?

Don't know $\mathcal{L}$? No problem!

What Now...?

Don't know $\mathcal{L}$? No problem!

Take data-generating process
$f_M : \boldsymbol{\theta} \rightarrow \mathbf{x}$.

What Now...?

Don't know $\mathcal{L}$? No problem!

Take data-generating process
$f_M : \boldsymbol{\theta} \rightarrow \mathbf{x}$.

Use neural network to learn $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}, M)$.

What Now...?

Don't know $\mathcal{L}$? No problem!

Take data-generating process
$f_M : \boldsymbol{\theta} \to \mathbf{x}$.

Use neural network to learn $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}, M)$.

This is **S**imulation-**B**ased **I**nference.

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:
  - Approximate likelihood $\tilde{\mathcal{L}}_r(\mathbf{x}|\theta)$ at round r

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:
  - Approximate likelihood $\tilde{\mathcal{L}}_r(\mathbf{x}|\theta)$ at round r
  - True likelihood $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:
  - Approximate likelihood $\tilde{\mathcal{L}}_r(\mathbf{x}|\theta)$ at round r
  - True likelihood $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$
  - $(x_i, \theta_i) \rightarrow$ (data, parameter) pair for simulation i

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:
  - Approximate likelihood $\tilde{\mathcal{L}}_r(\mathbf{x}|\theta)$ at round r
  - True likelihood $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$
  - $(x_i, \theta_i) \rightarrow$ (data, parameter) pair for simulation i
- Train a neural network to learn approximate likelihood.

$$\text{Loss function} = -\frac{1}{N}\sum_{i=1}^{N} \ln \tilde{\mathcal{L}}_r(x_i|\theta_i)$$

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:
  - Approximate likelihood $\tilde{\mathcal{L}}_r(\mathbf{x}|\theta)$ at round r
  - True likelihood $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$
  - $(x_i, \theta_i) \rightarrow$ (data, parameter) pair for simulation i
- Train a neural network to learn approximate likelihood.

$$\begin{aligned}\text{Loss function} &= -\frac{1}{N}\sum_{i=1}^{N} \ln \tilde{\mathcal{L}}_r(x_i|\theta_i) \\ &\approx -\mathbb{E}_{P(\boldsymbol{\theta},\mathbf{x})}\left[\ln \tilde{\mathcal{L}}_r\right]\end{aligned}$$

An Overview of SBI

- Part of a broad family, e.g. ABC, neural density estimation, etc.
- Focus $\rightarrow$ NLE, NPE.
- Amortised vs. **sequential** inference.
- Notation:
  - Approximate likelihood $\tilde{\mathcal{L}}_r(\mathbf{x}|\theta)$ at round r
  - True likelihood $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$
  - $(x_i, \theta_i) \rightarrow$ (data, parameter) pair for simulation i
- Train a neural network to learn approximate likelihood.

$$\begin{aligned}\text{Loss function} &= -\frac{1}{N}\sum_{i=1}^{N} \ln \tilde{\mathcal{L}}_r(x_i|\theta_i) \\ &\approx -\mathbb{E}_{P(\boldsymbol{\theta},\mathbf{x})}\left[\ln \tilde{\mathcal{L}}_r\right] \\ &= \mathbb{E}_{\pi(\boldsymbol{\theta})}\left[D_{\mathrm{KL}}(\mathcal{L}\|\tilde{\mathcal{L}}_r)\right] + \text{const.}\end{aligned}$$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

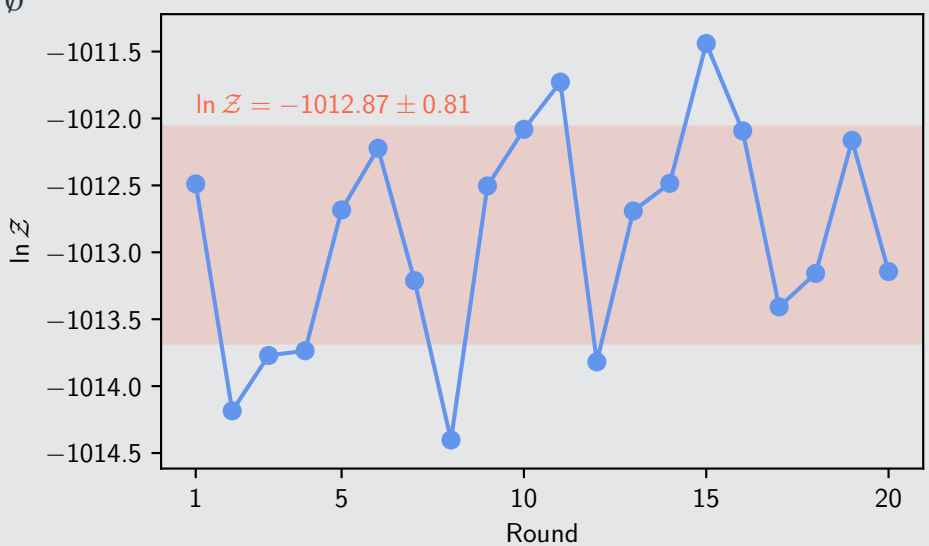
Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$



The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

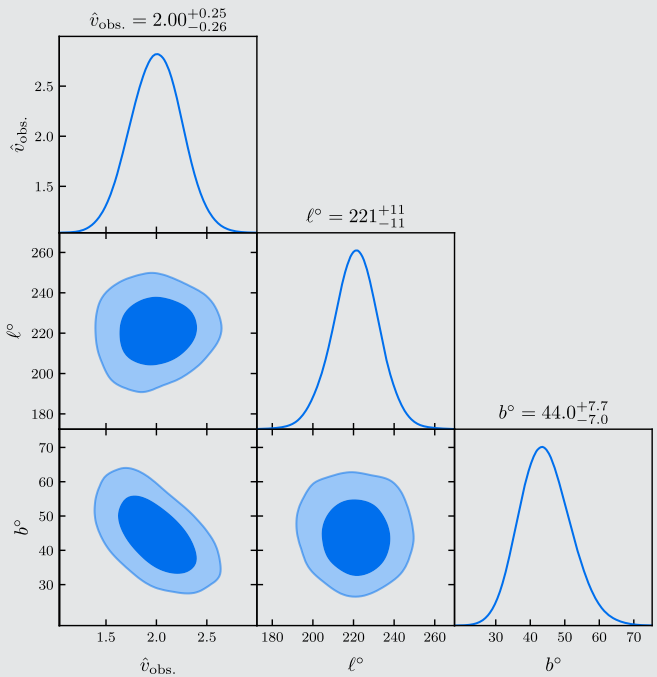
Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$



The Nuts & Bolts

Algorithm: Sequential Neural Likelihood Estimator (SNLE)

Source: Papamakarios+19, Oayda+26
Input: observed data $\mathbf{x}_0$, simulator f_M, prior $\pi(\boldsymbol{\theta})$, budget N and # rounds R
Output: approximate likelihood $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

---

set $P_0(\boldsymbol{\theta} \mid \mathbf{x}_0) = \pi(\boldsymbol{\theta})$ and $\mathbf{D}, \boldsymbol{\Theta}, \mathcal{Z}_{\text{all}} = \emptyset$

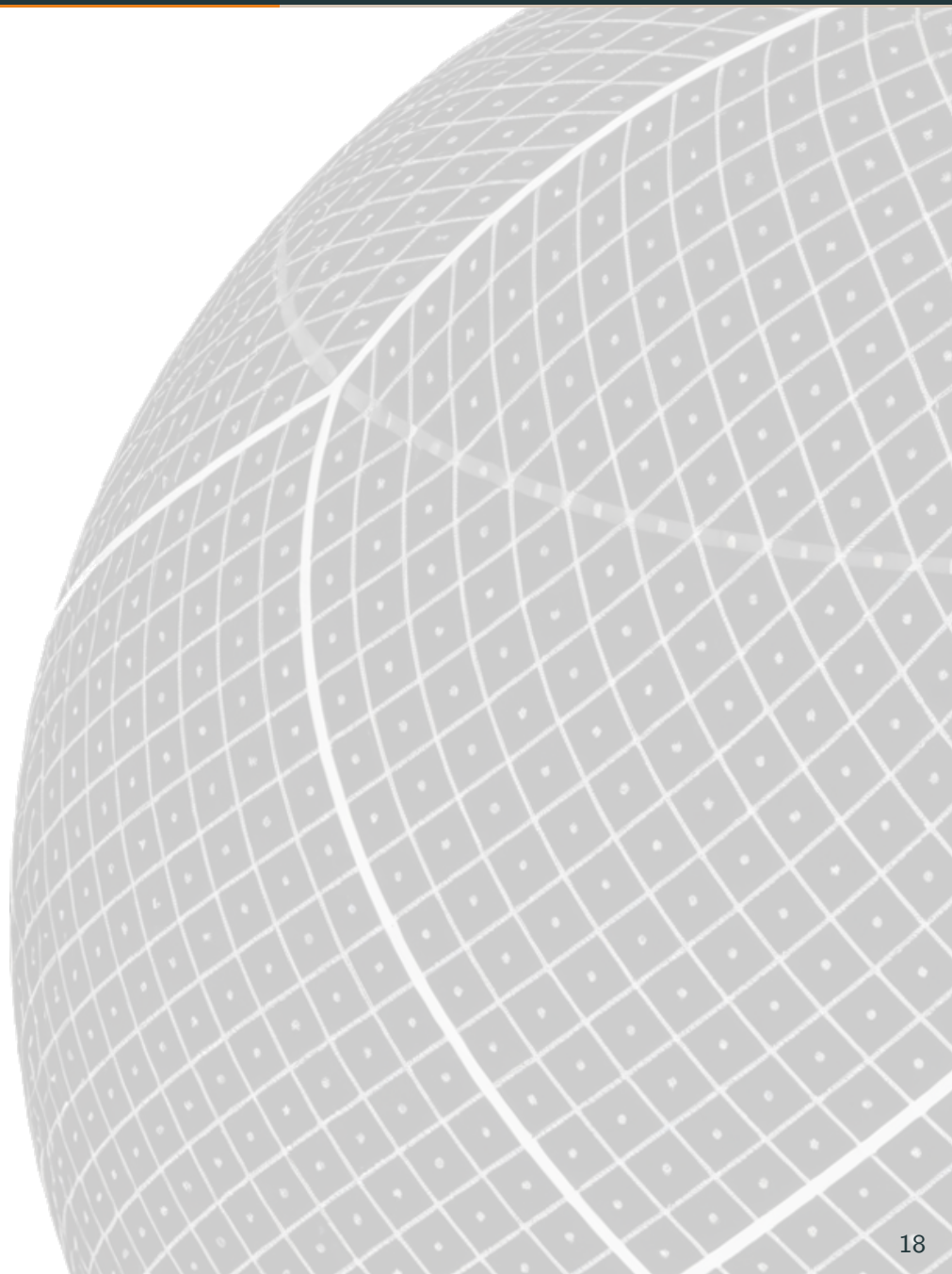
for r in rounds:
  for n in sims per round:
    sample $\boldsymbol{\theta}_n \hookleftarrow P_{r-1}(\boldsymbol{\theta} \mid \mathbf{x}_0)$
    simulate $\mathbf{x}_n = f_M(\boldsymbol{\theta}_n)$
    add $(\mathbf{x}_n, \boldsymbol{\theta}_n)$ to $(\mathbf{D}, \boldsymbol{\Theta})$
  train $\widetilde{\mathcal{L}}_r(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\mathbf{D}, \boldsymbol{\Theta})$
  obtain $\mathcal{Z}_r$ via NS, add to $\mathcal{Z}_{\text{all}}$
  set $P_r(\boldsymbol{\theta} \mid \mathbf{x}_0) \propto \widetilde{\mathcal{L}}_r(\mathbf{x}_0 \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})$

return $\widetilde{\mathcal{L}}_R(\mathbf{x} \mid \boldsymbol{\theta})$

Validation

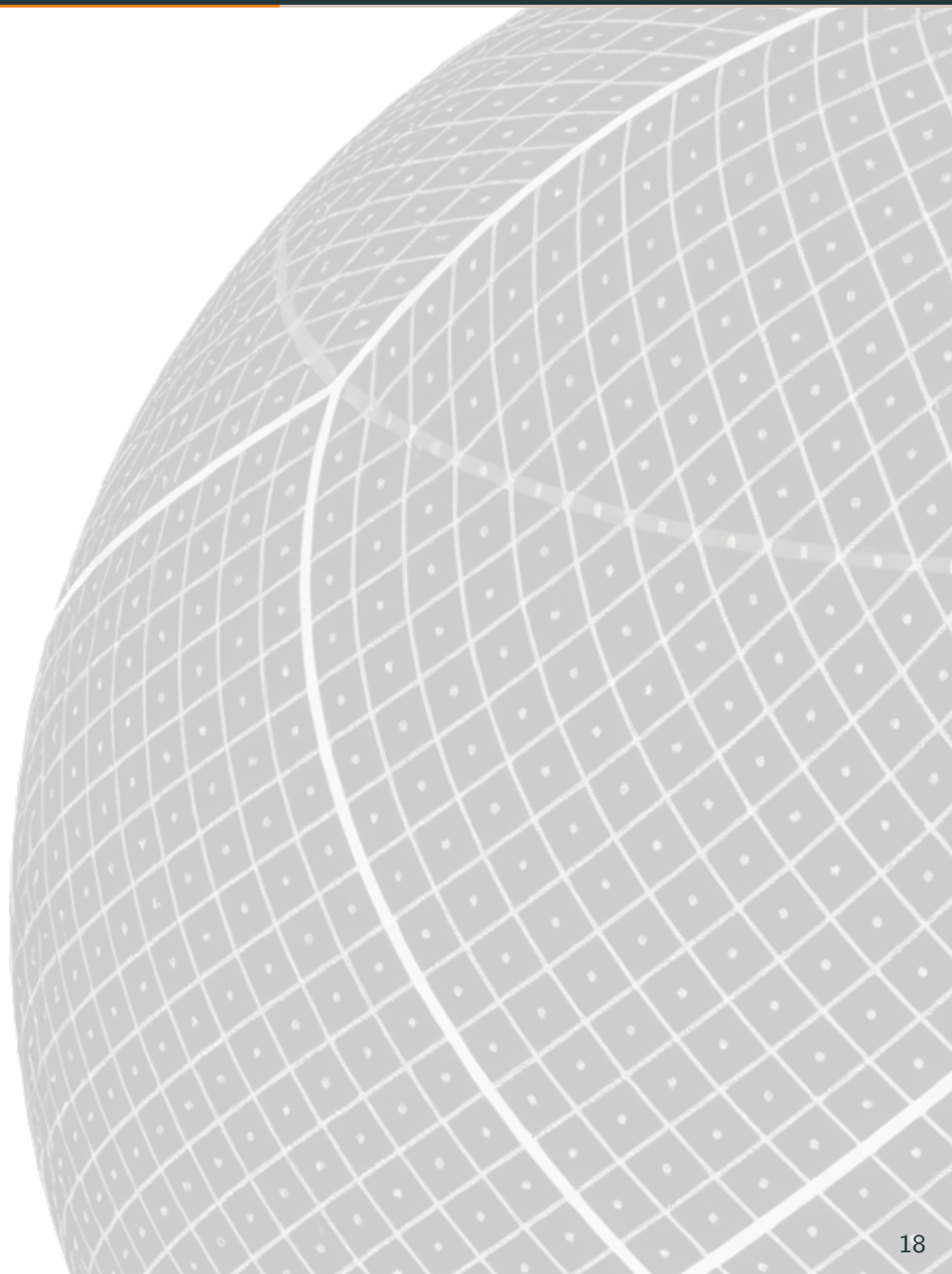
Challenge: Dimensionality

- $N_{\text{side}} = 4 \rightarrow 49\,152$ pixels.
- Model a $d = 49\,152$ base distribution? No.
- We want:
  - A way to $\downarrow$ data dim.
  - Make sure inferences don't change.
- $\mathcal{Z}(\mathbf{z}|M) \neq \mathcal{Z}(\mathbf{x}|M)$.



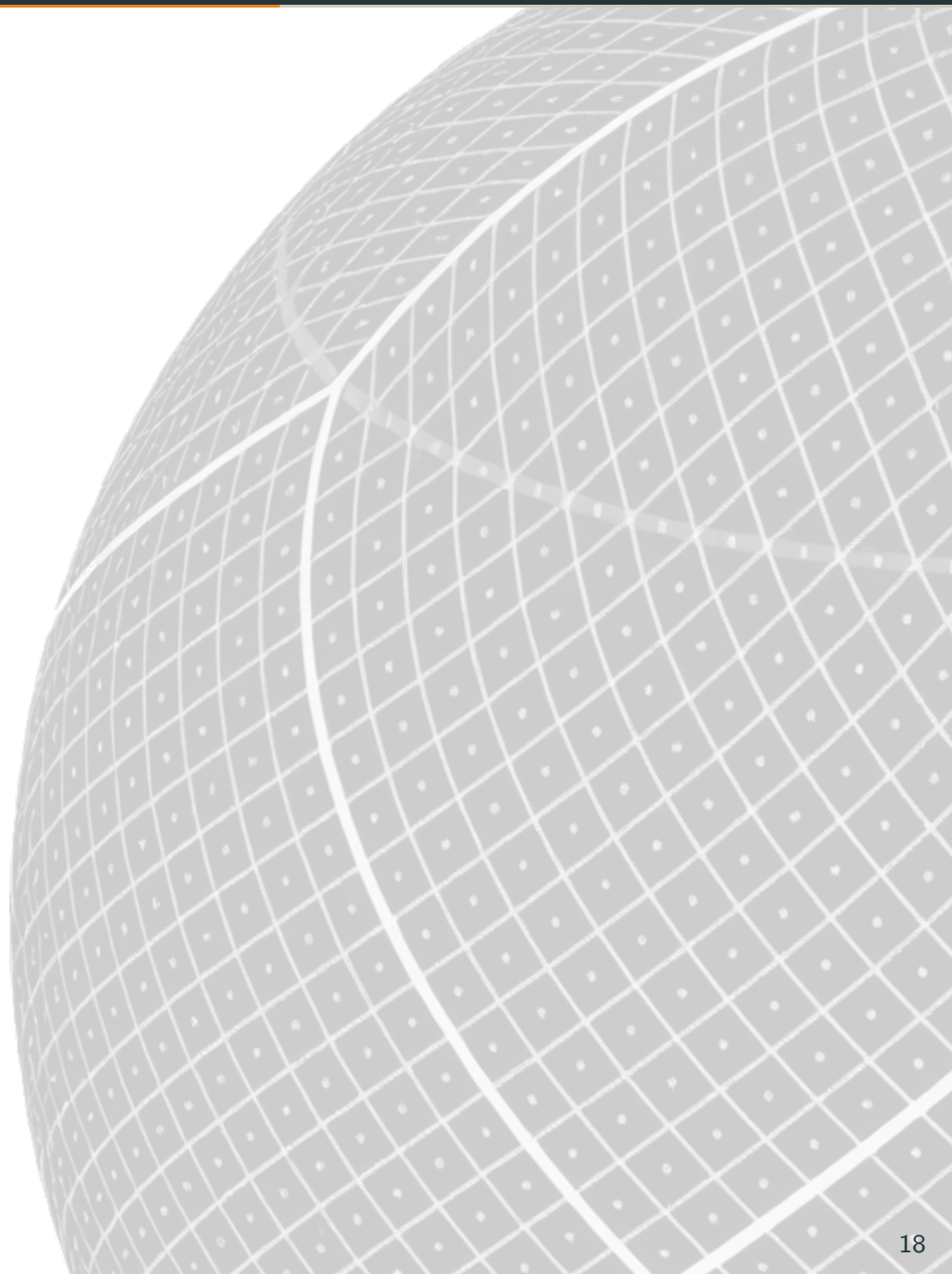
Challenge: Dimensionality

- $N_{\text{side}} = 4 \rightarrow 49\,152$ pixels.
- Model a $d = 49\,152$ base distribution? No.
- We want:
  - A way to $\downarrow$ data dim.
  - Make sure inferences don't change.
- $\mathcal{Z}(\mathbf{z}|M) \neq \mathcal{Z}(\mathbf{x}|M)$.



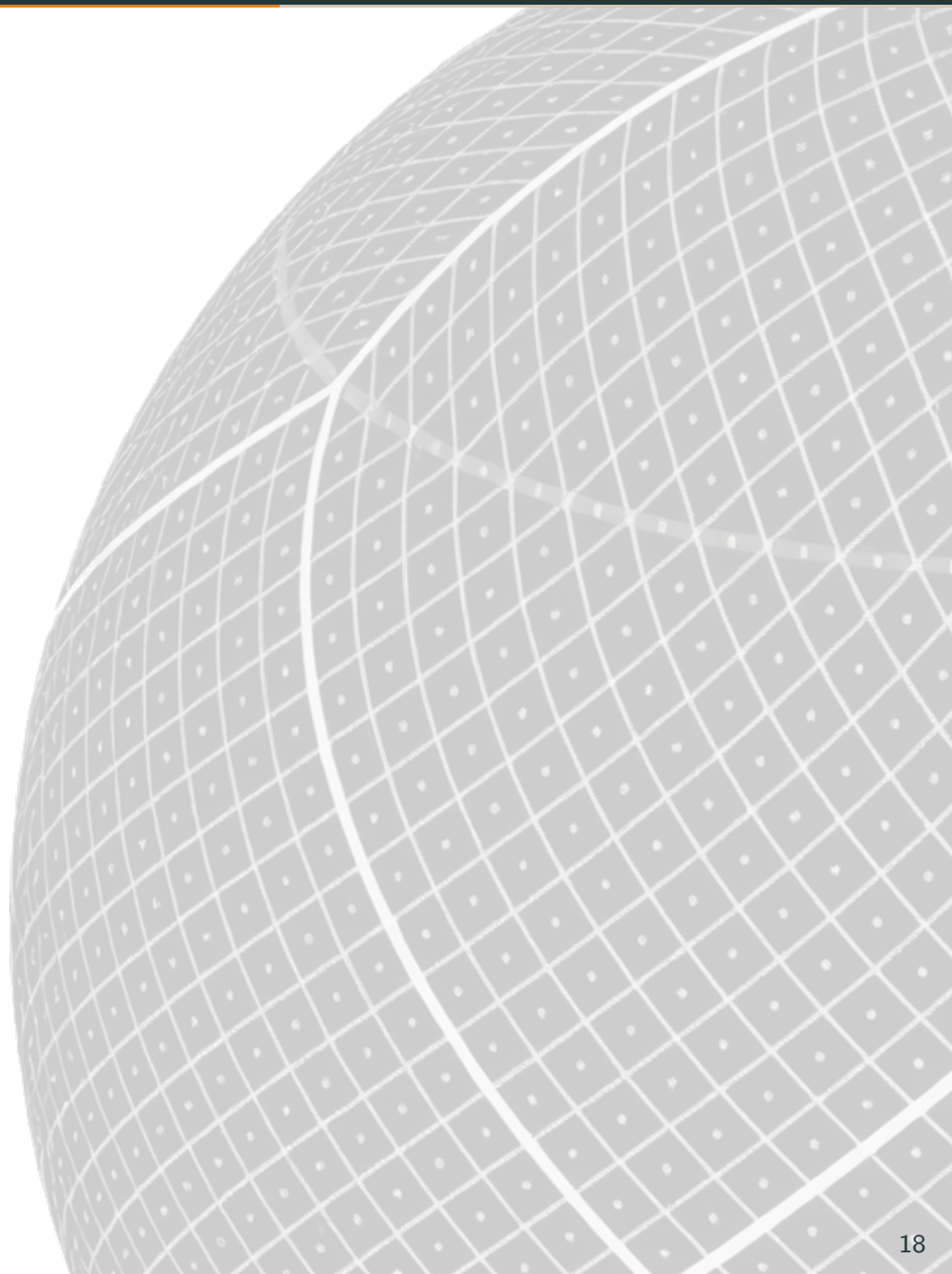
Challenge: Dimensionality

- $N_{\text{side}} = 4 \rightarrow 49\,152$ pixels.
- Model a $d = 49\,152$ base distribution? No.
- We want:
  - A way to $\downarrow$ data dim.
  - Make sure inferences don't change.
- $\mathcal{Z}(\mathbf{z}|M) \neq \mathcal{Z}(\mathbf{x}|M)$.



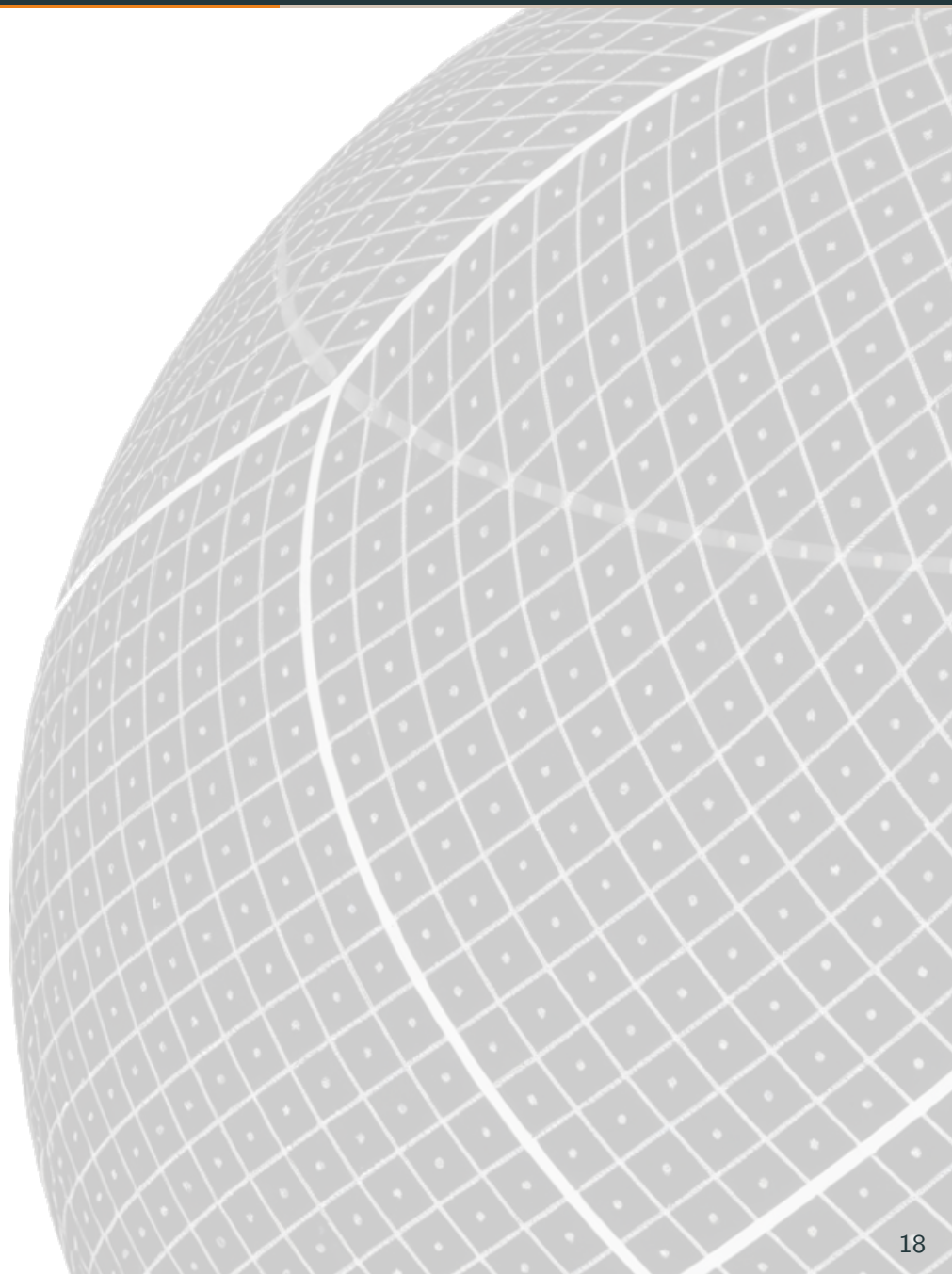
Challenge: Dimensionality

- $N_{\text{side}} = 4 \rightarrow 49\,152$ pixels.
- Model a $d = 49\,152$ base distribution? No.
- We want:
  - A way to $\downarrow$ data dim.
  - Make sure inferences don't change.
- $\mathcal{Z}(\mathbf{z}|M) \neq \mathcal{Z}(\mathbf{x}|M)$.



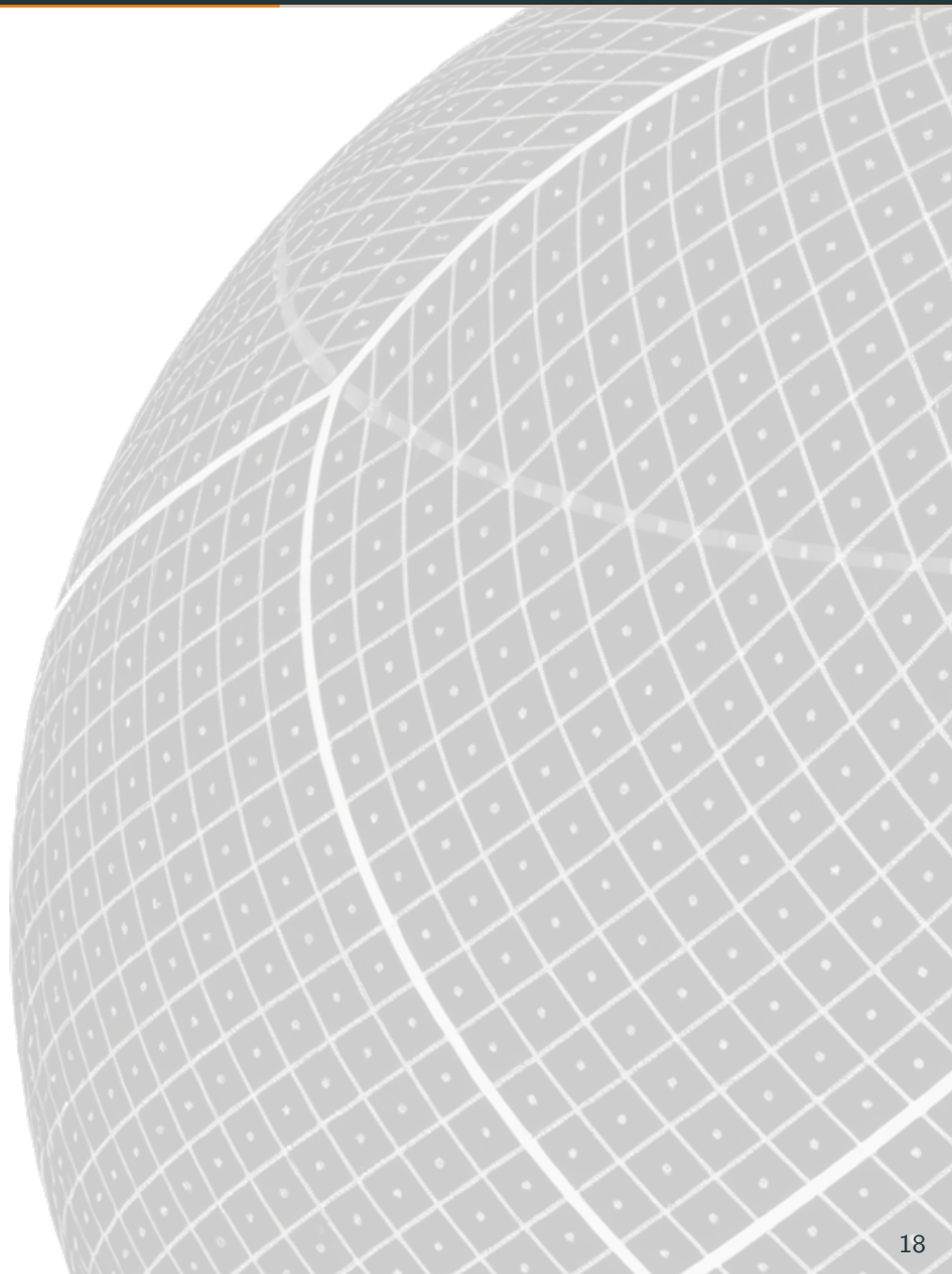
Challenge: Dimensionality

- $N_{\text{side}} = 4 \rightarrow 49\,152$ pixels.
- Model a $d = 49\,152$ base distribution? No.
- We want:
  - A way to $\downarrow$ data dim.
  - Make sure inferences don't change.
- $\mathcal{Z}(\mathbf{z}|M) \neq \mathcal{Z}(\mathbf{x}|M)$.



Challenge: Dimensionality

- $N_{\text{side}} = 4 \rightarrow 49\,152$ pixels.
- Model a $d = 49\,152$ base distribution? No.
- We want:
  - A way to $\downarrow$ data dim.
  - Make sure inferences don't change.
- $\mathcal{Z}(\mathbf{z}|M) \neq \mathcal{Z}(\mathbf{x}|M)$.



Downscaling

Validation

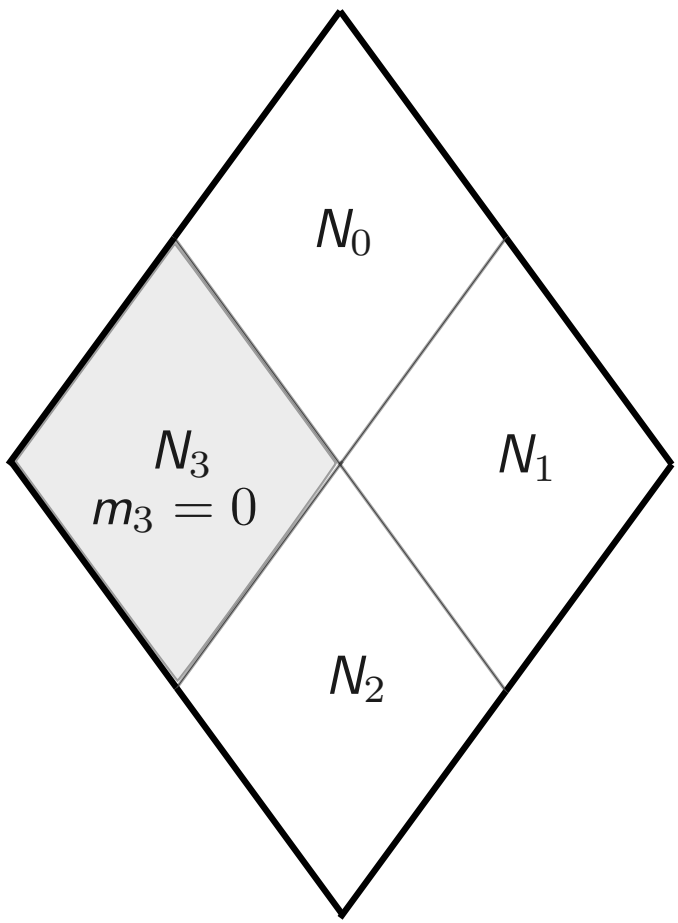
- SBI likelihood $\approx$ analytic likelihood
- $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$

Downscaling

Validation

- ➢ SBI likelihood $\approx$ analytic likelihood
- ➢ $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$

- Suppose $N_i \sim \text{Pois}(\lambda_i)$.

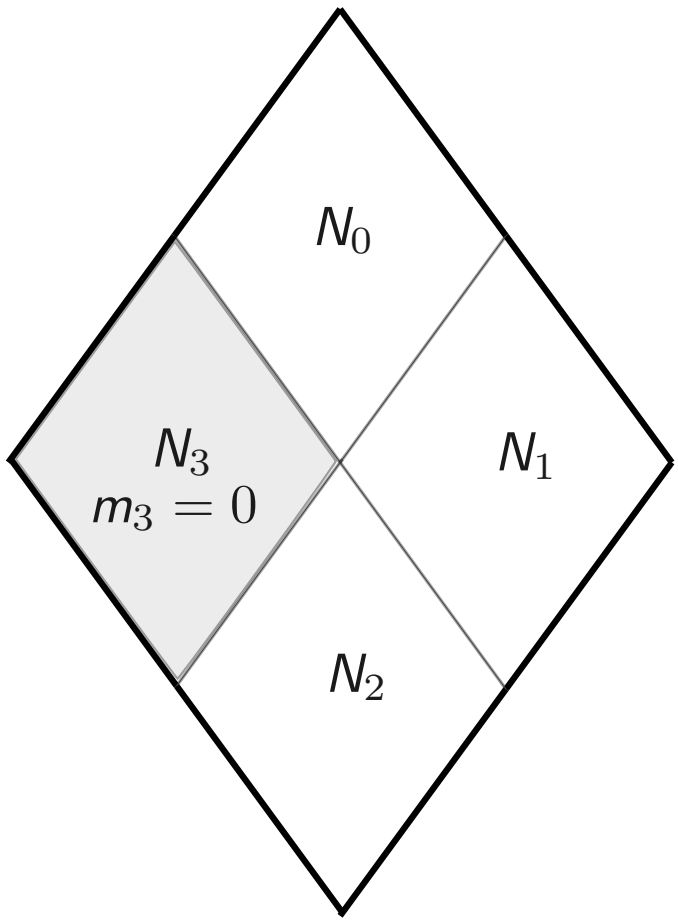


Downscaling

Validation

- ➢ SBI likelihood $\approx$ analytic likelihood
- ➢ $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$

- Suppose $N_i \sim \text{Pois}(\lambda_i)$.
- Then $N^{\text{super}} = \sum_{i=0}^{3} N_i$.

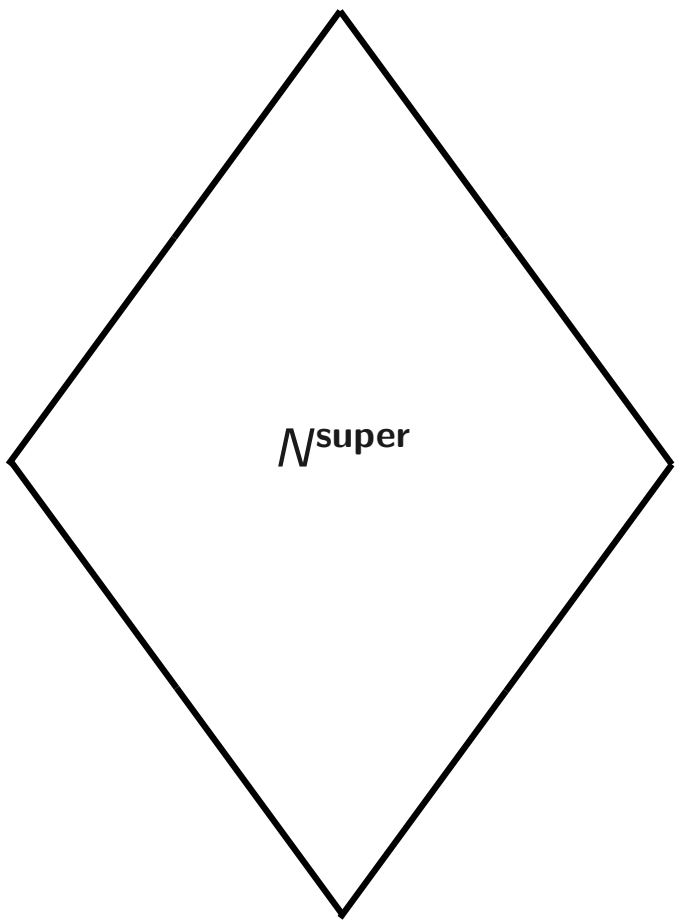


Downscaling

Validation

- ➢ SBI likelihood $\approx$ analytic likelihood
- ➢ $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$

- Suppose $N_i \sim \text{Pois}(\lambda_i)$.
- Then $N^{\text{super}} = \sum_{i=0}^{3} N_i$.
- $N_{\text{super}} \sim \text{Pois}(\lambda = \sum_{i=0}^{3} \lambda_i)$.

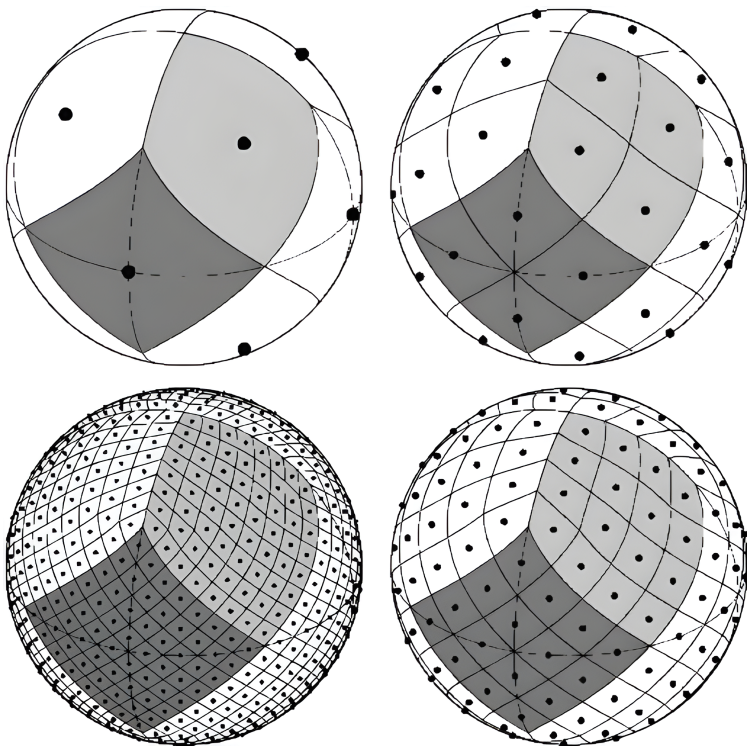


Downscaling

Validation

- SBI likelihood $\approx$ analytic likelihood
- $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$

- Suppose $N_i \sim \text{Pois}(\lambda_i)$.
- Then $N^{\text{super}} = \sum_{i=0}^{3} N_i$.
- $N_{\text{super}} \sim \text{Pois}(\lambda = \sum_{i=0}^{3} \lambda_i)$.
- Chain a series of downscales to get to a low resolution output.

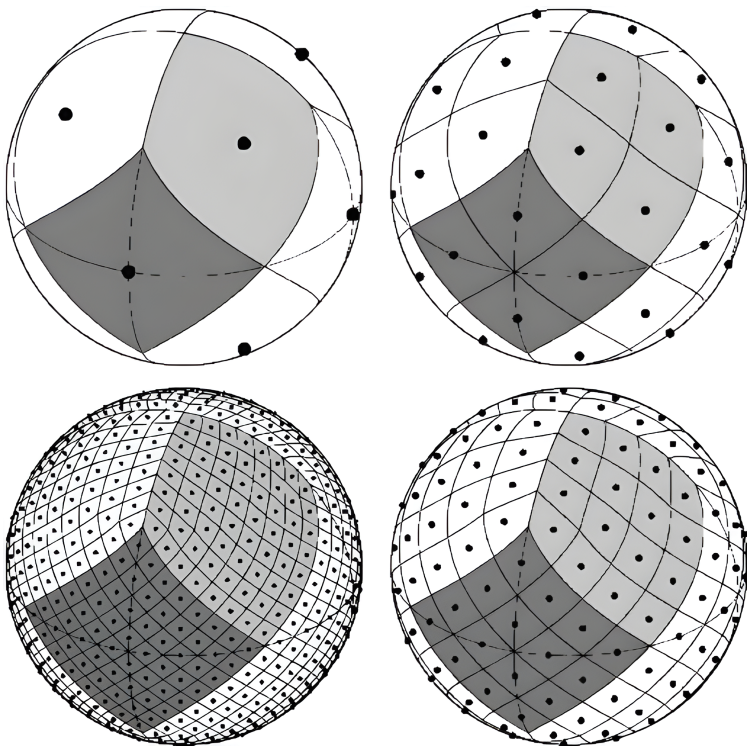


Downscaling

Validation

- SBI likelihood $\approx$ analytic likelihood
- $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$

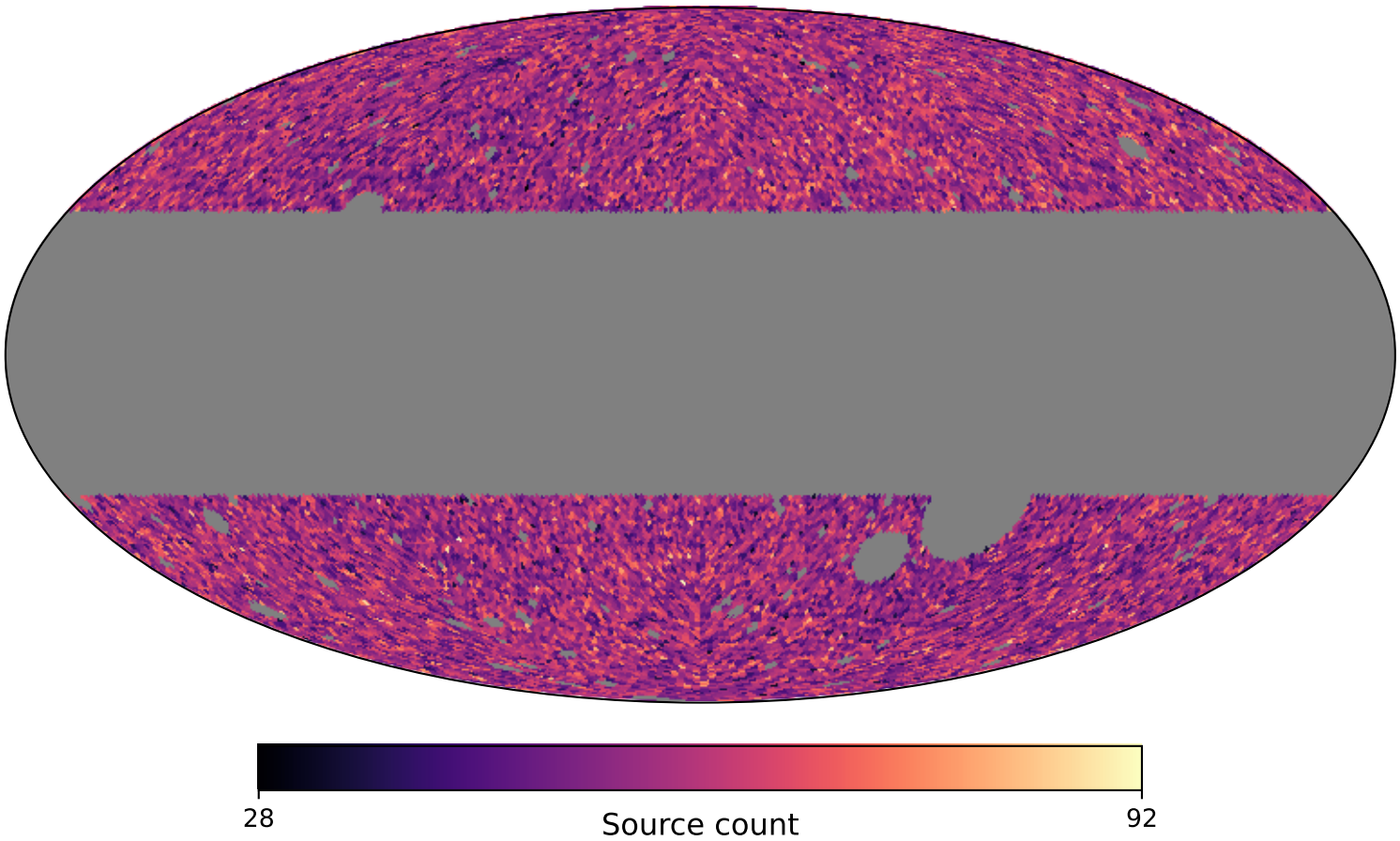
- Suppose $N_i \sim \text{Pois}(\lambda_i)$.
- Then $N^{\text{super}} = \sum_{i=0}^{3} N_i$.
- $N_{\text{super}} \sim \text{Pois}(\lambda = \sum_{i=0}^{3} \lambda_i)$.
- Chain a series of downscales to get to a low resolution output.
- Likelihood is Poissonian.



Downscaling Example

Validation

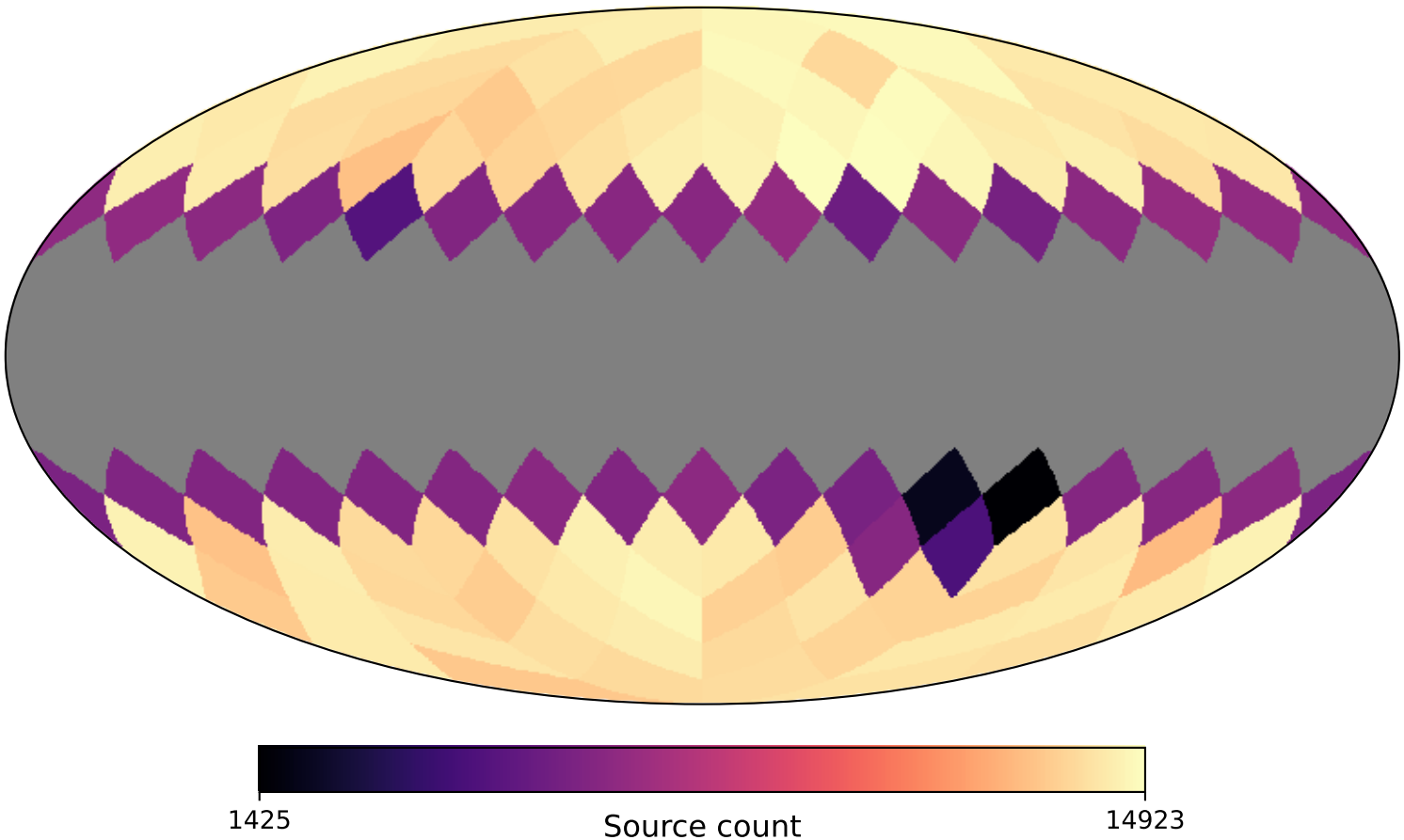
- SBI likelihood $\approx$ analytic likelihood
- $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$



Downscaling Example

Validation

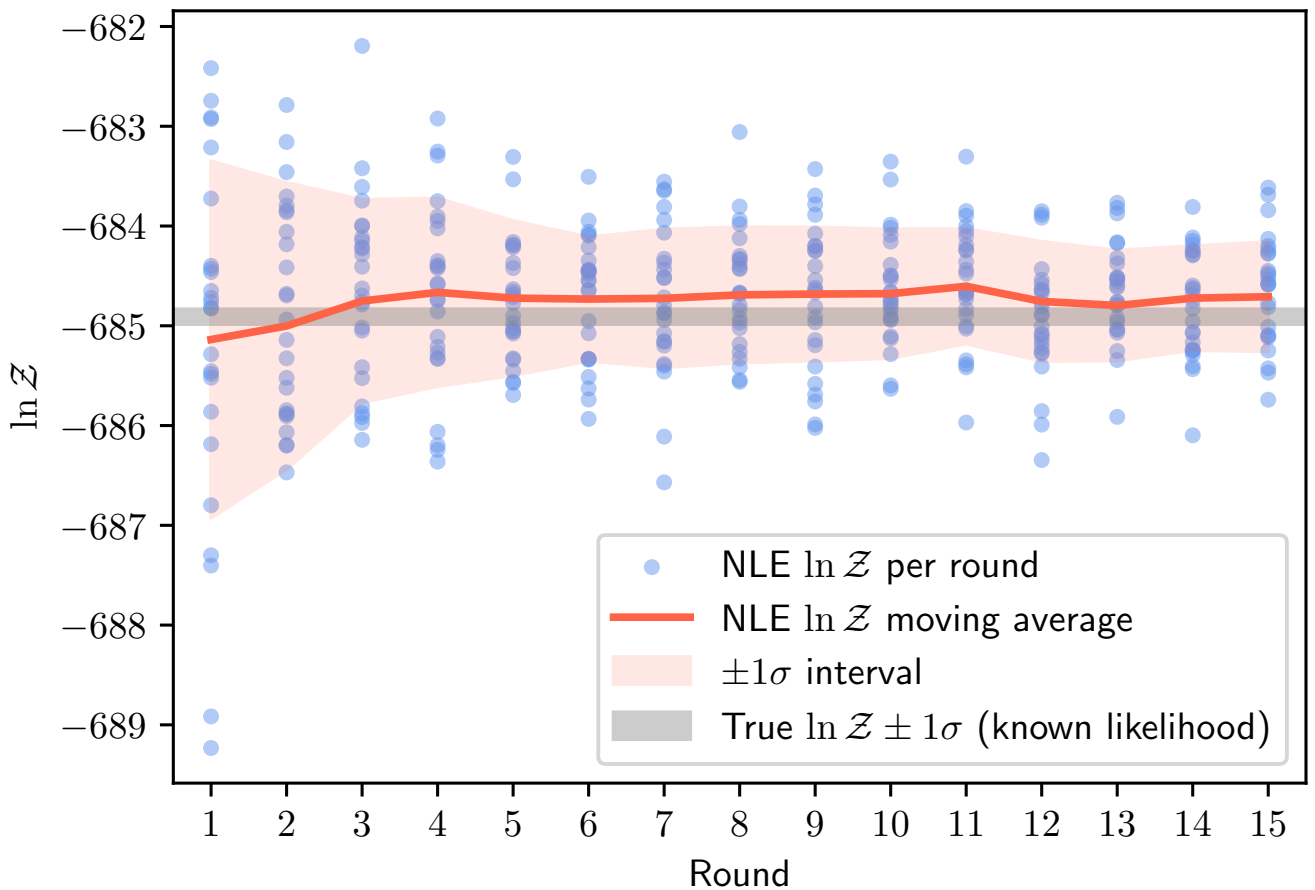
- SBI likelihood $\approx$ analytic likelihood
- $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$



Consistency Check — Evidence

Validation

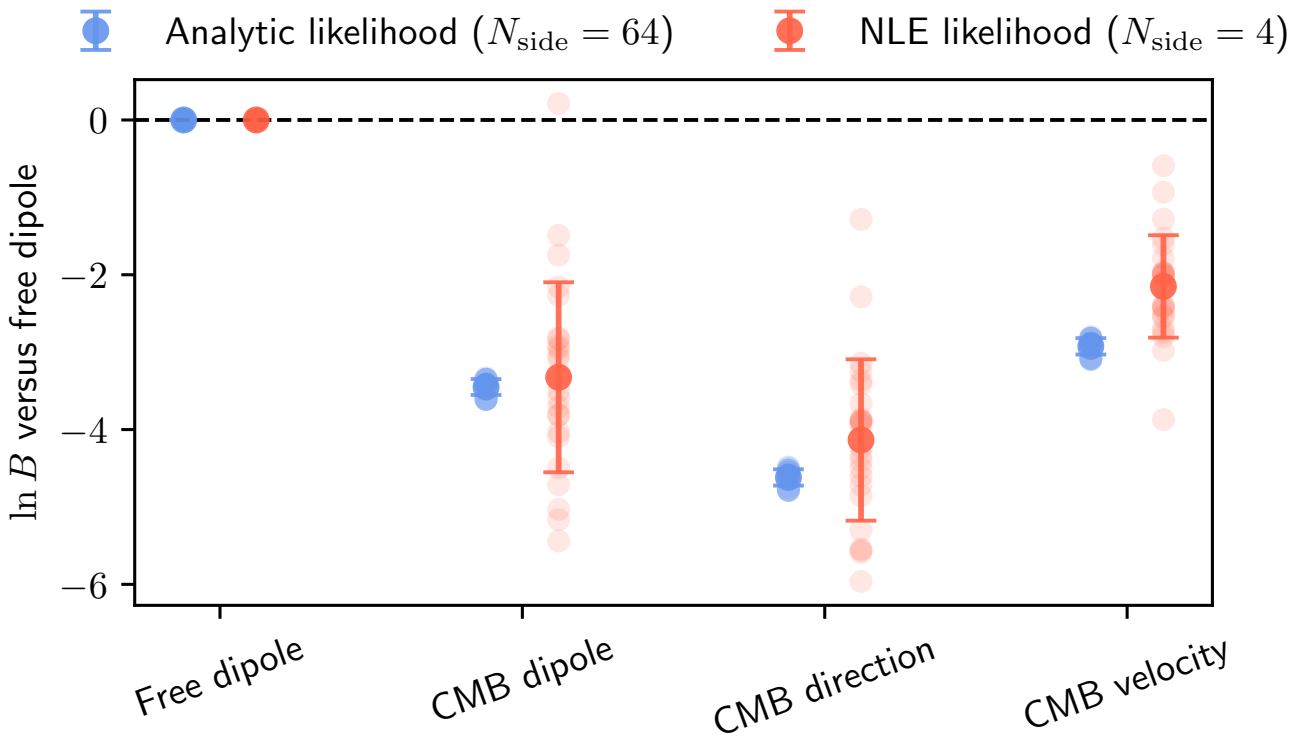
- SBI likelihood ≈ analytic likelihood
- $N_{\text{side}} = 4 \longleftrightarrow N_{\text{side}} = 64$



Consistency Check — Evidence

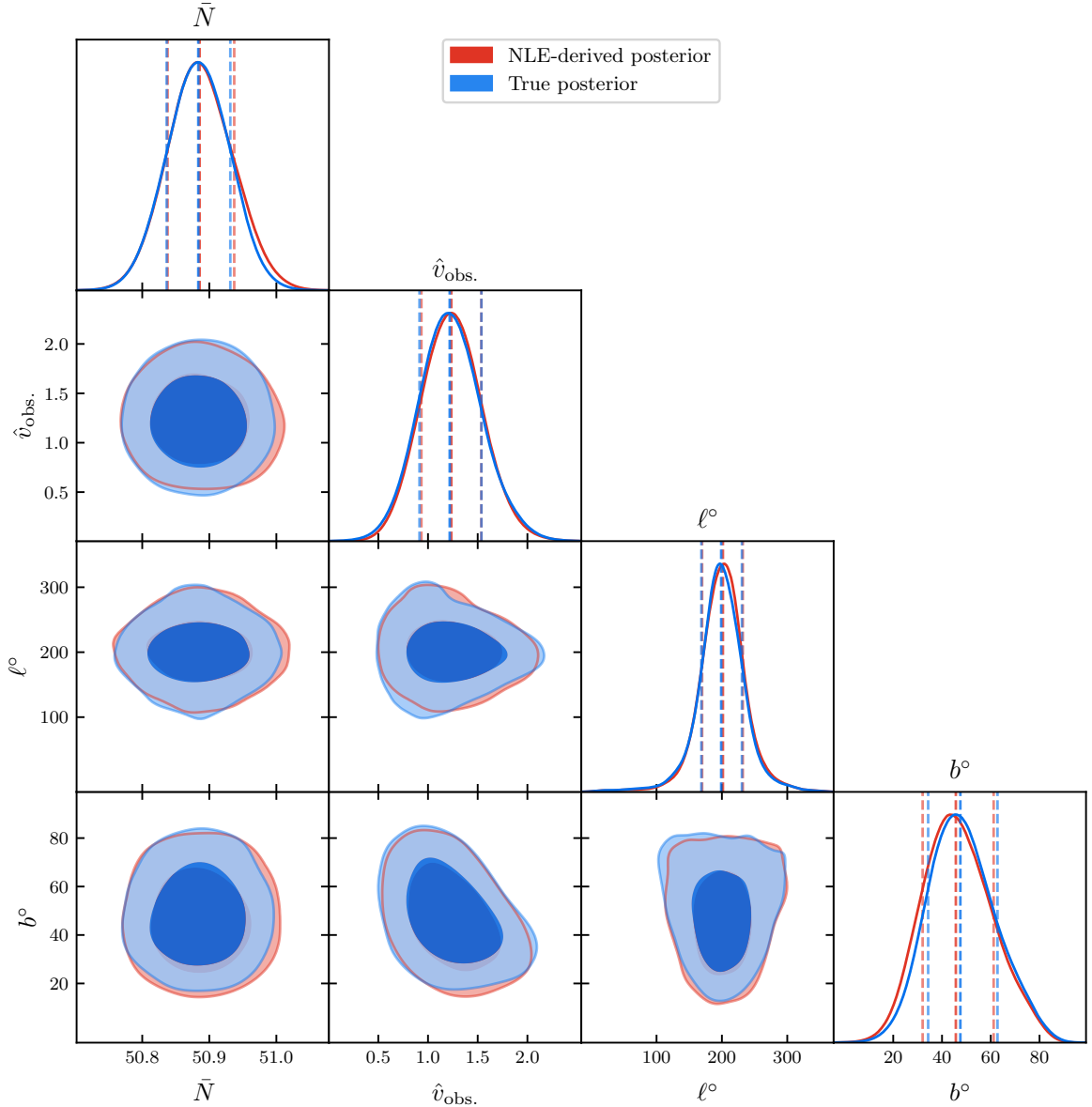
Validation

- SBI likelihood $\approx$ analytic likelihood
- $N_{\text{side}} = 4 \iff N_{\text{side}} = 64$



Consistency Check — Posteriors

SBI-derived posterior versus ground truth.



Oayda & Lewis (2026) MNRAS 546(4)

Results

Bayes Factors

- $\hat{v}_{\text{obs.}}$ Observer speed

Bayes Factors

- $\hat{v}_{\text{obs.}}$ Observer speed
- (l°, b°) Direction of motion

Bayes Factors

• $\hat{v}_{\text{obs.}}$	Observer speed
• (l°, b°)	Direction of motion
• $\log_{10} N_{\text{init.}}$	Initial count of sources

Bayes Factors

• $\hat{v}_{\text{obs.}}$	Observer speed
• (l°, b°)	Direction of motion
• $\log_{10} N_{\text{init.}}$	Initial count of sources
• $\sigma^2_{\text{WX, final}} = \sigma^2_{\text{WX}} + \eta_{\text{extra}}\sigma^2_{\text{WX}}$	'Extra error'

Bayes Factors

- $\hat{v}_{\text{obs.}}$ — Observer speed
- (l°, b°) — Direction of motion
- $\log_{10} N_{\text{init.}}$ — Initial count of sources
- $\sigma^2_{\text{WX, final}} = \sigma^2_{\text{WX}} + \eta_{\text{extra}}\sigma^2_{\text{WX}}$ — 'Extra error'
  - $\Delta m_{\text{WX}} \hookleftarrow \mathcal{G}(0, \sigma^2)$ — Gaussian

Oayda & Lewis (2026) MNRAS 546(4)

Bayes Factors

- $\hat{v}_{\text{obs.}}$ — Observer speed
- (l°, b°) — Direction of motion
- $\log_{10} N_{\text{init.}}$ — Initial count of sources
- $\sigma^2_{\text{WX, final}} = \sigma^2_{\text{WX}} + \eta_{\text{extra}}\sigma^2_{\text{WX}}$ — 'Extra error'
  - $\Delta m_{\text{WX}} \hookleftarrow \mathcal{G}(0, \sigma^2)$ — Gaussian
  - $\Delta m_{\text{WX}} \hookleftarrow t(0, \sigma^2, \xi)$ — Student's t

Bayes Factors

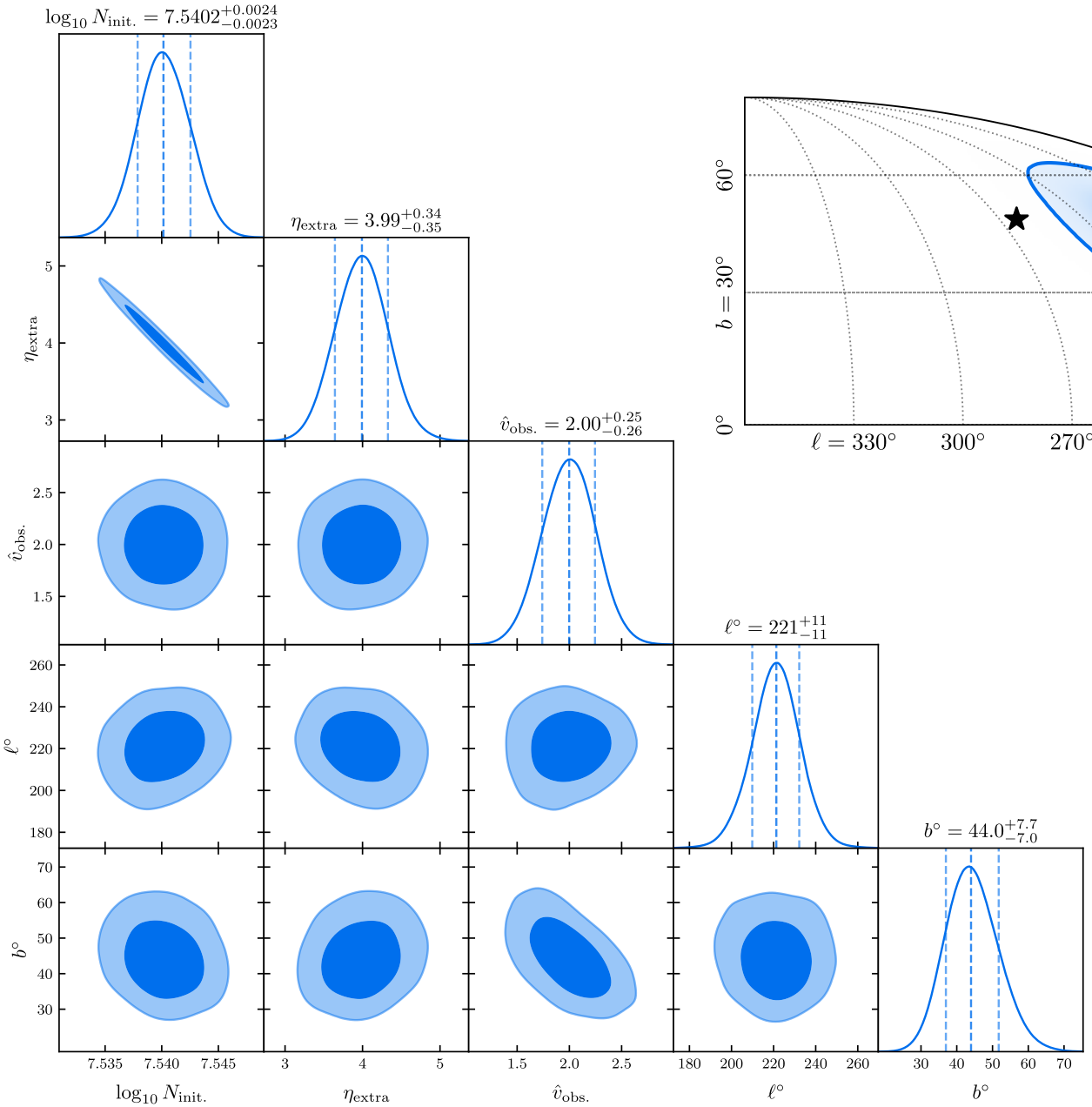
- $\hat{v}_{\text{obs.}}$ — Observer speed
- (l°, b°) — Direction of motion
- $\log_{10} N_{\text{init.}}$ — Initial count of sources
- $\sigma^2_{\text{WX, final}} = \sigma^2_{\text{WX}} + \eta_{\text{extra}} \sigma^2_{\text{WX}}$ — 'Extra error'
  - $\Delta m_{\text{WX}} \hookleftarrow \mathcal{G}(0, \sigma^2)$ — Gaussian
  - $\Delta m_{\text{WX}} \hookleftarrow t(0, \sigma^2, \xi)$ — Student's t

Model	**ln B**
Dipole from Secrest+21	2.8 ± 0.9
Dipole from Dam+23	1.9 ± 0.8
Free dipole, extra error, Gaussian	0 ± 0
CMB direction, free velocity	-2.4 ± 0.9
CMB velocity & direction	-8.3 ± 0.9
Free dipole, no extra error	-113.6 ± 6.4

Table 1: Bayes factors w.r.t. fiducial model.

Oayda & Lewis (2026) MNRAS 546(4)

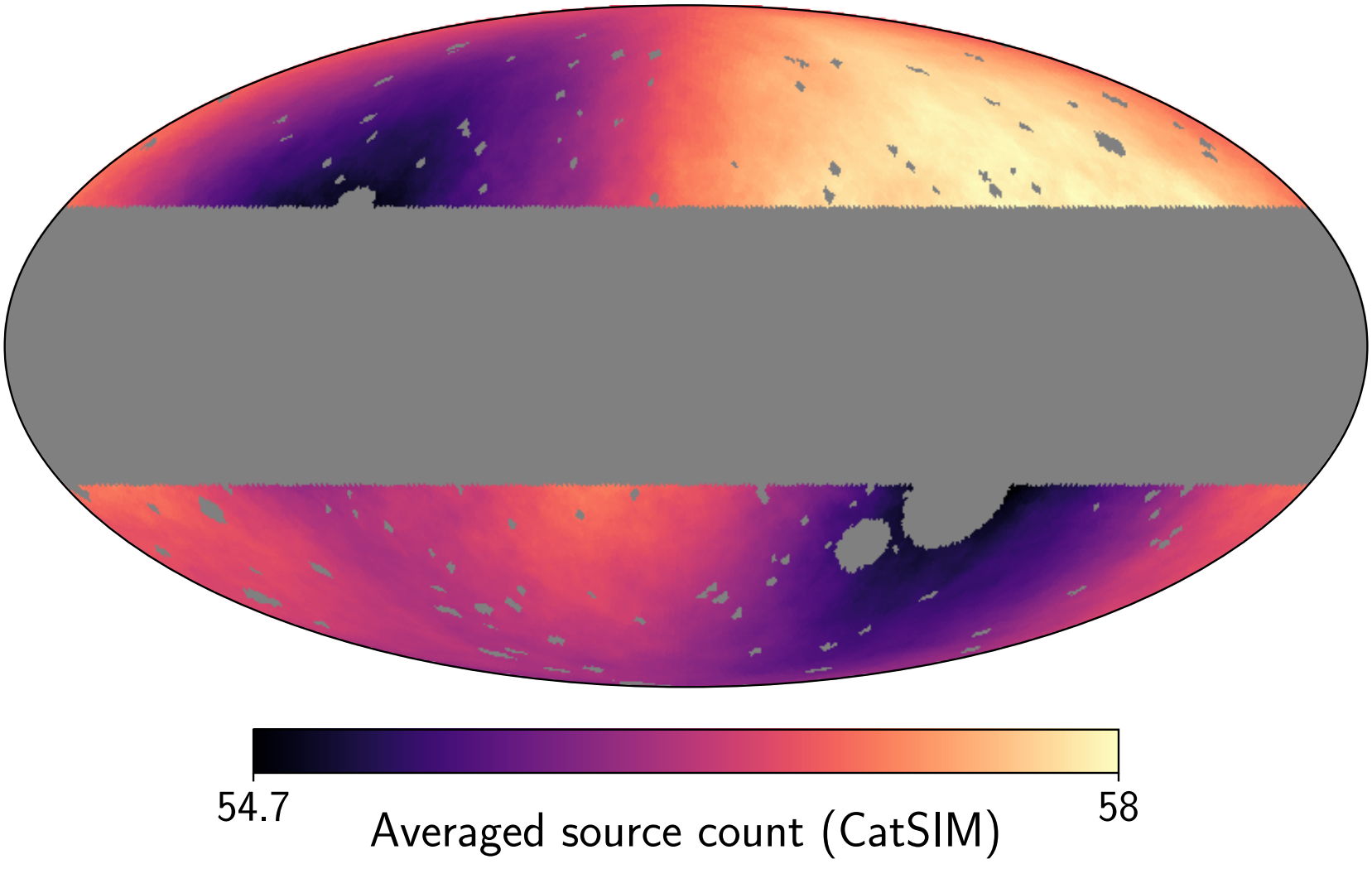
Posterior Distribution



Fiducial model

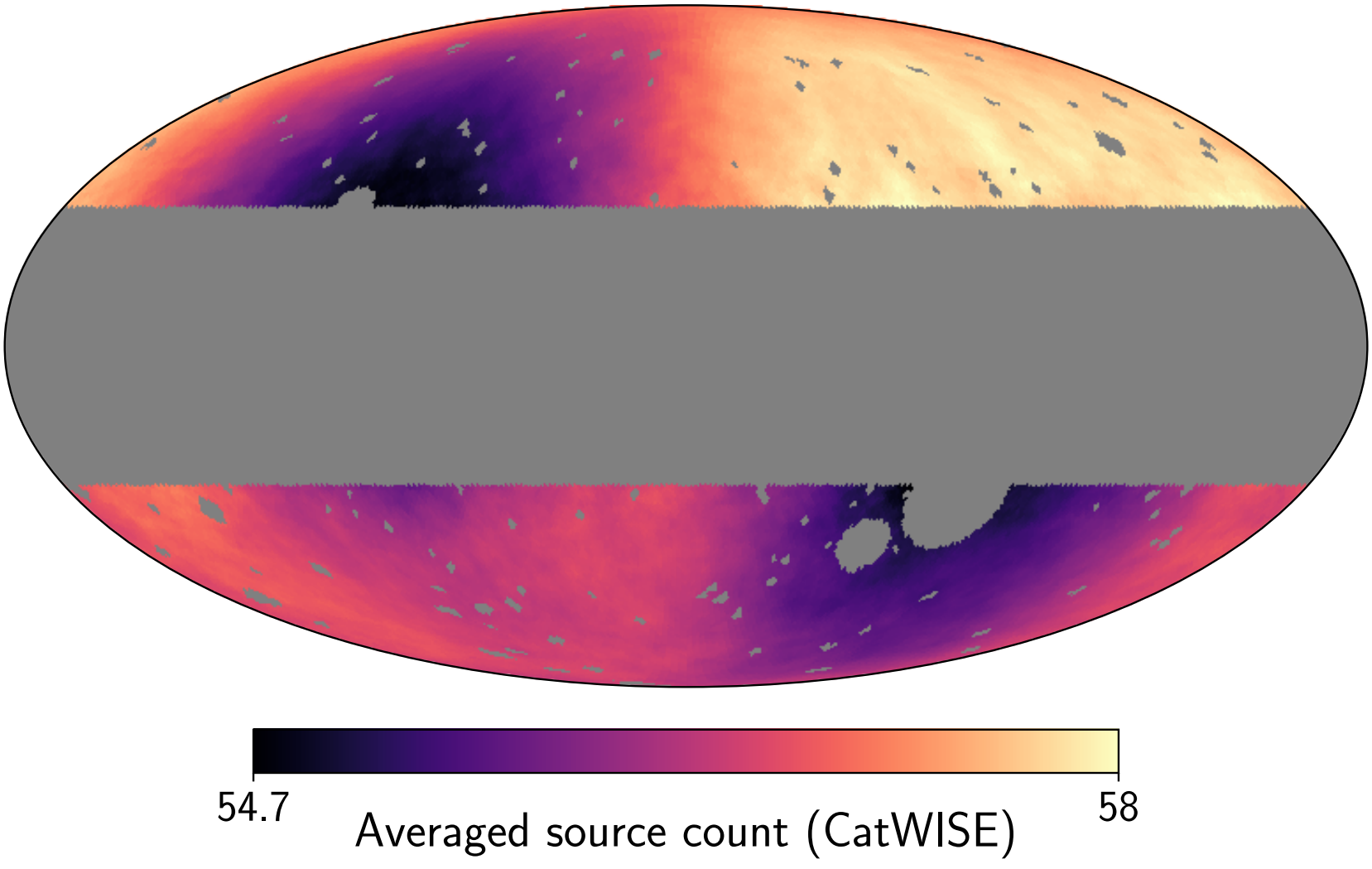
Oayda & Lewis (2026) MNRAS 546(4)

Posterior Predictive



Oayda & Lewis (2026) MNRAS 546(4)

Posterior Predictive



Oayda & Lewis (2026) MNRAS 546(4)

Where To Go Next?

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **546,** 1–16 (2026) https://doi.org/10.1093/mnras/stag248
Advance Access publication 2026 February 9

Wising up to CatWISE: using simulation-based inference to interpret the ecliptic bias and confirm the cosmic dipole excess

Oliver T. Oayda ★ and Geraint F. Lewis
Sydney Institute for Astronomy, School of Physics A28, The University of Sydney, SydneyNSW 2006,Australia

- SBI → generalisable to other datasets.
- Radio → flux calibration, resolution, etc.
- Optical → Galactic extinction, scanning law (with colour cuts), etc.

Where To Go Next?

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **546,** 1–16 (2026) https://doi.org/10.1093/mnras/stag248
Advance Access publication 2026 February 9

Wising up to CatWISE: using simulation-based inference to interpret the ecliptic bias and confirm the cosmic dipole excess

Oliver T. Oayda ★ and Geraint F. Lewis

Sydney Institute for Astronomy, School of Physics A28, The University of Sydney, SydneyNSW 2006,Australia

- SBI → generalisable to other datasets.
- Radio → flux calibration, resolution, etc.
- Optical → Galactic extinction, scanning law (with colour cuts), etc.

Where To Go Next?

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **546,** 1–16 (2026) https://doi.org/10.1093/mnras/stag248
Advance Access publication 2026 February 9

Wising up to CatWISE: using simulation-based inference to interpret the ecliptic bias and confirm the cosmic dipole excess

Oliver T. Oayda ★ and Geraint F. Lewis
Sydney Institute for Astronomy, School of Physics A28, The University of Sydney, SydneyNSW 2006,Australia

- SBI → generalisable to other datasets.
- Radio → flux calibration, resolution, etc.
- Optical → Galactic extinction, scanning law (with colour cuts), etc.

Where To Go Next?

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **546,** 1–16 (2026) https://doi.org/10.1093/mnras/stag248
Advance Access publication 2026 February 9

Wising up to CatWISE: using simulation-based inference to interpret the ecliptic bias and confirm the cosmic dipole excess

Oliver T. Oayda ★ and Geraint F. Lewis
Sydney Institute for Astronomy, School of Physics A28, The University of Sydney, SydneyNSW 2006,Australia

- SBI → generalisable to other datasets.
- Radio → flux calibration, resolution, etc.
- Optical → Galactic extinction, scanning law (with colour cuts), etc.

The dipole tension is sticking around!

- **Describe and account for ecliptic bias** ✓

- **Describe and account for ecliptic bias** ✓
- **Measure cosmic dipole** ✓

- **Describe and account for ecliptic bias** ✓
- **Measure cosmic dipole** ✓
- **Leverage power of simulations and SBI** ✓

- Describe and account for ecliptic bias ✓
- Measure cosmic dipole ✓
- Leverage power of simulations and SBI ✓
- What is the future for SBI & the dipole ?

Appendix

Normalising Flows

- **Goal:** Describe a d-dimensional base distribution $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$.
- We use a series of Masked Autoregressive Flows (MAFs).

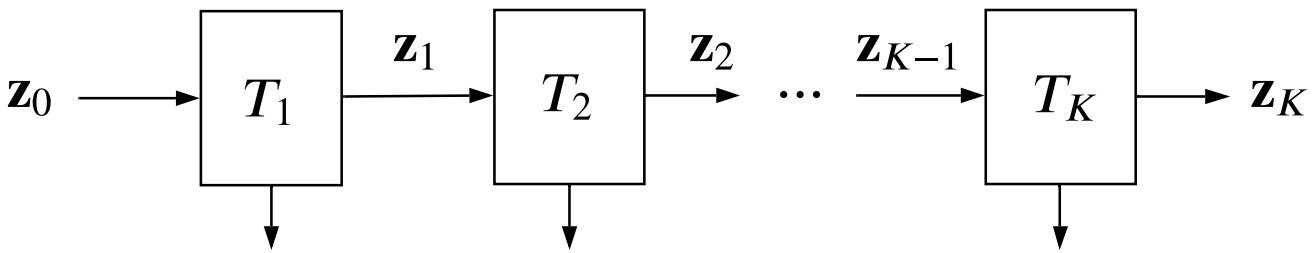
Normalising Flows

- **Goal:** Describe a d-dimensional base distribution $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$.
- We use a series of Masked Autoregressive Flows (MAFs).
- **Flow:** Transform a simple base distribution to a more complex one.

Normalising Flows

- **Goal:** Describe a d-dimensional base distribution $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$.
- We use a series of Masked Autoregressive Flows (MAFs).
- **Flow:** Transform a simple base distribution to a more complex one.

Composing individual bijections (from Papamakarios+19).

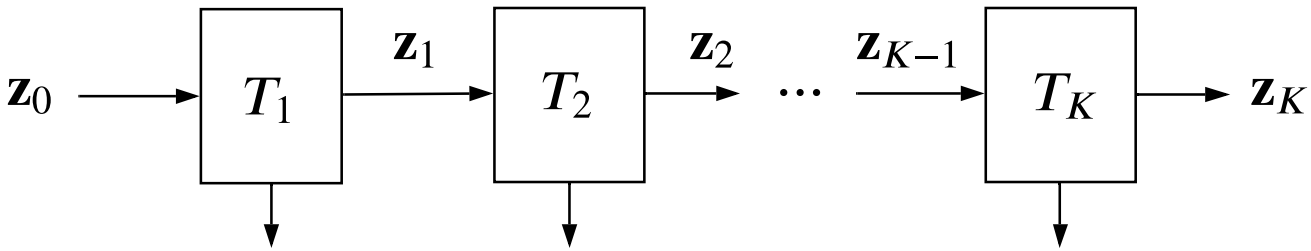


$$\log|\det J_{T_1}(\mathbf{z}_0)| + \log|\det J_{T_2}(\mathbf{z}_1)| + \cdots + \log|\det J_{T_K}(\mathbf{z}_{K-1})| = \log|\det J_T(\mathbf{z}_0)|$$

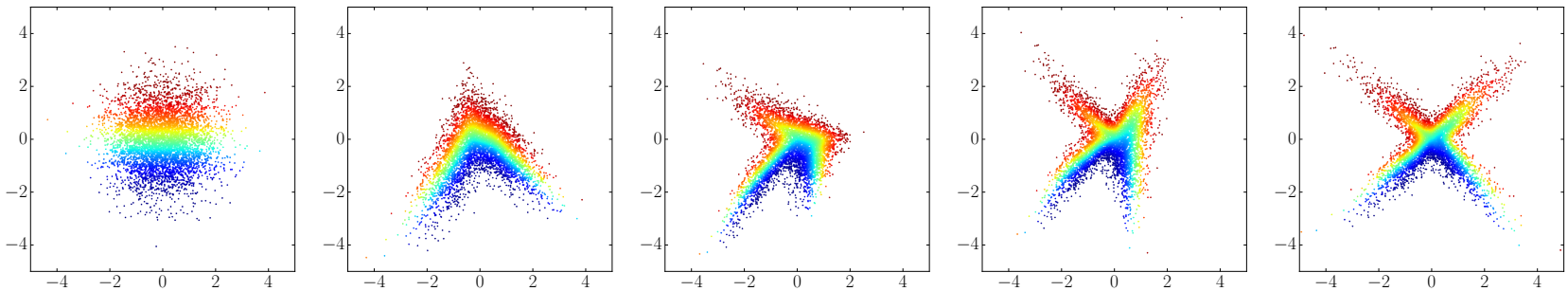
Normalising Flows

- **Goal:** Describe a d-dimensional base distribution $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})$.
- We use a series of Masked Autoregressive Flows (MAFs).
- **Flow:** Transform a simple base distribution to a more complex one.

Composing individual bijections (from Papamakarios+19).



$$\log|\det J_{T_1}(\mathbf{z}_0)| + \log|\det J_{T_2}(\mathbf{z}_1)| + \cdots + \log|\det J_{T_K}(\mathbf{z}_{K-1})| = \log|\det J_T(\mathbf{z}_0)|$$



Validation: SNLE vs. SNLE