I Can See Clearly Now the Dust is Gone: Mapping Galactic Dark Matter in 3D with Modern Machine Learning and the Gaia Space Telescope

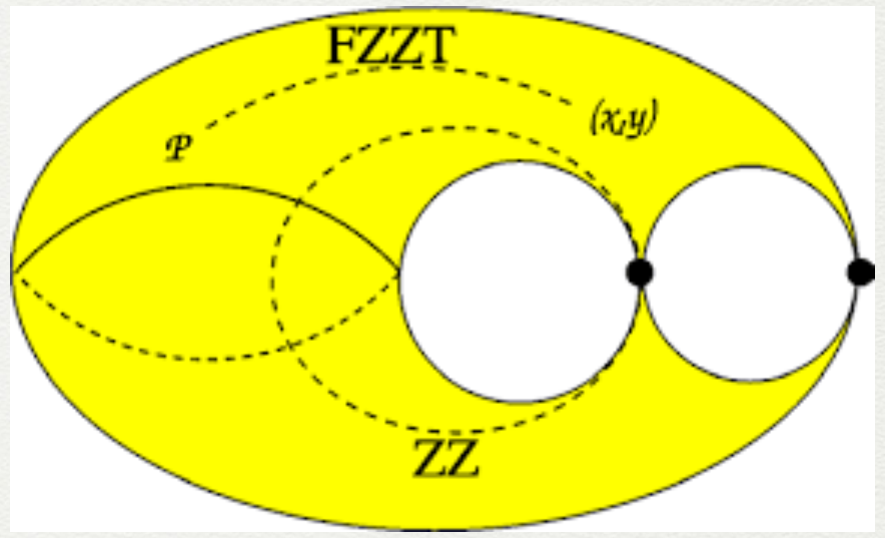
LSPC Grenoble Seminar

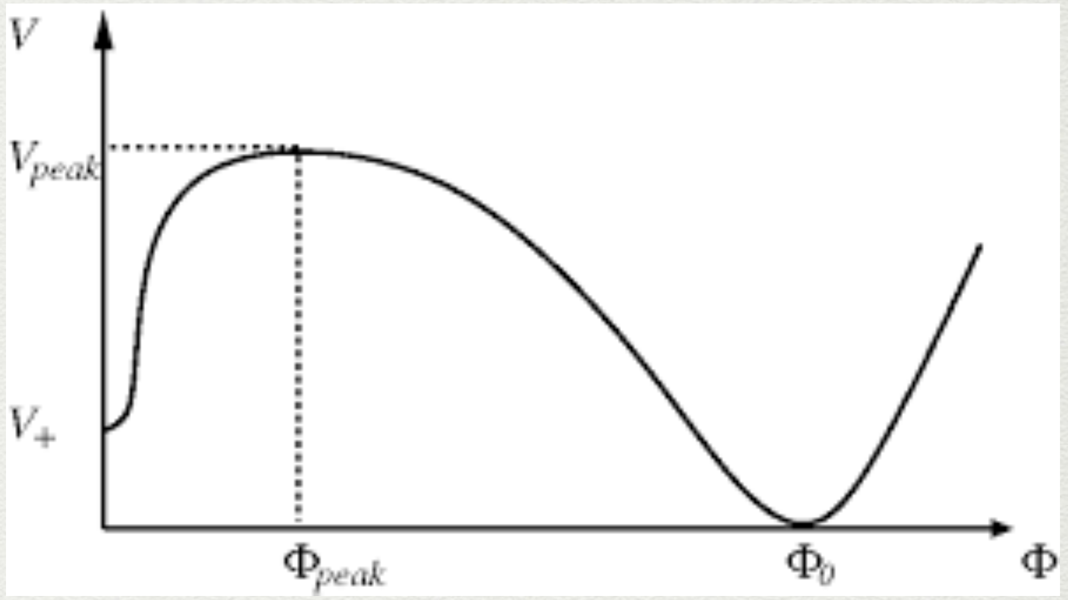
David Shih
September 22, 2025

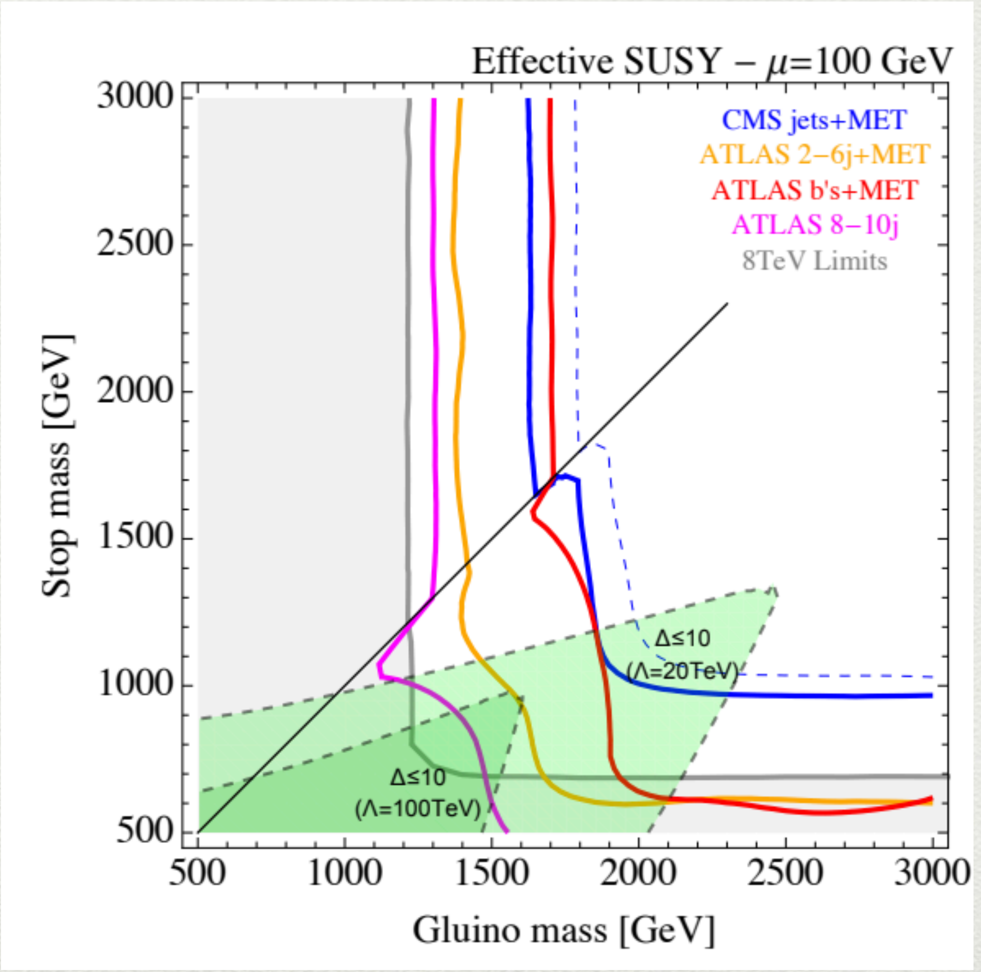
RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

A bit about myself

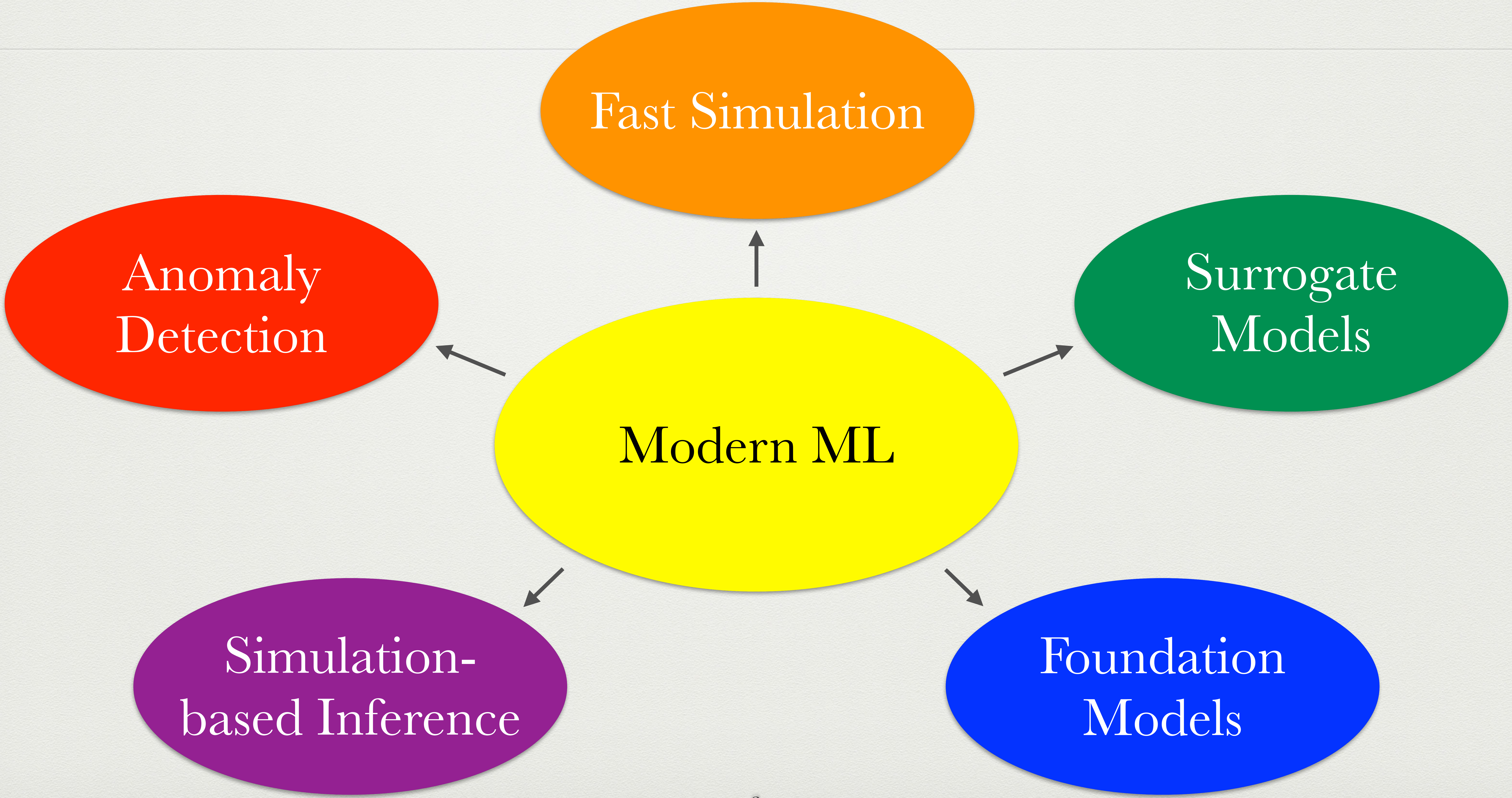
- Faculty at Rutgers since 2010
- Previous work:
    - formal theory (matrix models and 2d quantum gravity, black holes, CFTs)
    - phenomenology (SUSY breaking, gauge mediation, collider pheno, naturalness, recasting)
- Since 2018: increasingly focused on applying machine learning to fundamental physics





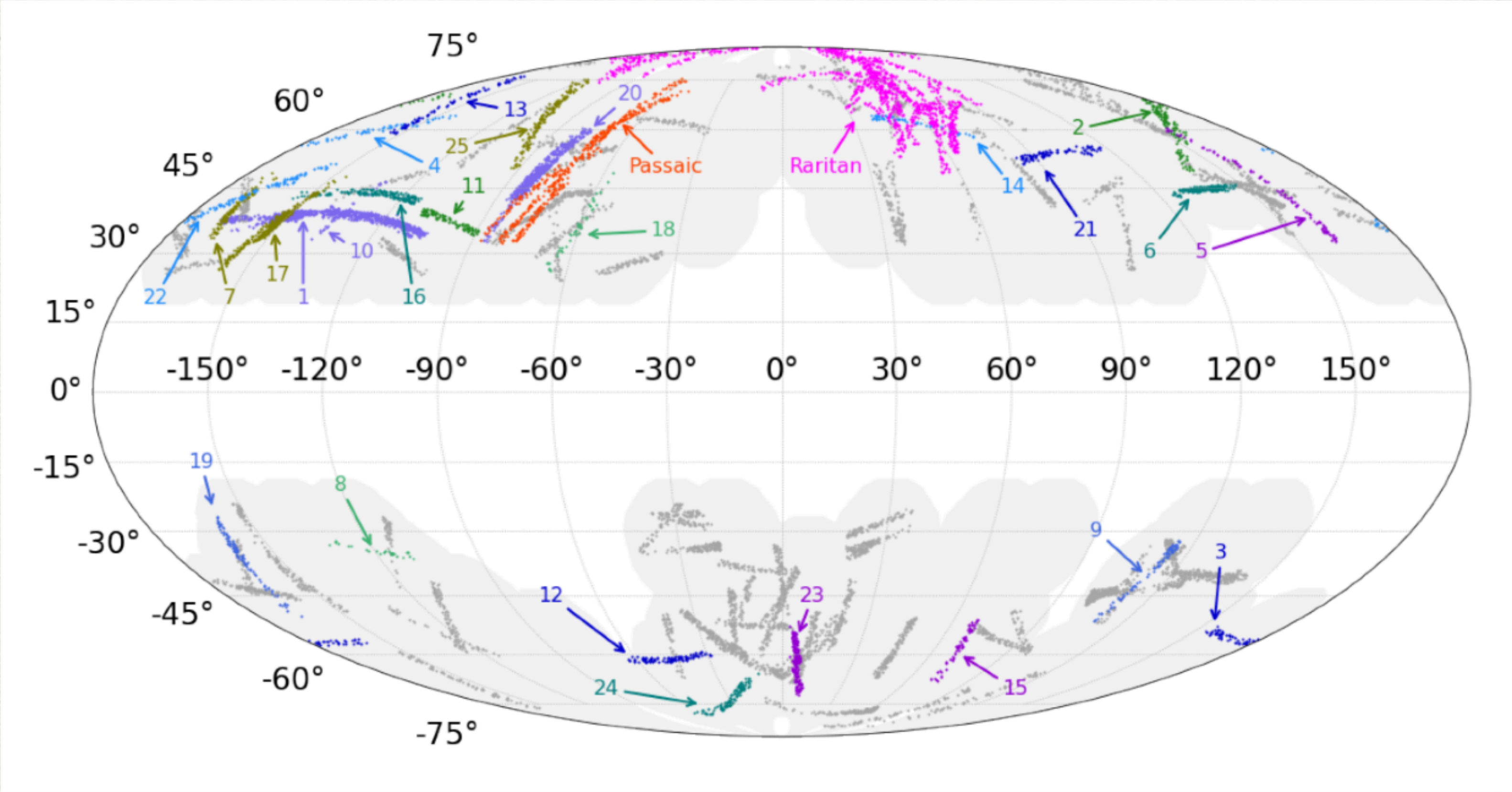


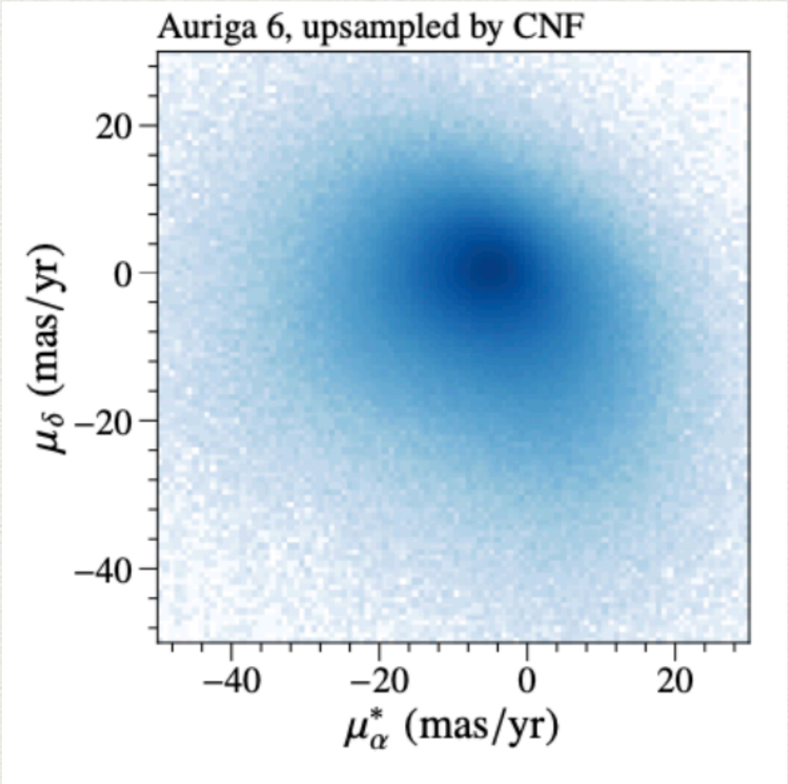
My research on ML for LHC physics



ML for astro/cosmo

- I am also increasingly branching out into ML for astro/cosmo, including a number of projects with Gaia data:
  - Via Machinae [2104.12789, 2303.01529, 2509.08064]: model-agnostic search for stellar streams, using resonant anomaly detection methods originally developed for LHC, discovered ~100 stream candidates, awaiting follow-up confirmation
  - GalaxyFlow [2211.11765]: using generative models to "upsample" cosmological simulations (star particles) to produce smooth and faithful Gaia mock catalogs (stars)
  - ClearPotential [2205.01129, 2305.13358, 2412.14236, 2510.xxxxx]: new technique using modern ML (normalizing flows) to measure local Galactic potential, acceleration field, and mass density field; first application to Gaia data





ML for astro/cosmo

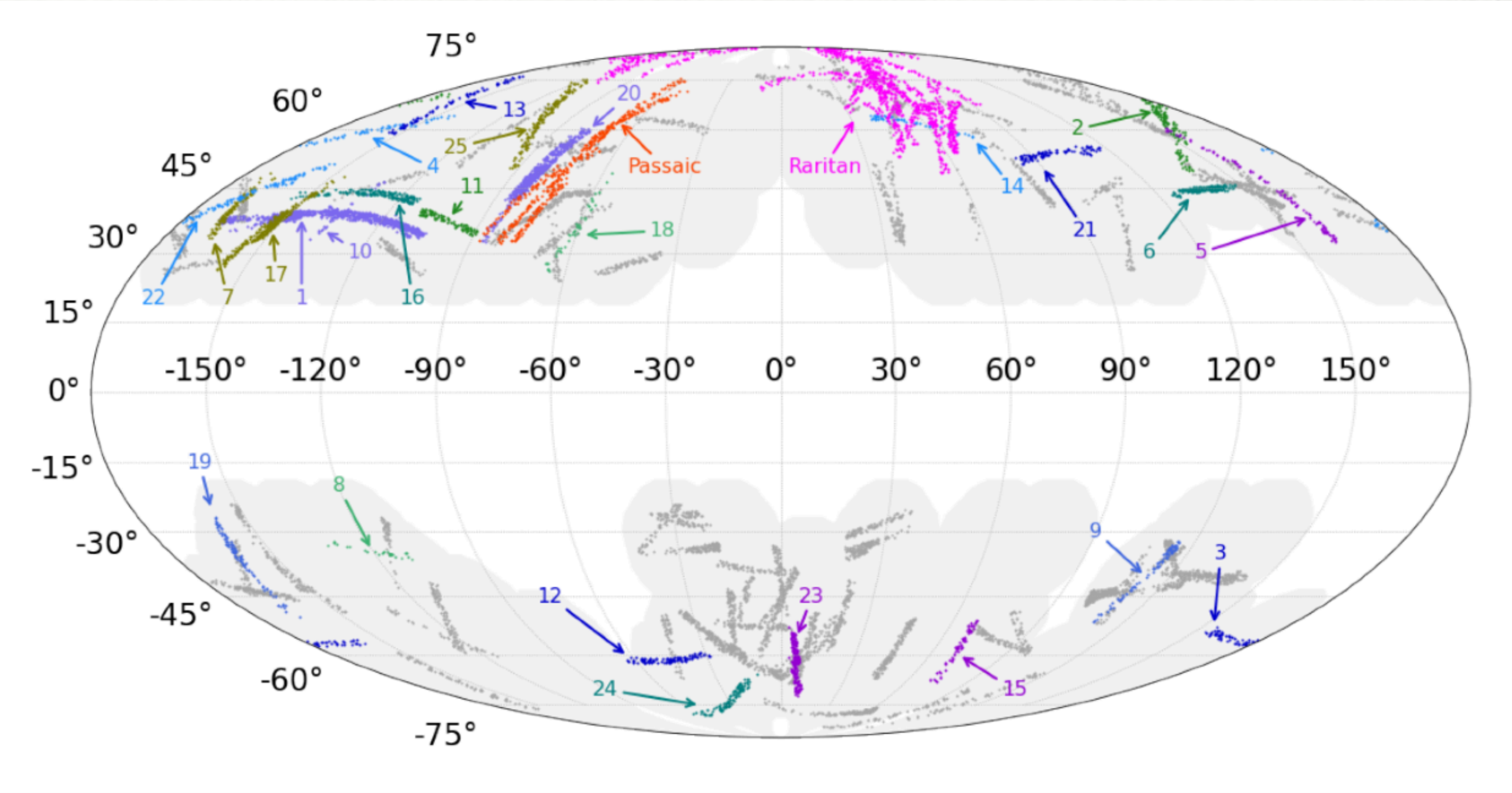
- I am also increasingly branching out into ML for astro/cosmo, including a number of projects with Gaia data:

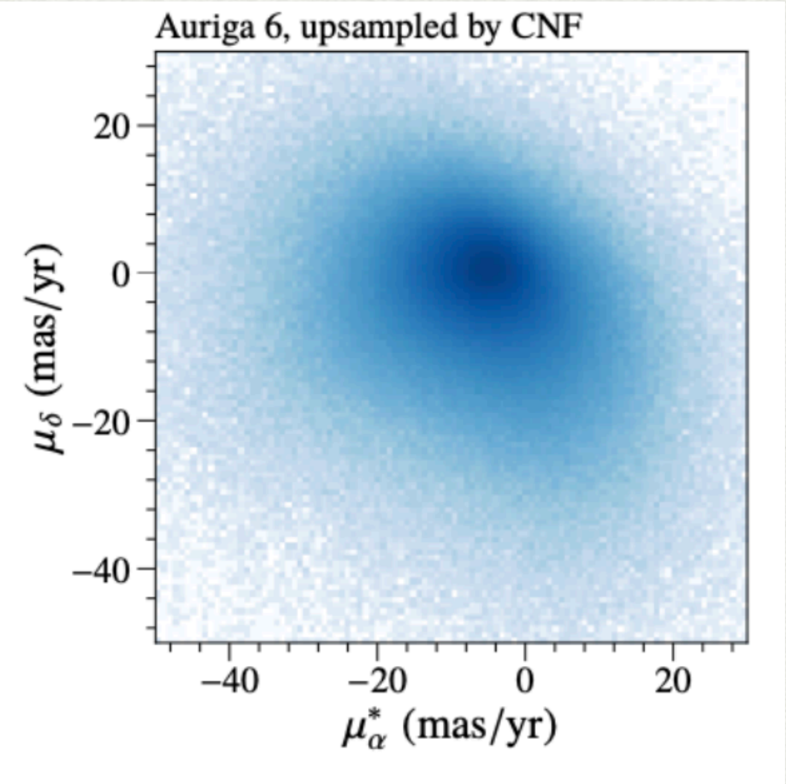
    - Via Machinae [2104.12789, 2303.01529, 2509.08064]: model-agnostic search for stellar streams, using resonant anomaly detection methods originally developed for LHC, discovered ~100 stream candidates, awaiting follow-up confirmation

    - GalaxyFlow [2211.11765]: using generative models to “upsample” cosmological simulations (star particles) to produce smooth and faithful Gaia mock catalogs (stars)

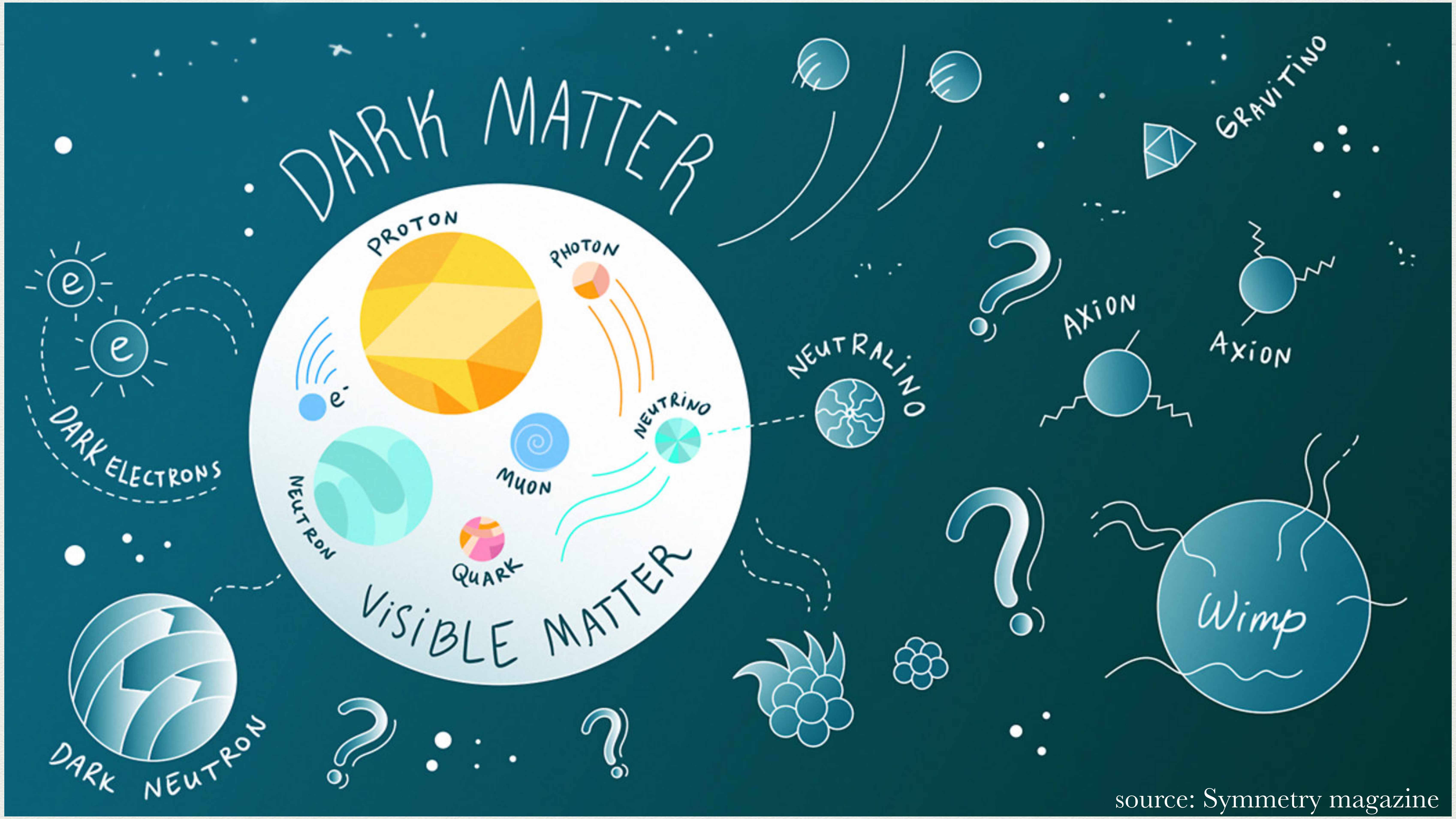
    - ClearPotential [2205.01129, 2305.13358, 2412.14236, 2510.xxxxx]: new technique using modern ML (normalizing flows) to measure local Galactic potential, acceleration field, and mass density field; first application to Gaia data

THIS TALK



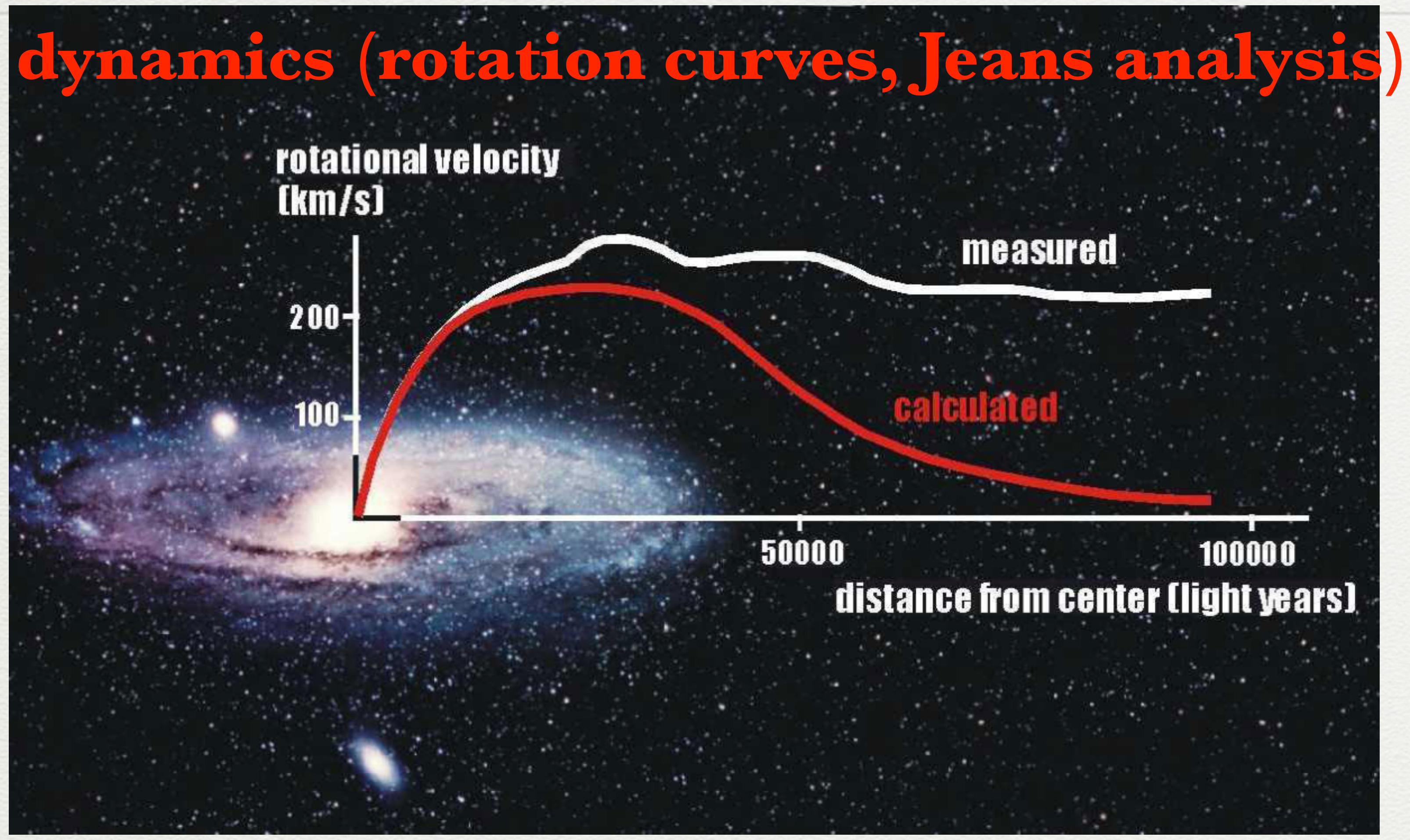


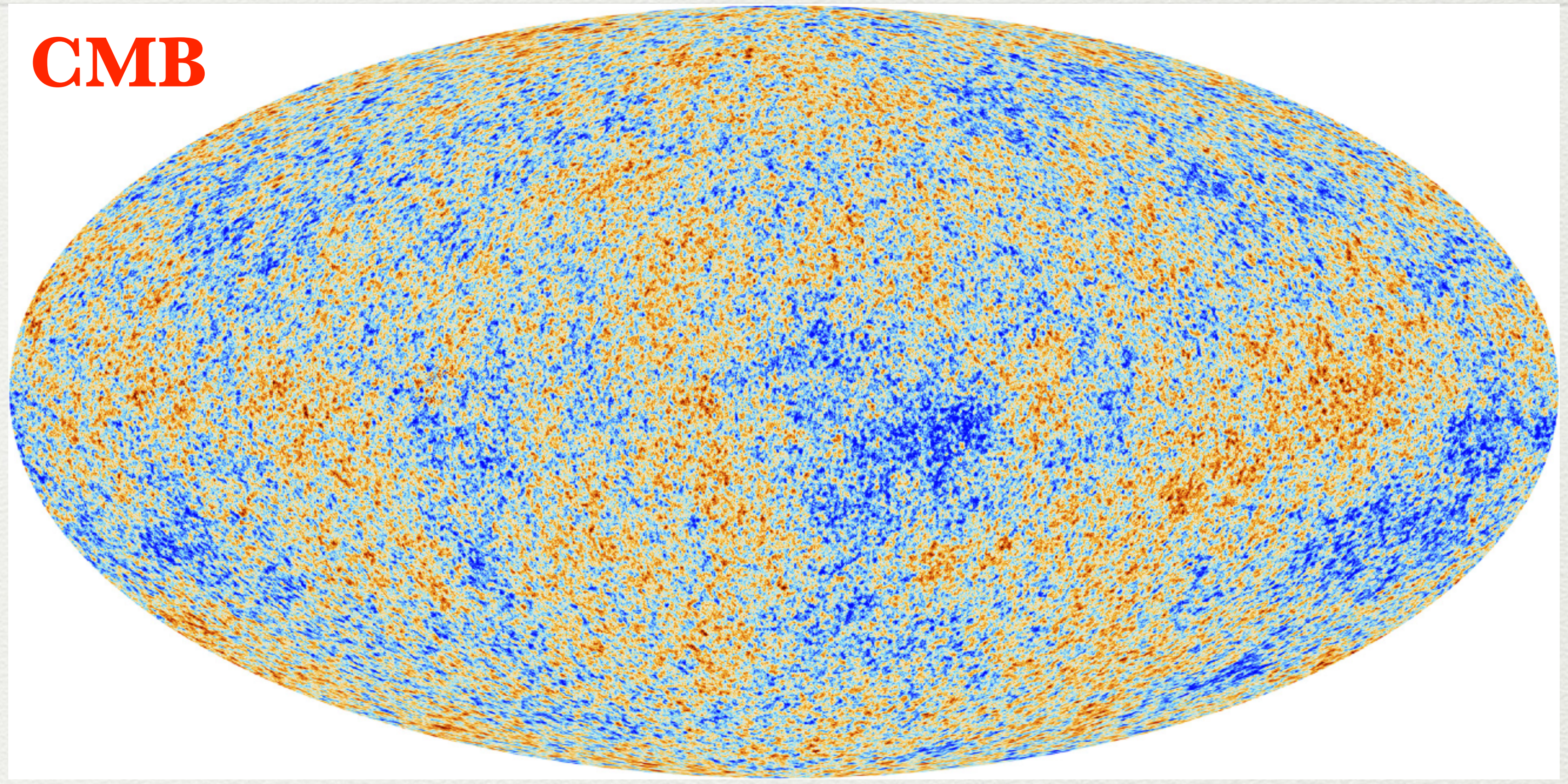
Motivation: Dark Matter

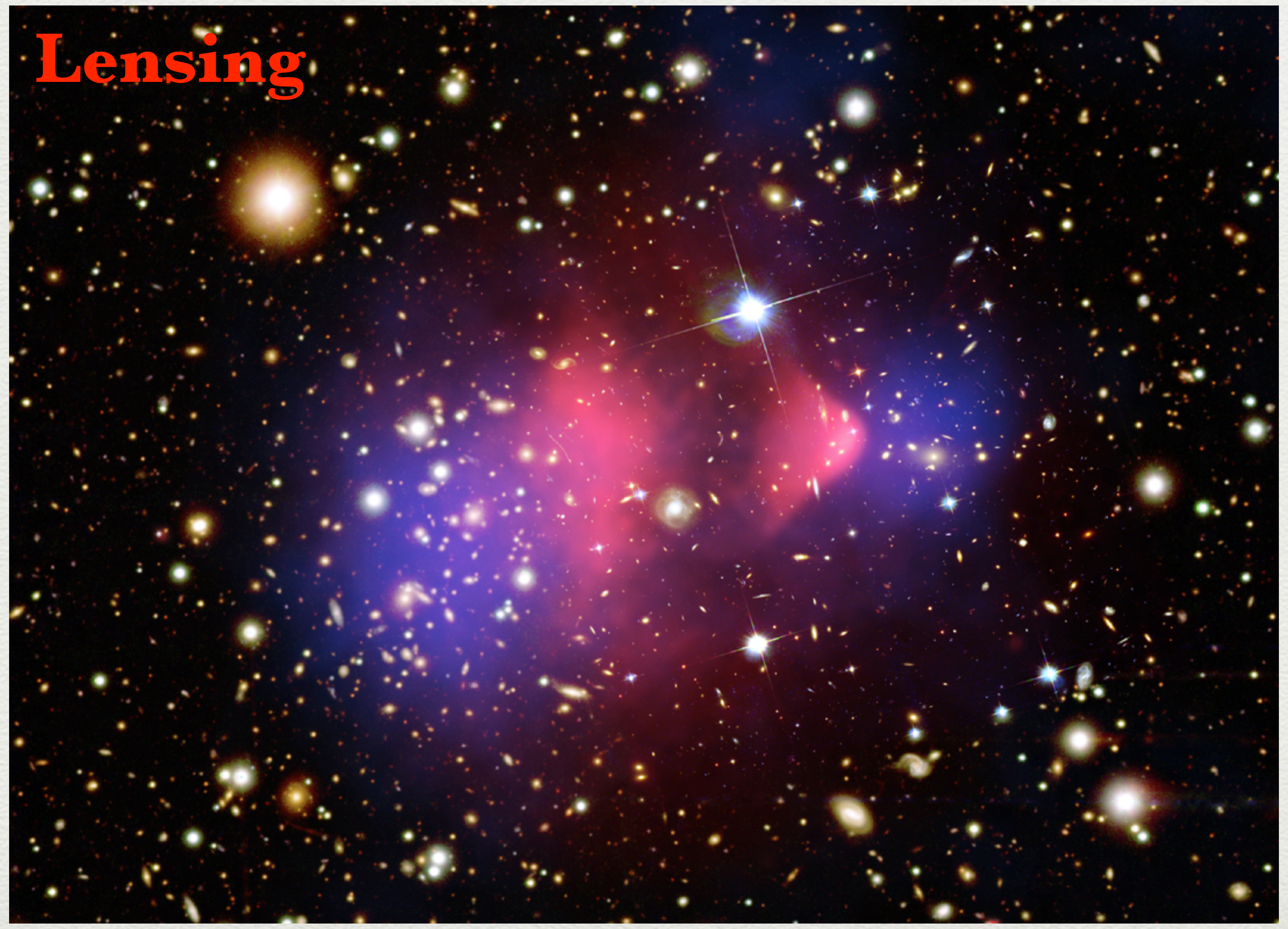


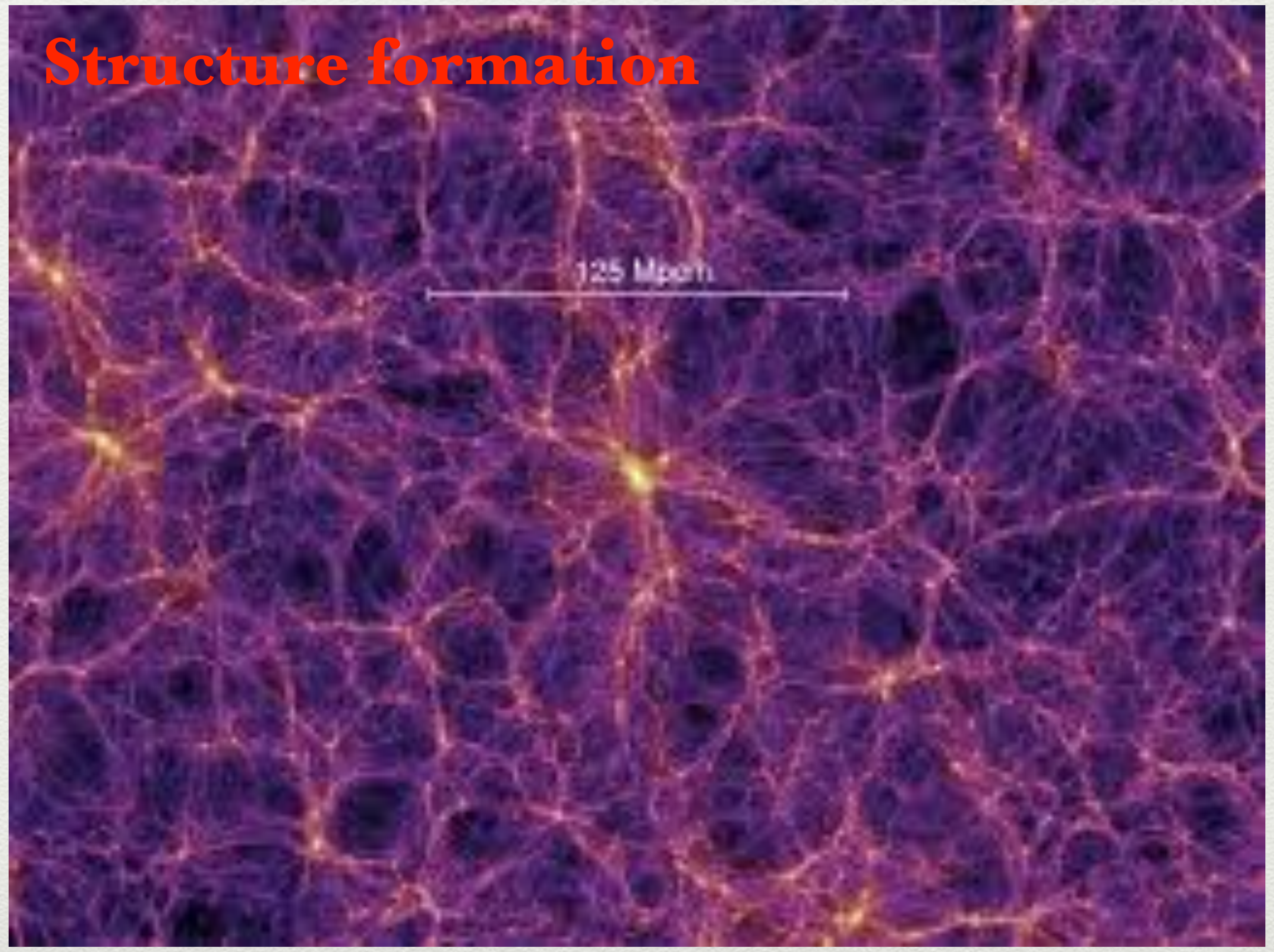
Dark matter is one of the greatest mysteries of our time

Dark Matter: Overwhelming evidence

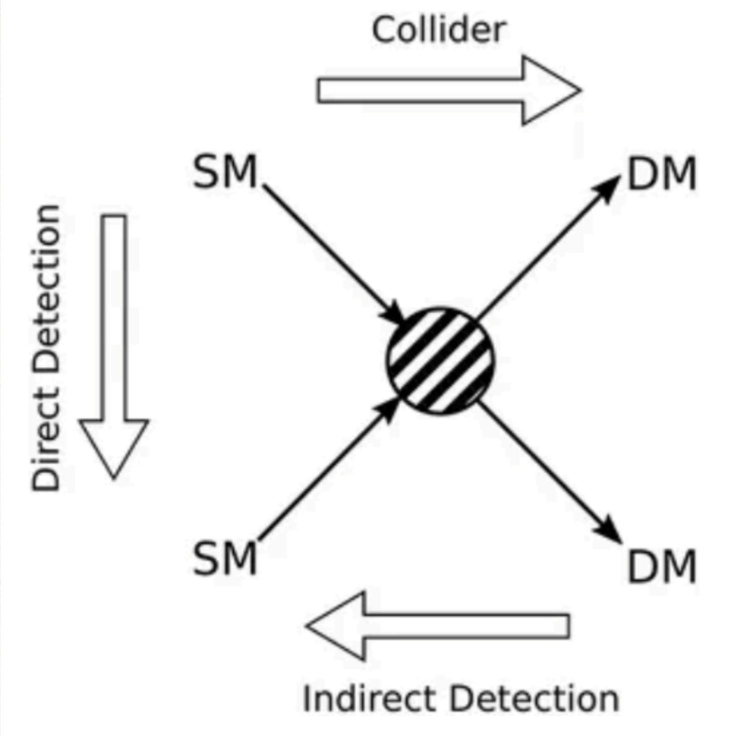


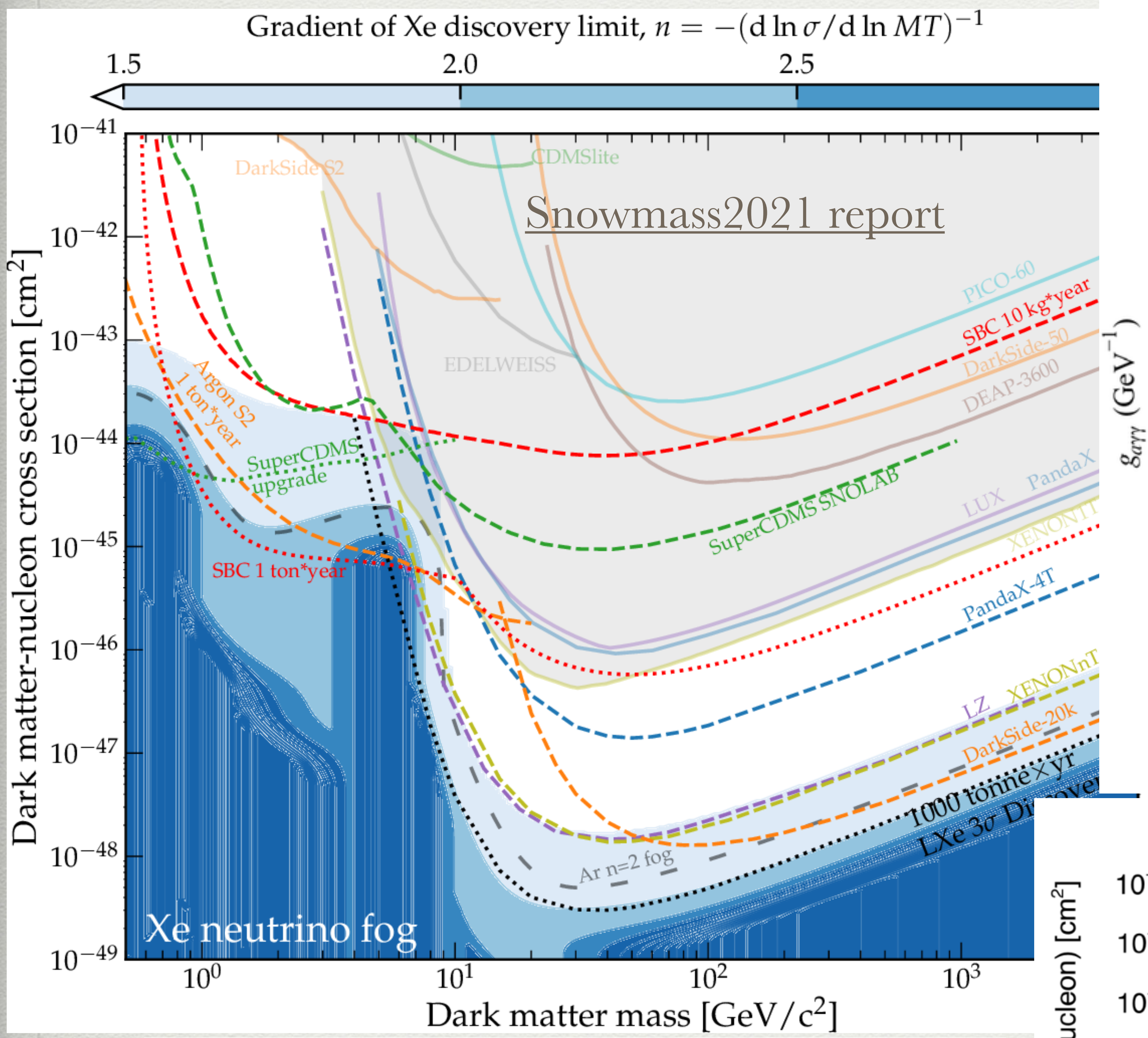


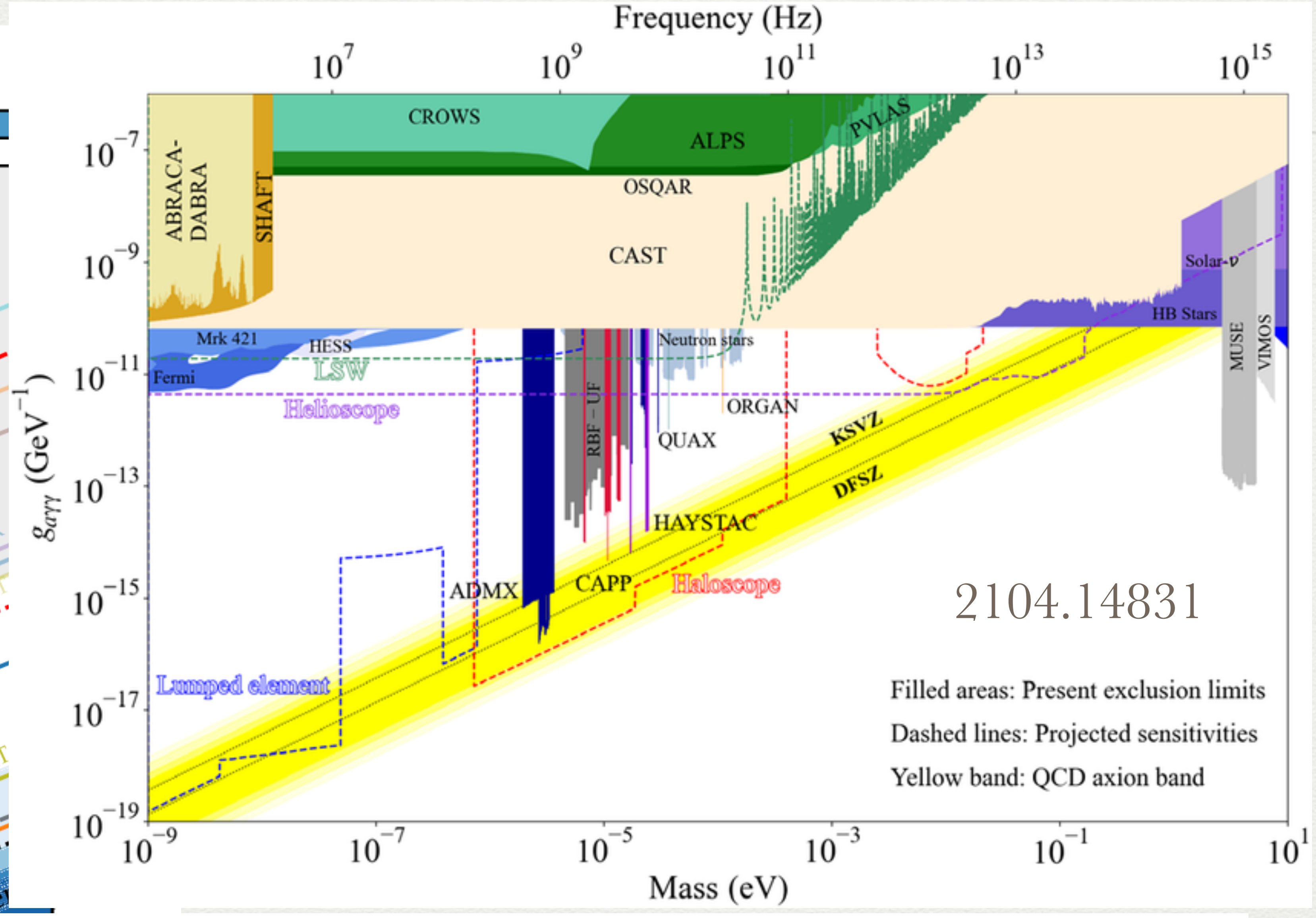


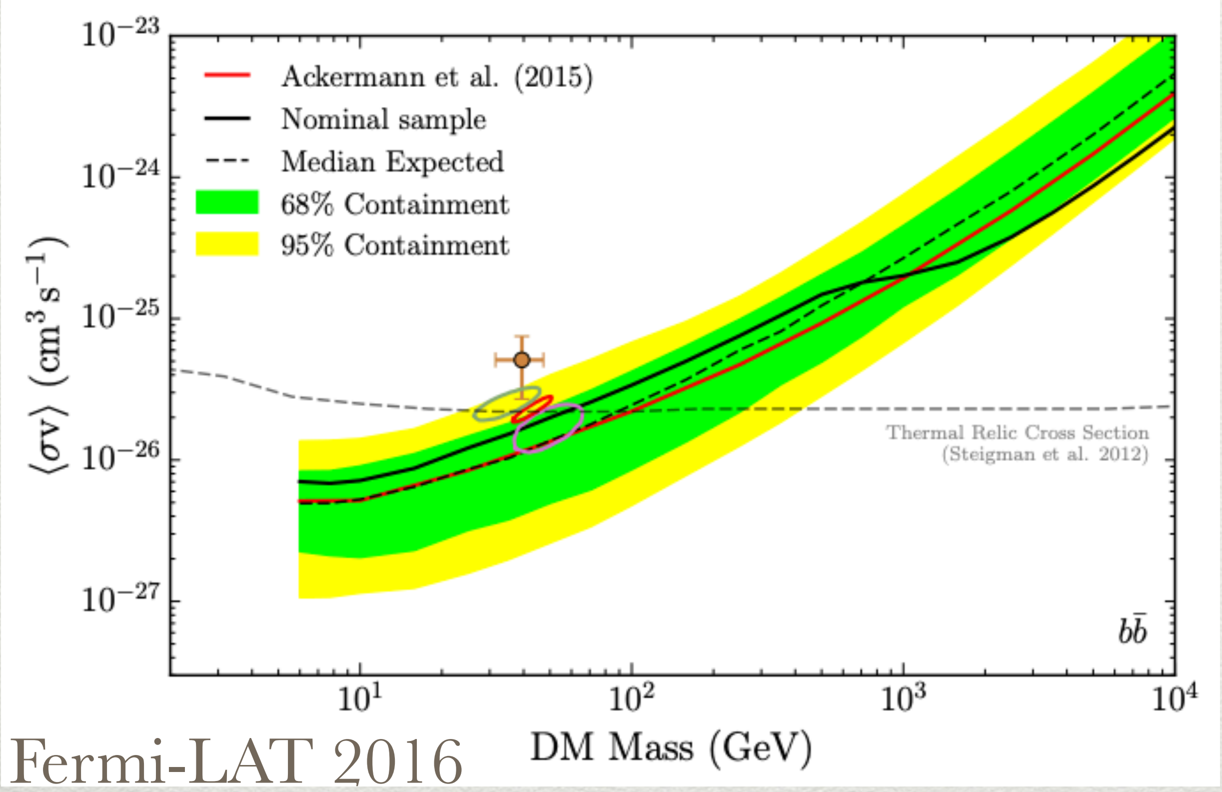


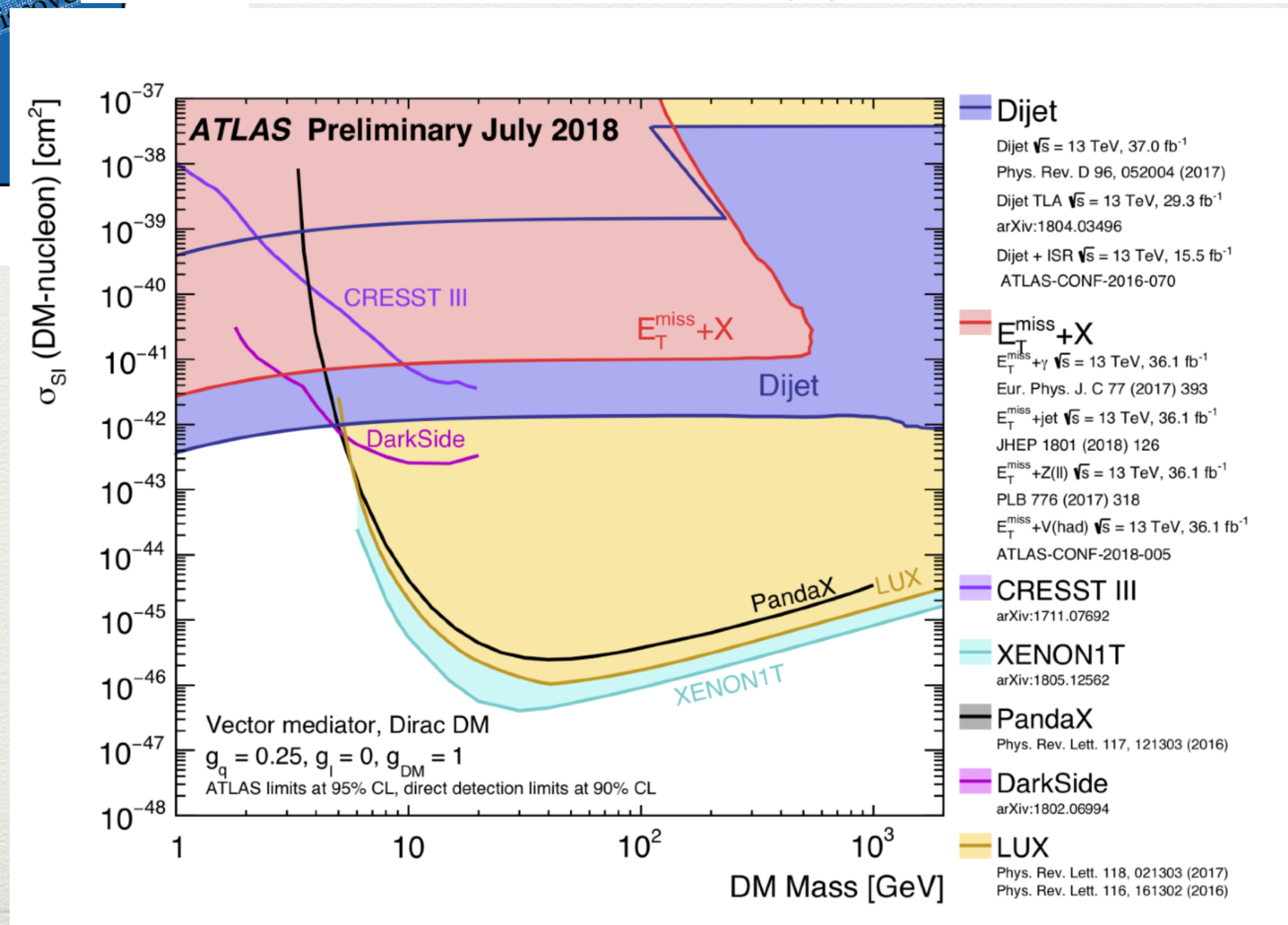
Searching for Dark Matter











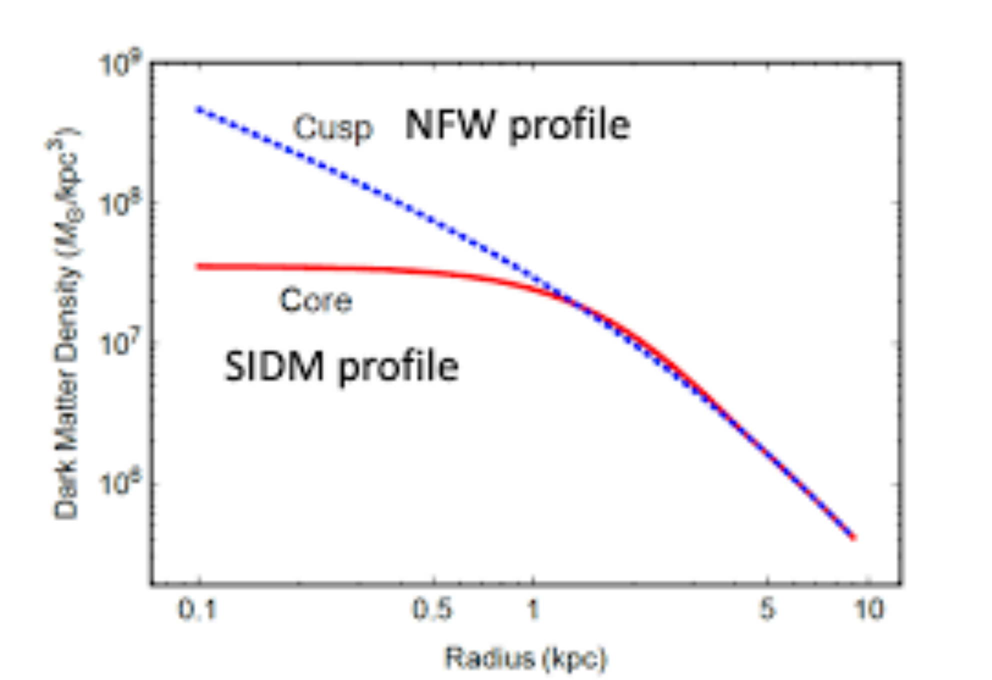
Yet the particle nature of dark matter remains elusive.

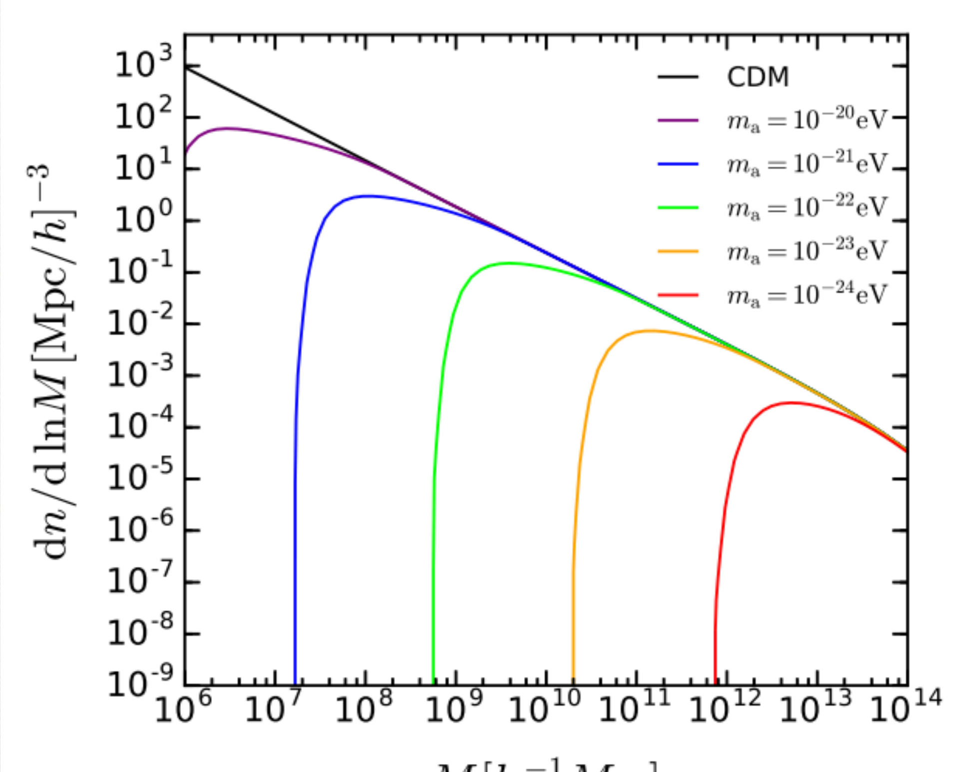
Many direct and indirect searches, but so far no additional evidence of dark matter interactions beyond gravity.

Dark matter from local dynamics

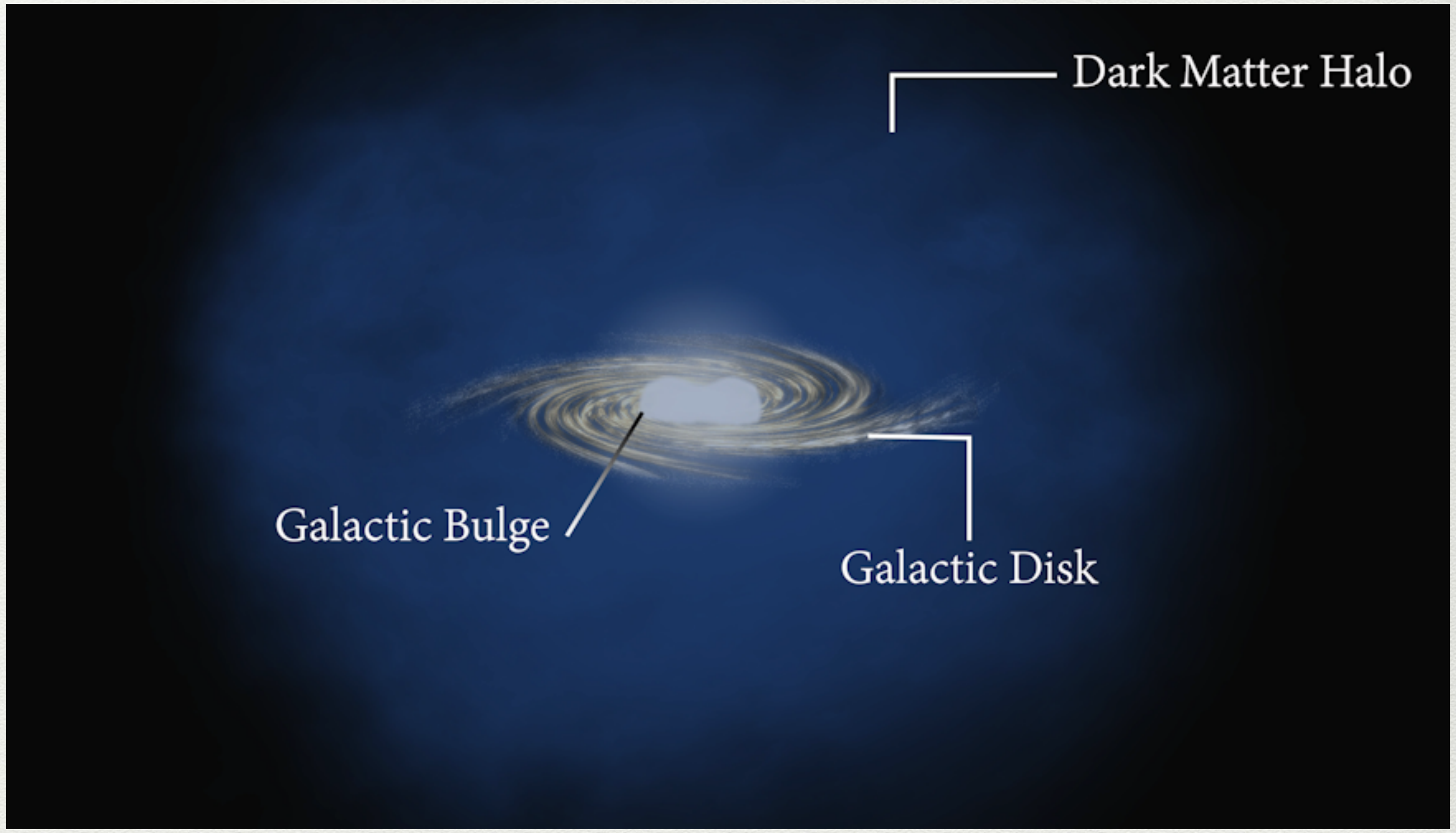
Local Galactic dynamics remain one of the best astrophysical probes of dark matter

- Local dark matter density important input to direct detection
- Density profile and substructure sensitive to DM properties





from 1608.02575



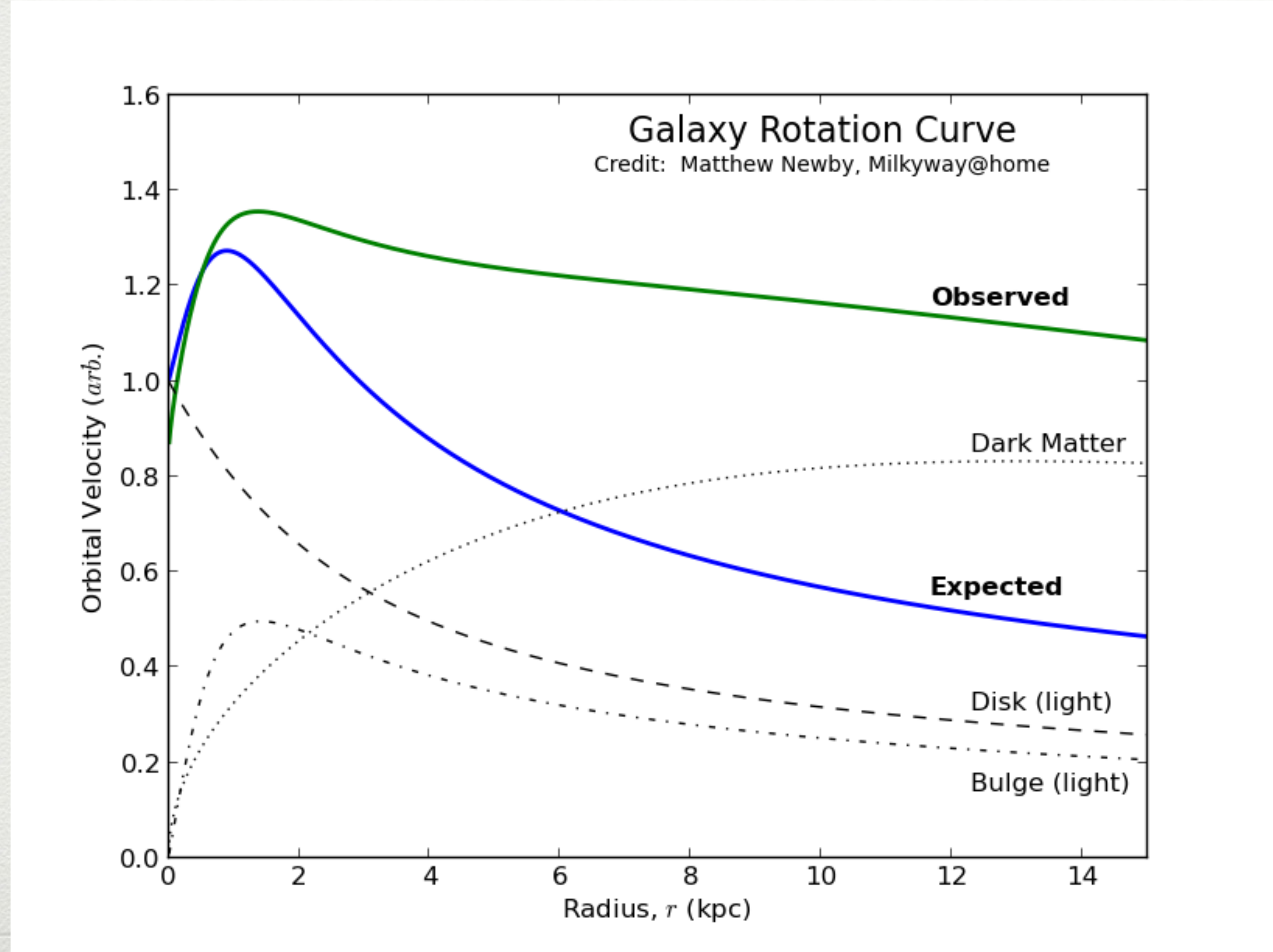
The Milky Way is a dark matter laboratory!

Dark matter from local dynamics

General idea: use kinematics of tracer stars to infer Galactic potential

Two main approaches

- Rotation curves: fit parametric potentials to circular orbits of disk stars



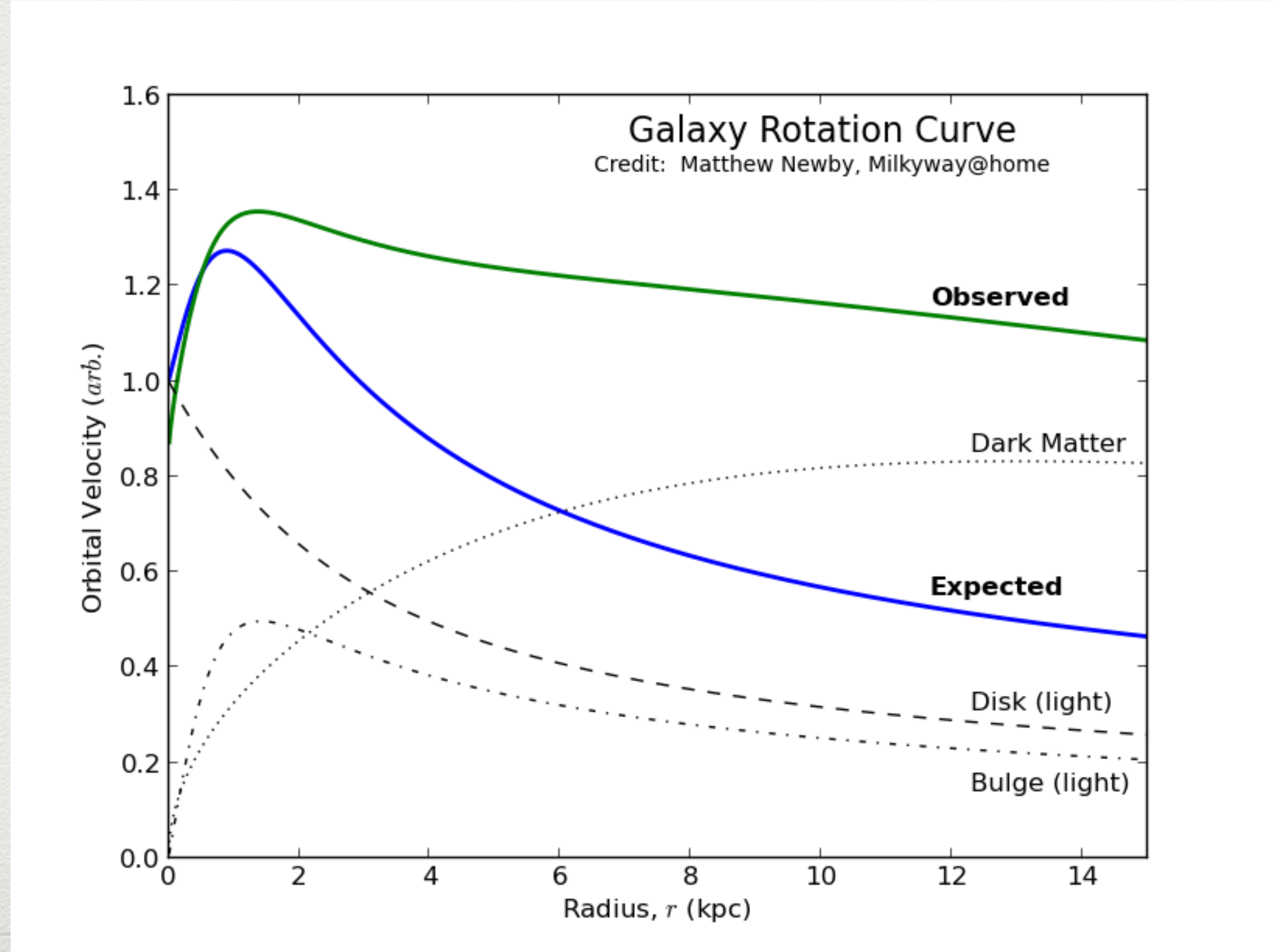
- Statistical properties: infer potential from statistics of tracer stars

Dark matter from local dynamics

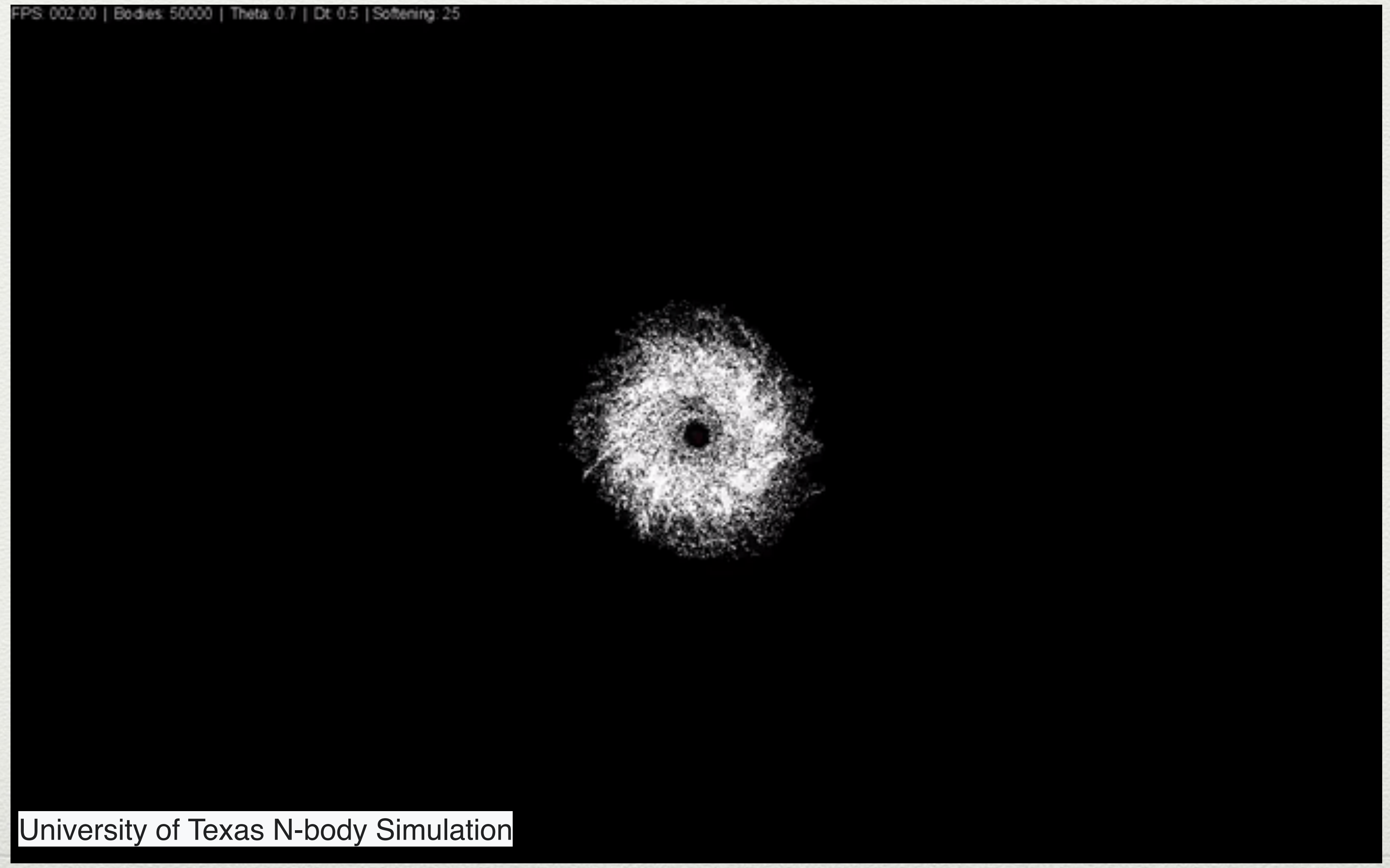
General idea: use kinematics of tracer stars to infer Galactic potential

Two main approaches

- Rotation curves: fit parametric potentials to circular orbits of disk stars



- Statistical properties: infer potential from statistics of tracer stars

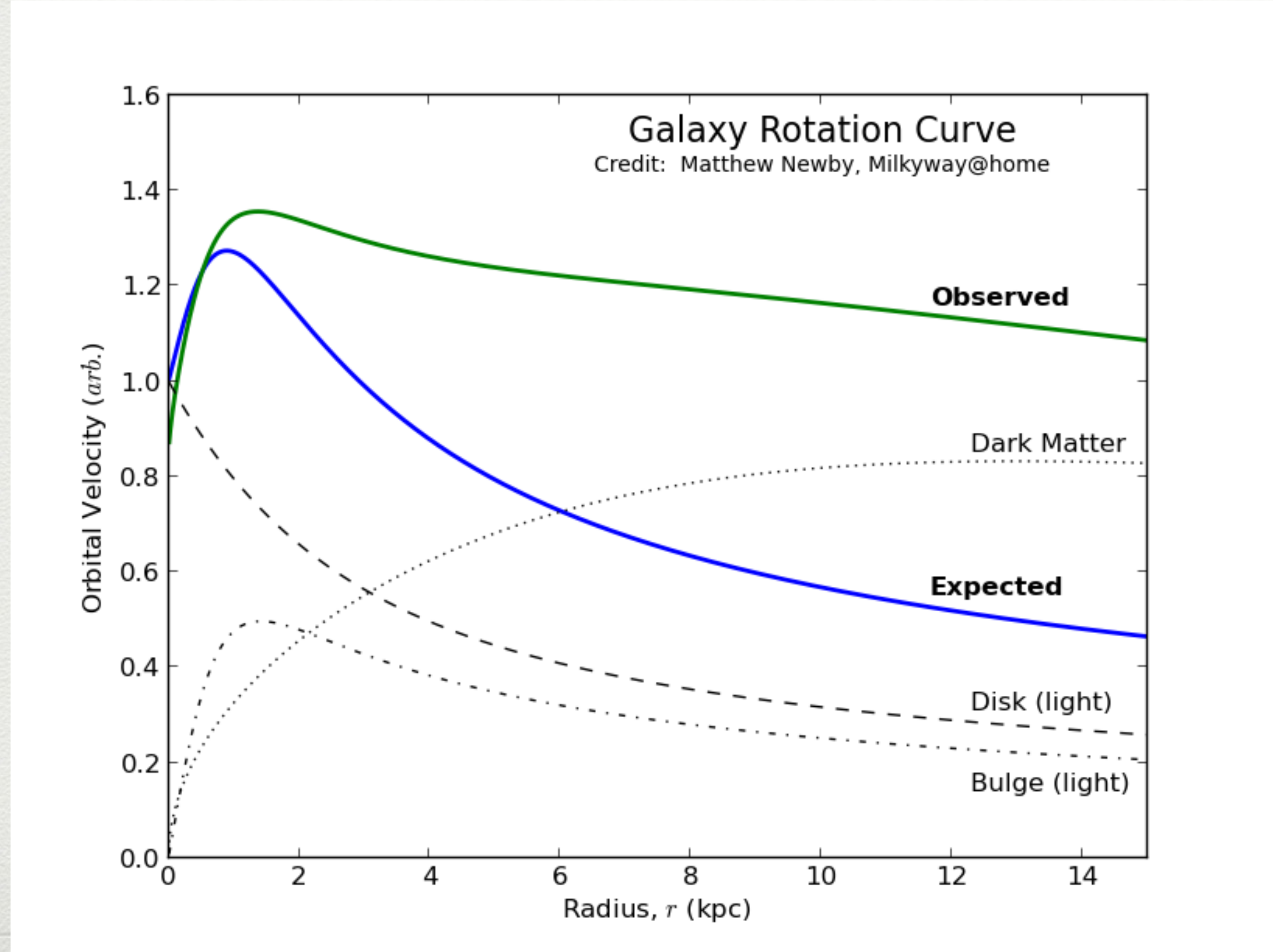


Dark matter from local dynamics

General idea: use kinematics of tracer stars to infer Galactic potential

<u>Two main approaches</u>

- Rotation curves: fit parametric potentials to circular orbits of disk stars



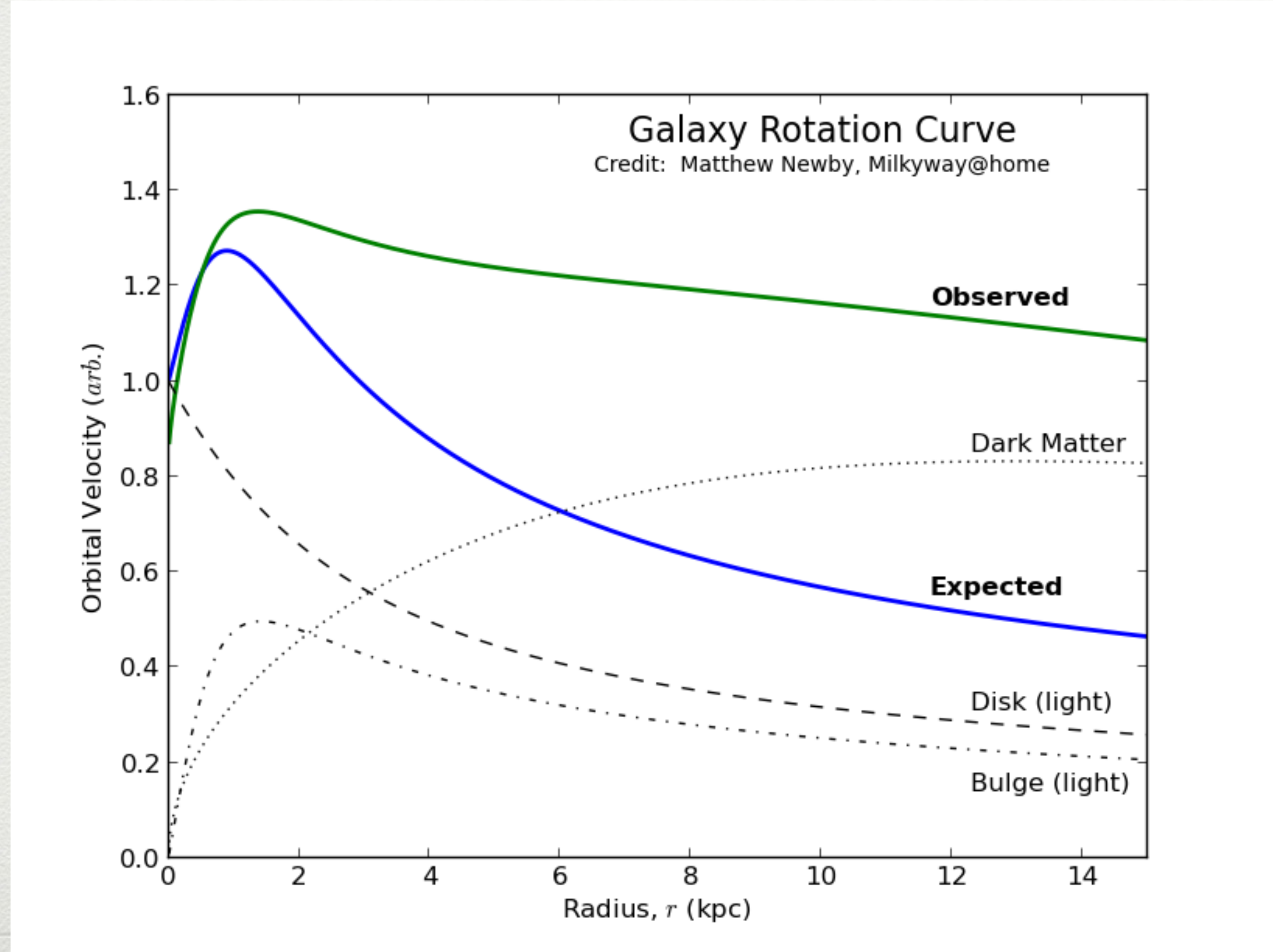
- **Statistical properties: infer potential from statistics of tracer stars**

Dark matter from local dynamics

General idea: use kinematics of tracer stars to infer Galactic potential

<u>Two main approaches</u>

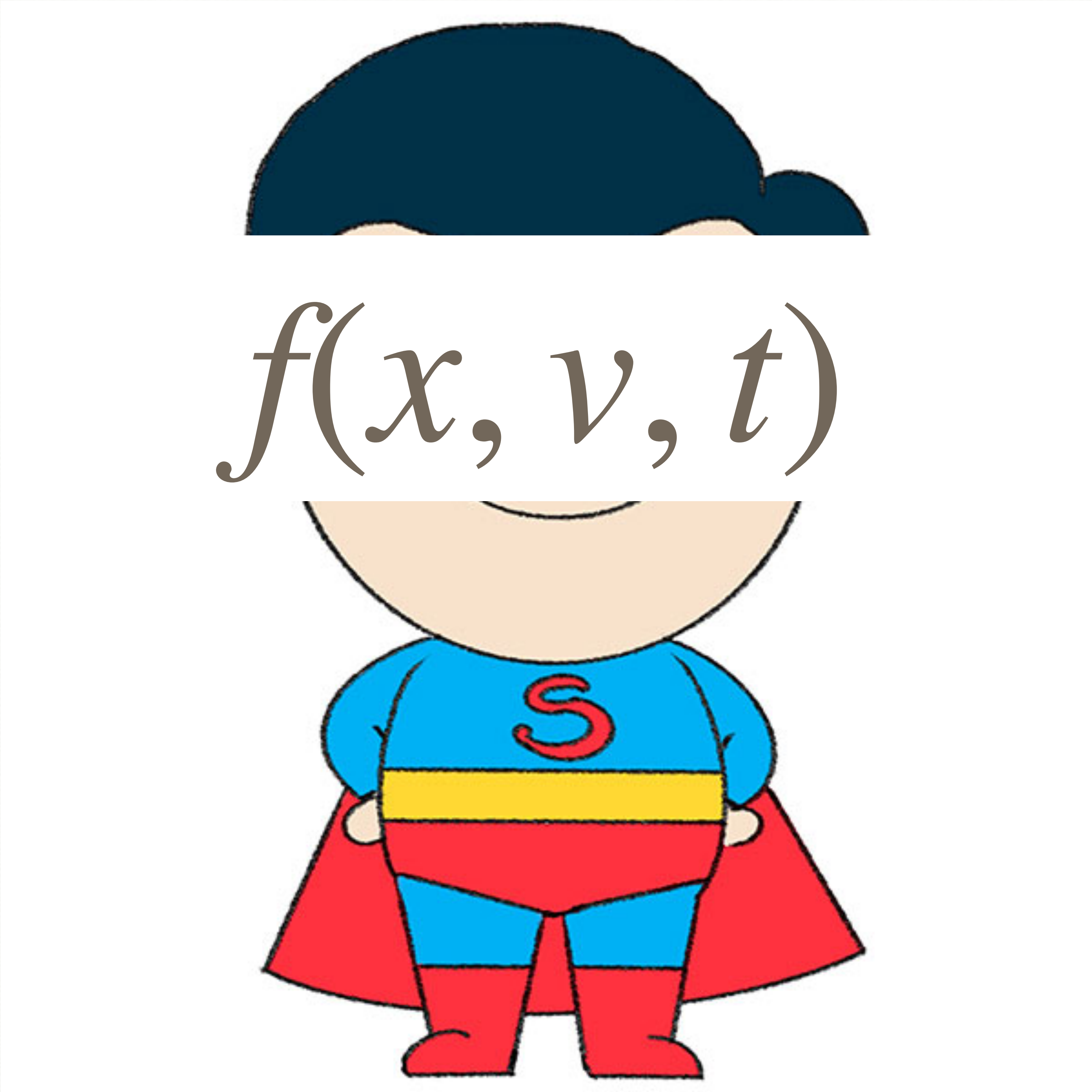
- Rotation curves: fit parametric potentials to circular orbits of disk stars



- **Statistical properties: infer potential from statistics of tracer stars**

Dark matter from stellar phase space

The hero of the story:



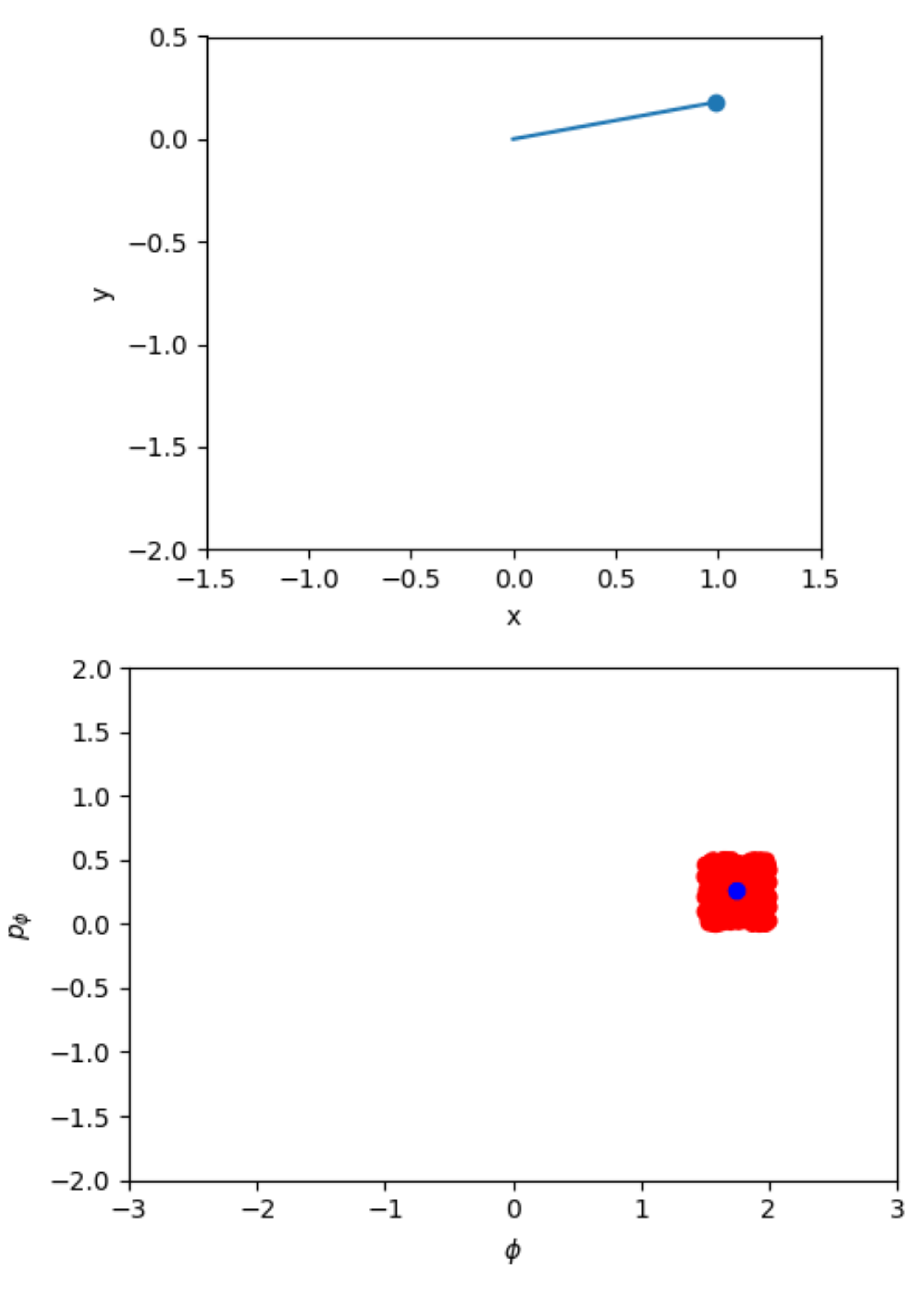
$$f(x, v, t)$$

Stellar phase space density (PSD)

Liouville's Theorem and the Collisionless Boltzmann Equation

For a particle moving under a conservative (frictionless) force, phase space volumes are constant in time

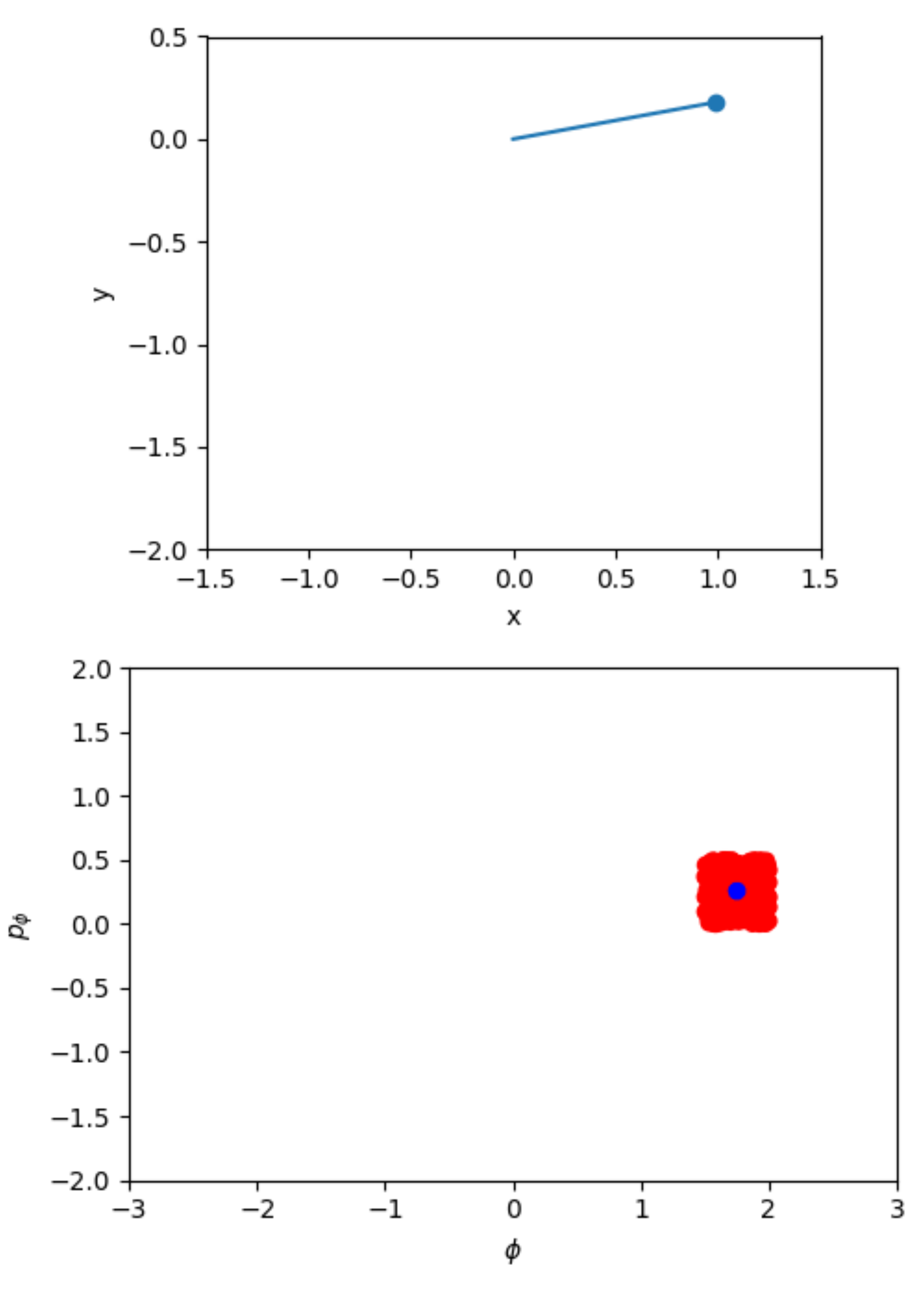
$$\frac{dV_{phase\ space}}{dt} = 0$$



Liouville's Theorem and the Collisionless Boltzmann Equation

For a particle moving under a conservative (frictionless) force, phase space volumes are constant in time

$$\frac{dV_{phase\ space}}{dt} = 0$$

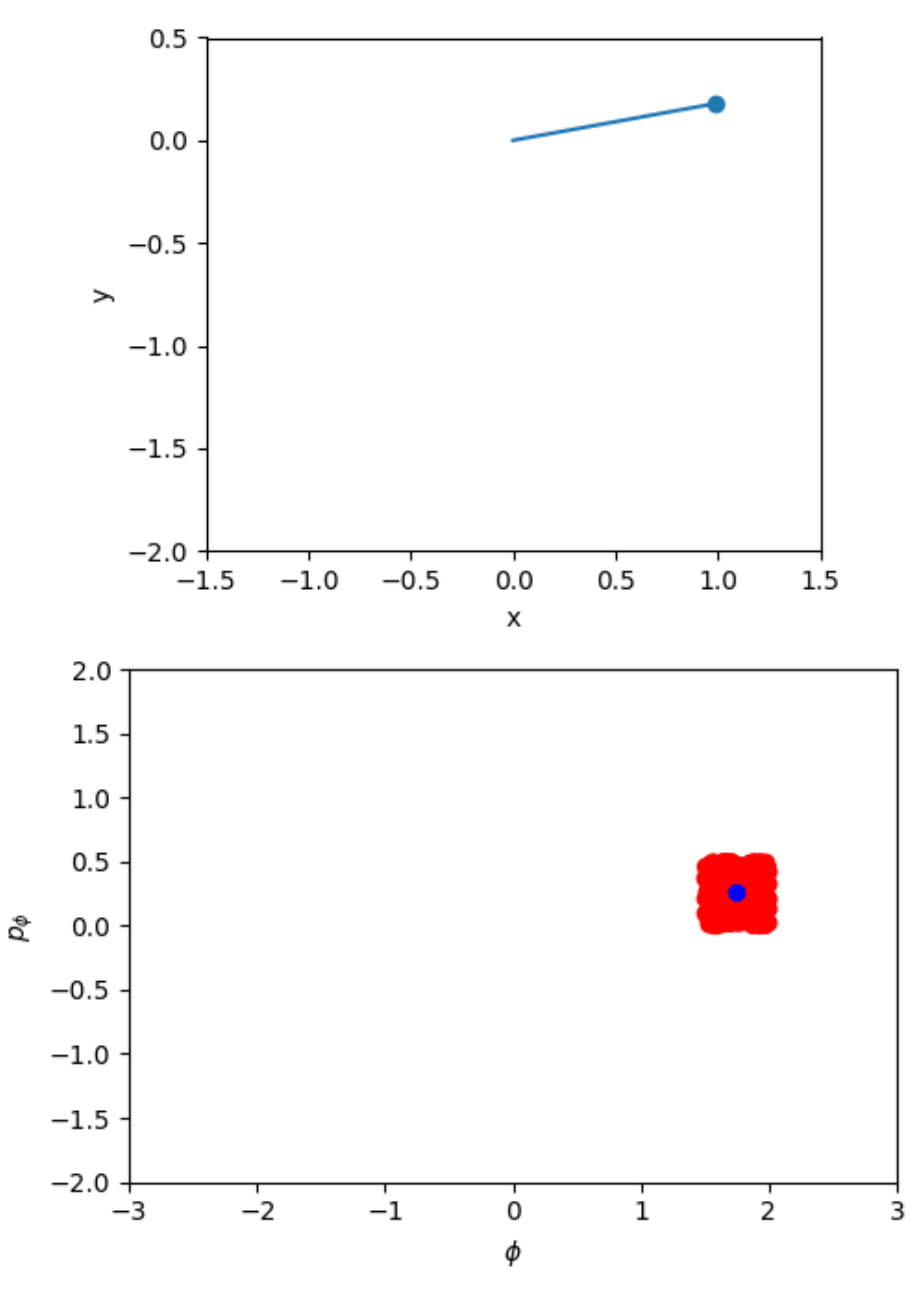


Liouville's Theorem and the Collisionless Boltzmann Equation

For a particle moving under a conservative (frictionless) force, phase space volumes are constant in time

$$\frac{dV_{phase\ space}}{dt} = 0$$

$$dV_{phase\ space} = f(x, v, t)dxdv$$



Liouville's Theorem and the Collisionless Boltzmann Equation

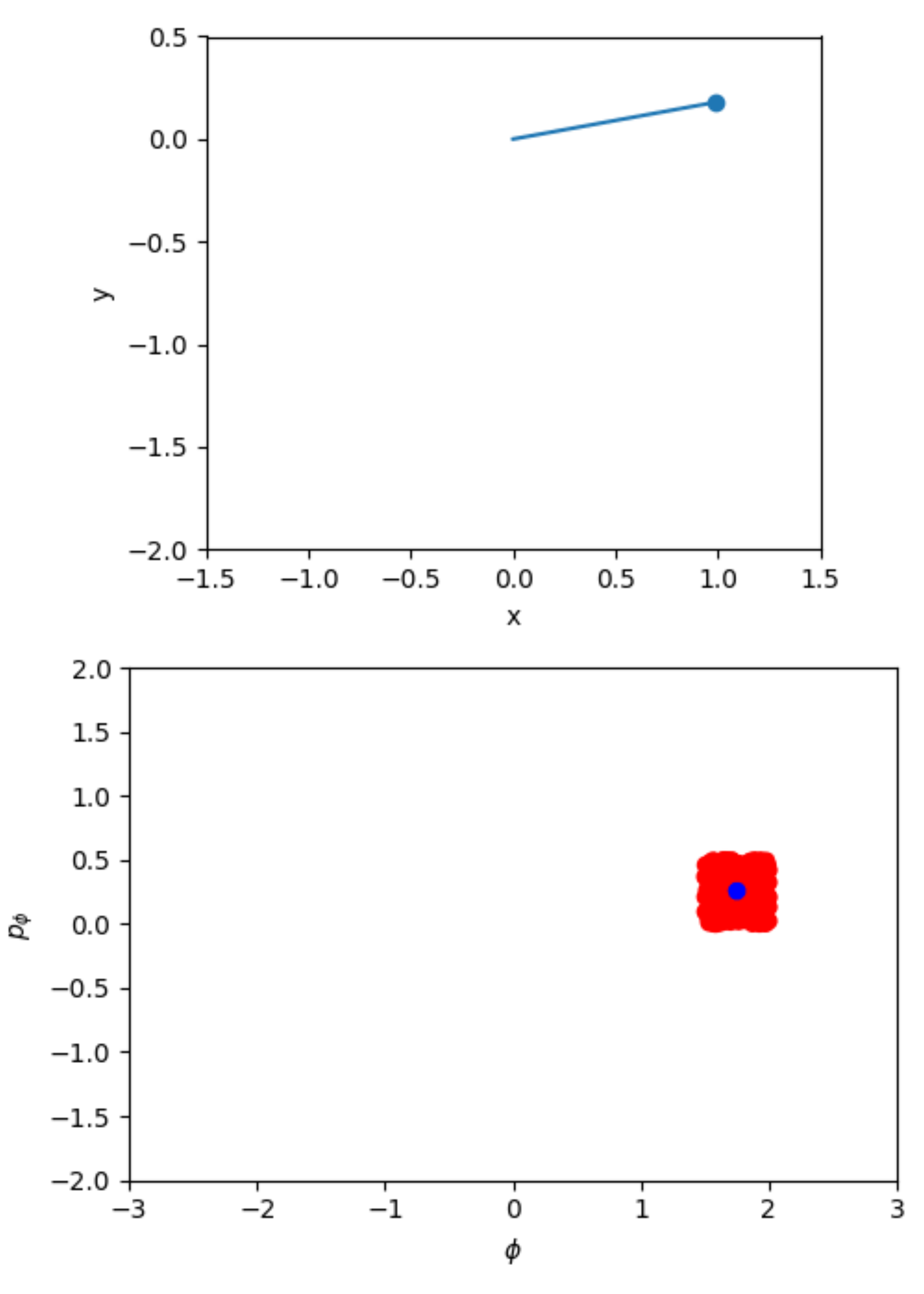
For a particle moving under a conservative (frictionless) force, phase space volumes are constant in time

$$\frac{dV_{phase\ space}}{dt} = 0$$

$$dV_{phase\ space} = f(x, v, t)dxdv$$

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} = 0$$

Collisionless Boltzmann Equation



Galactic potential from stellar phase space

- Stars are well described by CBE
- Old stars should be mostly in **dynamical equilibrium**

Galactic potential from stellar phase space

- Stars are well described by CBE
- Old stars should be mostly in **dynamical equilibrium**

$$\frac{\partial f}{\partial t} = 0$$

$$f(x, v, t) = f(x, v)$$

Galactic potential from stellar phase space

- Stars are well described by CBE
- Old stars should be mostly in **dynamical equilibrium**

$$\frac{\partial f}{\partial t} = 0$$

$$f(x, v, t) = f(x, v)$$

$$\boxed{\vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} = 0}$$

Galactic potential from stellar phase space

- Stars are well described by CBE
- Old stars should be mostly in **dynamical equilibrium**

$$\frac{\partial f}{\partial t} = 0$$

$$f(x, v, t) = f(x, v)$$

$$\boxed{\vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} = 0}$$

Stellar accelerations

Galactic potential from stellar phase space

$$\boxed{\vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} = 0}$$

- So if we know stellar phase space density, can infer their accelerations!
- From accelerations we can obtain gravitational potential and mass density

$$\vec{a}(x) = -\nabla\Phi(x) \qquad \nabla^2\Phi = -\nabla \cdot \vec{a} = 4\pi G\rho$$

Jeans Analysis

- Previously, 6d phase space density not learnable from data, nor its derivatives
- Instead: **Jeans Equations**

Jeans Analysis

- Previously, 6d phase space density not learnable from data, nor its derivatives
- Instead: **Jeans Equations**

$$0 = \int d^3v \, v_j \left(v_i \frac{\partial f}{\partial x_i} + a_i(x) \frac{\partial f}{\partial v_i} \right)$$

Jeans Analysis

- Previously, 6d phase space density not learnable from data, nor its derivatives
- Instead: **Jeans Equations**

$$0 = \int d^3v\, v_j \left(v_i \frac{\partial f}{\partial x_i} + a_i(x) \frac{\partial f}{\partial v_i} \right)$$

number density $$n(x) \equiv \int d^3v\, f(x, v)$$

Jeans Analysis

- Previously, 6d phase space density not learnable from data, nor its derivatives
- Instead: **Jeans Equations**

$$0 = \int d^3v\, v_j \left(v_i \frac{\partial f}{\partial x_i} + a_i(x) \frac{\partial f}{\partial v_i} \right)$$

number density $$n(x) \equiv \int d^3v\, f(x, v)$$

second velocity moments $$\langle v_i v_j \rangle(x) \equiv \frac{1}{n(x)} \int d^3v\, f(x, v)\, v_i v_j$$

Jeans Analysis

- Previously, 6d phase space density not learnable from data, nor its derivatives

- Instead: **Jeans Equations**

$$0 = \int d^3v\, v_j \left(v_i \frac{\partial f}{\partial x_i} + a_i(x) \frac{\partial f}{\partial v_i} \right)$$

$$= \frac{\partial}{\partial x_i} \Big(n(x) \langle v_i v_j \rangle (x) \Big) - a_j(x) n(x)$$

number density $n(x) \equiv \int d^3v\, f(x, v)$

second velocity moments $\langle v_i v_j \rangle (x) \equiv \frac{1}{n(x)} \int d^3v\, f(x, v)\, v_i v_j$

Jeans Analysis

- Previously, 6d phase space density not learnable from data, nor its derivatives

- Instead: **Jeans Equations**

$$0 = \int d^3v\, v_j \left(v_i \frac{\partial f}{\partial x_i} + a_i(x) \frac{\partial f}{\partial v_i} \right)$$

$$= \boxed{\frac{\partial}{\partial x_i} \Big(n(x) \langle v_i v_j \rangle (x) \Big) - a_j(x) n(x)}$$

number density $n(x) \equiv \int d^3v\, f(x, v)$

second velocity moments $\langle v_i v_j \rangle (x) \equiv \frac{1}{n(x)} \int d^3v\, f(x, v)\, v_i v_j$

Jeans Analysis

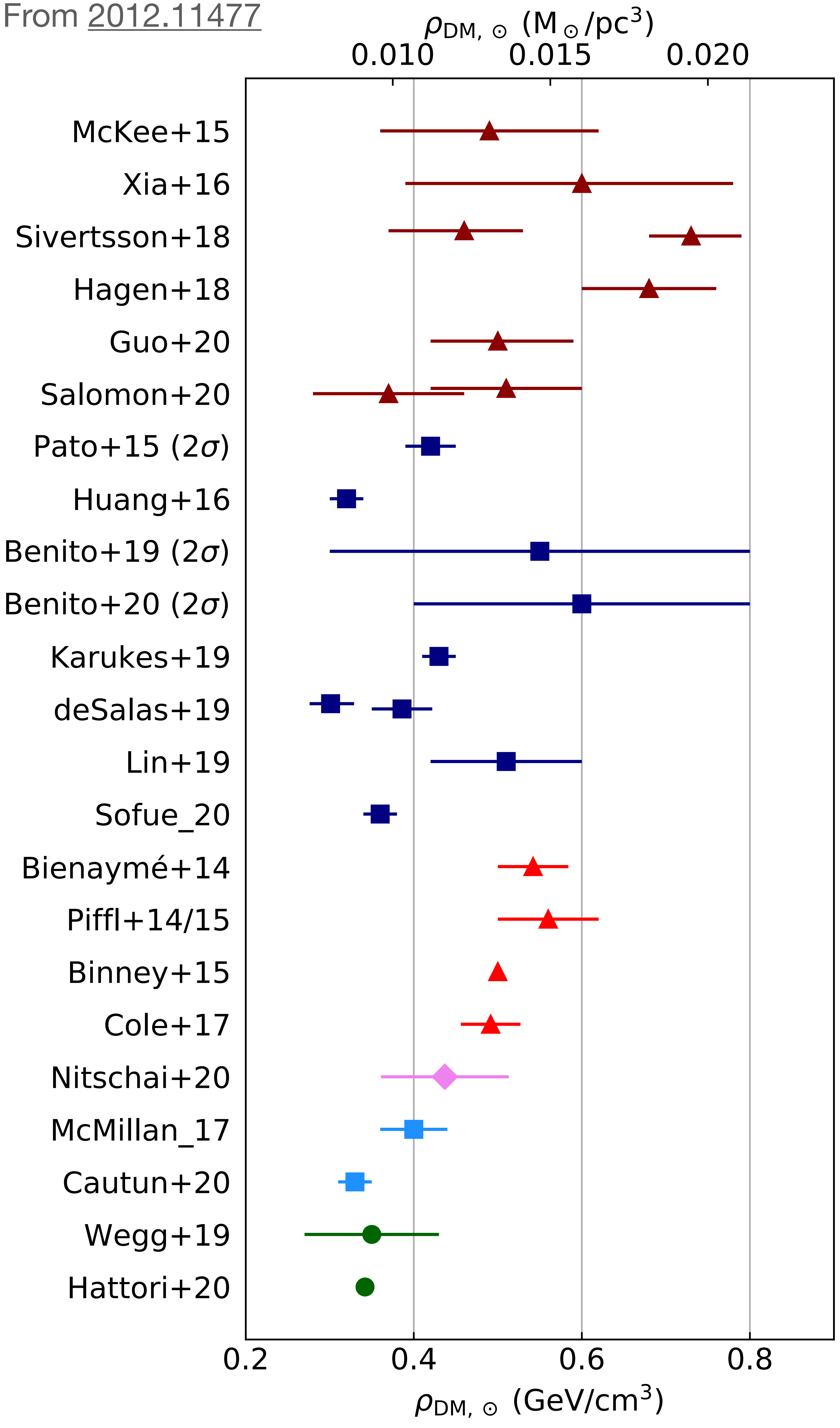
$$\boxed{0 = \frac{\partial}{\partial x_i}\Big(n(x)\langle v_i v_j\rangle(x)\Big) - a_j(x)n(x)}$$

- Number density and velocity moments can be estimated from 3d position bins
- Derivatives estimated from finite differences
- Often for enough statistics need to invoke additional assumptions of symmetries, eg azimuthal or north/south
- Finally, fit is often performed to a parametrized gravitational potential, eg

$$\Phi = \Phi_{NFW} + \Phi_{disk} + \Phi_{bulge} + \dots$$

- or just settle for determining ρ_{DM} at the solar location

From 2012.11477



A New Approach Powered by Modern ML

Green et al 2011.04673, 2205.02244; Naik et al 2112.07657; An et al 2106.05981; Kalda et al 2310.00040 2507.03742
Buckley, Lim, Putney & DS 2205.01129, 2305.13358, 2412.14236

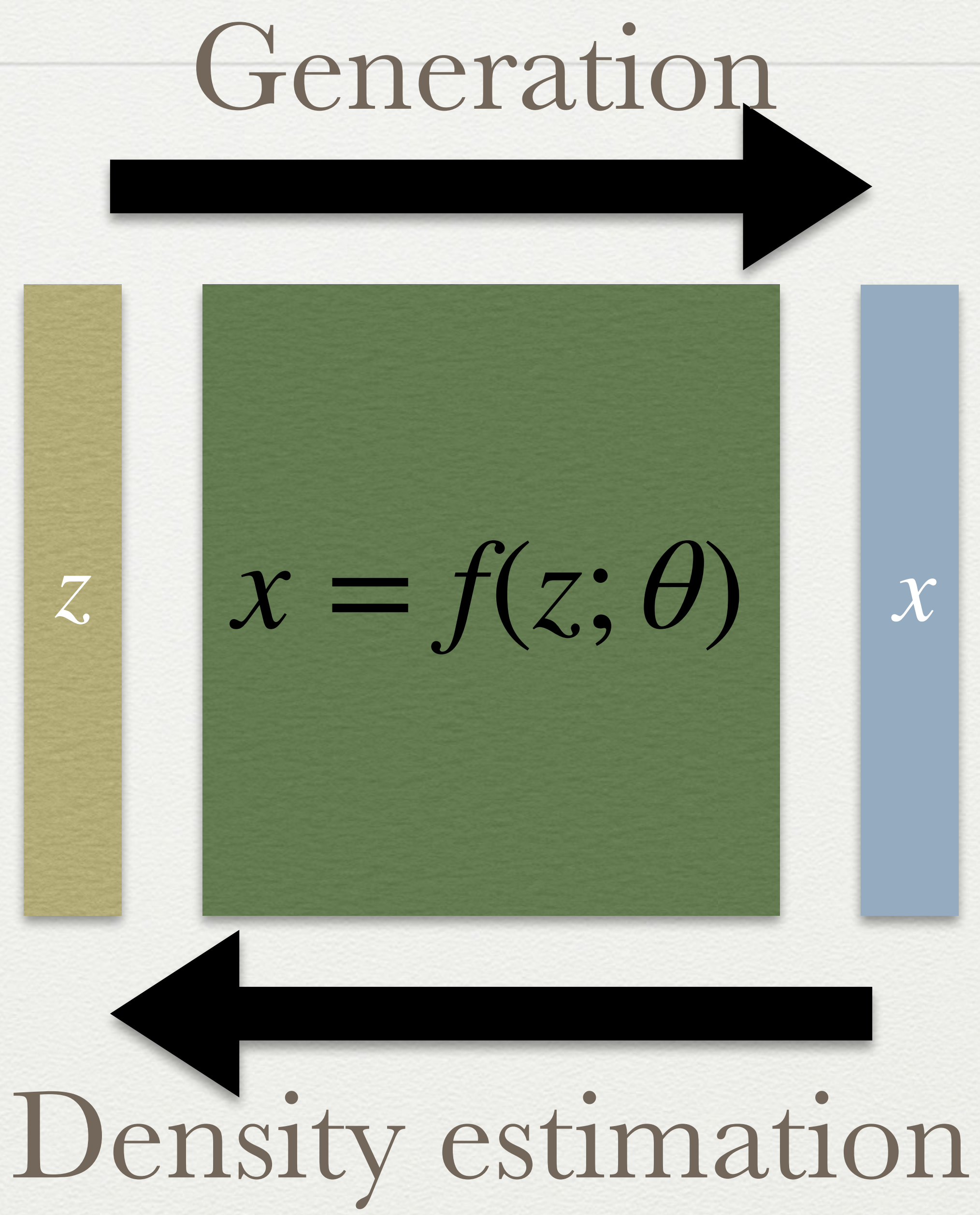
- With **modern ML** and the unprecedented coverage of **Gaia data**, can do much better than Jeans analysis
- Density estimation eg with normalizing flows can easily handle 6d phase space and its derivatives
- Neural networks can parametrize potential, acceleration and/or mass density

Fully data-driven, model-free, symmetry-free, unbinned measurement of Galactic potential!

Step 1: fit normalizing flows to Gaia data

Normalizing Flows

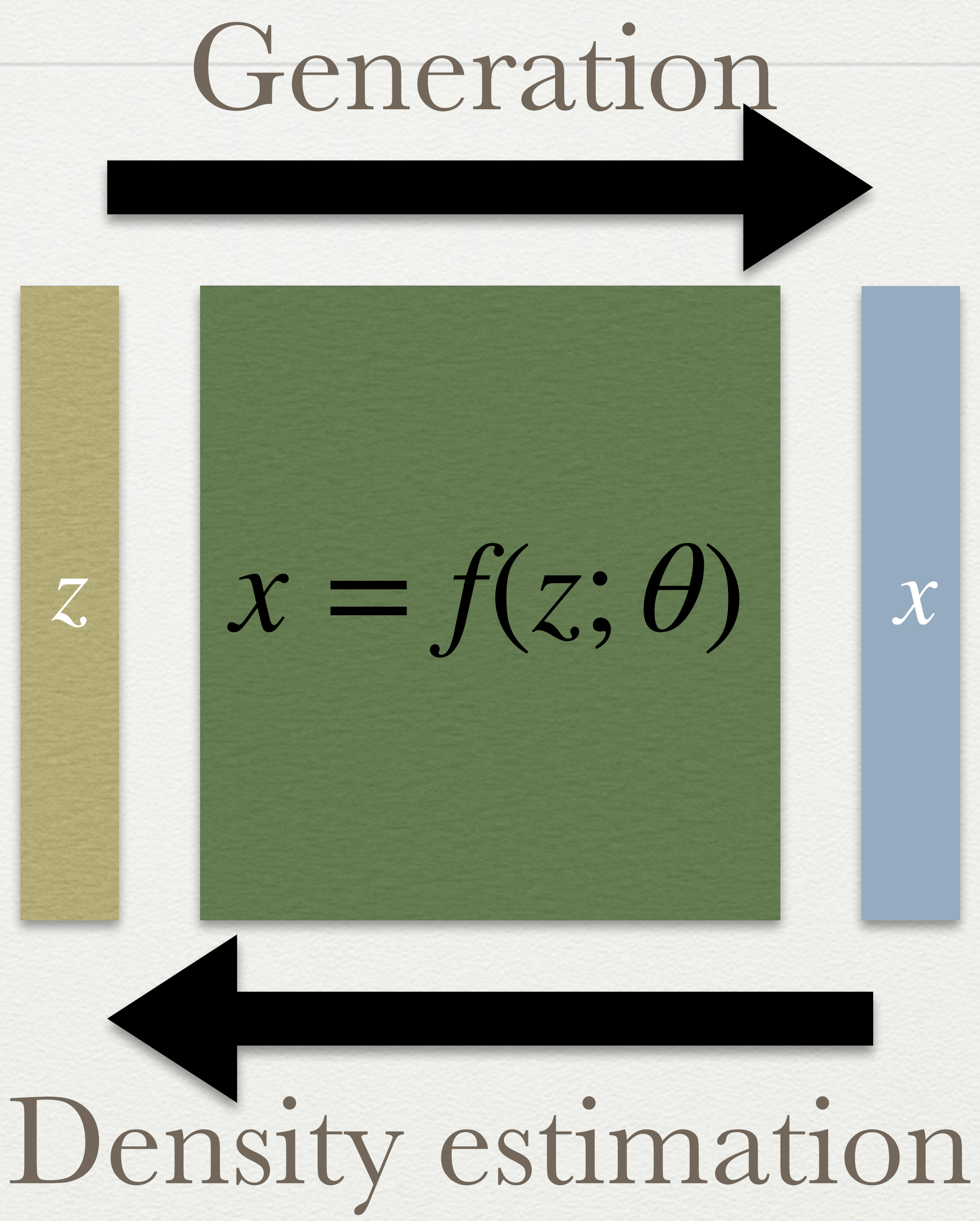
$z \sim p_{base}(z) = \mathcal{N}(0,1)^d$



$x \sim p_\theta(x)$

Normalizing Flows

$z \sim p_{base}(z) = \mathcal{N}(0,1)^d$

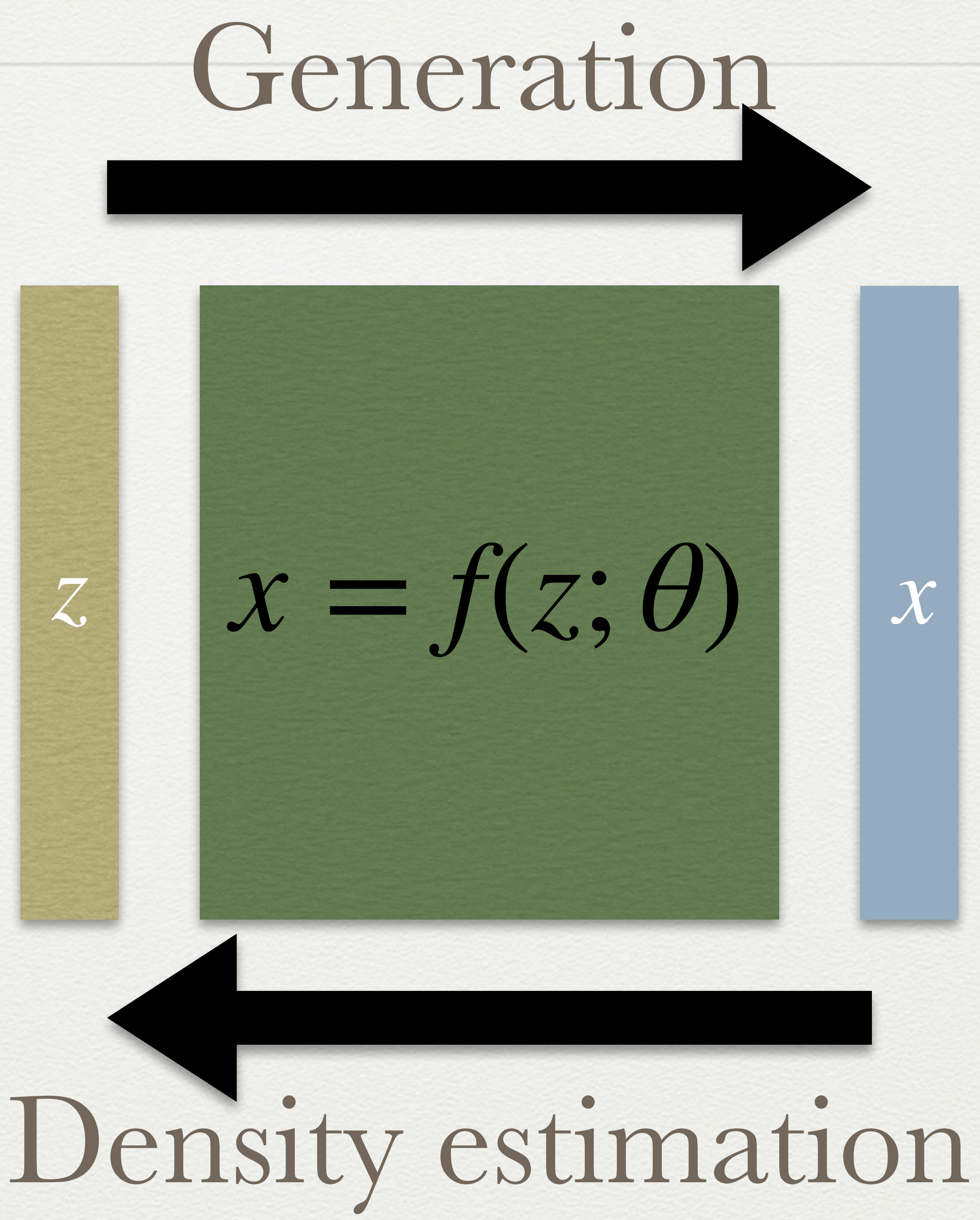


$x \sim p_\theta(x)$

Powerful class of density estimators that are also generative models

Normalizing Flows

$$z \sim p_{base}(z) = \mathcal{N}(0,1)^d$$



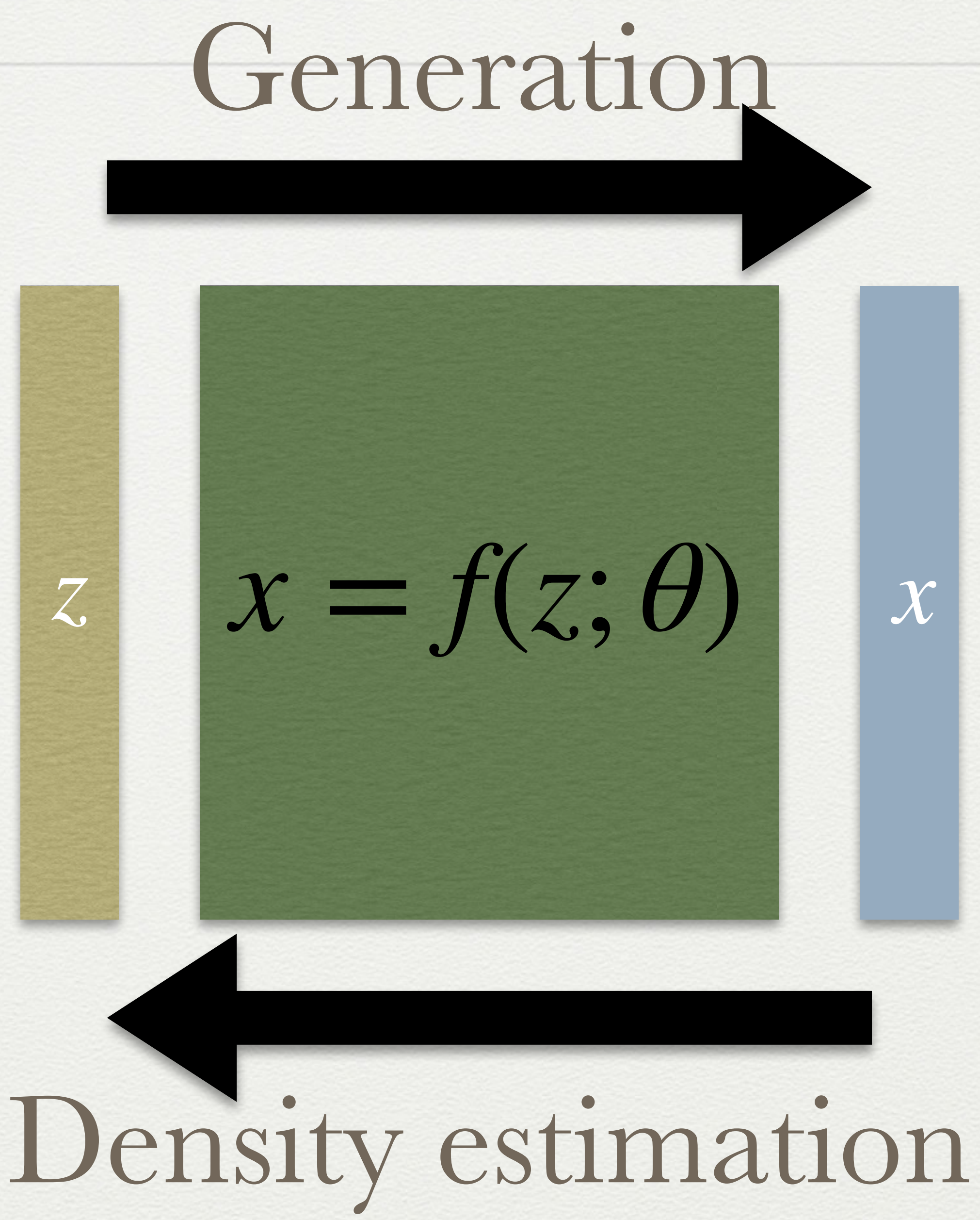
$$x \sim p_\theta(x)$$

Powerful class of density estimators that are also generative models

- Family of invertible maps parametrized by neural networks

Normalizing Flows

$$z \sim p_{base}(z) = \mathcal{N}(0,1)^d$$



$$x \sim p_\theta(x)$$

Powerful class of density estimators that are also generative models

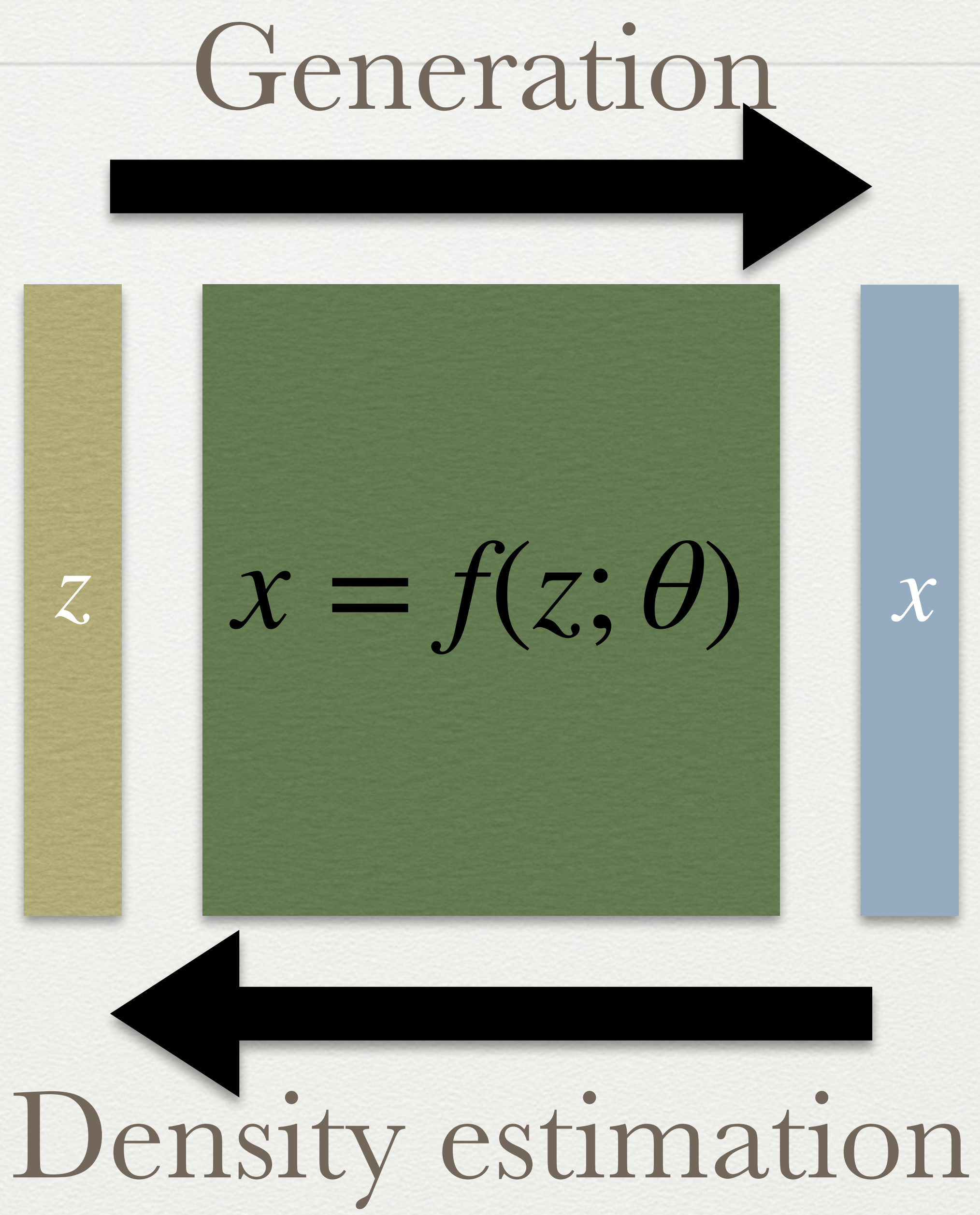
- Family of invertible maps parametrized by neural networks

$$p_\theta(x) = p_{base}(z = f_\theta(x)) \left| \frac{\partial z}{\partial x} \right|$$

(need tractable Jacobian!)

Normalizing Flows

$$z \sim p_{base}(z) = \mathcal{N}(0,1)^d$$



$$x \sim p_\theta(x)$$

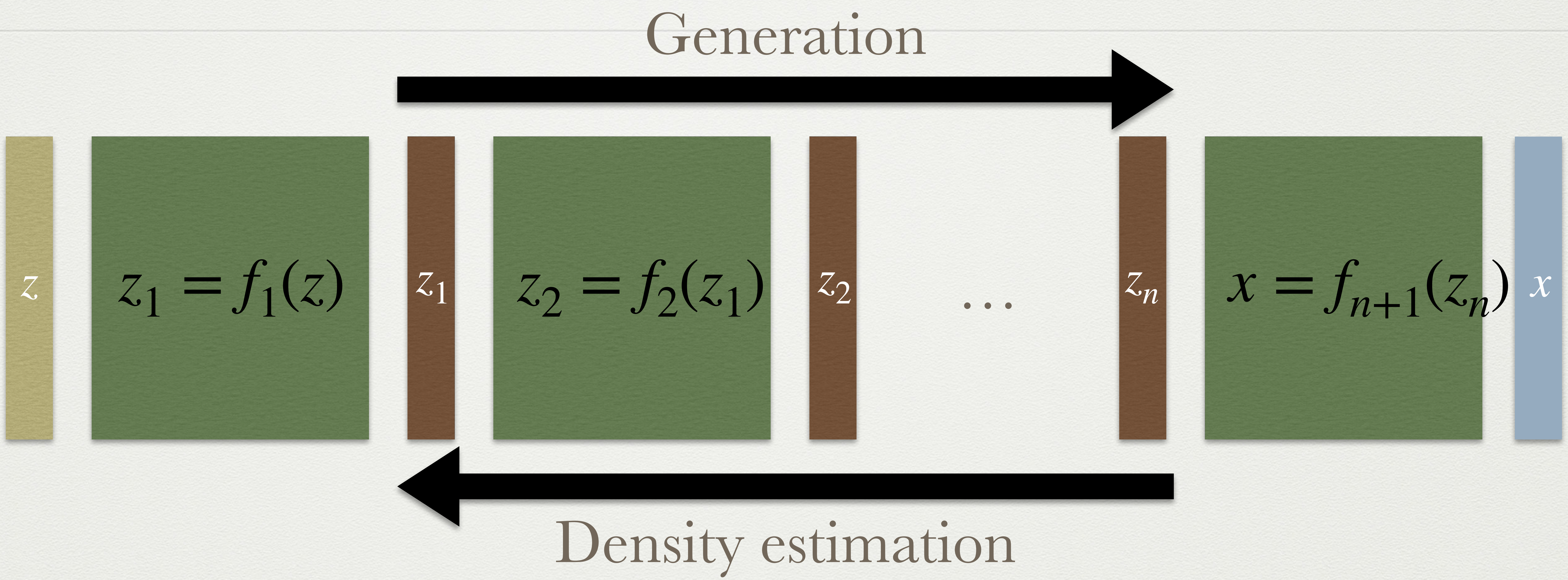
Powerful class of density estimators that are also generative models

- Family of invertible maps parametrized by neural networks $\quad p_\theta(x) = p_{base}(z = f_\theta(x))\left|\frac{\partial z}{\partial x}\right|$
- Train with maximum likelihood objective $\quad L = -\sum_{x_i \in data} \log p_\theta(x_i)$

(need tractable Jacobian!)

Normalizing Flows

$$z \sim p_{base}(z) = \mathcal{N}(0,1)^d$$



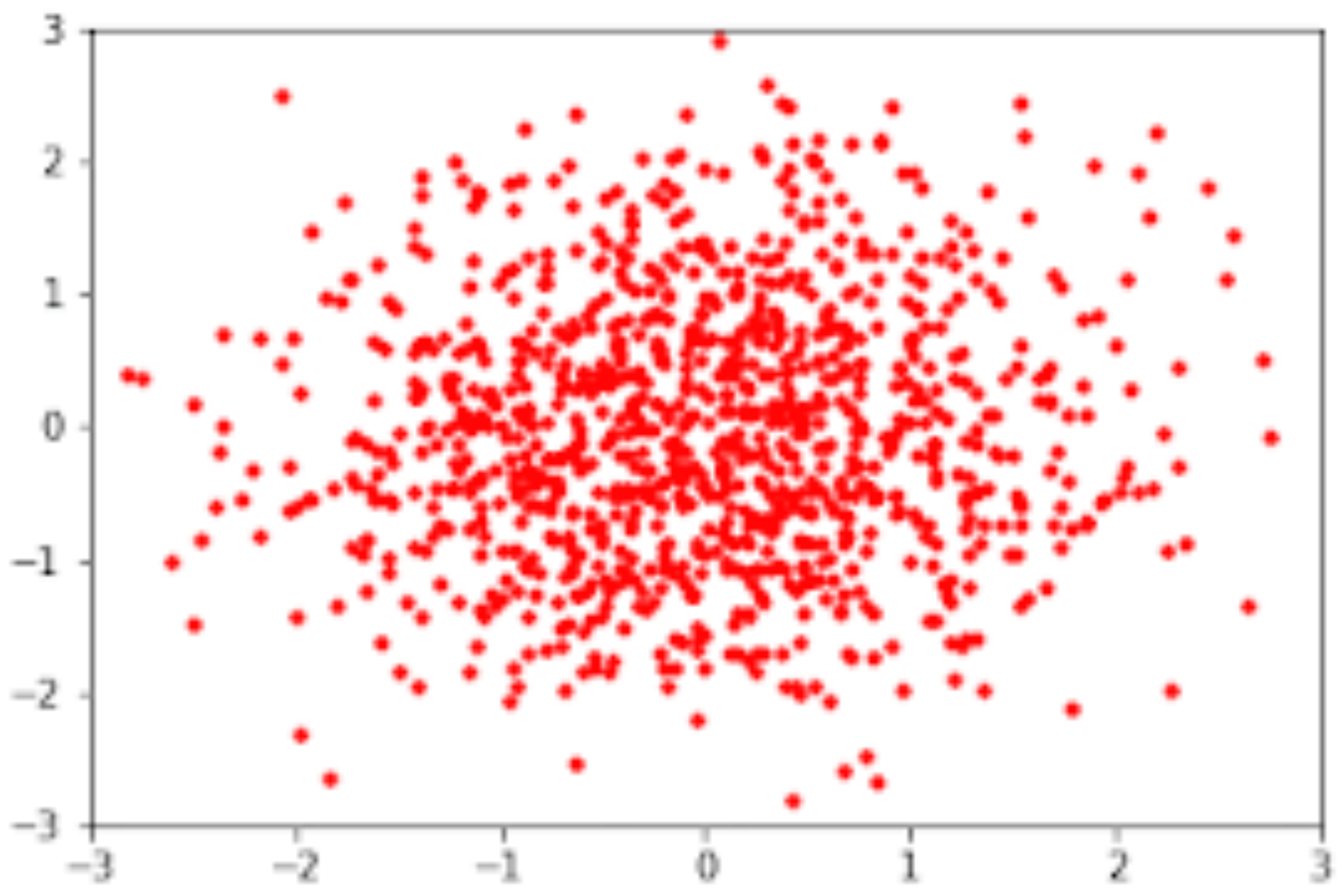
$$x \sim p(x)$$

Powerful class of density estimators that are also generative models

- Family of invertible maps parametrized by neural networks $\quad p_\theta(x) = p_{base}(z = f_\theta(x)) \left| \frac{\partial z}{\partial x} \right|$
- Train with maximum likelihood objective $\quad L = -\sum_{x_i \in data} \log p_\theta(x_i)$ **(need tractable Jacobian!)**
- **Compose multiple maps for greater expressivity**

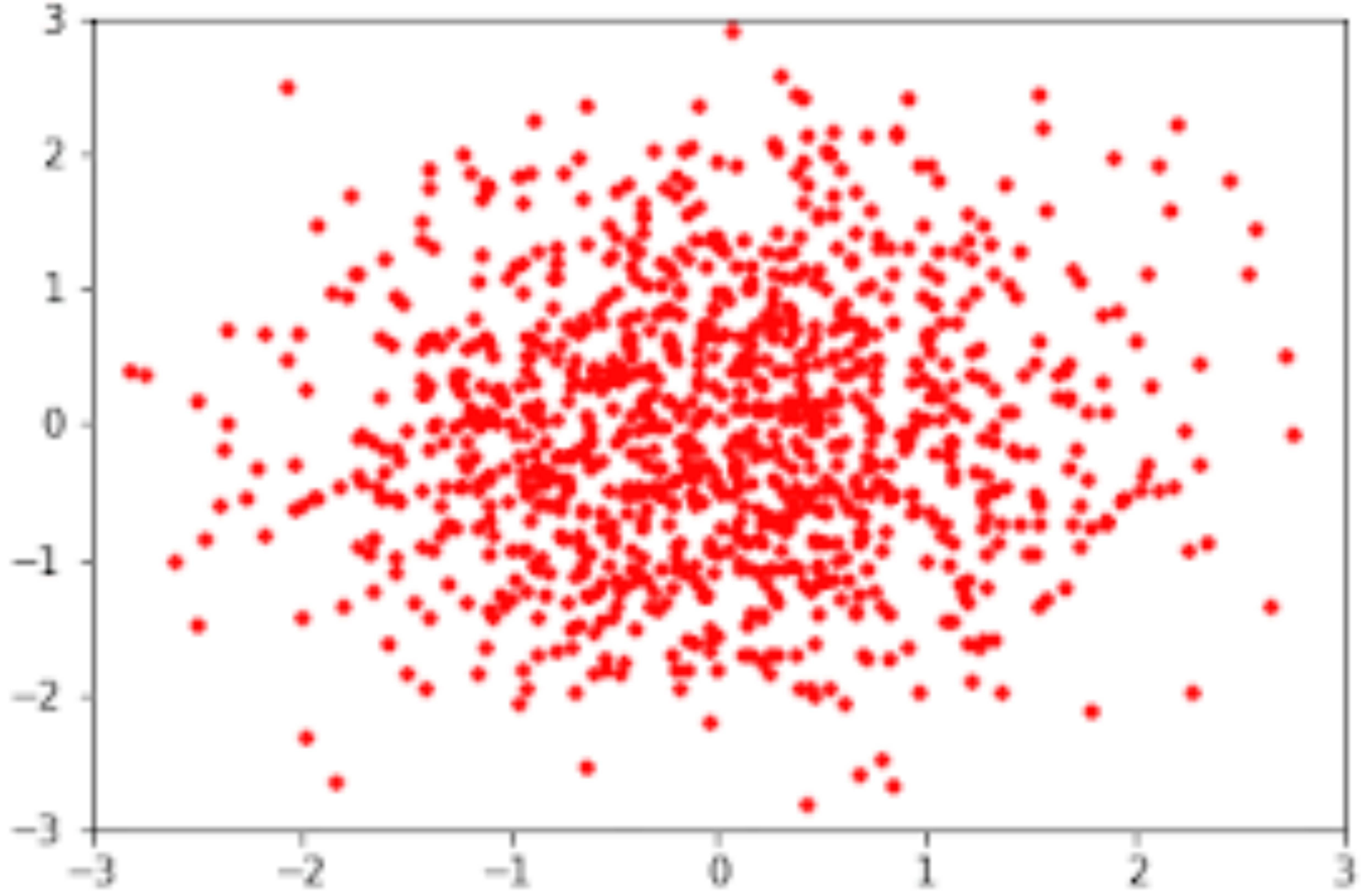
Example: Normalizing Flows

source: Eric Jang



Example: Normalizing Flows

source: Eric Jang



Our normalizing flows

- We used Masked Autoregressive Flows with affine transformations

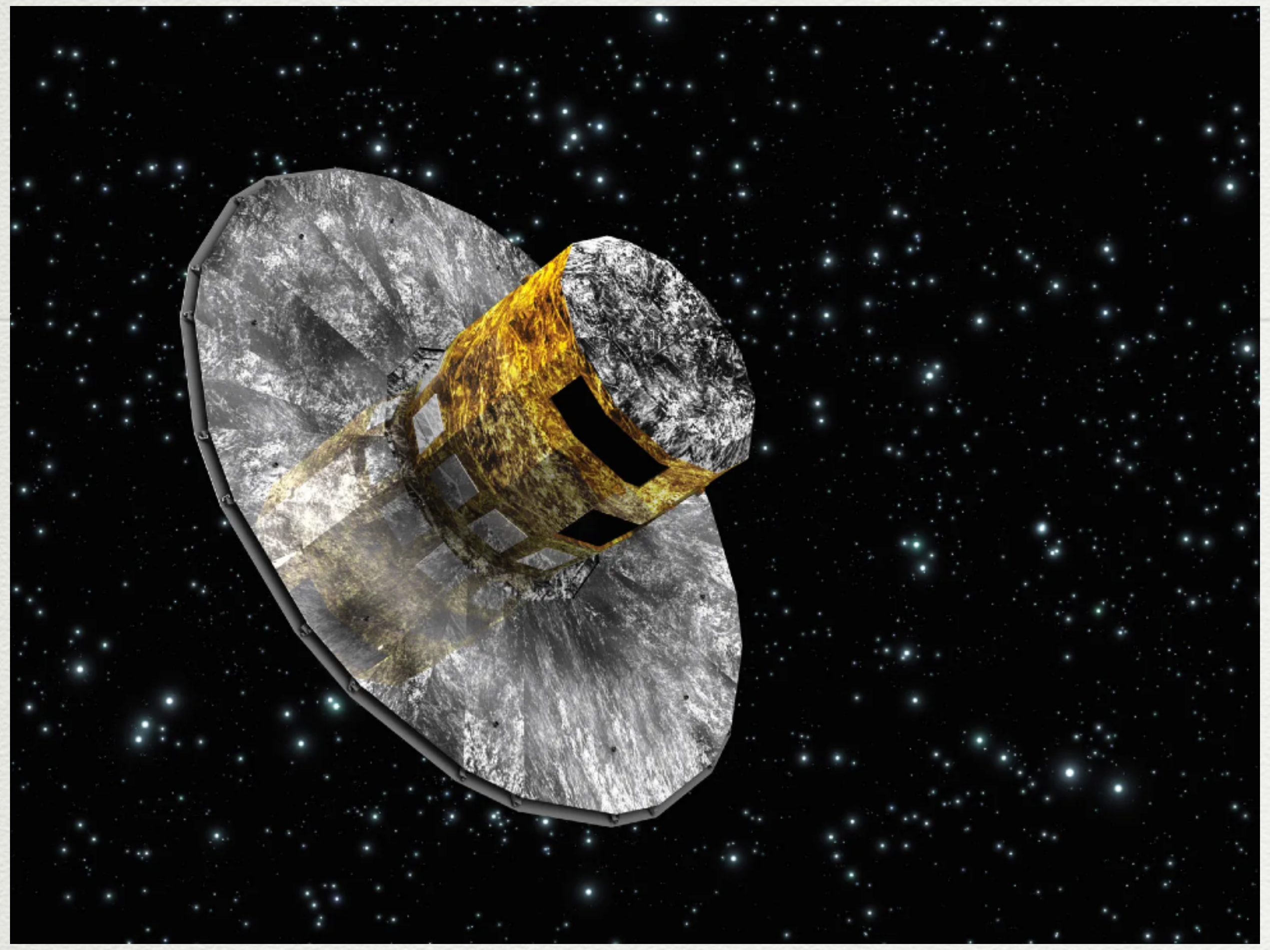
$$z_n = a_n(x_1, \ldots, x_{n-1})x_n + b_n(x_1, \ldots, x_{n-1})$$

MLPs

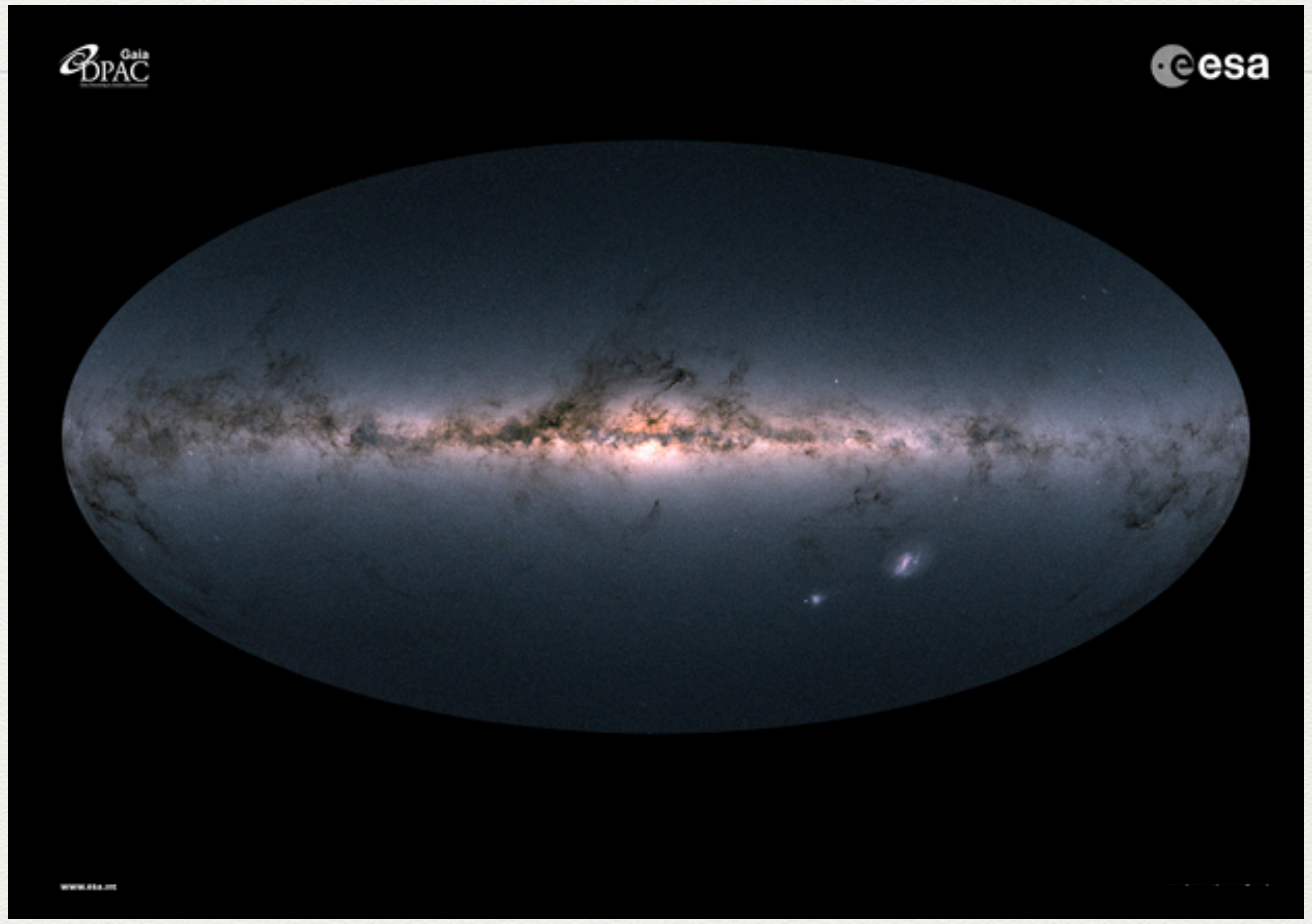
- Separate MAFs for $n(x)$ and $p(v\,|\,x)$ $[f(x, v) = p(v\,|\,x)n(x)]$

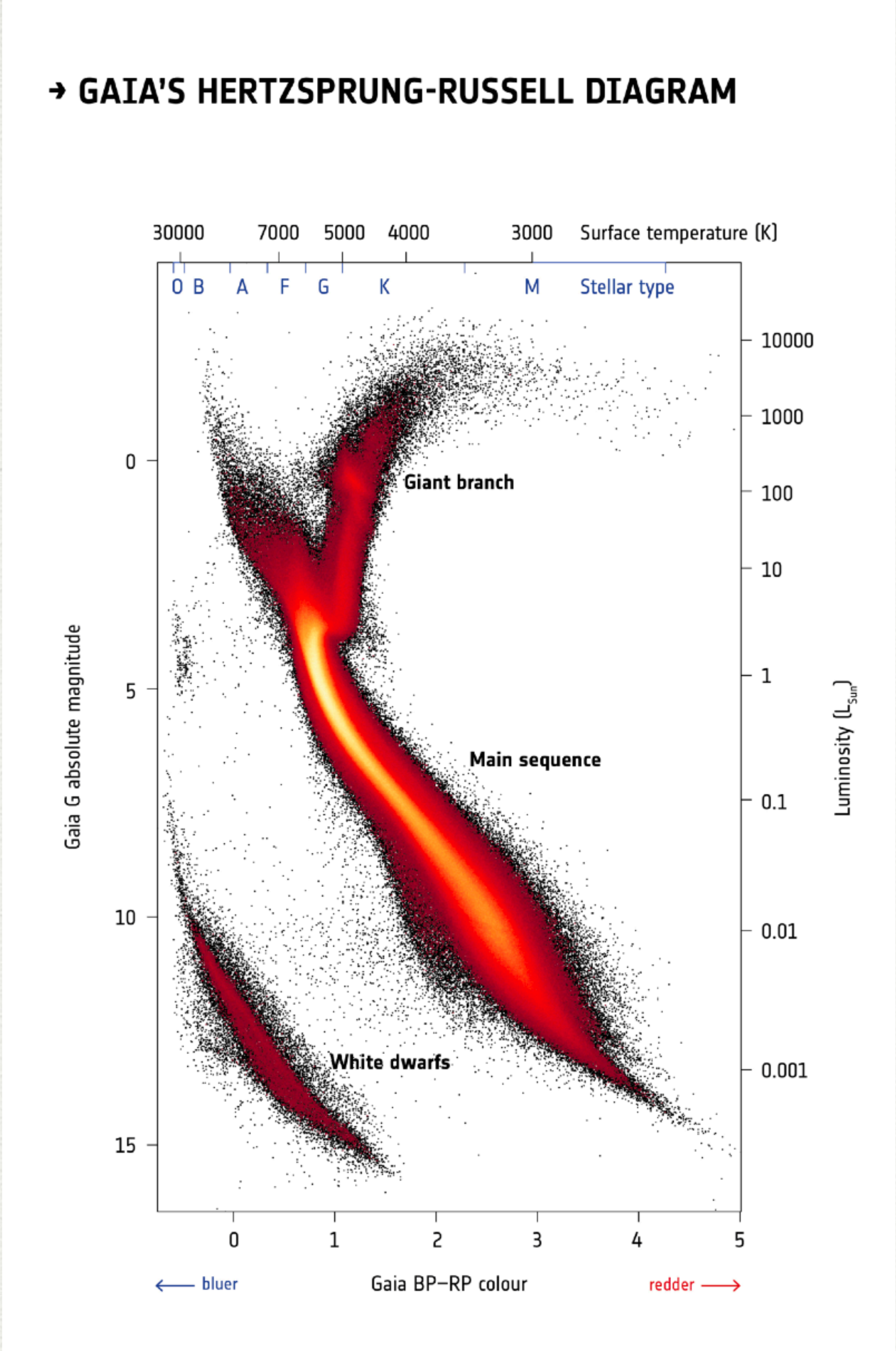
- Used smooth activations (GELU, GINT) instead of RELU for smooth derivatives

- Trained and ensemble-averaged 100 MAFs for more accurate PSDs



Gaia data

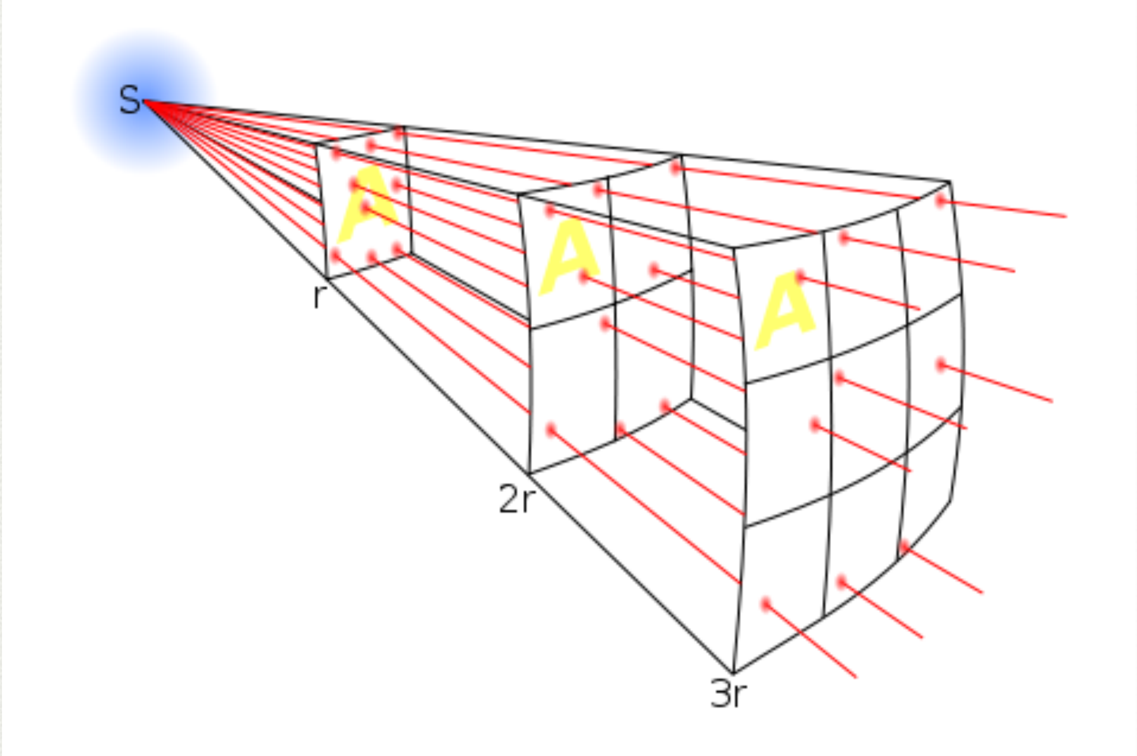




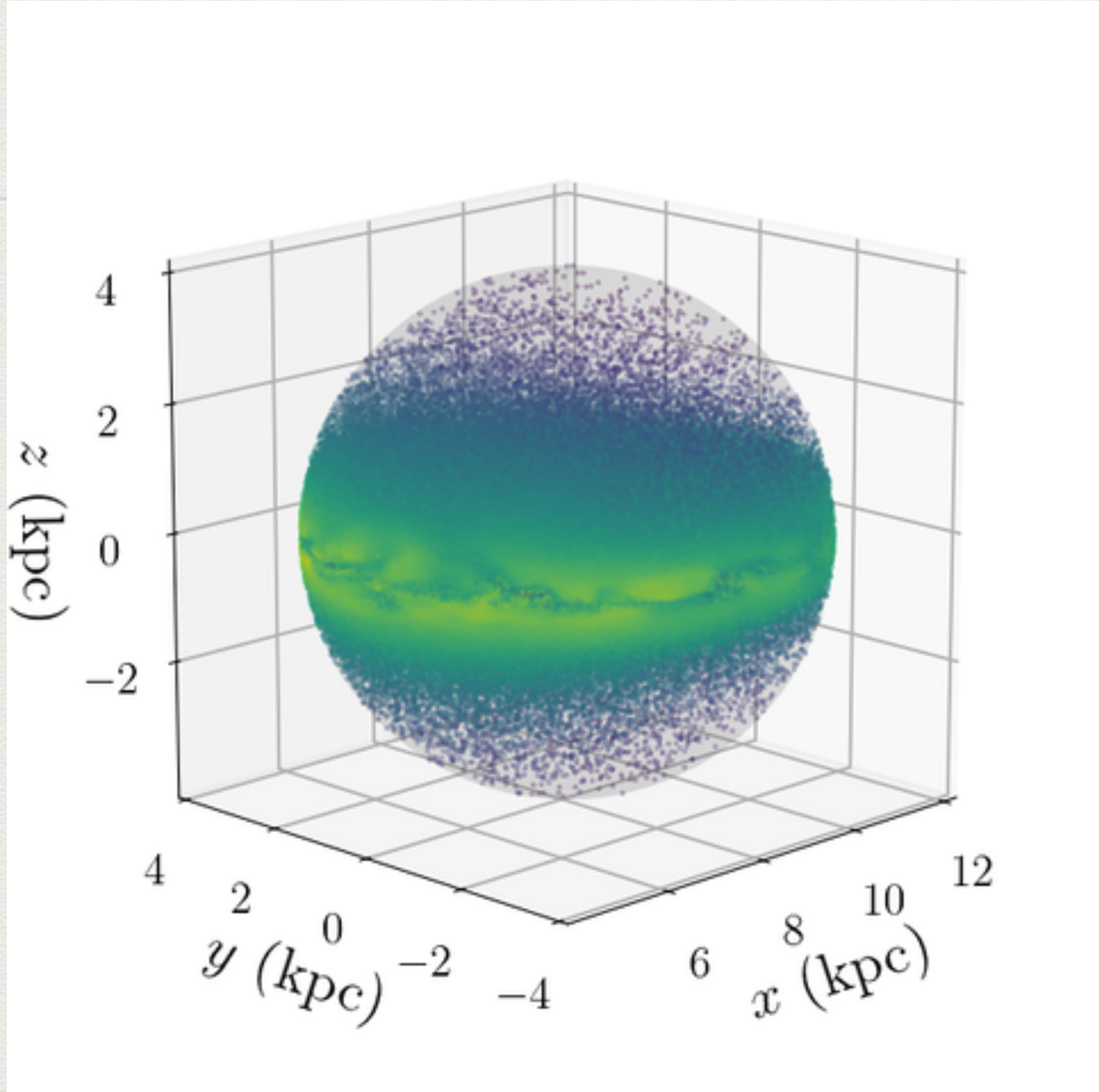
- Space telescope launched in 2013, currently on Data Release 3 (DR3)
- Angular positions, proper motions, color & magnitude of over **1 billion stars** in our Galaxy
- Distances and radial velocities for a smaller subset of nearby stars

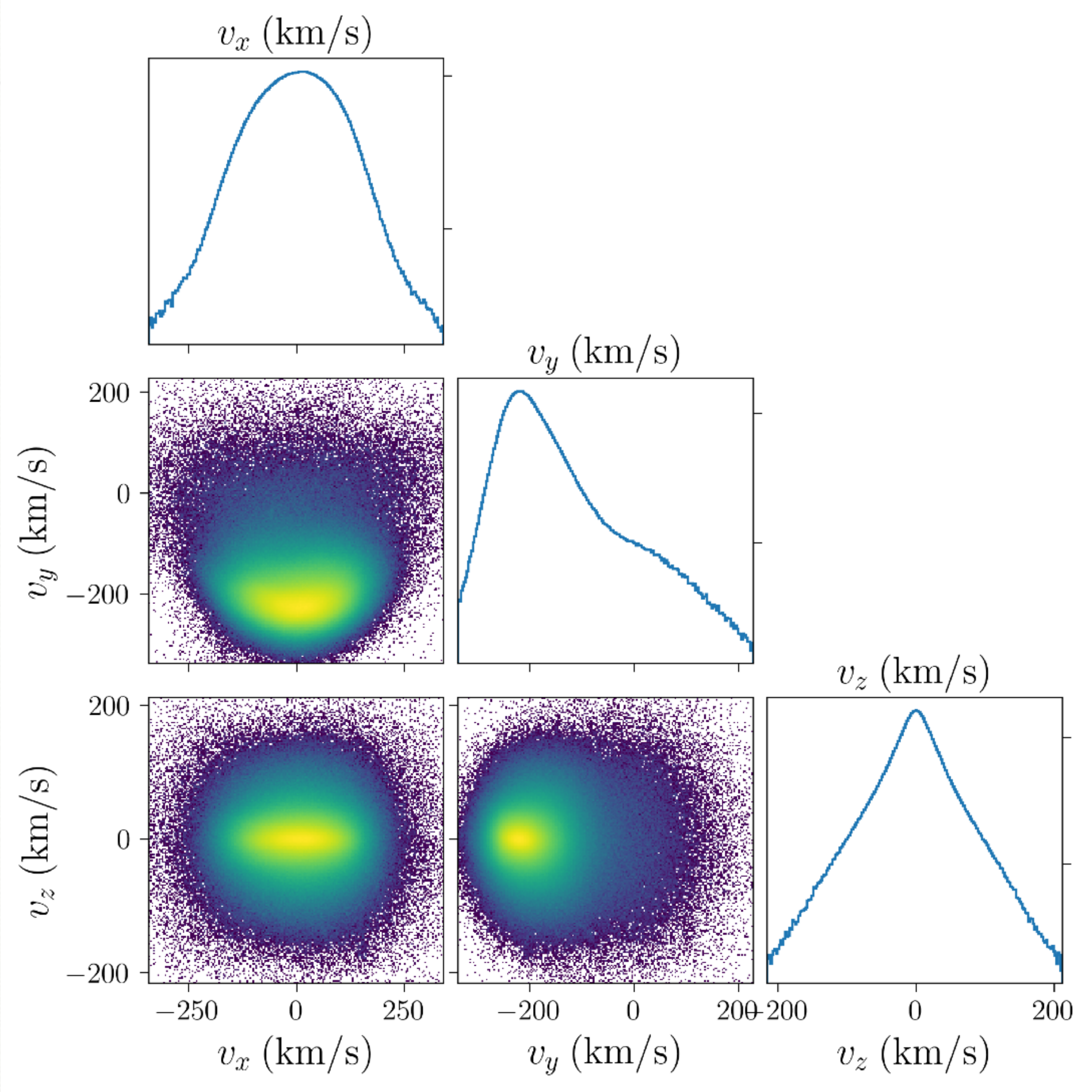
Gaia data

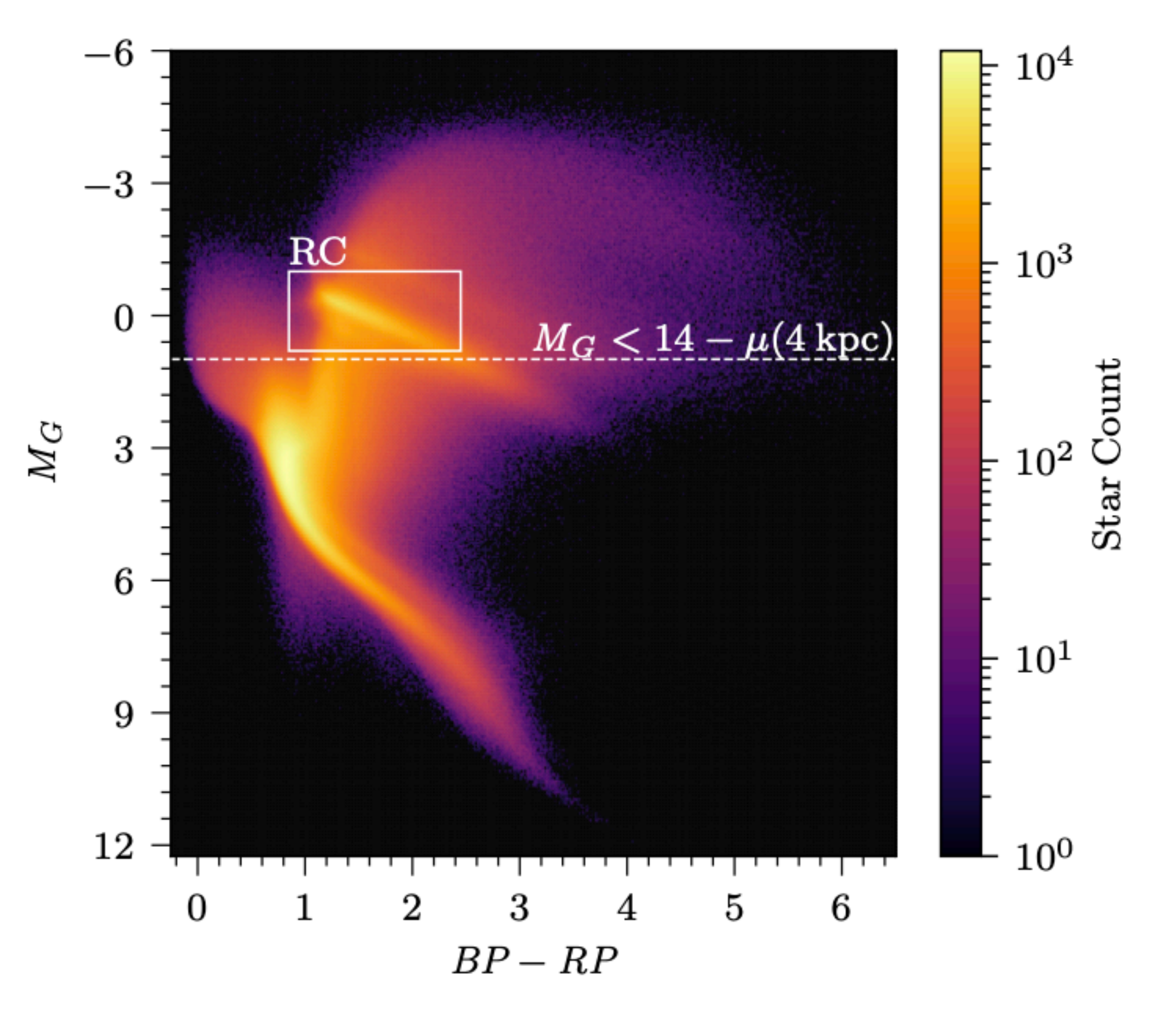
- ~35M stars in DR3 with 6d kinematics, located within ~4 kpc
- Want a **complete** and **unbiased** subset
- Gaia is complete up to observed magnitude $G_{obs} < 14$
- However, more distant stars are dimmer
- **For an unbiased subset, require all stars to be bright enough to have been observable anywhere within 4 kpc volume**



Gaia data



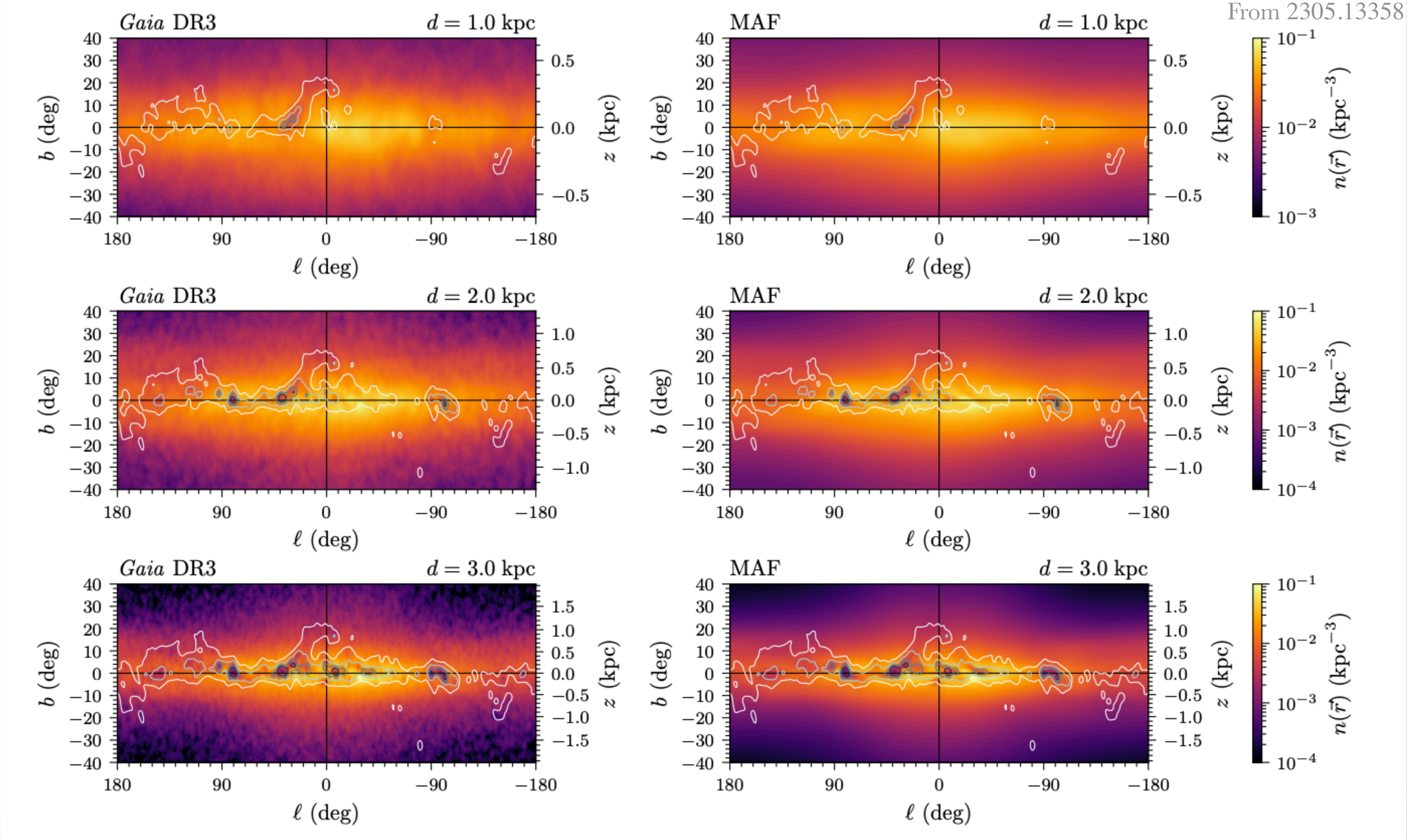




- **This results in 6M stars**
- Includes Red Clump stars (65%) — old, bright red giants believed to be in equilibrium, widely used already in Jeans analysis

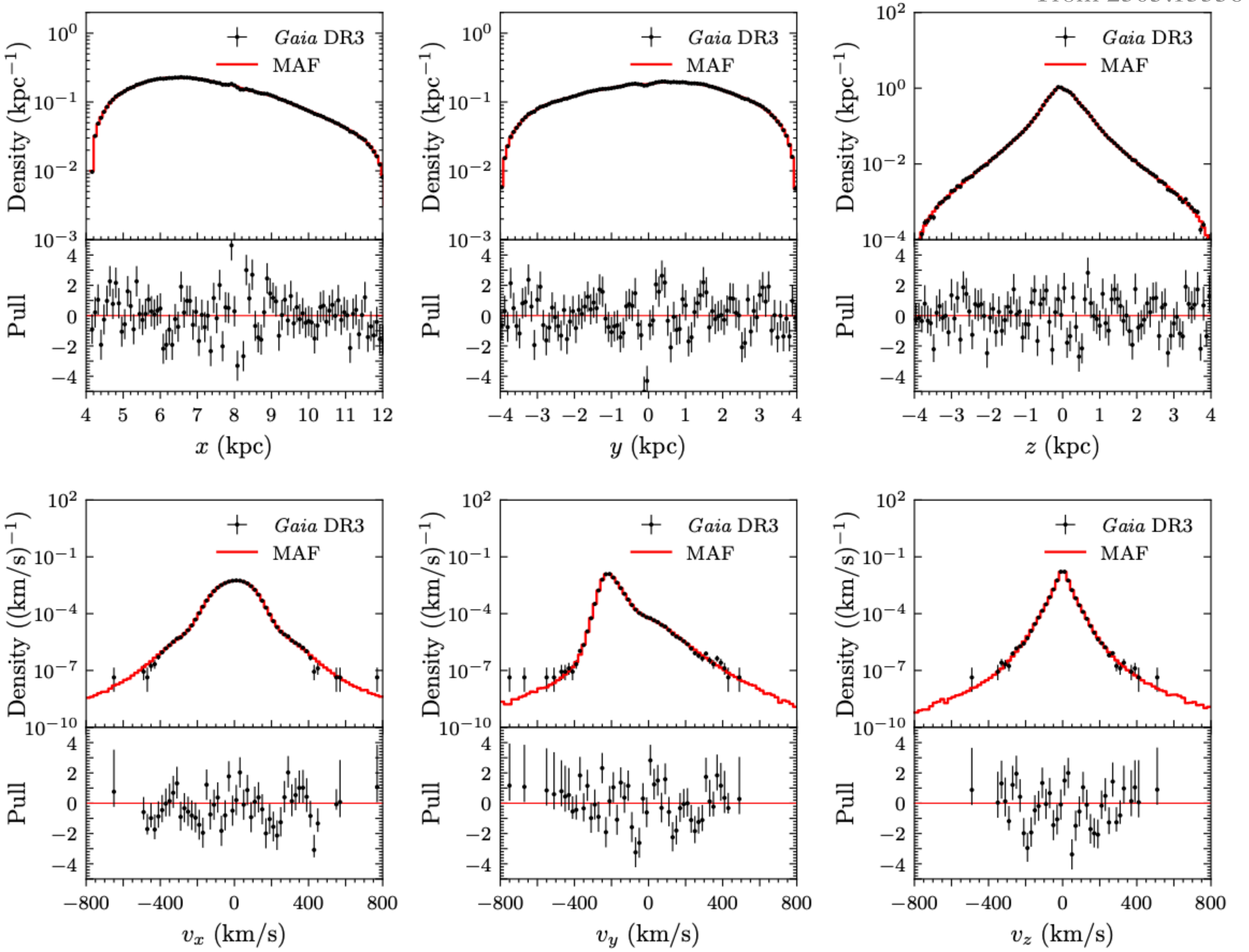
`Results: stellar PSD

From 2305.13358



Results: stellar PSD

From 2305.13358



Step 2: from PSD to Galactic mass density

Back to the CBE

$$\boxed{\vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} = 0}$$

- Just one scalar equation for three acceleration components.
Underdetermined system?

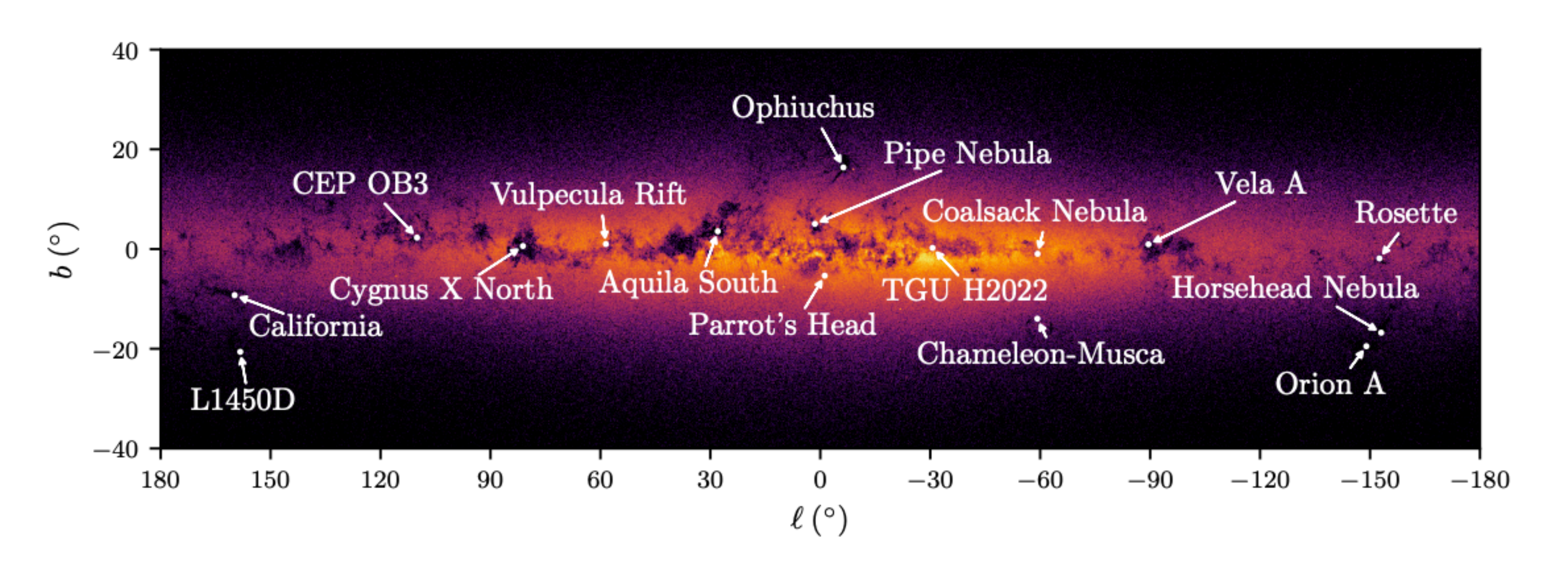
- But acceleration only depends on position! So actually many equations (one for each velocity) at the same position.
Highly overdetermined system!

- Many ways to solve. Our approach: minimize MSE weighted by phase space density

$$\sum_{\vec{v}\sim f(\vec{v}|\vec{x})} \left| \vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} \right|^2$$

Problem: interstellar dust

- But wait! What are all those blotches in the number density?
- Interstellar dust!
- Blocks many lines of sight
- Stars are reddened and dimmed — can fall out of the Gaia dataset



- Affects the completeness of the dataset, biases the phase space density

First version:

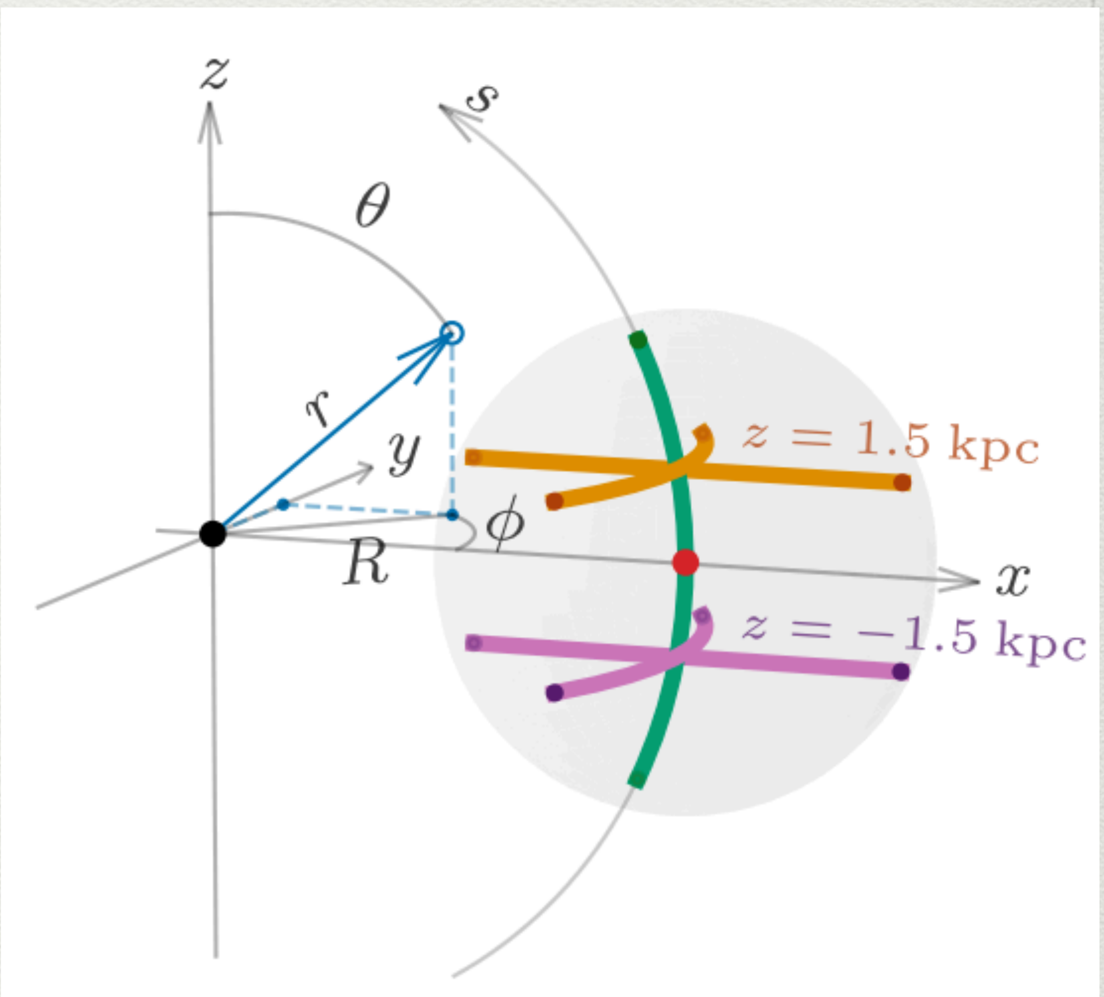
Mapping Dark Matter in the Milky Way using Normalizing Flows and Gaia DR3

Sung Hak Lim, Eric Putney, Matthew R. Buckley, David Shih (May 22, 2023)

e-Print: 2305.13358 [astro-ph.GA]

- In previous work, we chose specific lines of sight that are not blocked by dust

From 2305.13358



- Solved MSE for accelerations pointwise at each position

$$\sum_{\vec{v}\sim f(\vec{v}|\vec{x})} \left| \vec{v} \cdot \frac{\partial f}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f}{\partial \vec{v}} \right|^2$$

- Take another derivative (using kernel trick) to obtain $\rho(x)$

New version:

Mapping Dark Matter Through the Dust of the Milky Way Part I: Dust Correction and Phase Space Density

Eric Putney, David Shih, Sung Hak Lim, Matthew R. Buckley (Dec 18, 2024)

e-Print: 2412.14236 [astro-ph.GA]

- New idea: can simultaneously determine Galactic potential and "dust efficiency" function

$$f_{obs}(x, v) = \epsilon(x)\, f_{corr}(x, v)$$

- **Key point: dust efficiency doesn't depend on star's velocity, just its position!**

- Dust-corrected PSD should satisfy equilibrium CBE

$$\vec{v} \cdot \frac{\partial f_{corr}}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f_{corr}}{\partial \vec{v}} = 0$$

New version:

Mapping Dark Matter Through the Dust of the Milky Way Part I: Dust Correction and Phase Space Density

Eric Putney, David Shih, Sung Hak Lim, Matthew R. Buckley (Dec 18, 2024)

e-Print: 2412.14236 [astro-ph.GA]

- New idea: can simultaneously determine Galactic potential and "dust efficiency" function

$$f_{obs}(x, v) = \epsilon(x)\, f_{corr}(x, v)$$

- **Key point: dust efficiency doesn't depend on star's velocity, just its position!**

- Dust-corrected PSD should satisfy equilibrium CBE

$$\vec{v} \cdot \frac{\partial f_{corr}}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f_{corr}}{\partial \vec{v}} = 0 \qquad \Rightarrow$$

New version:

Mapping Dark Matter Through the Dust of the Milky Way Part I: Dust Correction and Phase Space Density

Eric Putney, David Shih, Sung Hak Lim, Matthew R. Buckley (Dec 18, 2024)

e-Print: 2412.14236 [astro-ph.GA]

- New idea: can simultaneously determine Galactic potential and "dust efficiency" function

$$f_{obs}(x, v) = \epsilon(x)\, f_{corr}(x, v)$$

- **Key point: dust efficiency doesn't depend on star's velocity, just its position!**

- Dust-corrected PSD should satisfy equilibrium CBE

$$\vec{v} \cdot \frac{\partial f_{corr}}{\partial \vec{x}} + \vec{a}(x) \cdot \frac{\partial f_{corr}}{\partial \vec{v}} = 0 \quad \Rightarrow \quad \boxed{\vec{v} \cdot \frac{\partial \log f_{obs}}{\partial \vec{x}} - \vec{v} \cdot \nabla \log \epsilon - \nabla \Phi \cdot \frac{\partial \log f_{obs}}{\partial \vec{v}} = 0}$$

New version:

Mapping Dark Matter Through the Dust of the Milky Way Part I: Dust Correction and Phase Space Density

Eric Putney, David Shih, Sung Hak Lim, Matthew R. Buckley (Dec 18, 2024)

e-Print: 2412.14236 [astro-ph.GA]

$$\boxed{\vec{v} \cdot \frac{\partial \log f_{obs}}{\partial \vec{x}} - \vec{v} \cdot \nabla \log \epsilon - \nabla \Phi \cdot \frac{\partial \log f_{obs}}{\partial \vec{v}} = 0}$$

- Different velocity dependence — still overdetermined system!
- Can simultaneously determine $\Phi(x)$ and $\epsilon(x)$

New version:

Mapping Dark Matter Through the Dust of the Milky Way Part I: Dust Correction and Phase Space Density

Eric Putney, David Shih, Sung Hak Lim, Matthew R. Buckley (Dec 18, 2024)

e-Print: 2412.14236 [astro-ph.GA]

$$\boxed{\vec{v} \cdot \frac{\partial \log f_{obs}}{\partial \vec{x}} - \vec{v} \cdot \nabla \log \epsilon - \nabla \Phi \cdot \frac{\partial \log f_{obs}}{\partial \vec{v}} = 0}$$

- Different velocity dependence — still overdetermined system!
- Can simultaneously determine $\Phi(x)$ and $\epsilon(x)$

$$L(\theta, \varphi) = \sum_{(x,v)\sim f(x,v)} \left| \vec{v} \cdot \frac{\partial \log f_{obs}}{\partial \vec{x}} - \vec{v} \cdot \nabla \log \epsilon_\theta(x) - \nabla \Phi_\varphi(x) \cdot \frac{\partial \log f_{obs}}{\partial \vec{v}} \right|^2$$

New version:

Mapping Dark Matter Through the Dust of the Milky Way Part I: Dust Correction and Phase Space Density

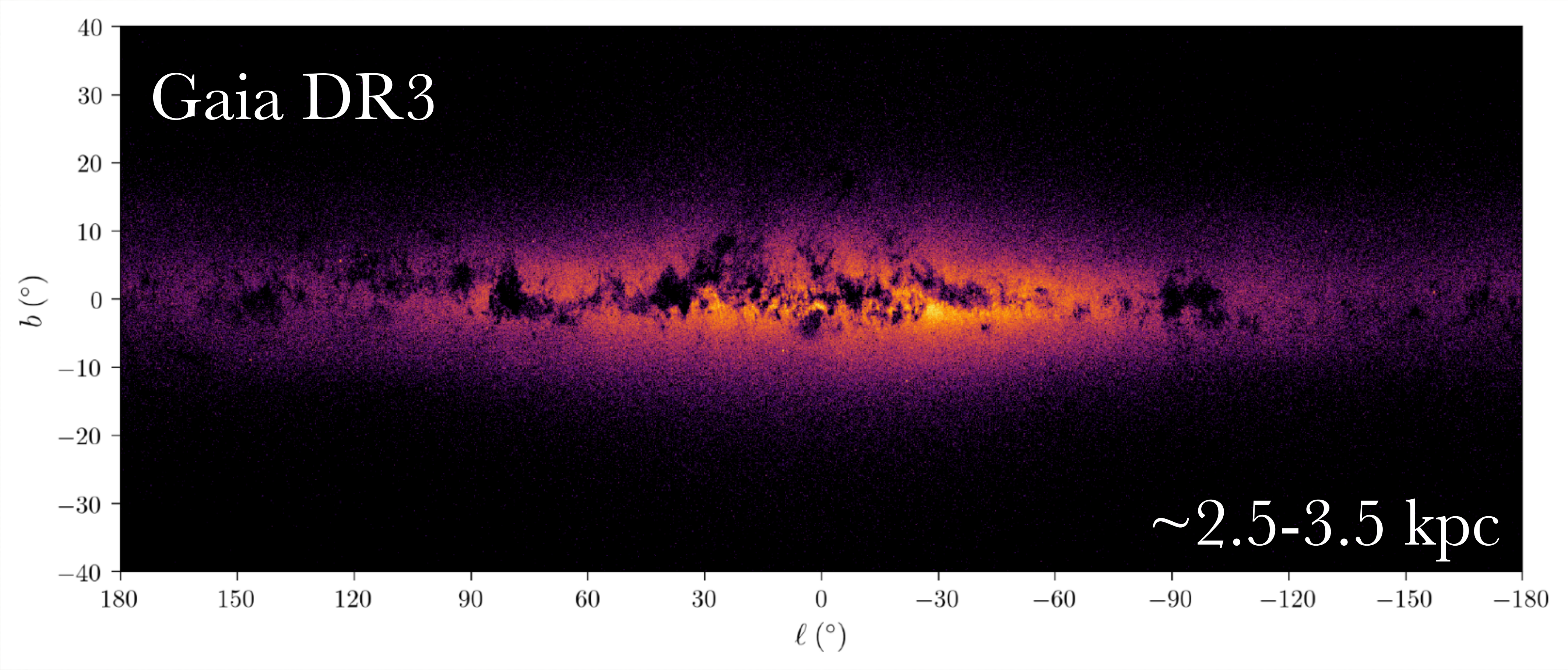
Eric Putney, David Shih, Sung Hak Lim, Matthew R. Buckley (Dec 18, 2024)

e-Print: 2412.14236 [astro-ph.GA]

$$L(\theta, \varphi) = \sum_{(x,v)\sim f(x,v)} \left| \vec{v} \cdot \frac{\partial \log f_{obs}}{\partial \vec{x}} - \vec{v} \cdot \nabla \log \epsilon_\theta(x) - \nabla \Phi_\varphi(x) \cdot \frac{\partial \log f_{obs}}{\partial \vec{v}} \right|^2$$

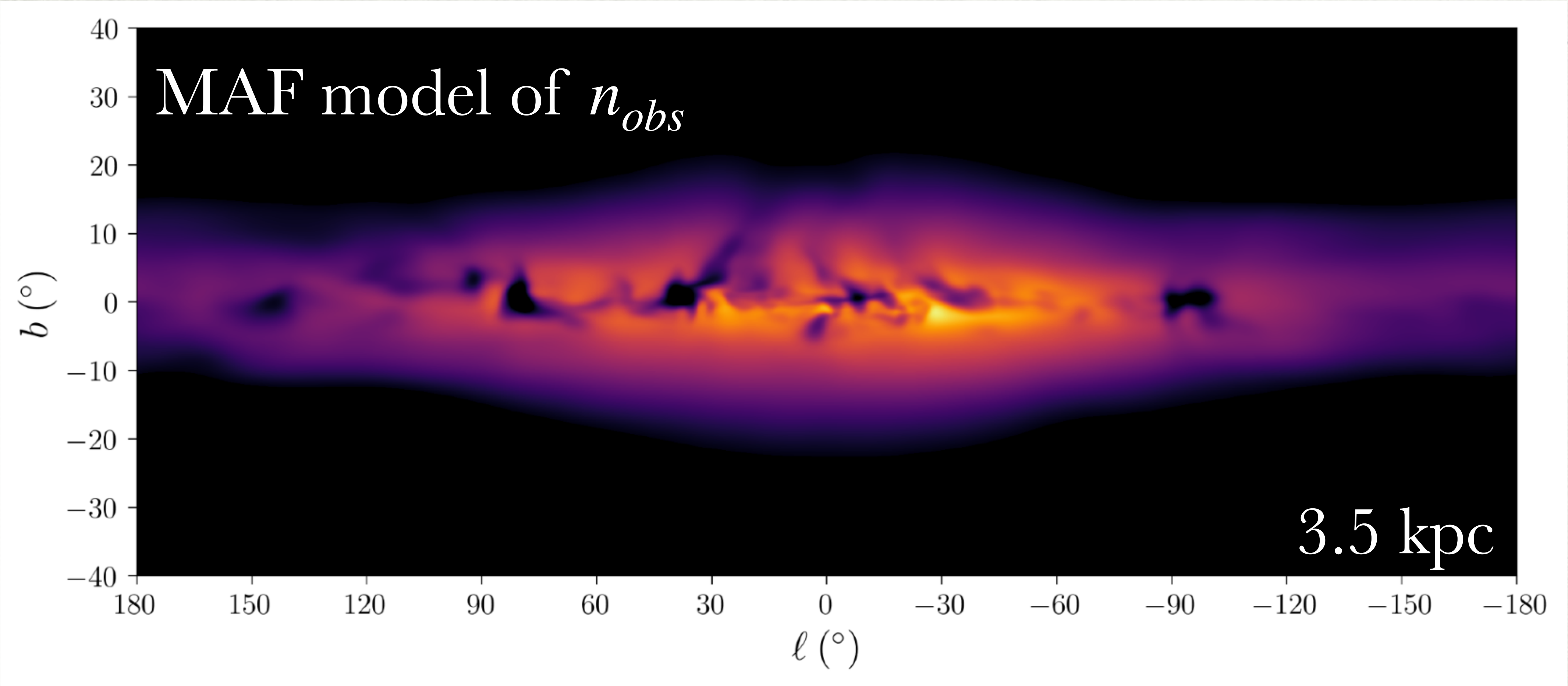
- Parametrize both $\epsilon(x)$ and $\Phi(x)$ with NNs (simple MLPs)
- Advantages
  - Much more computationally efficient compared to previous pointwise approach
  - More physical — guarantees curl-free accelerations

Results: Dust-Corrected PSD



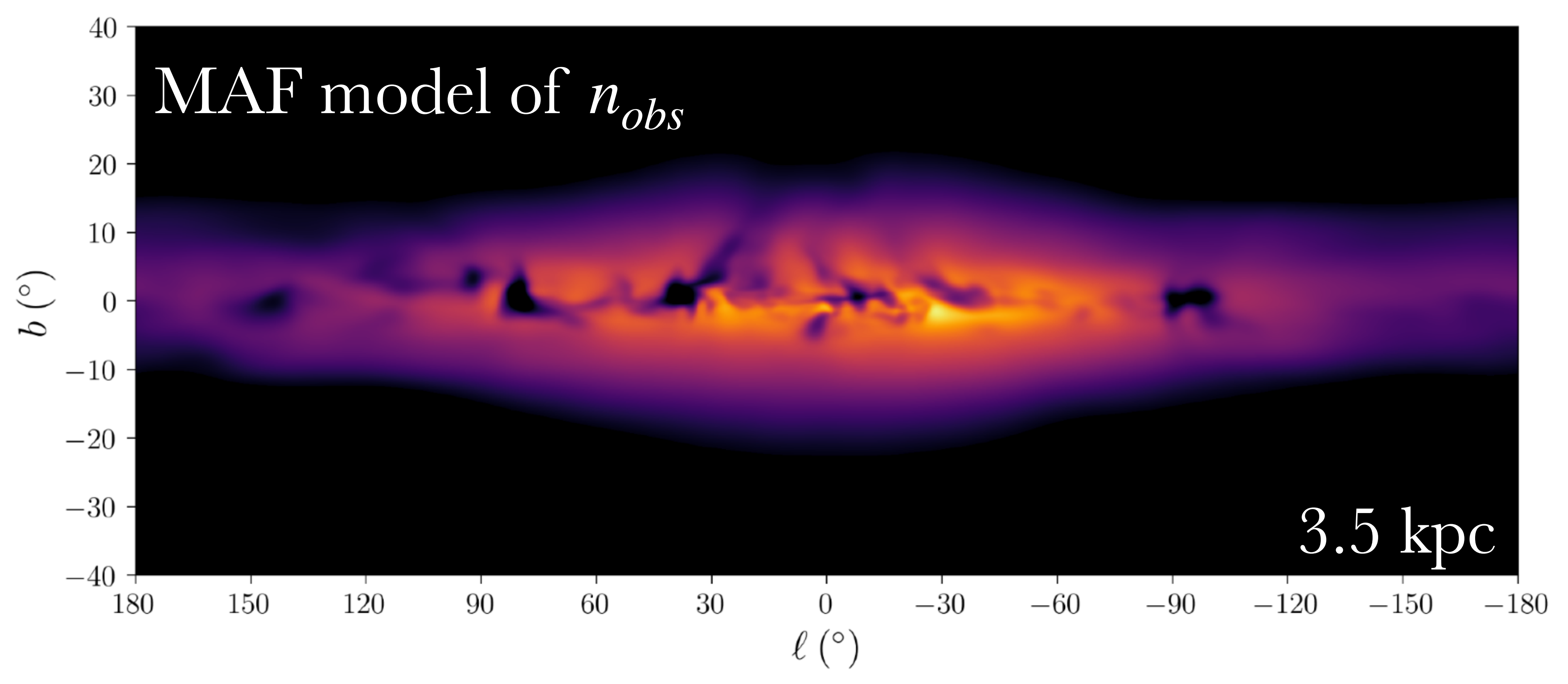
Observed data is biased by dust

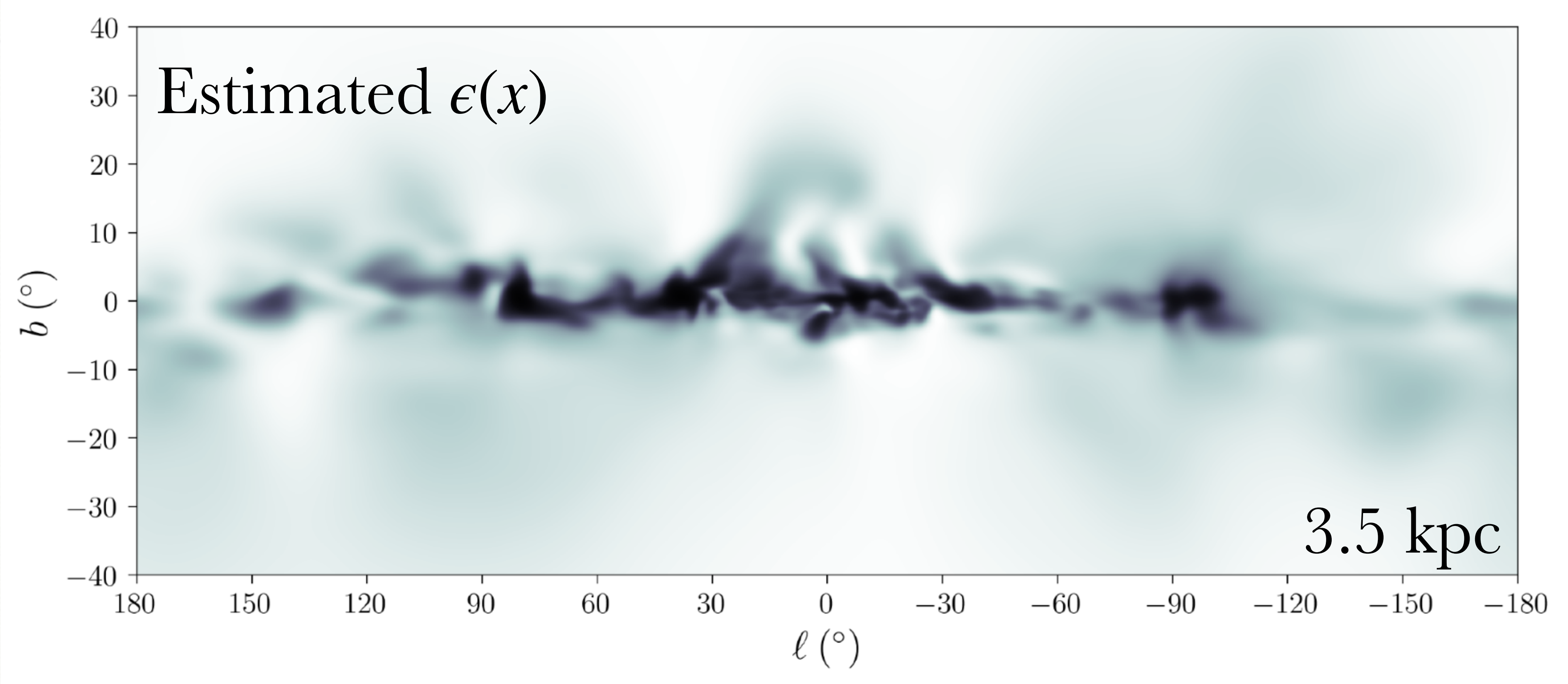
Results: Dust-Corrected PSD



Learned PSD is biased by dust

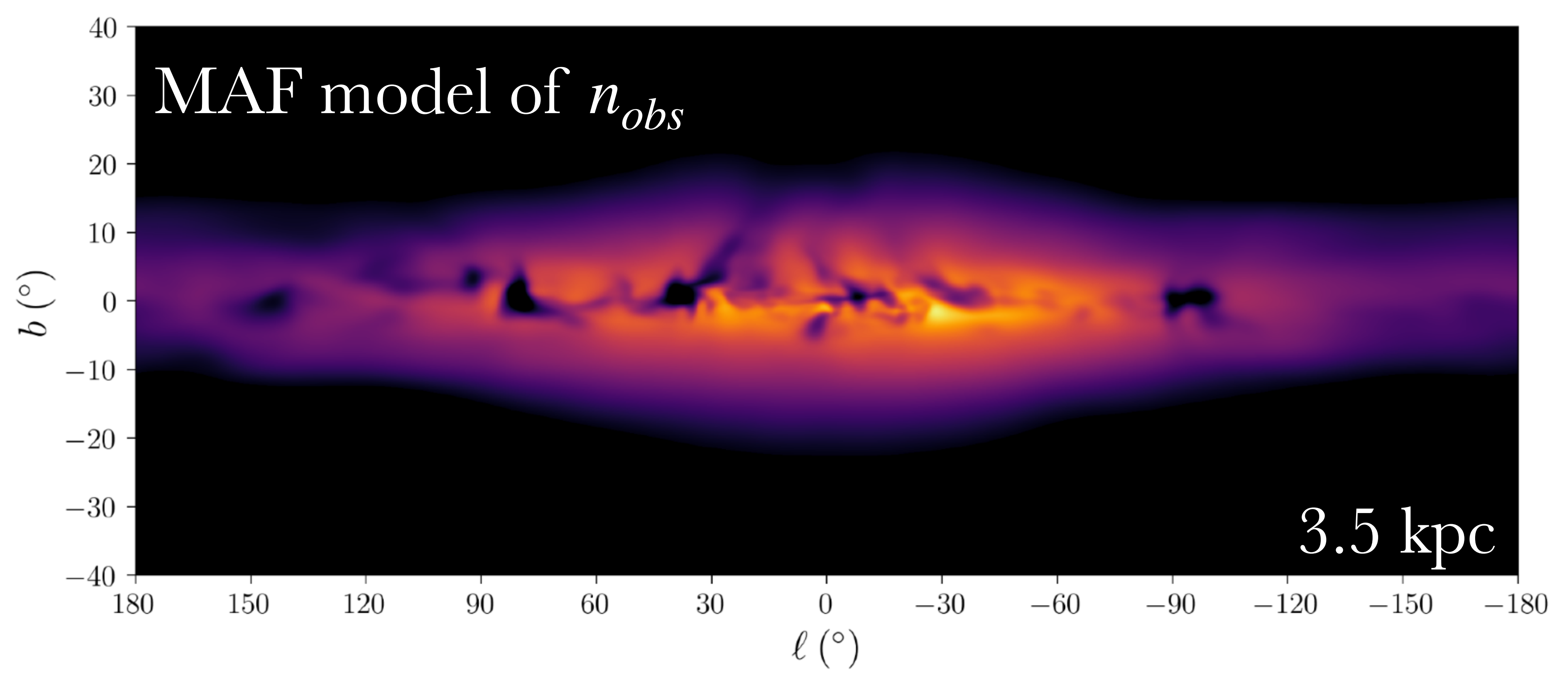
Results: Dust-Corrected PSD

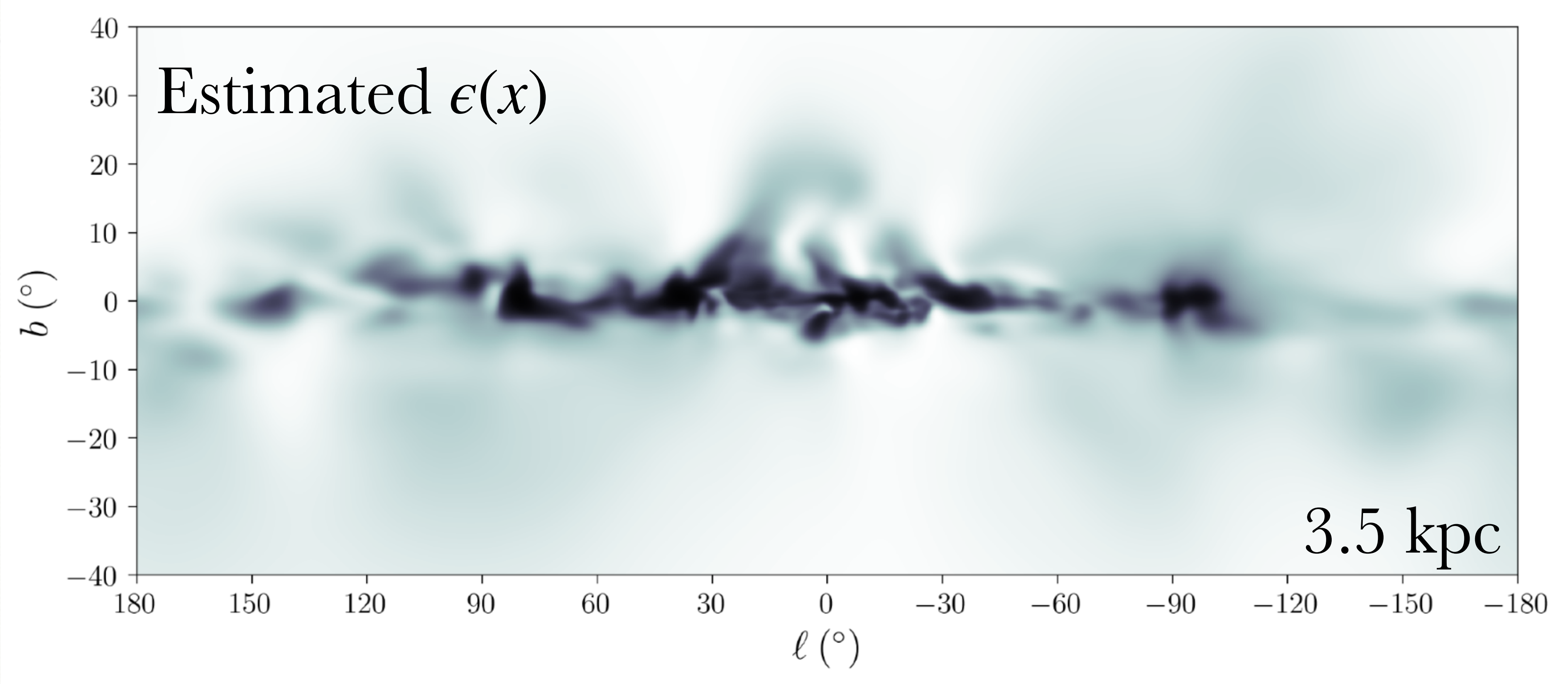




Estimate $\epsilon(x)$ from CBE

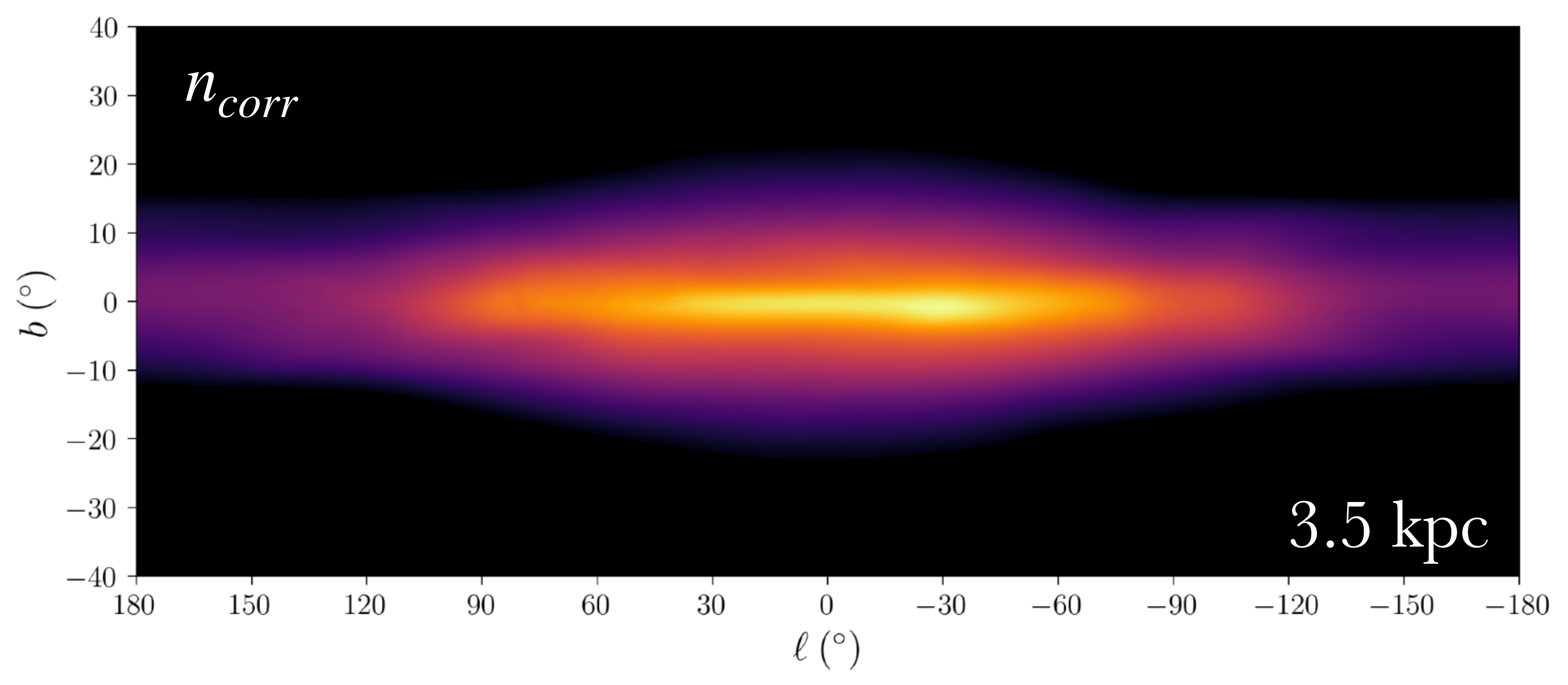
Results: Dust-Corrected PSD



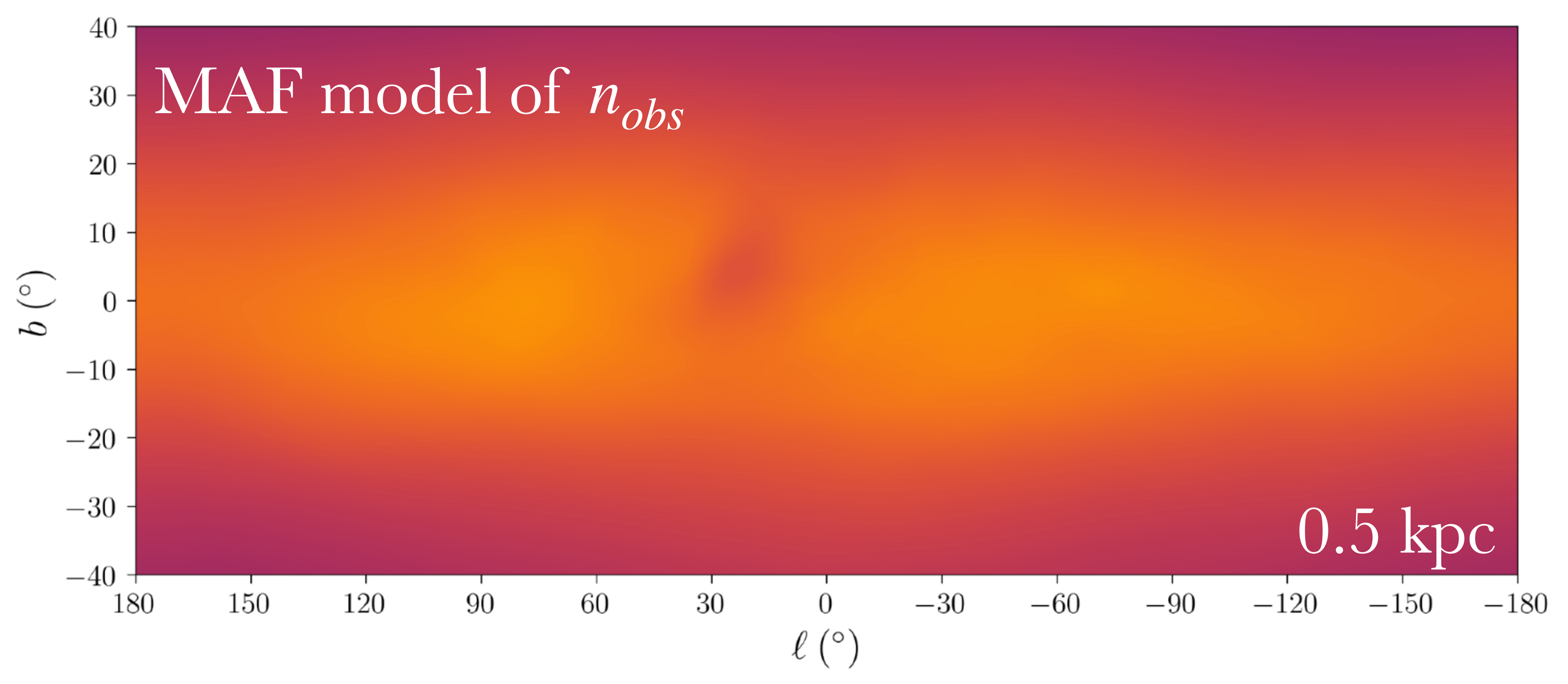


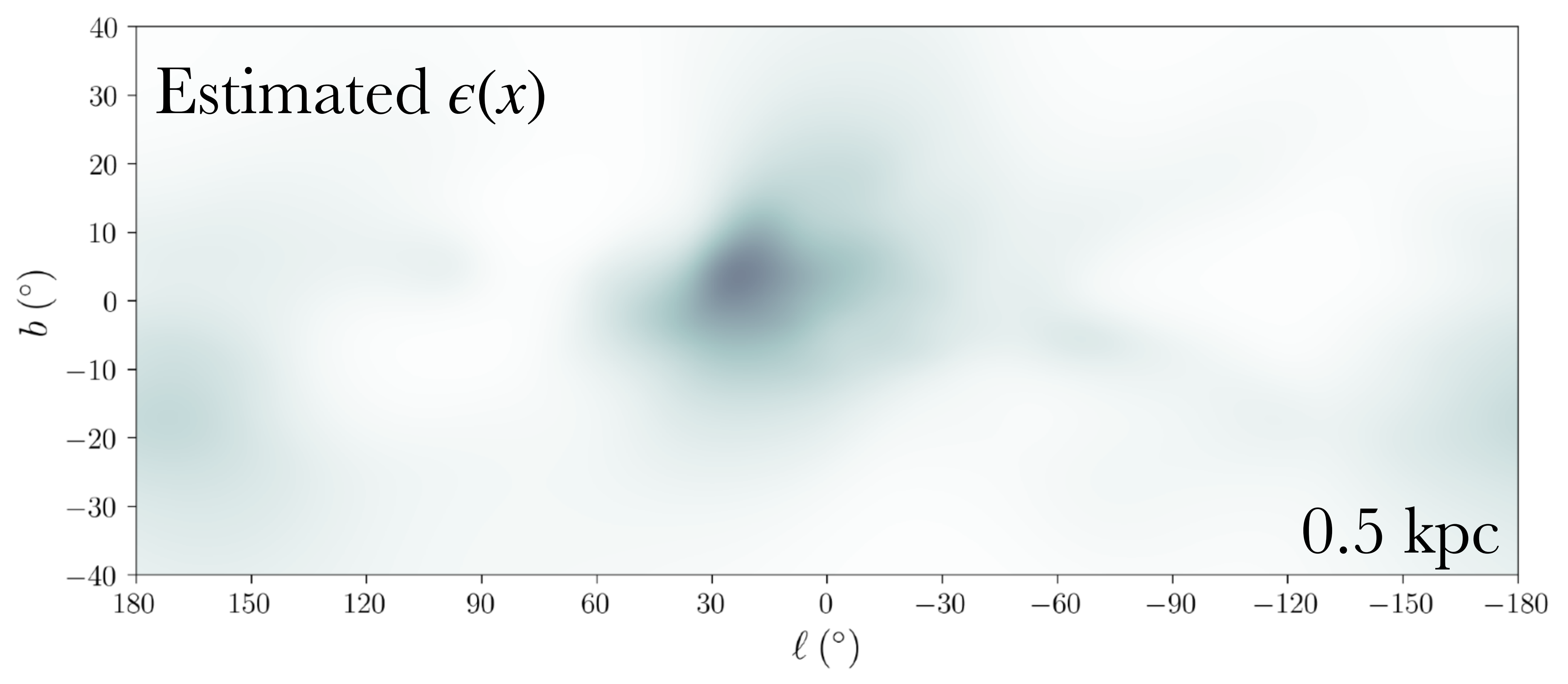
Corrected density

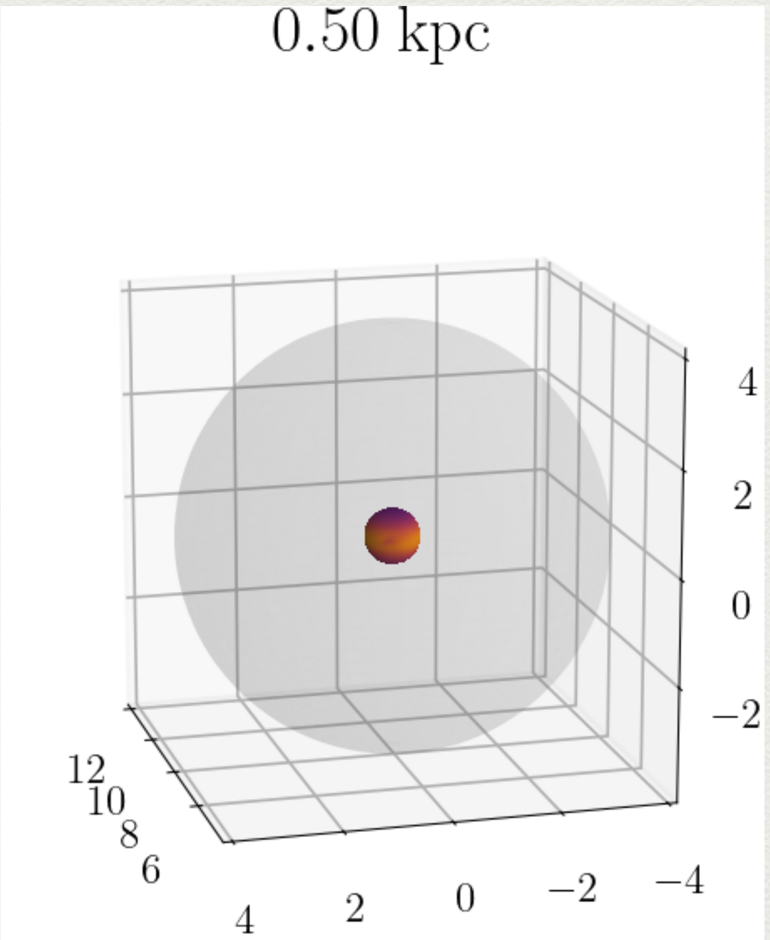
$n_{corr}(x) = n_{obs}(x)/\epsilon(x)$

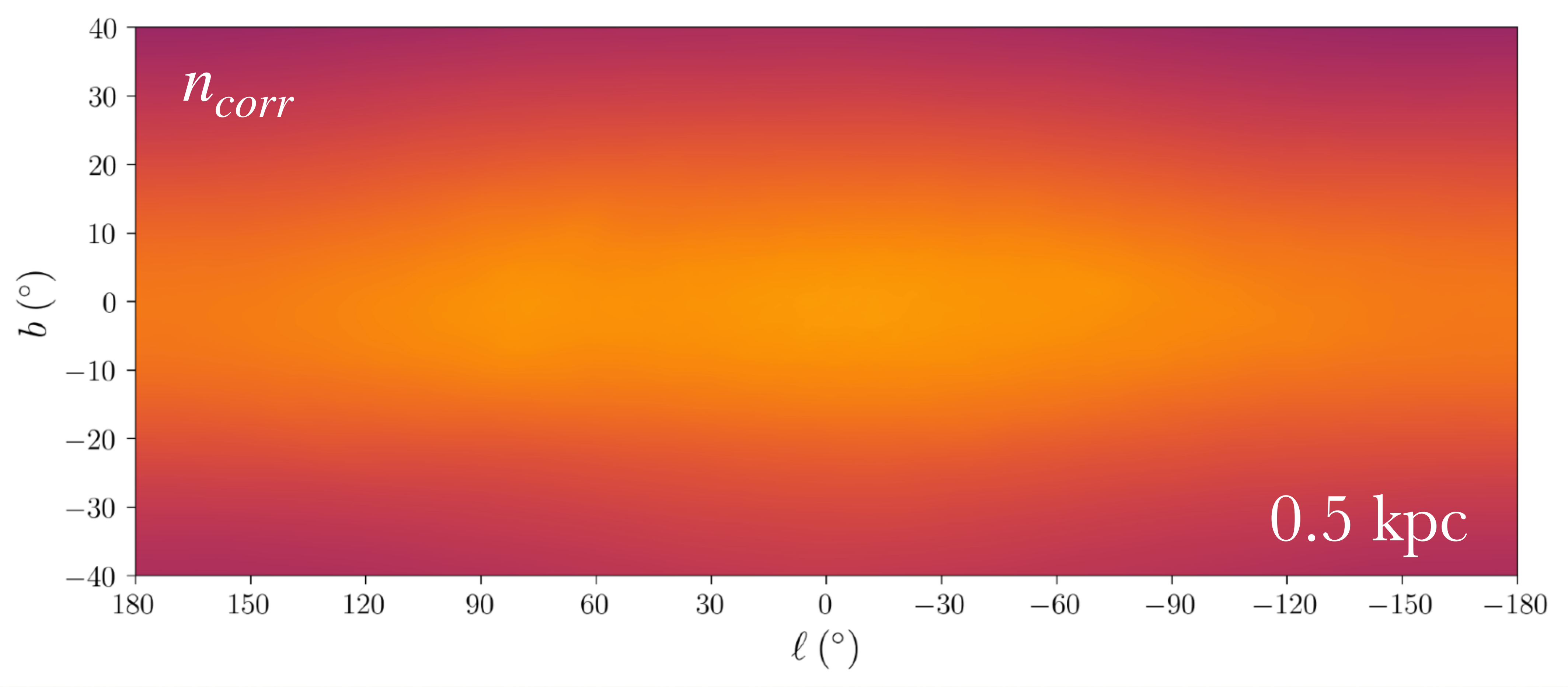


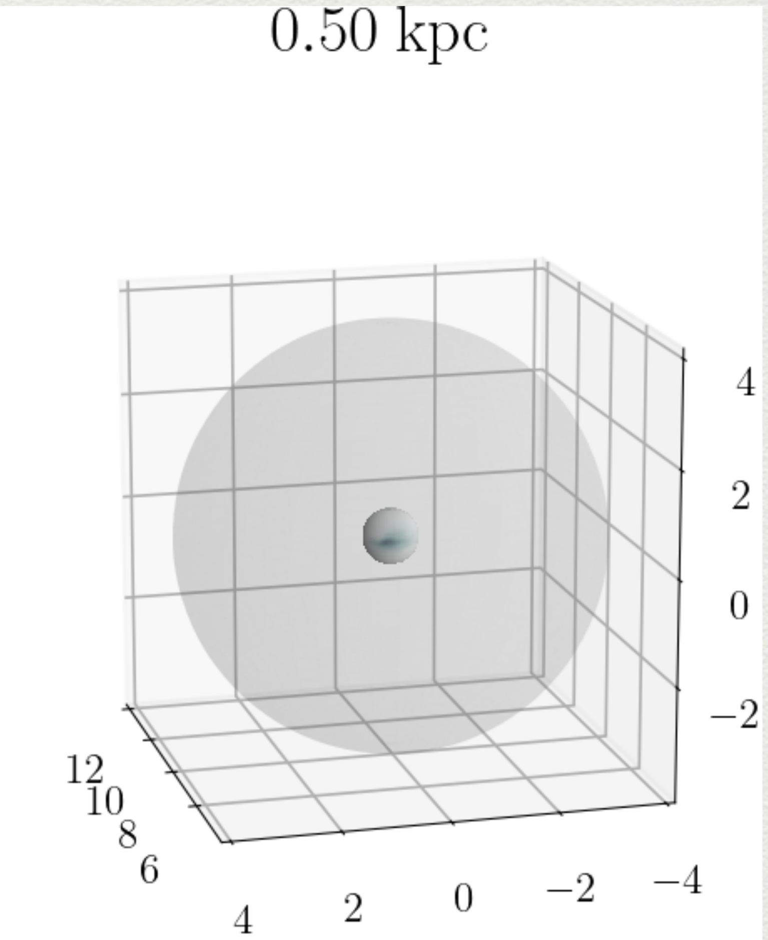
Results: Dust-Corrected PSD



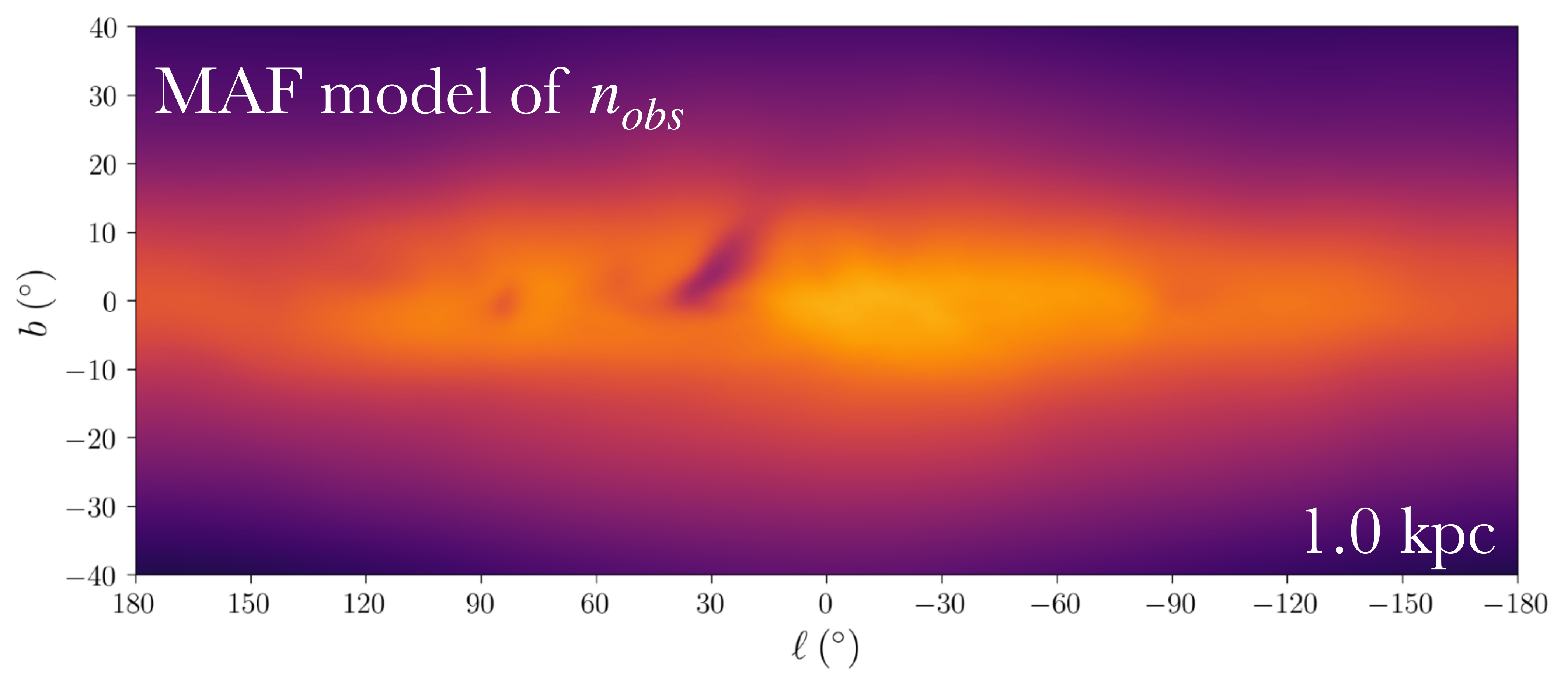


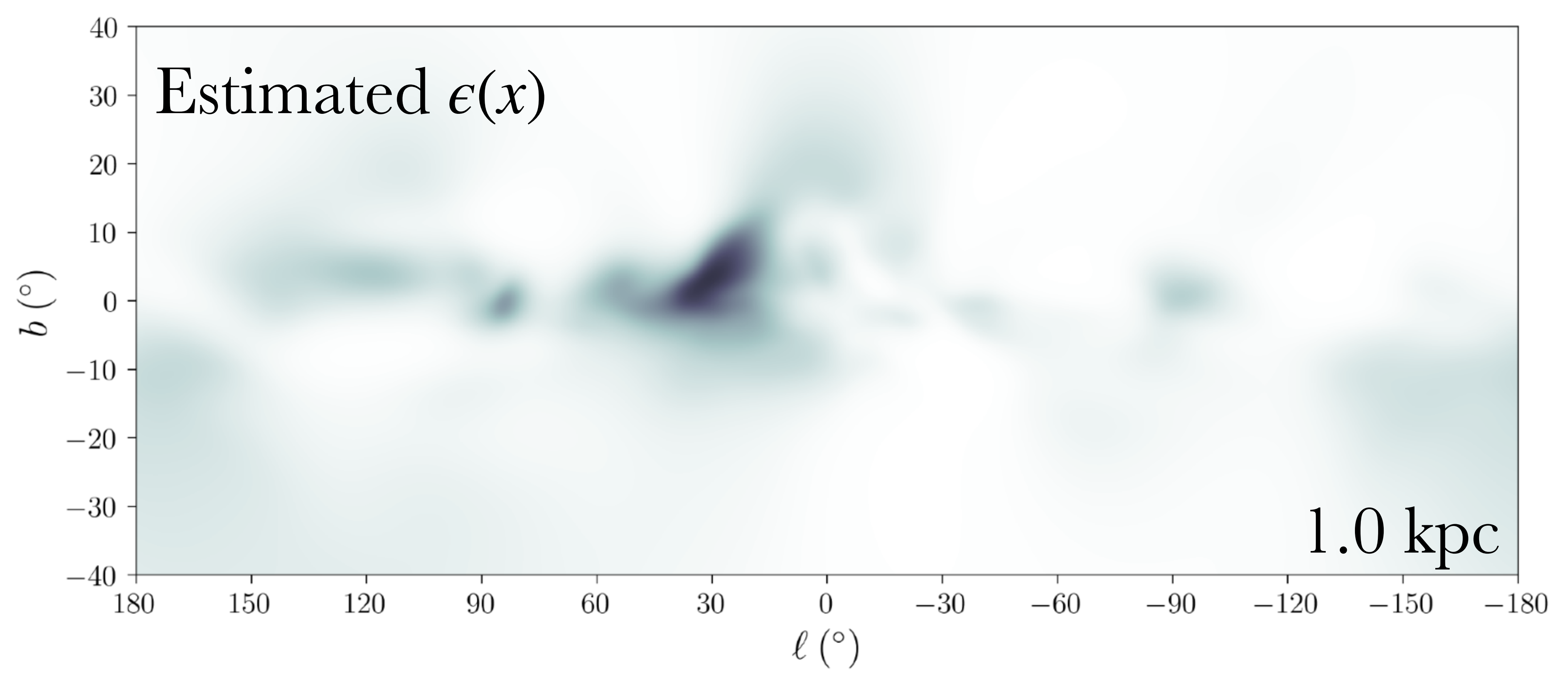


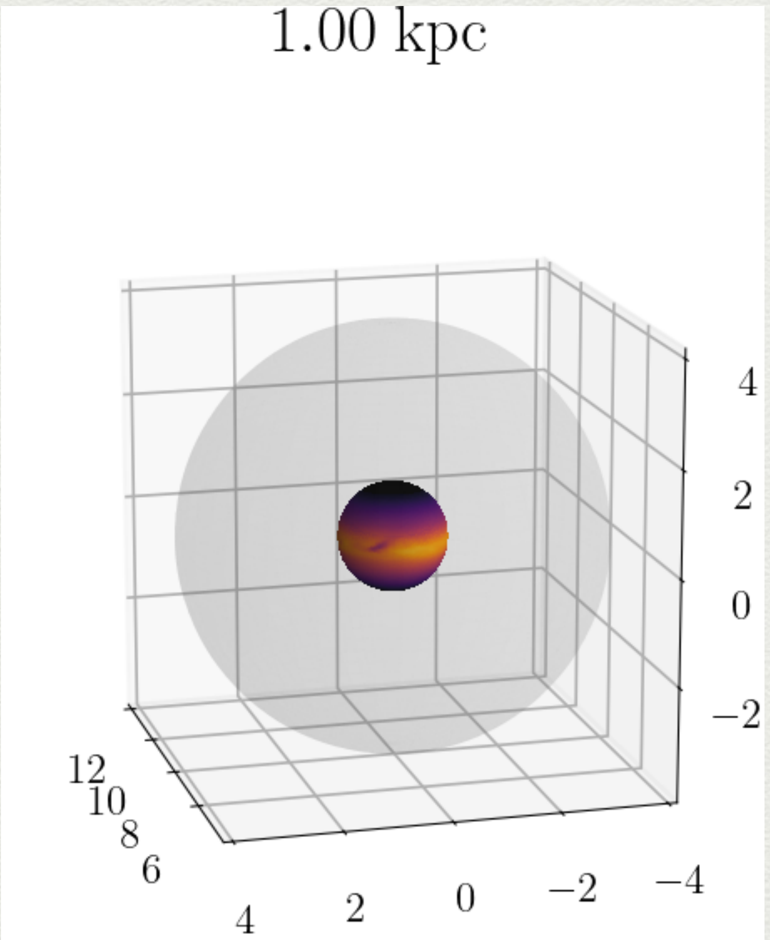


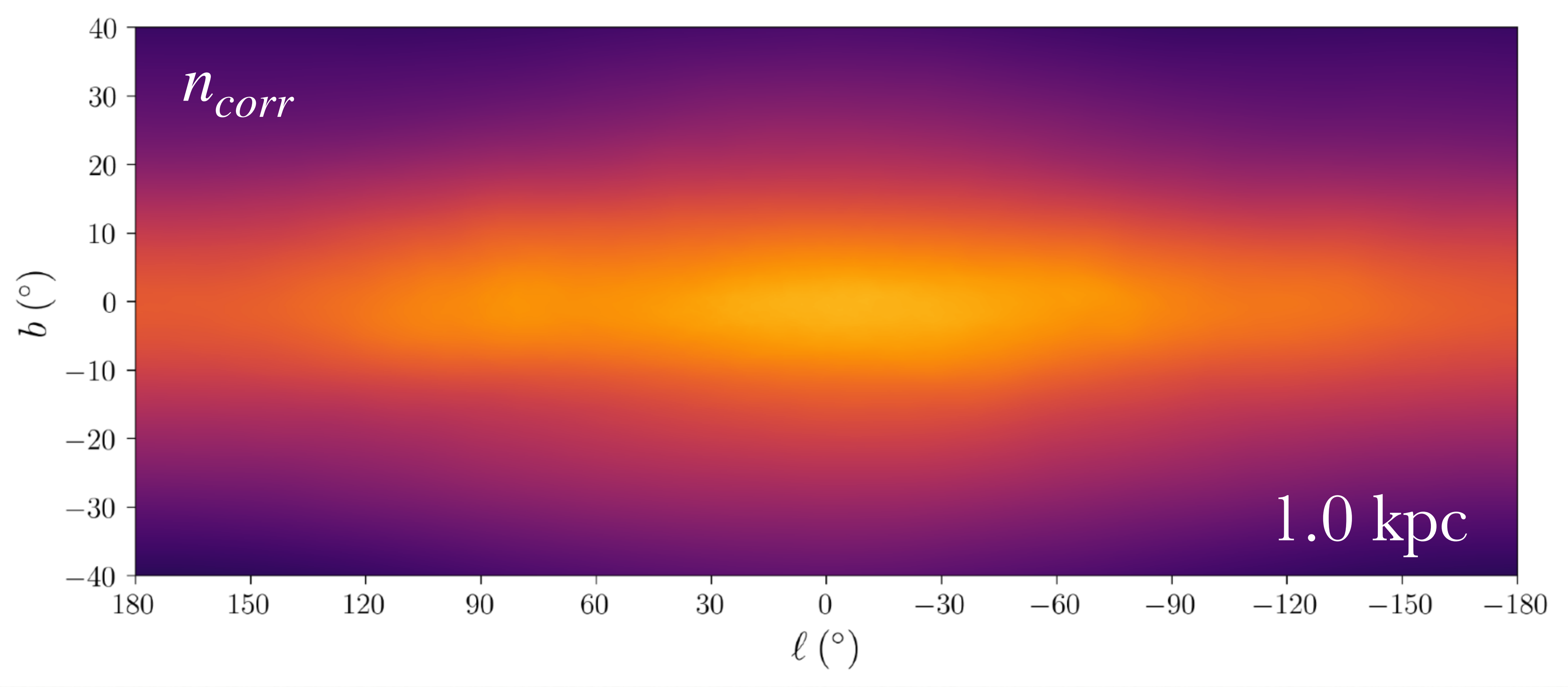


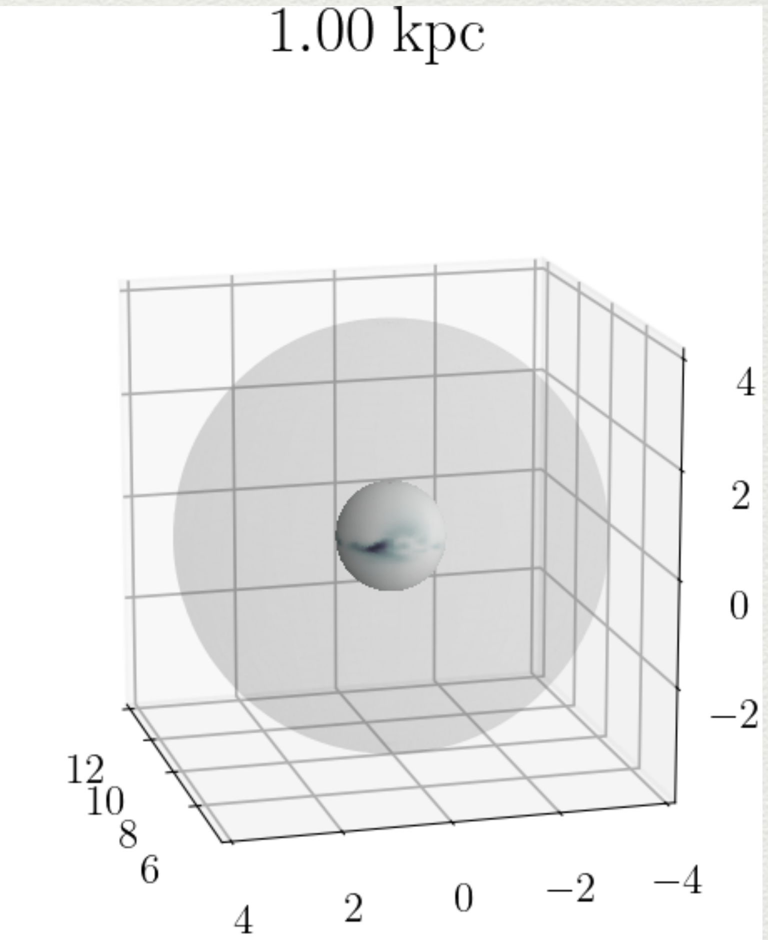
Results: Dust-Corrected PSD



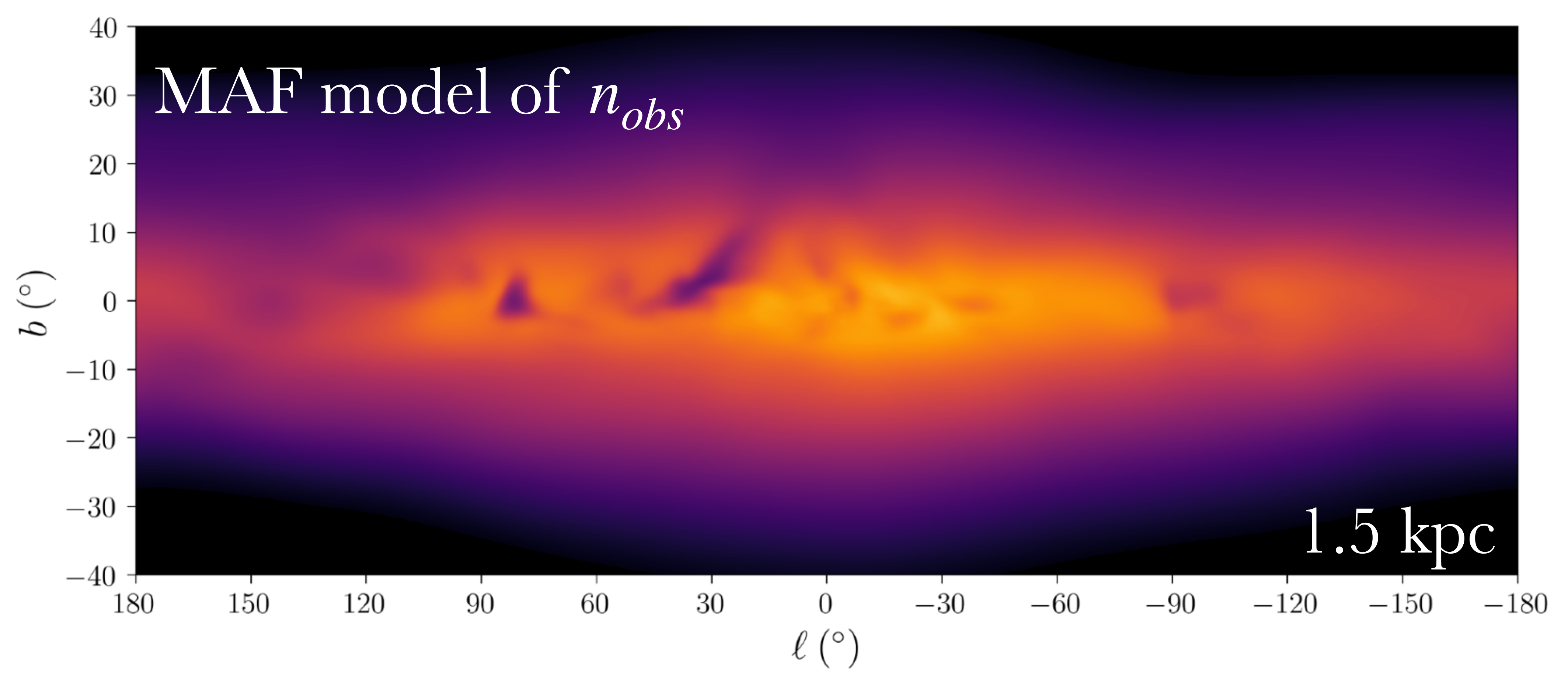


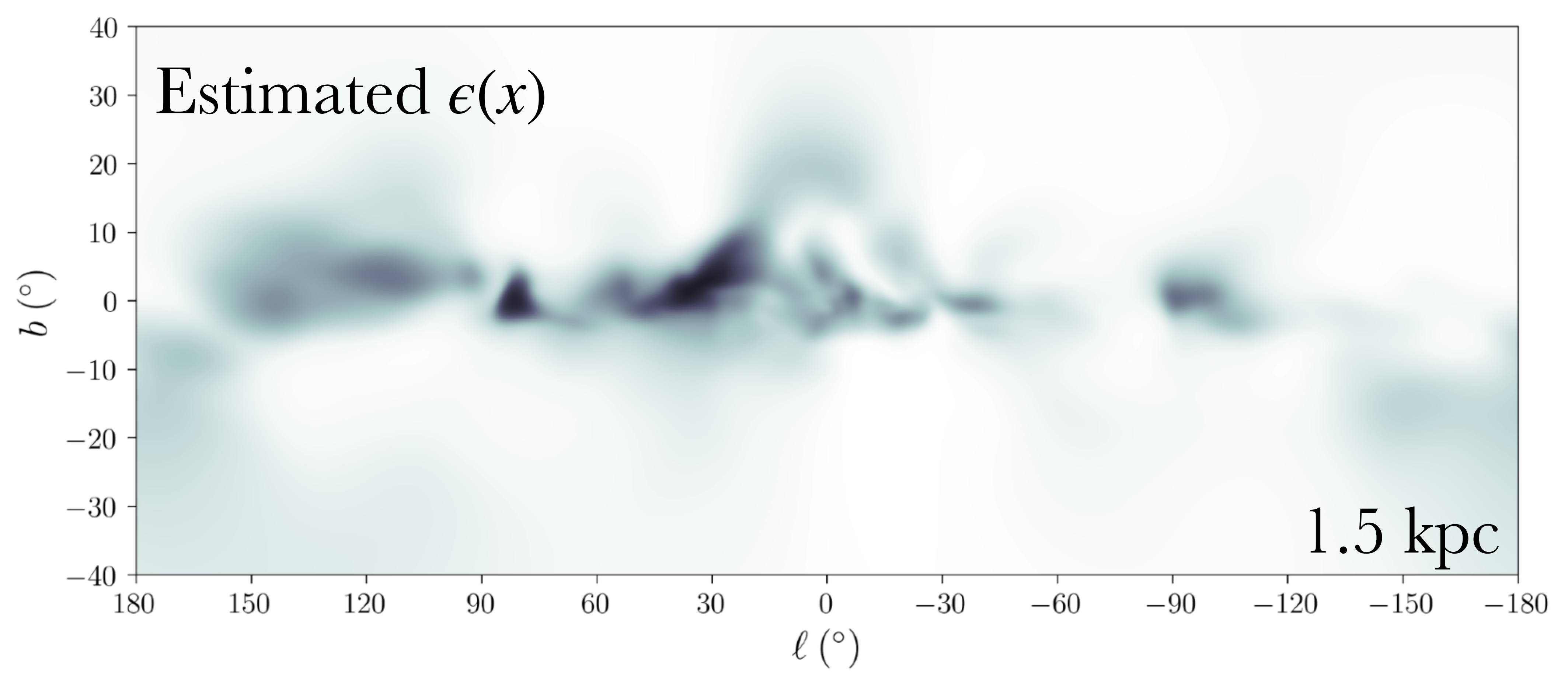


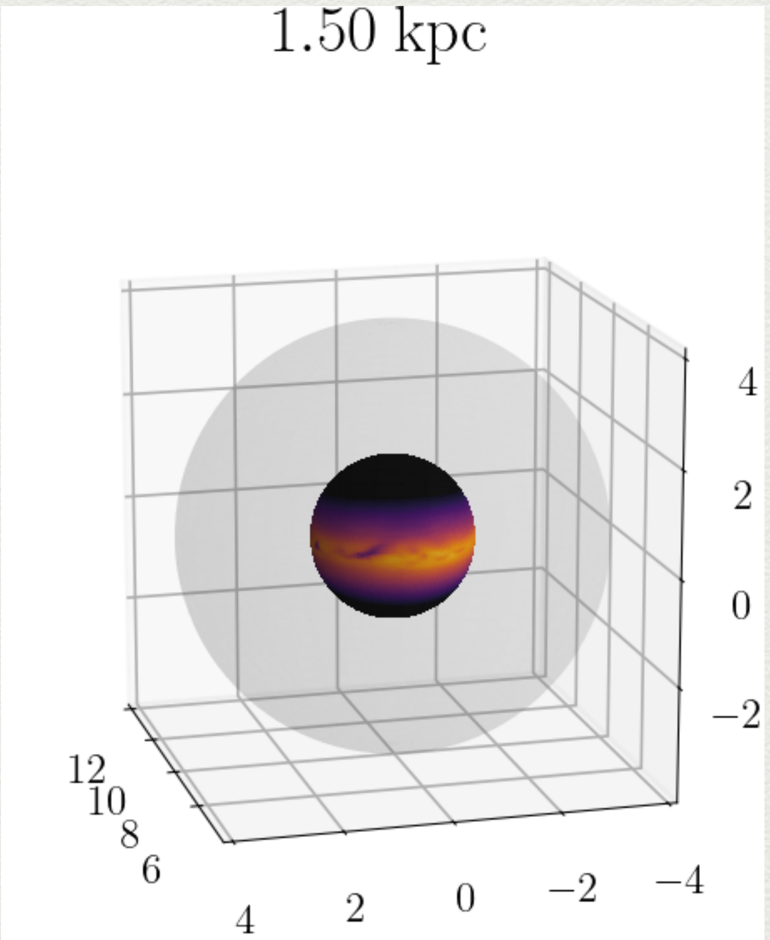


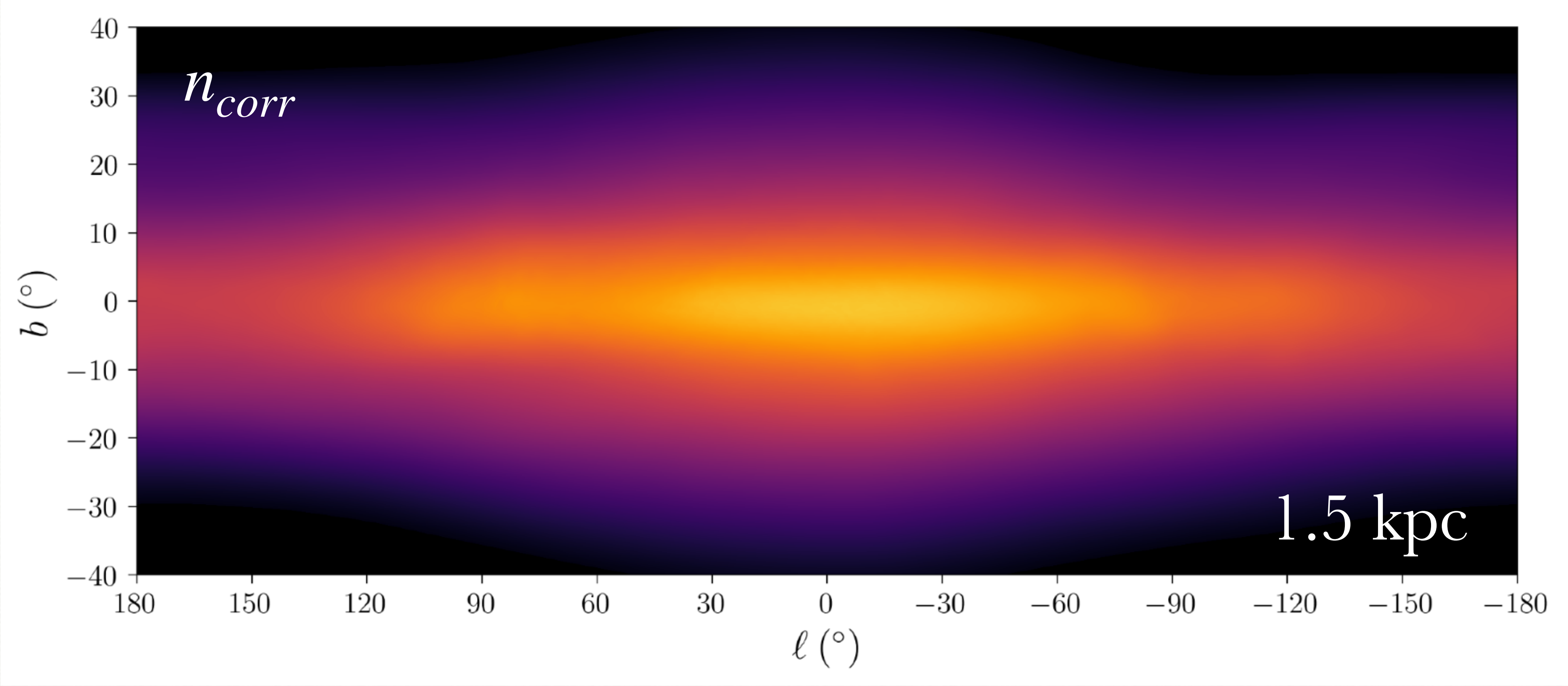


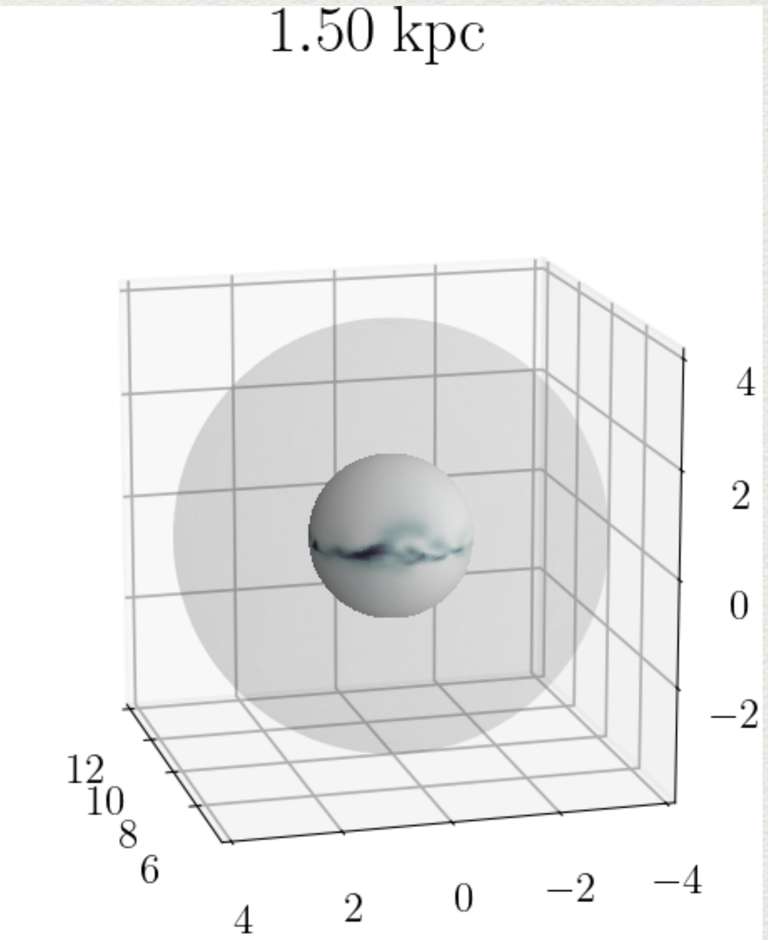
Results: Dust-Corrected PSD



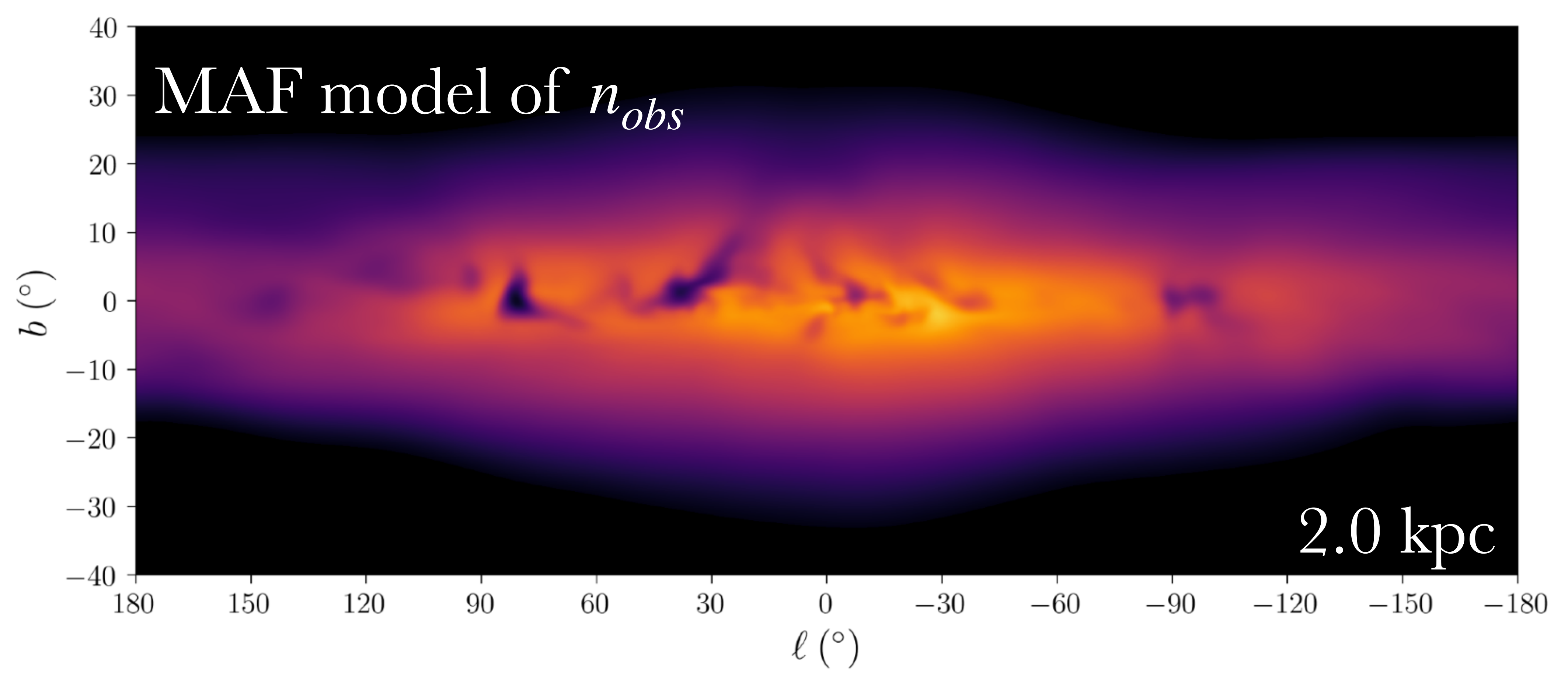


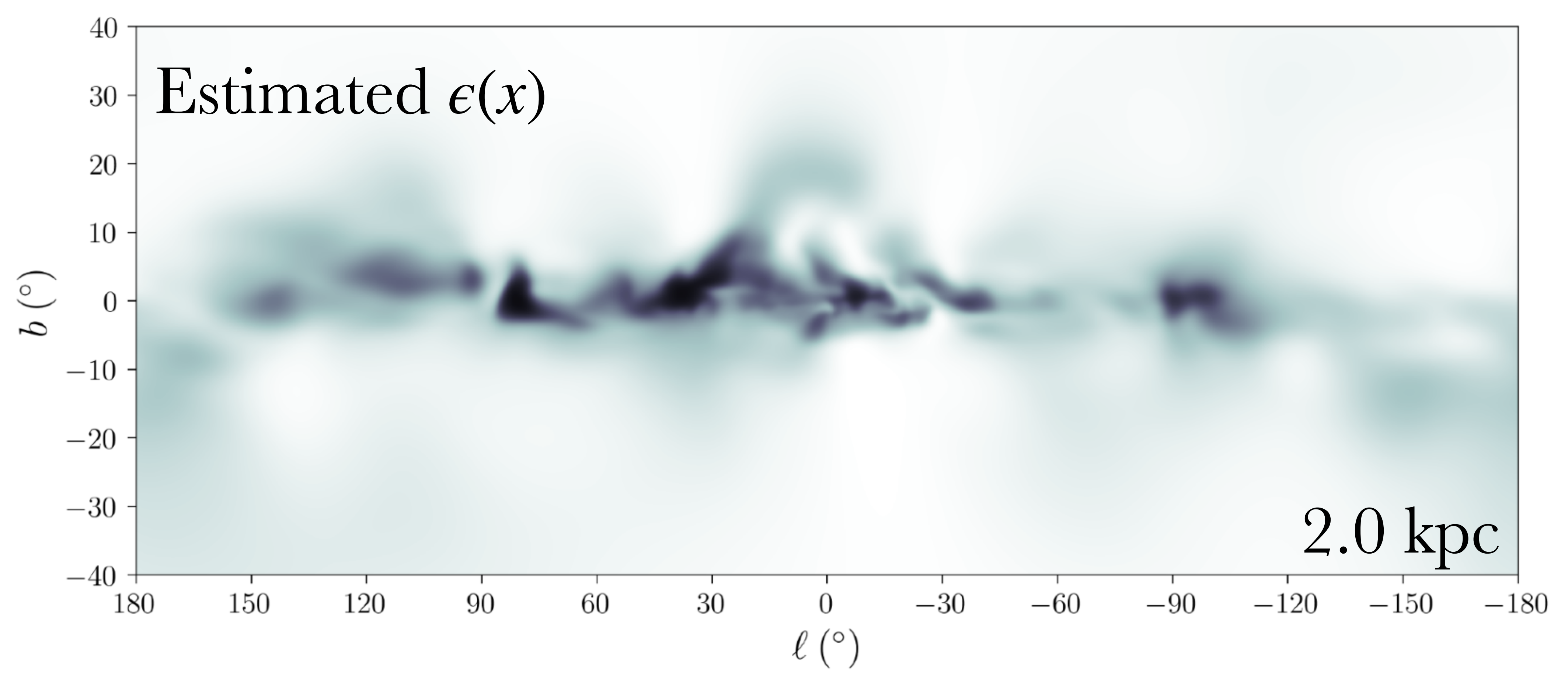


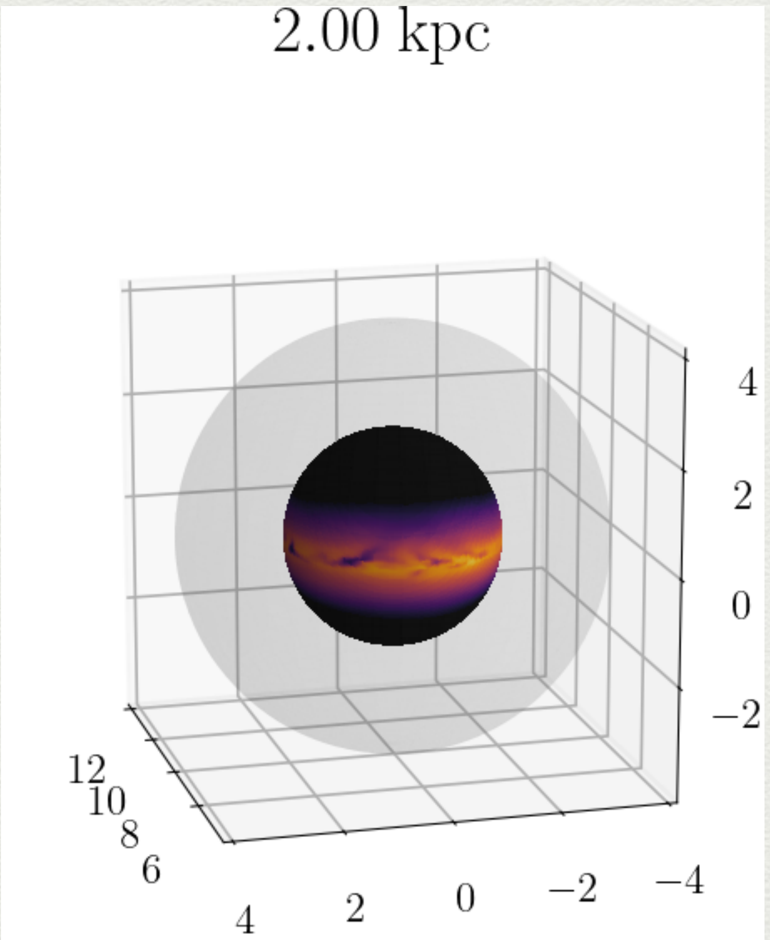


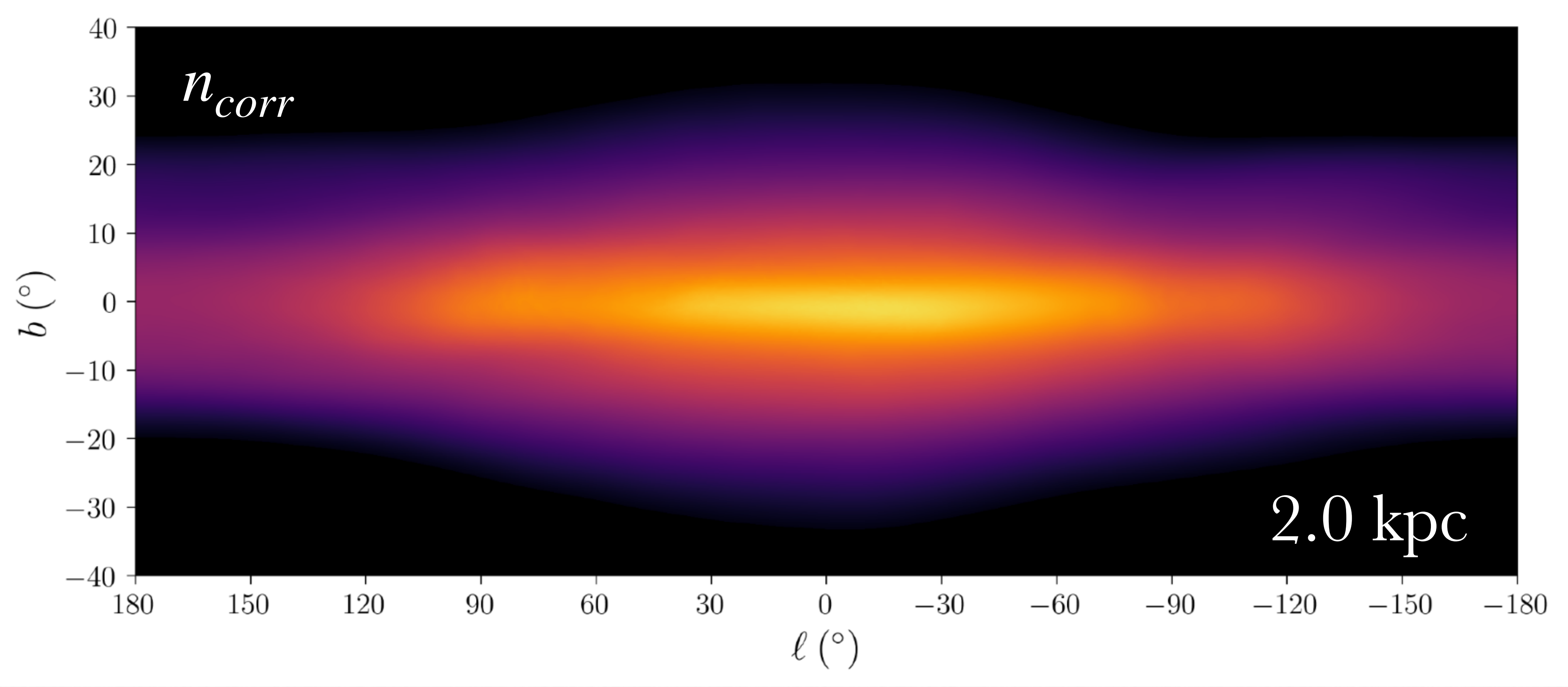


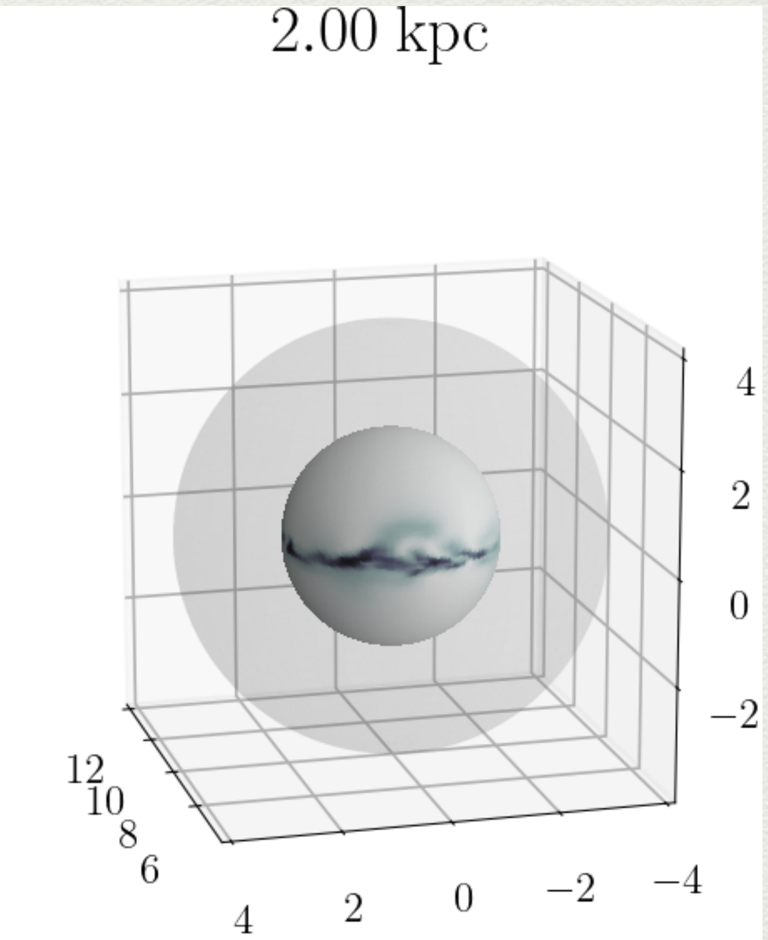
Results: Dust-Corrected PSD



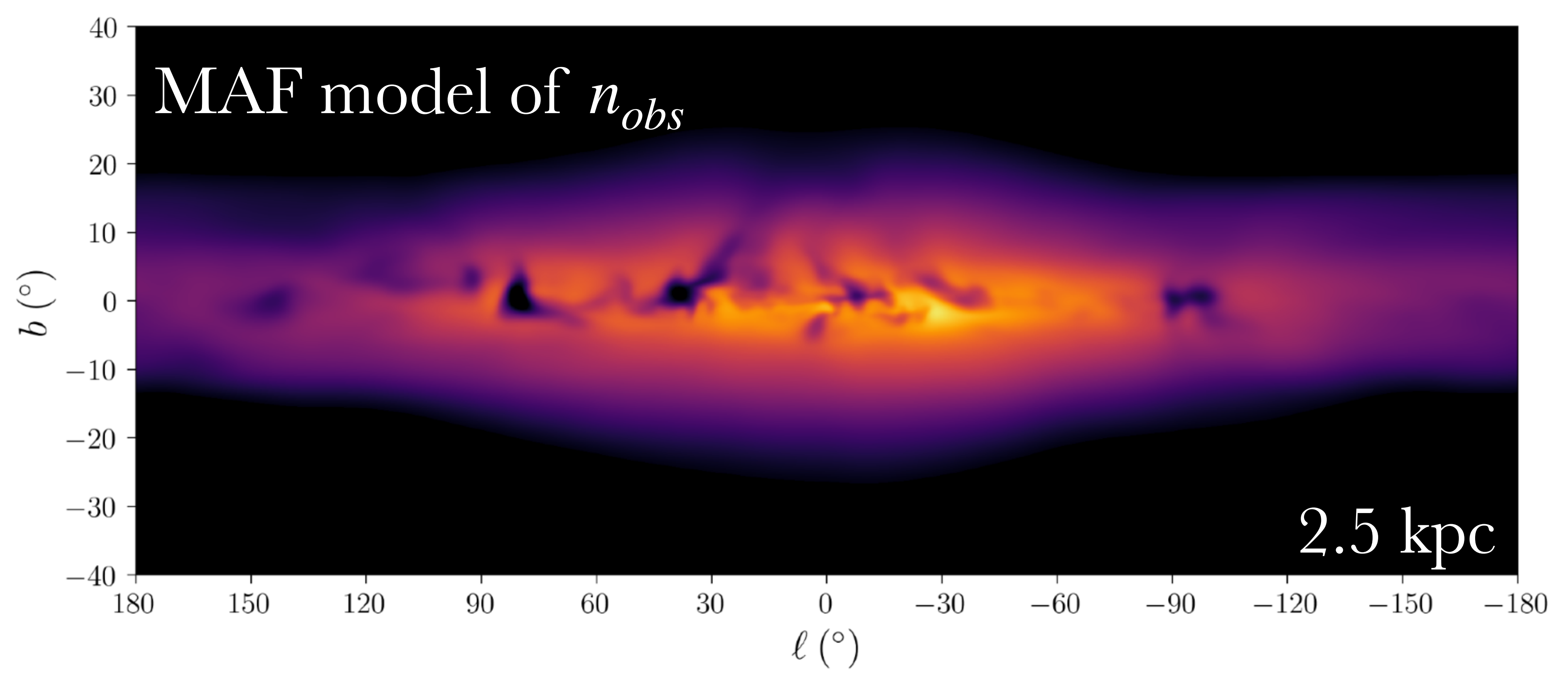


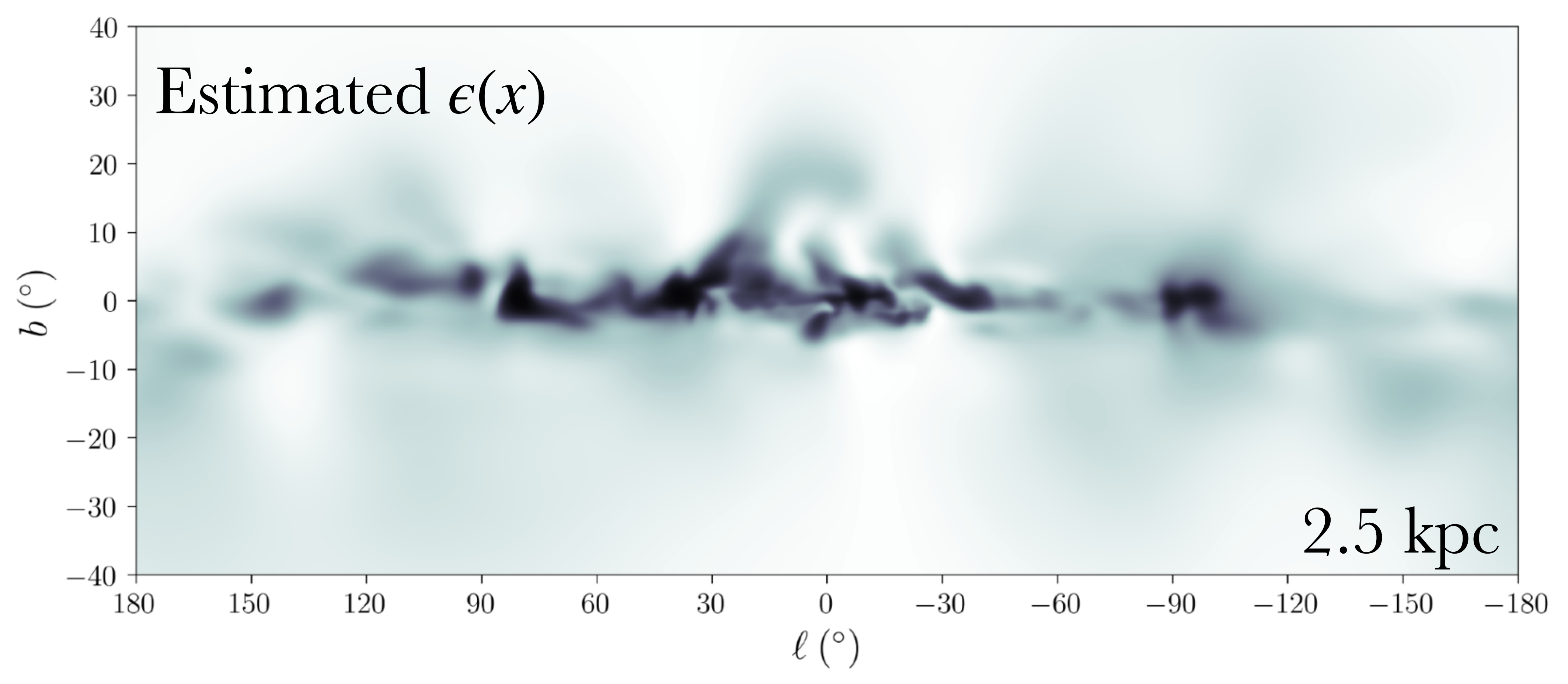


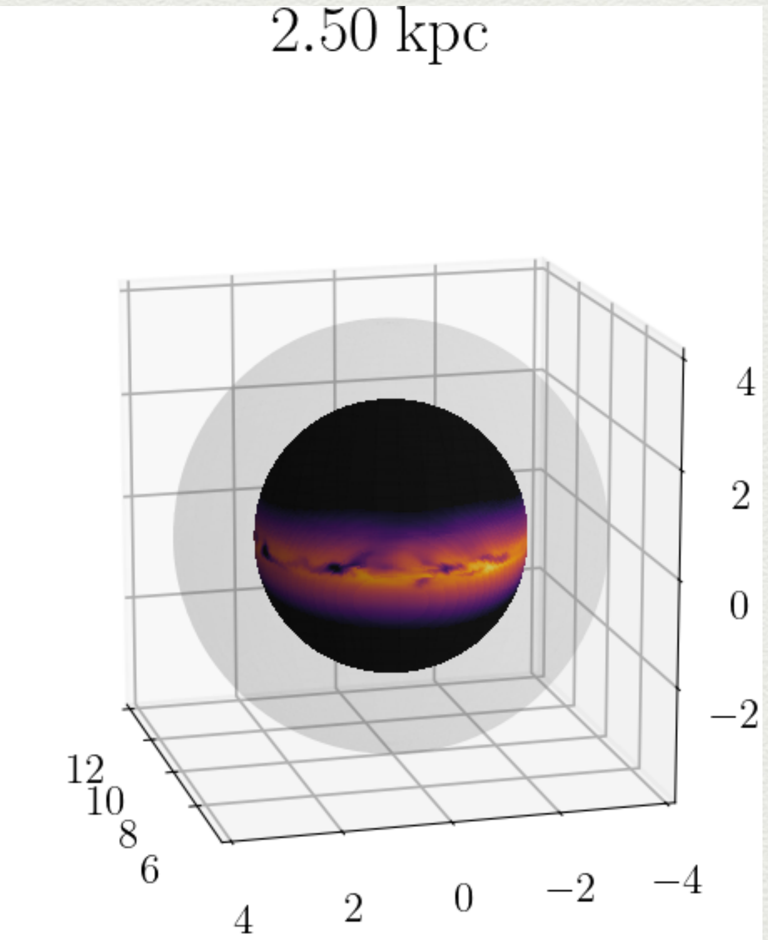


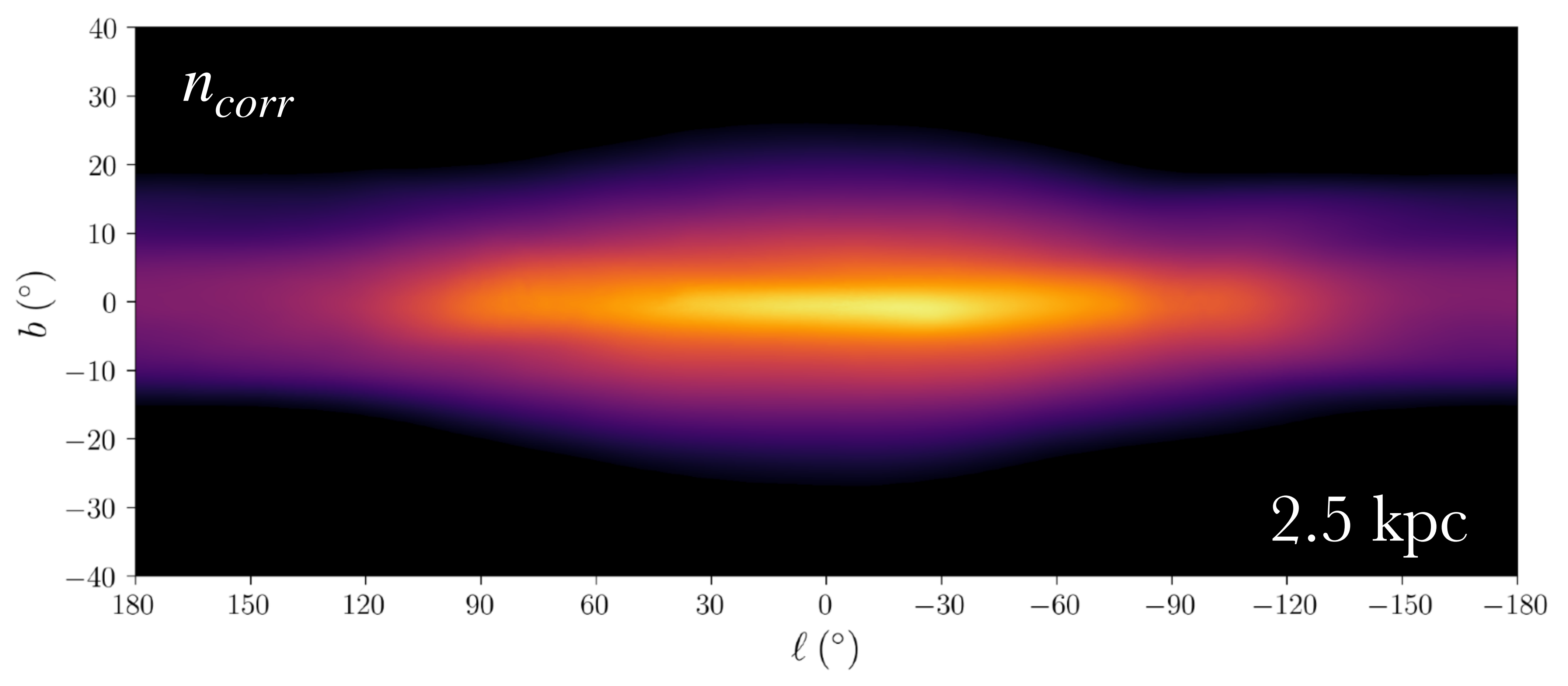


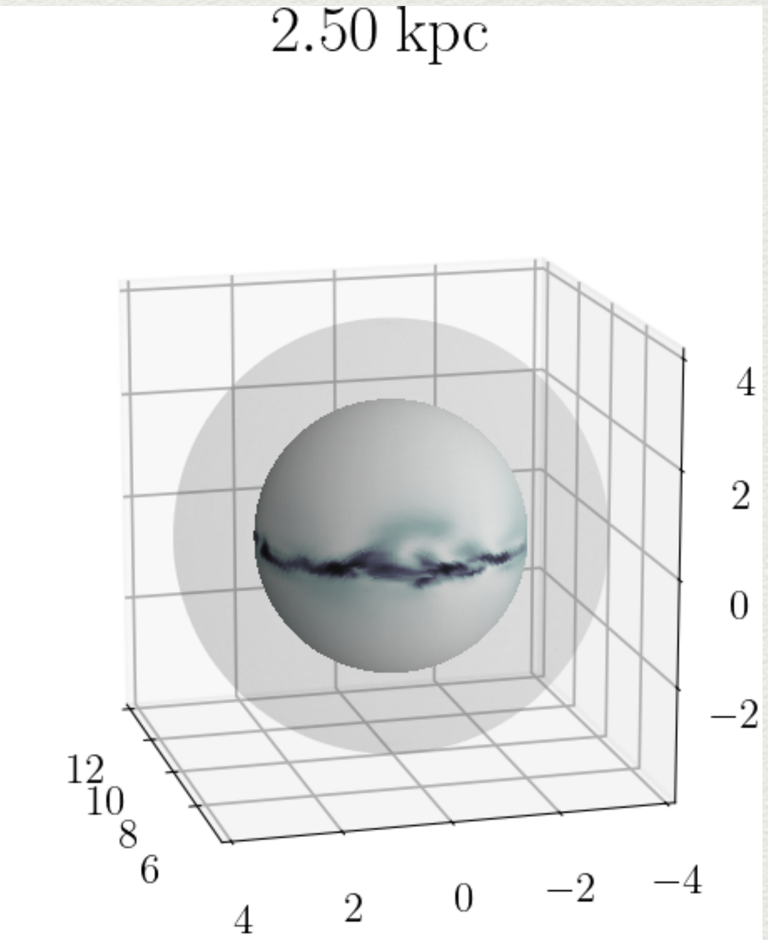
Results: Dust-Corrected PSD



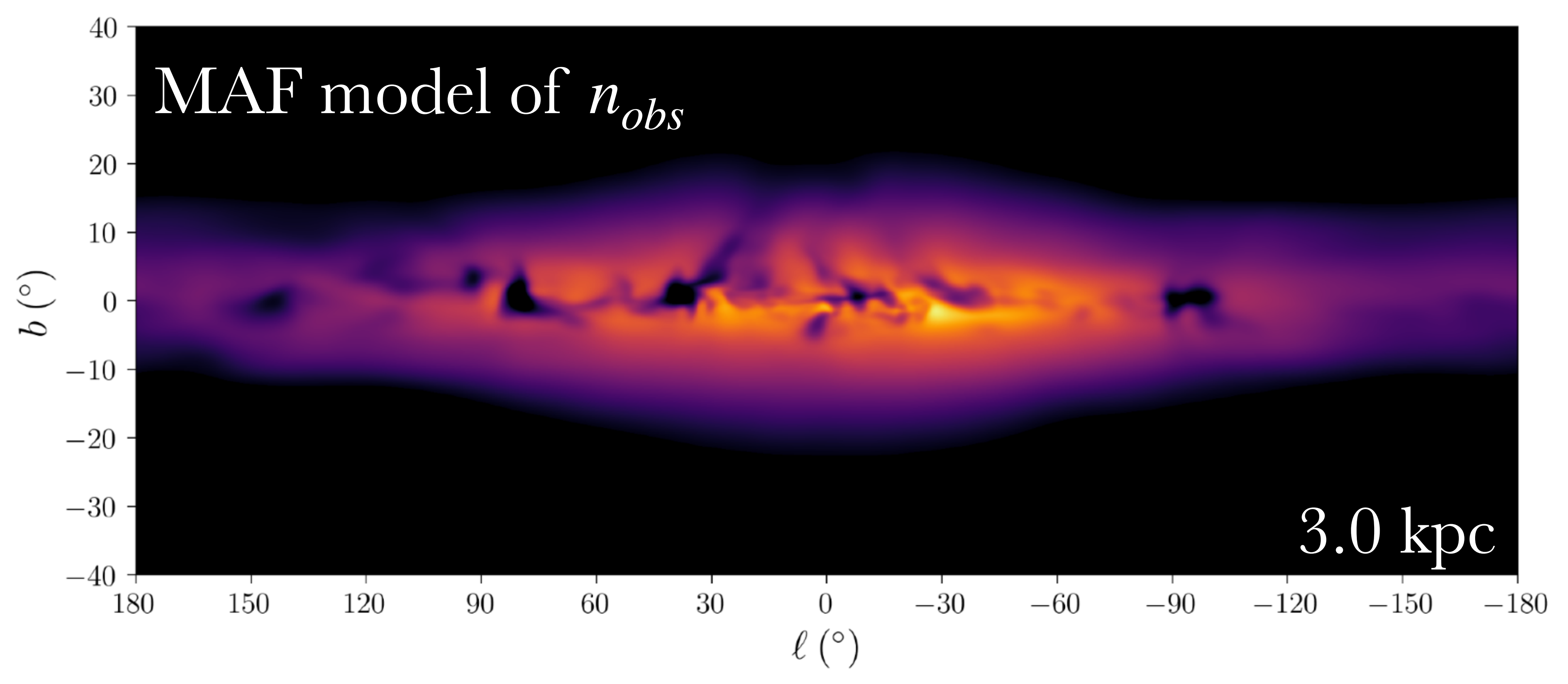


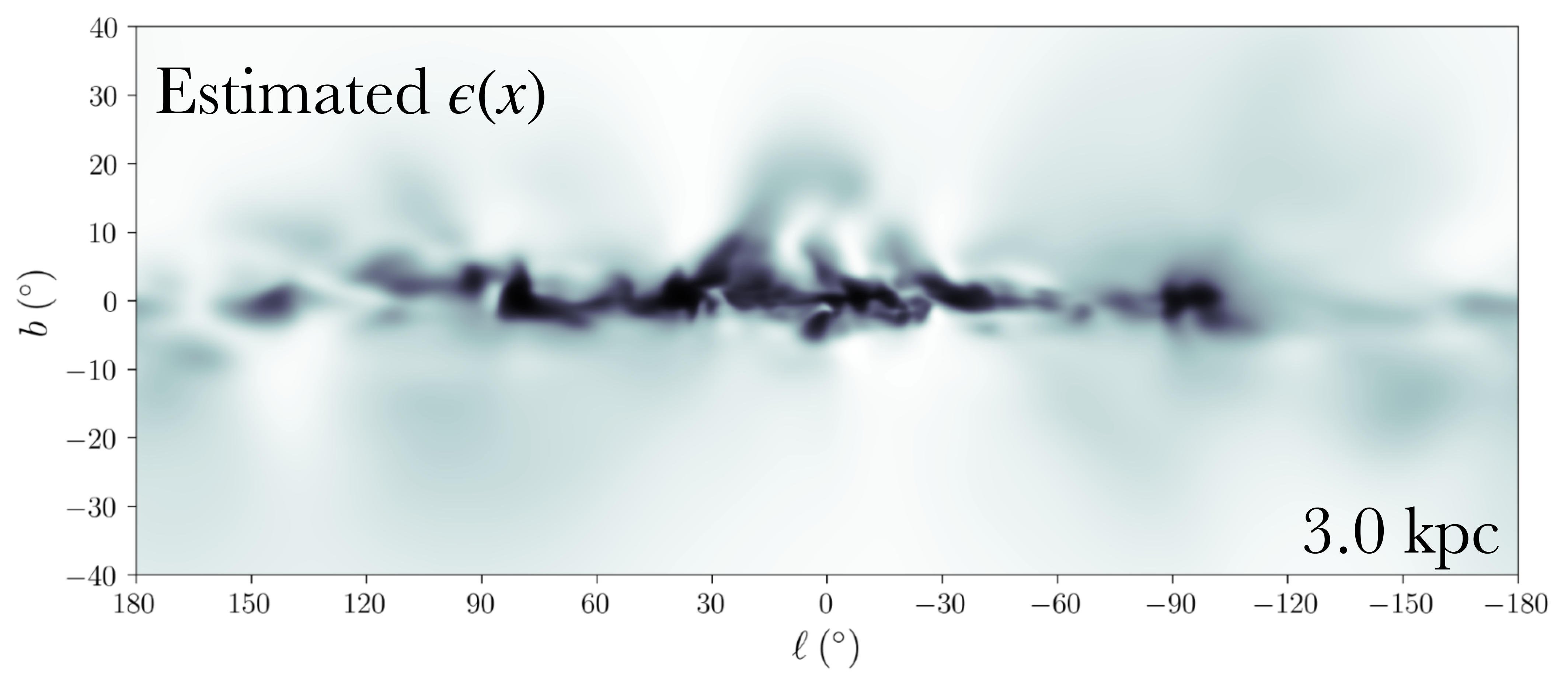


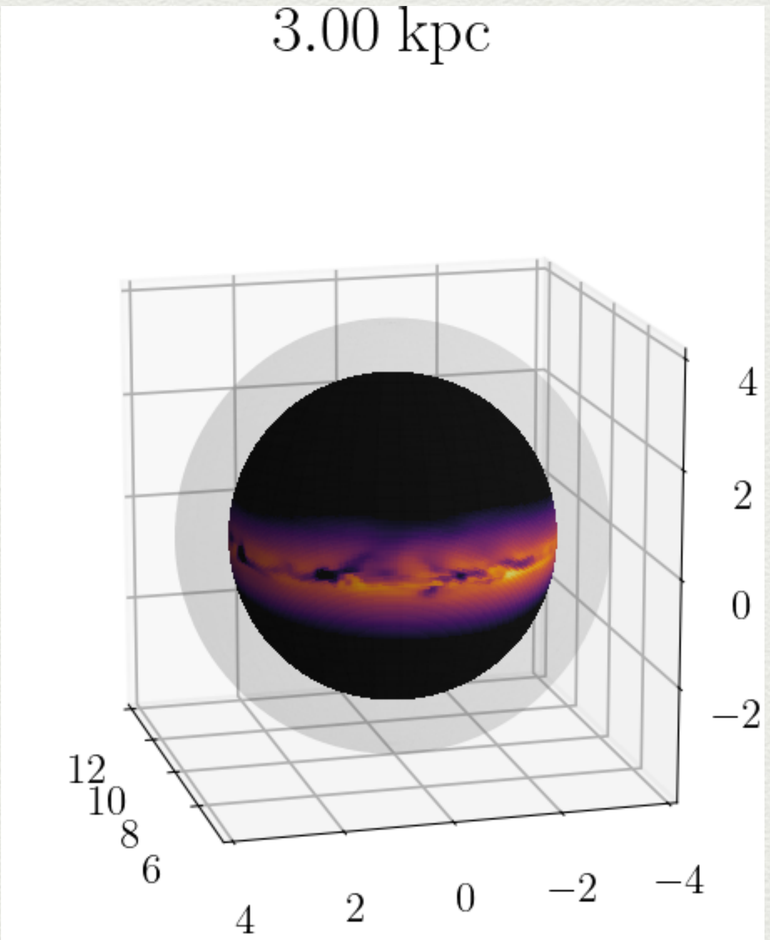


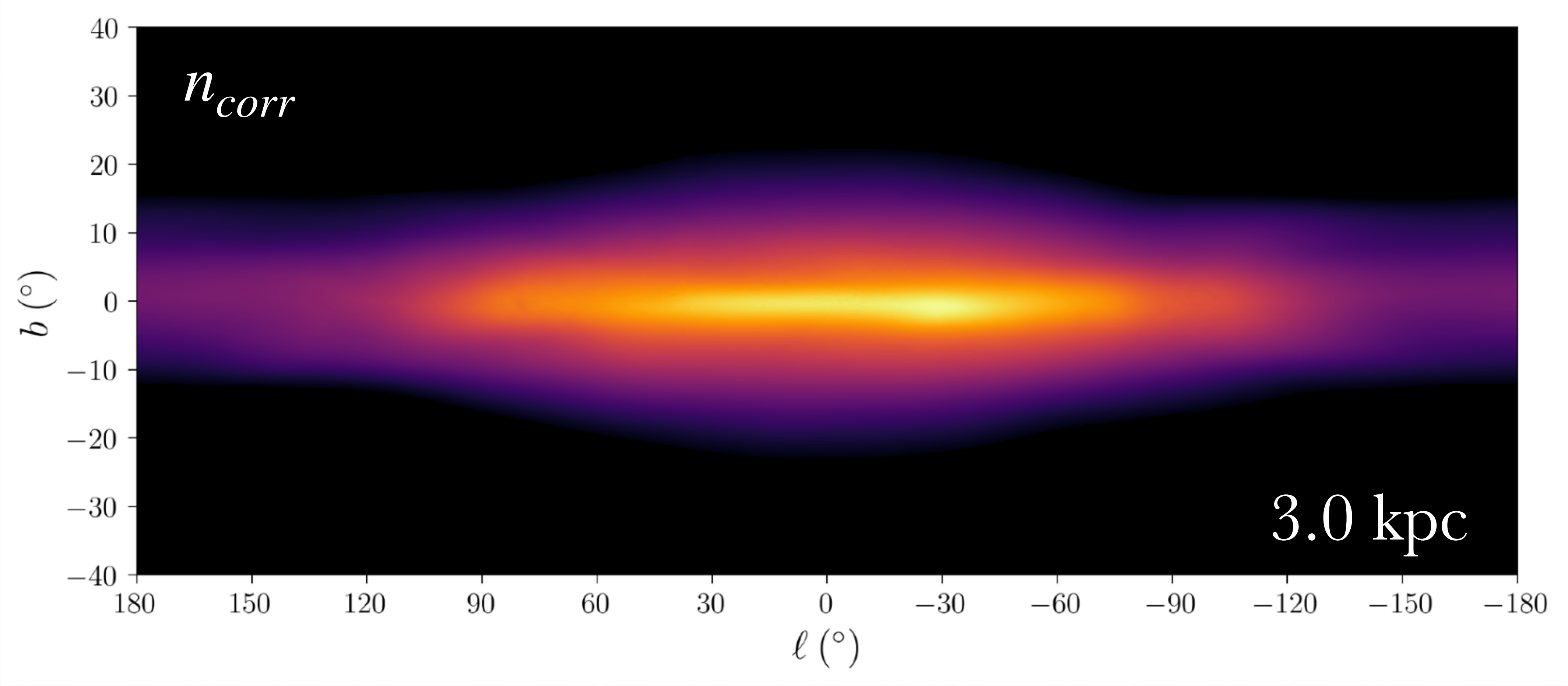


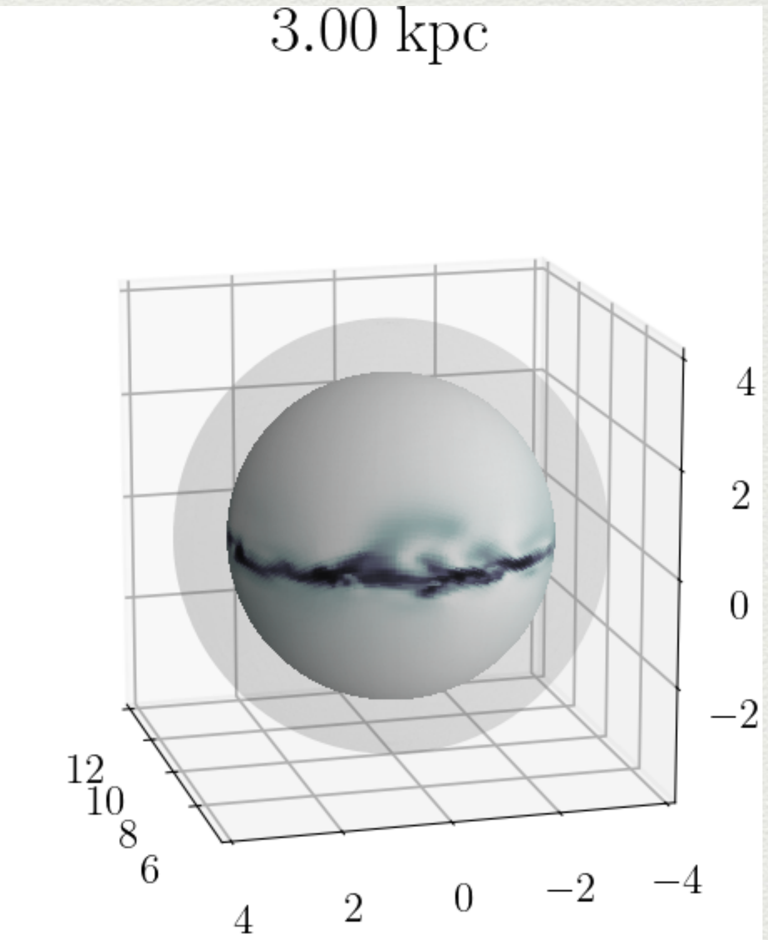
Results: Dust-Corrected PSD



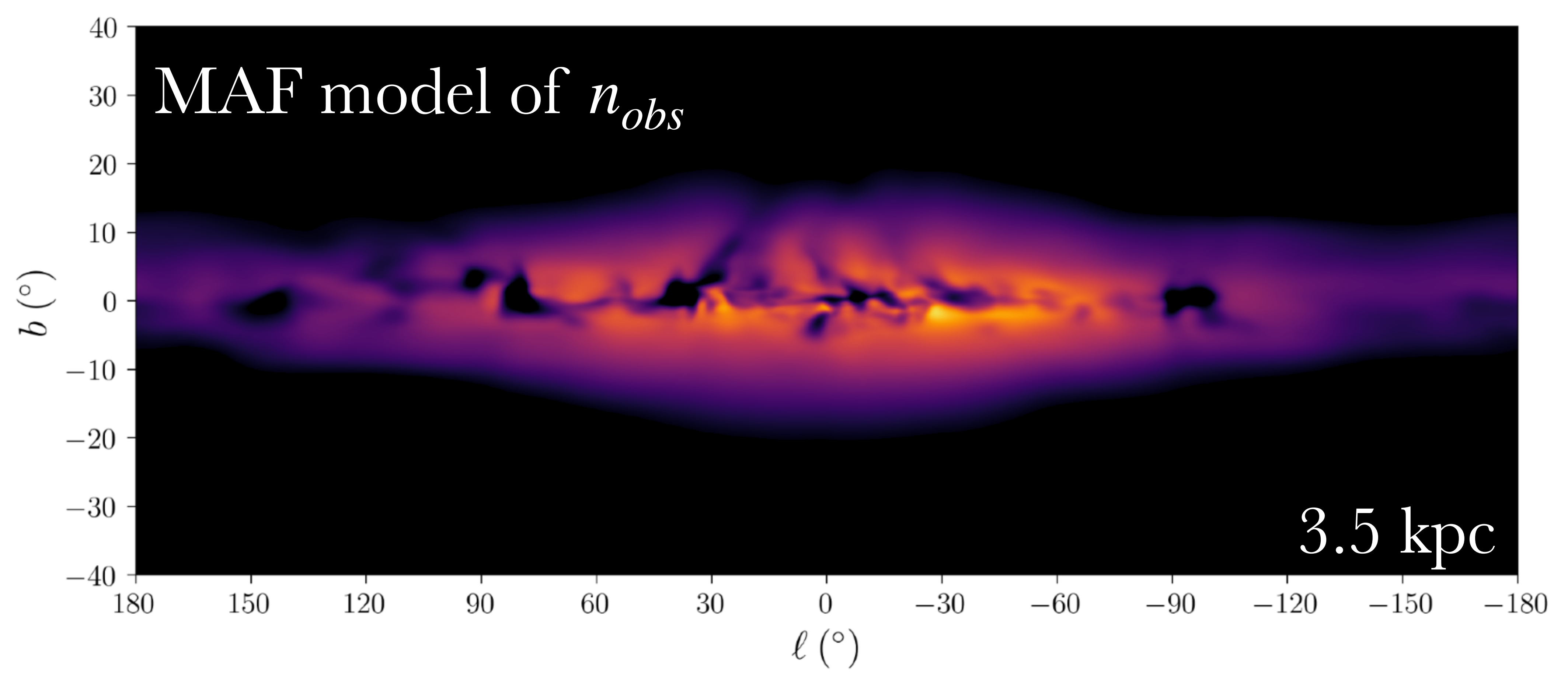


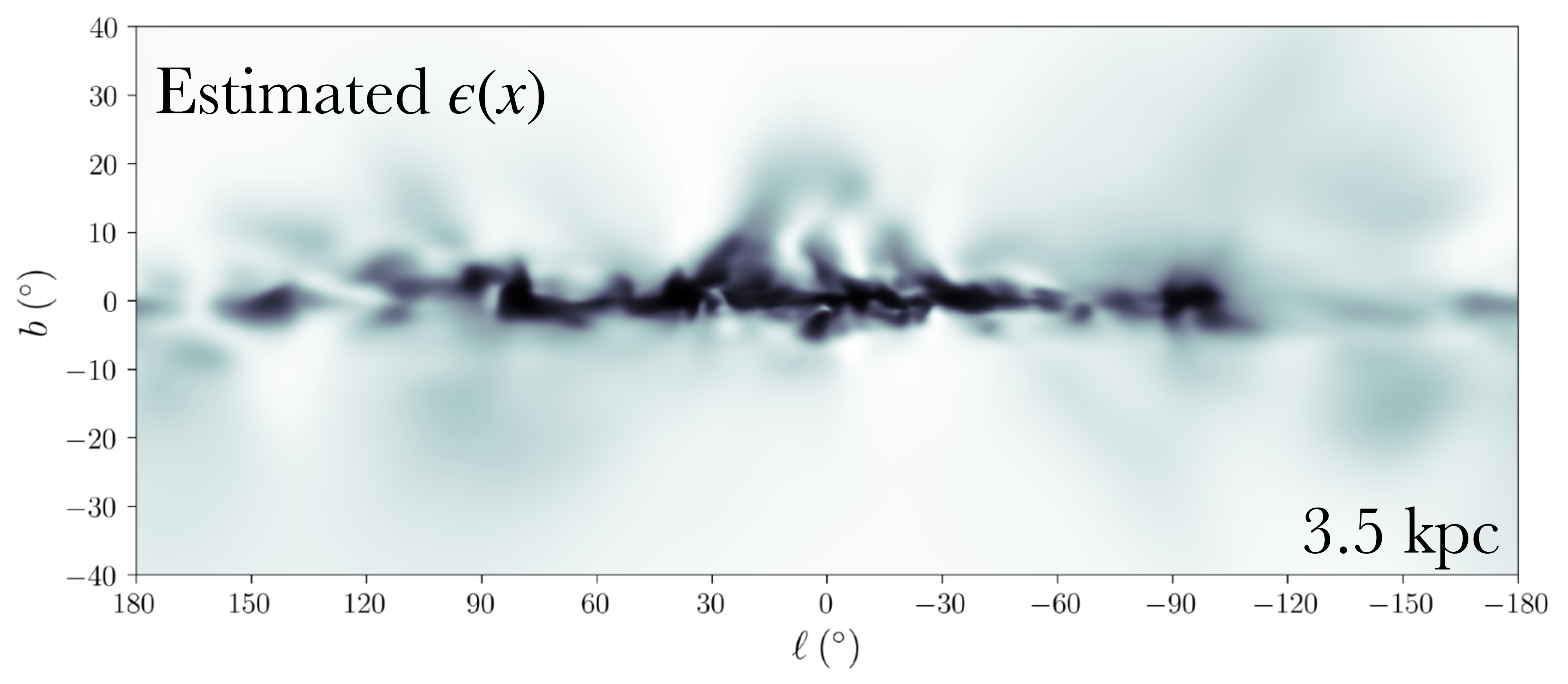


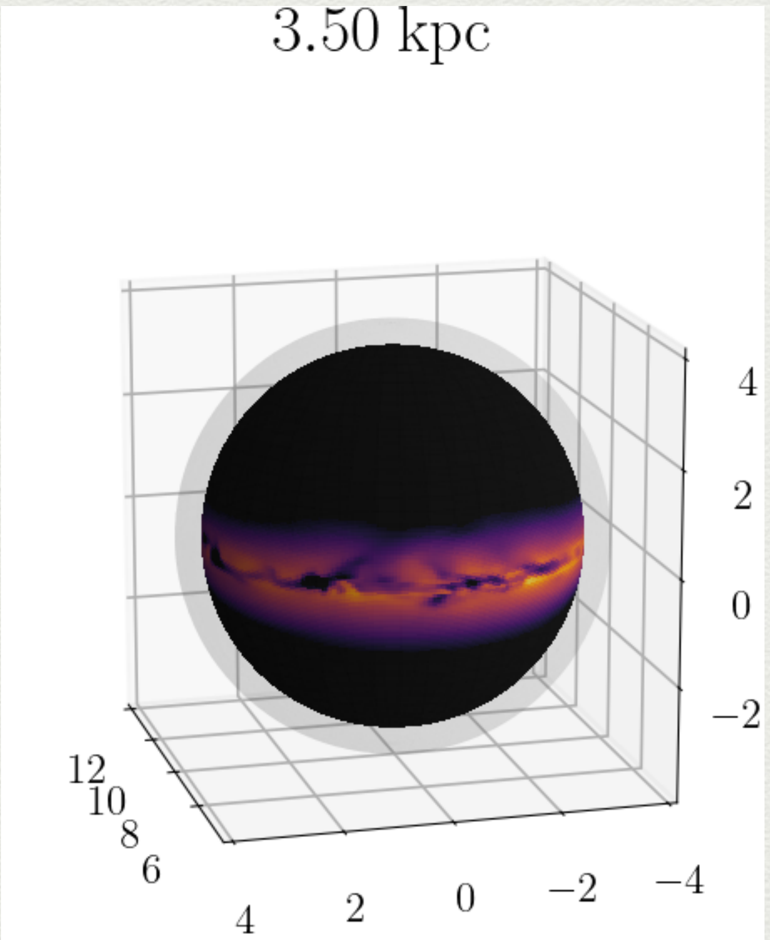


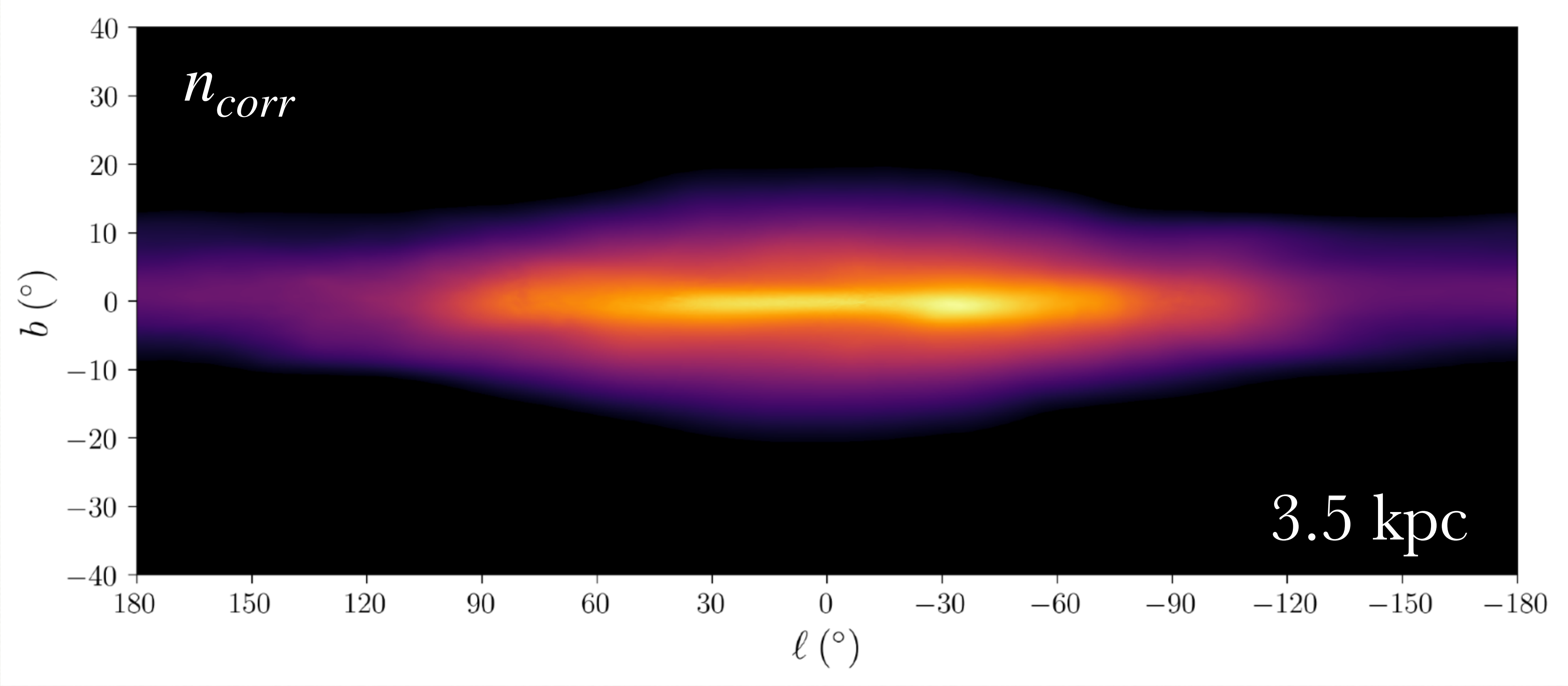


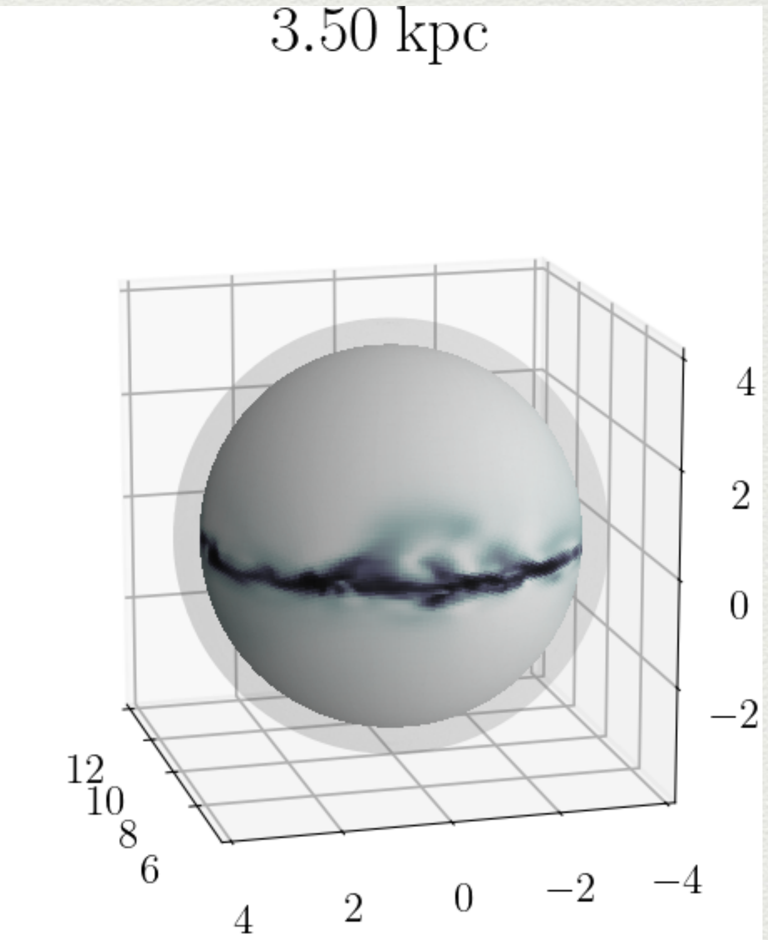
Results: Dust-Corrected PSD





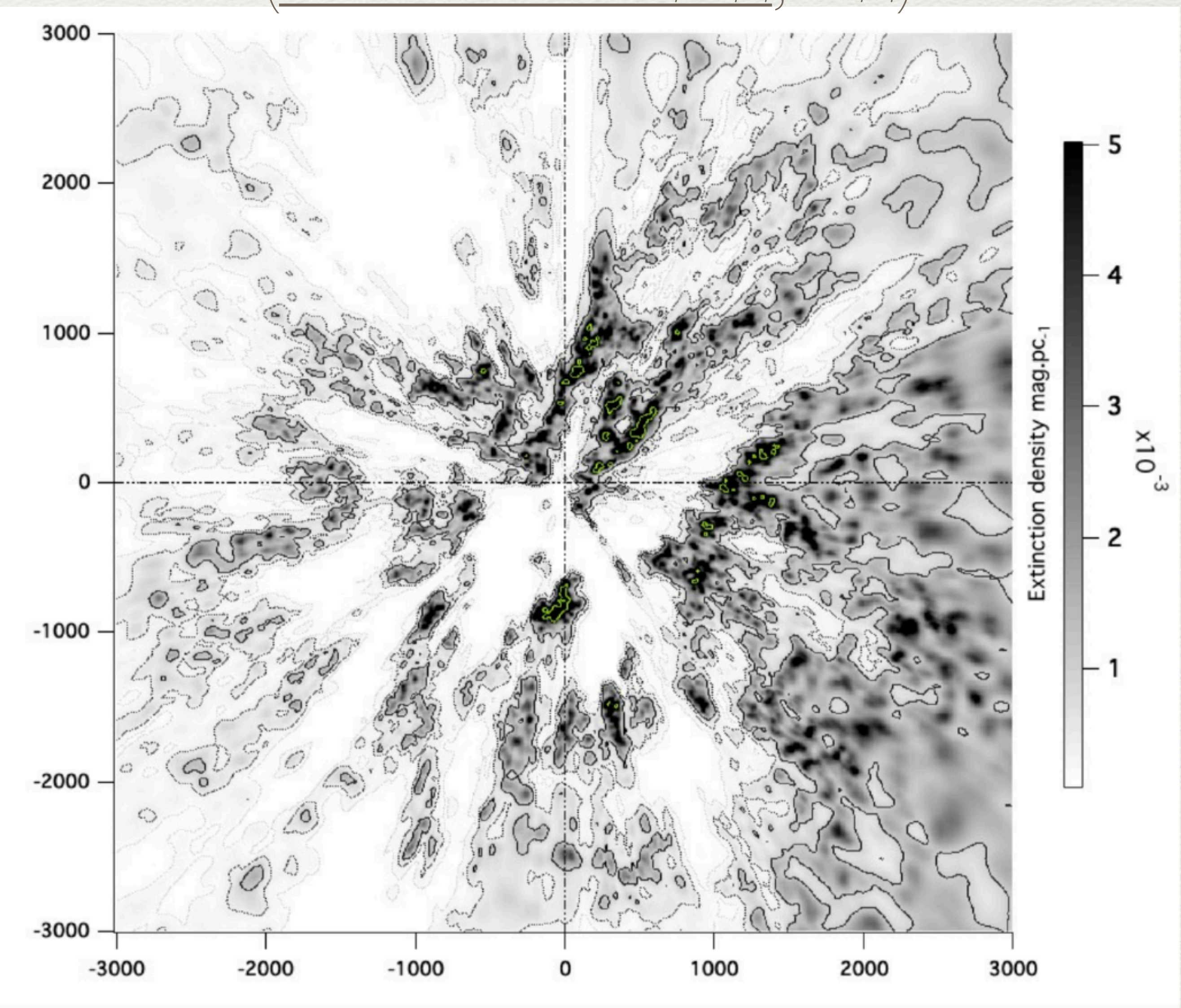






Validating dust efficiency map

(Lallement et al 2022, L22)

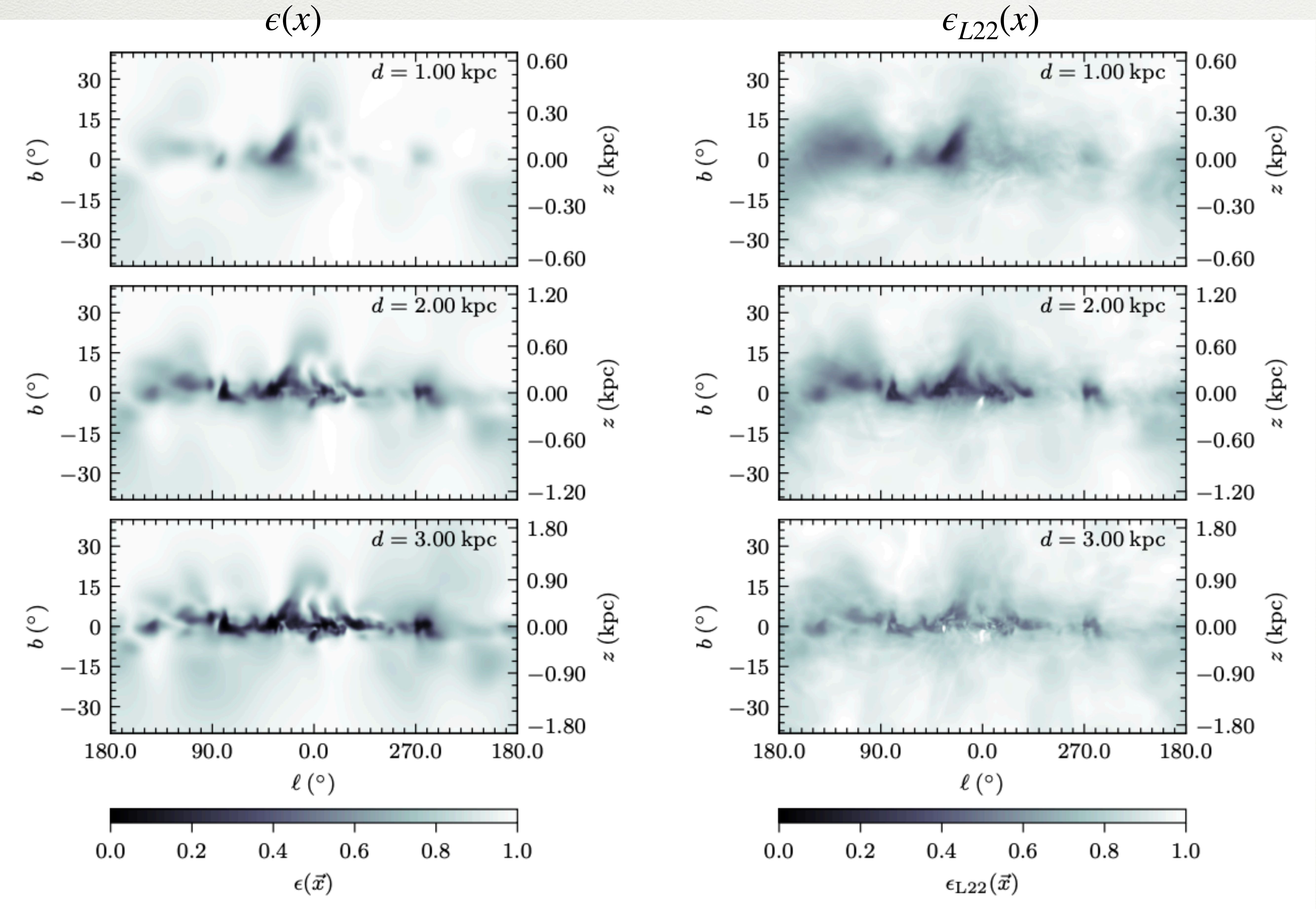


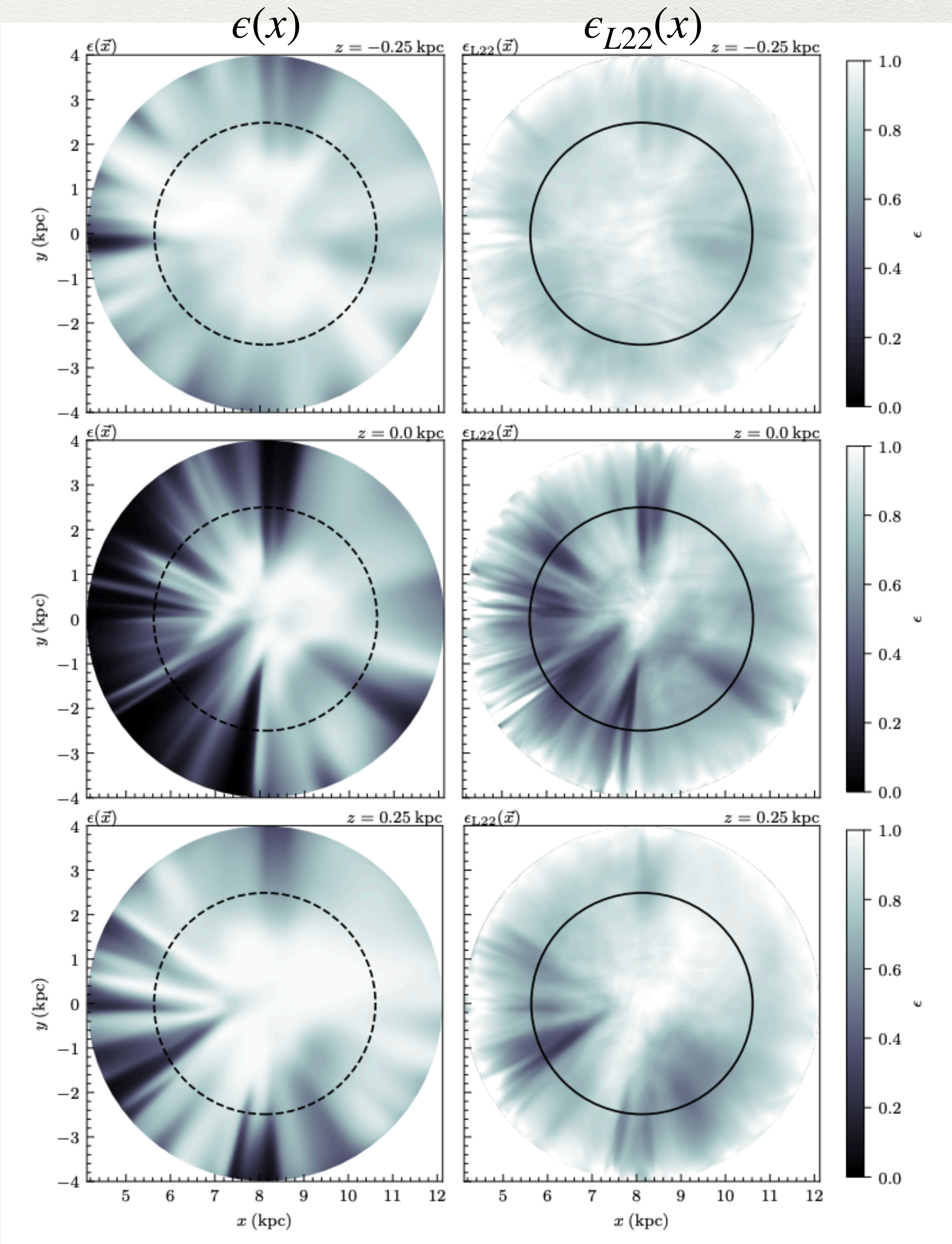
- How do we know that $\epsilon(x)$ is really correct?
- Astronomers have constructed extremely detailed and precise 3d dust maps using Gaia and other data
- However, these quantify how much an *observed* star is dimmed and reddened by dust.
- They cannot by themselves tell us about stars that are lost due to dust

Validating dust efficiency map

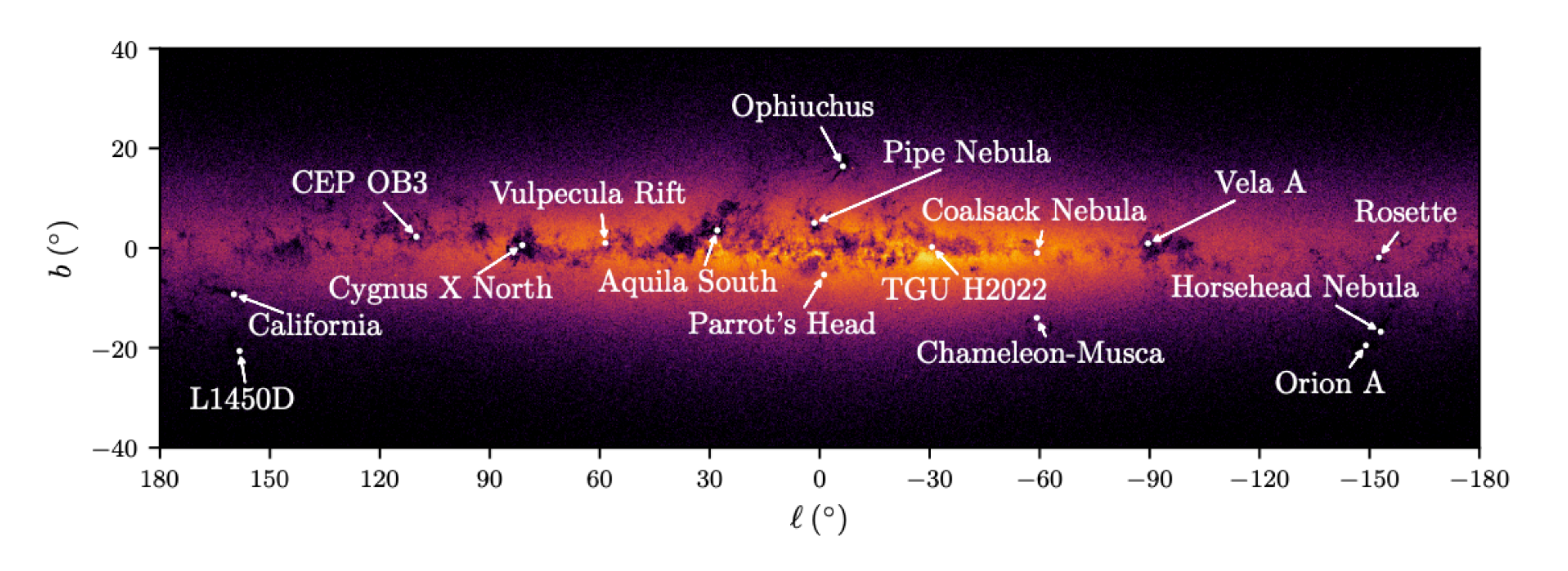
- From L22 dust map we can construct an alternative efficiency map as follows:
  - Within ~2.5 kpc, Gaia is expected to be complete even accounting for dust extinction
  - Using L22 dust-corrected magnitudes, re-apply 4 kpc magnitude cut
  - Train flow on surviving stars to learn $n_{L22}(x)$
  - Compute $\epsilon_{L22}(x) = n_{L22}(x)/n_{obs}(x)$

Validating dust efficiency map

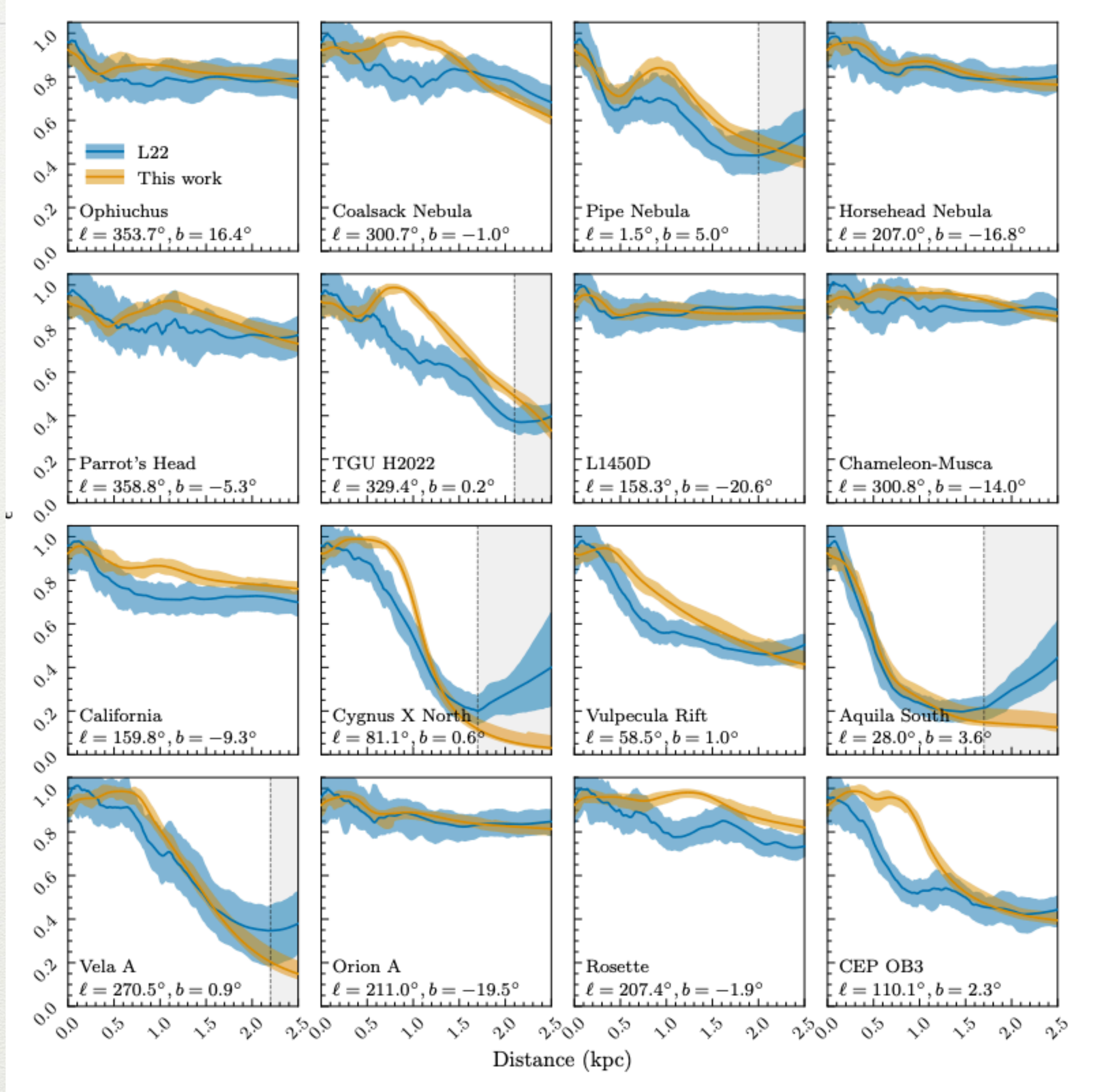




Validating dust efficiency map

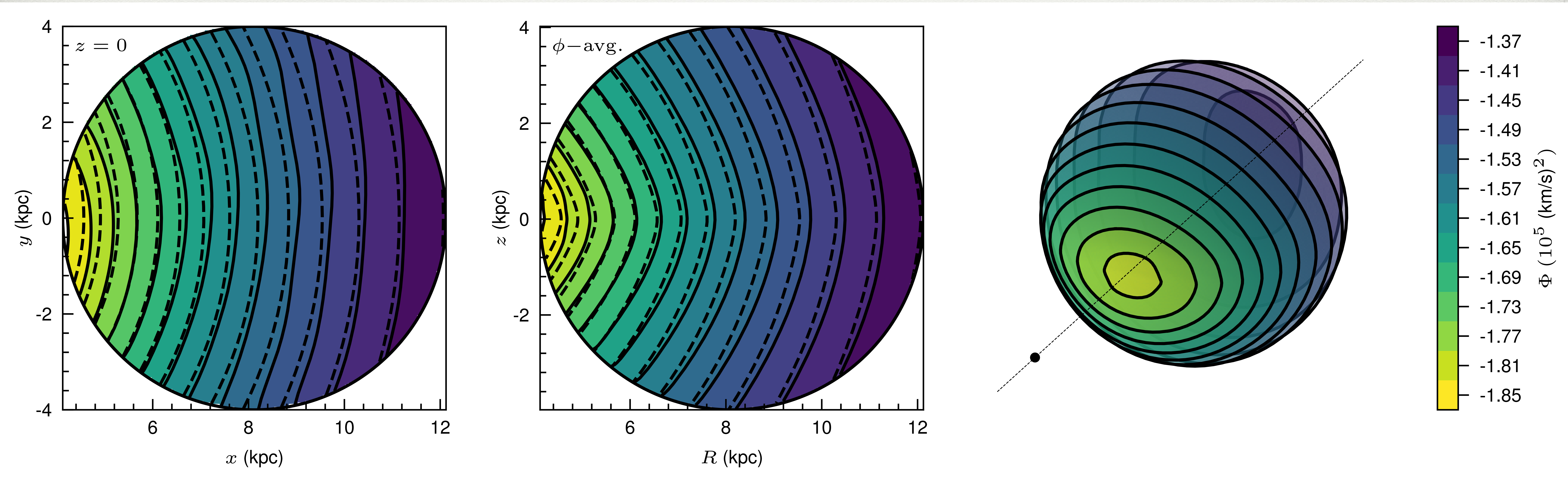


(Mostly) reproduce known dust clouds!



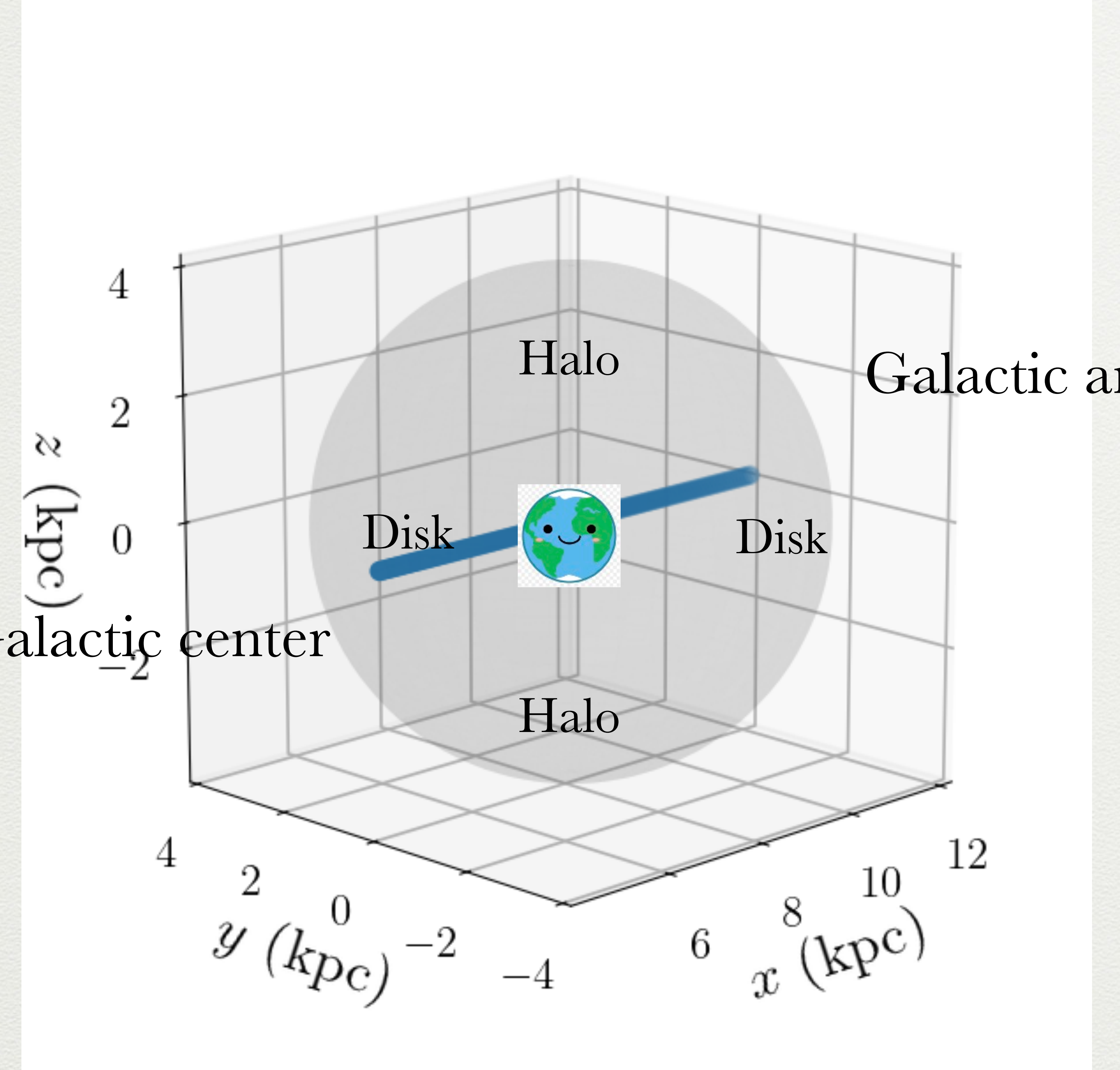
Results: Potential

Solid: our $\Phi(x)$
Dashed: MWPotential2014 $\Phi(x)$



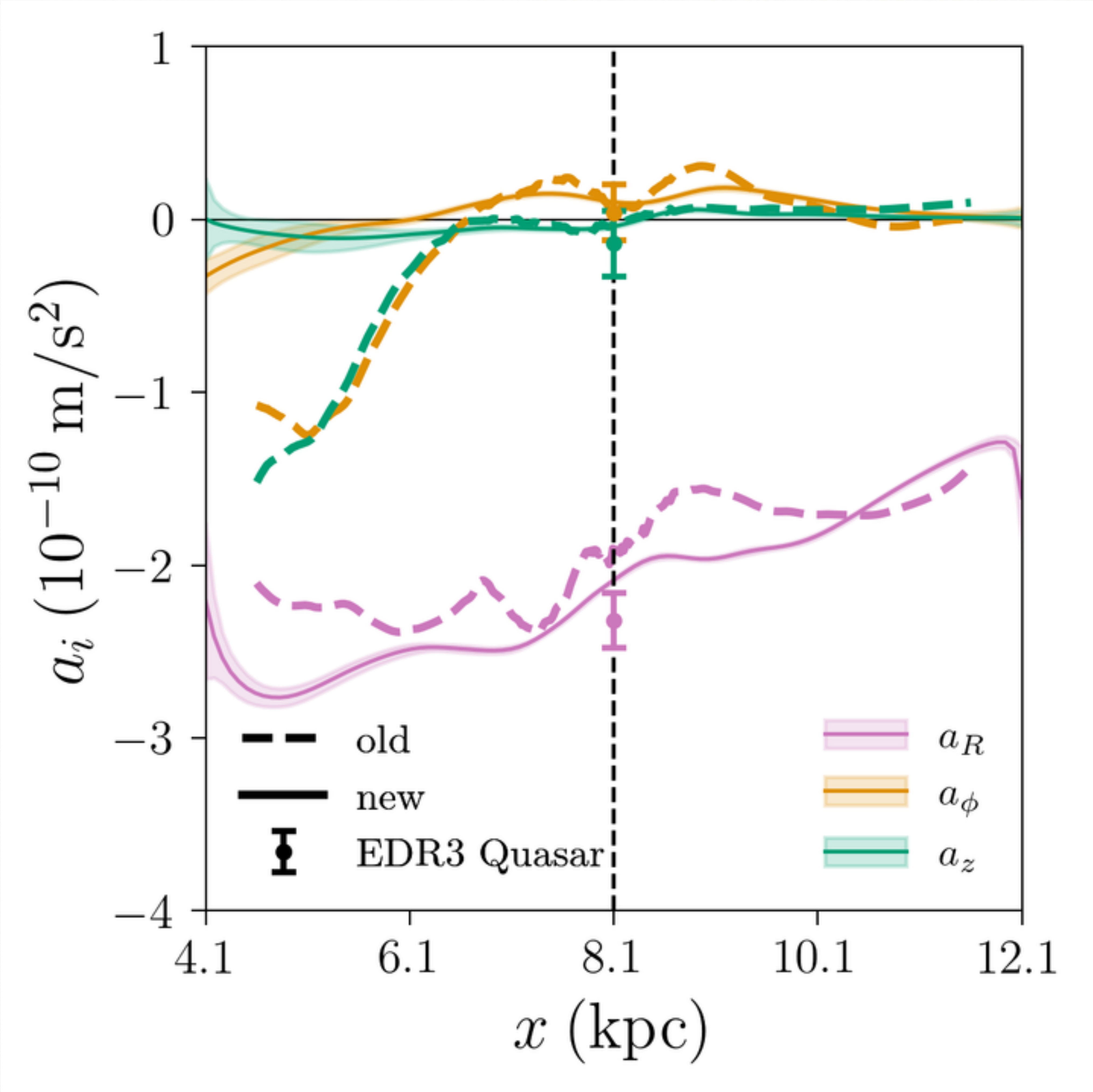
- Generally good agreement with gala's MWPotential2014!
- Can recover smooth and reasonable potential even in highly dust-obscured disk

Results: accelerations

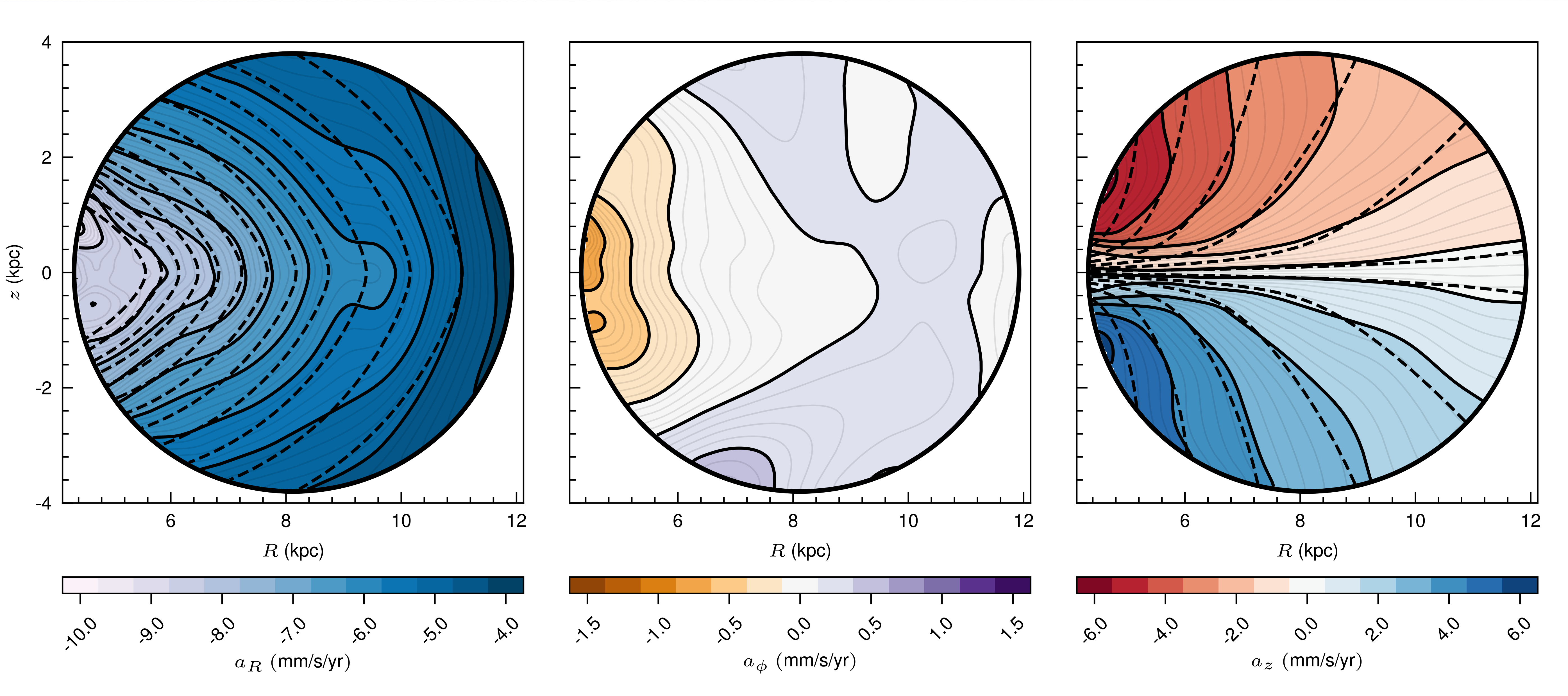


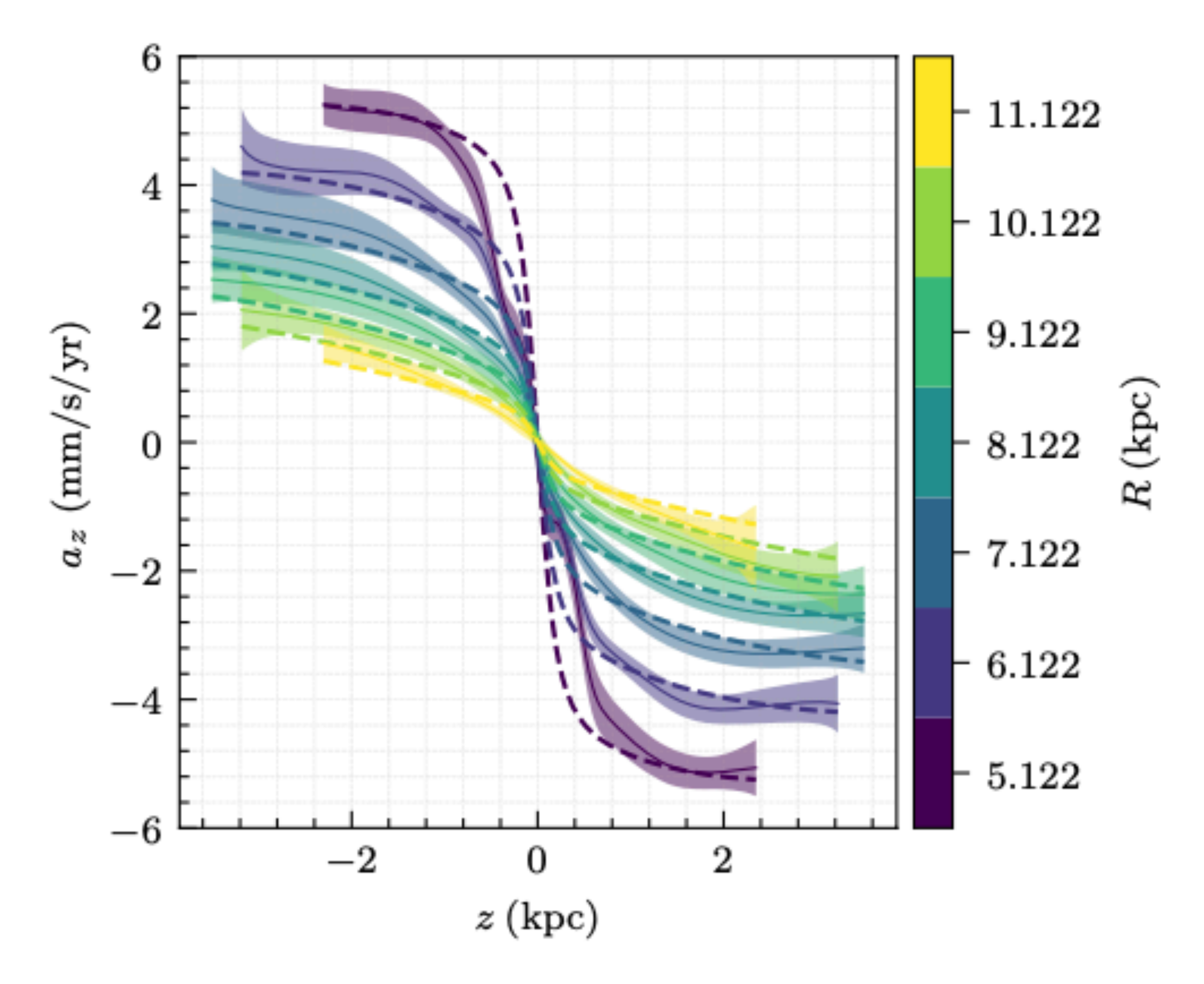
Can now access accelerations in the disk!

← Galactic center (dust clouds thicken)



Results: accelerations



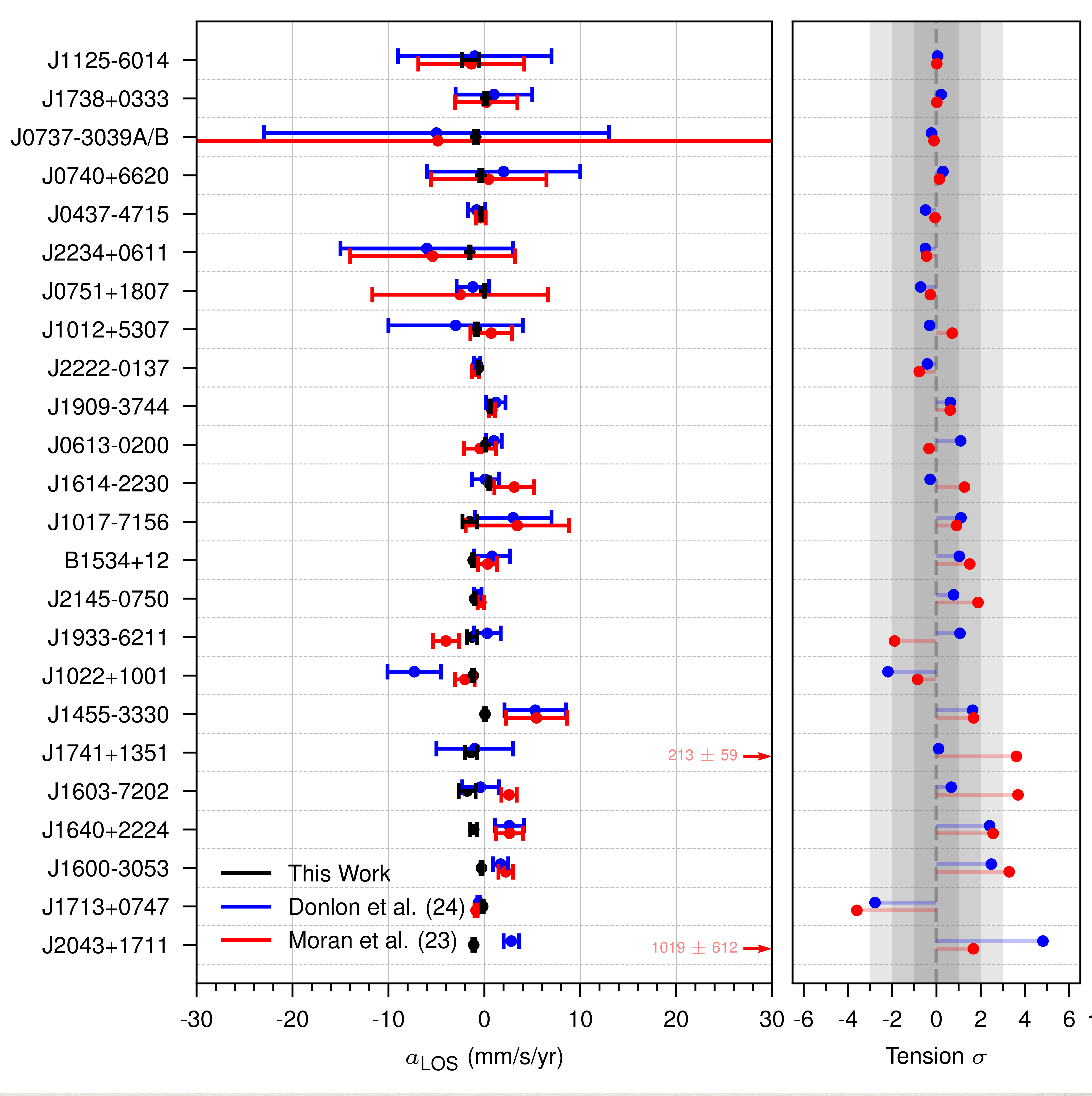


Excellent overall agreement with MWPotential2014

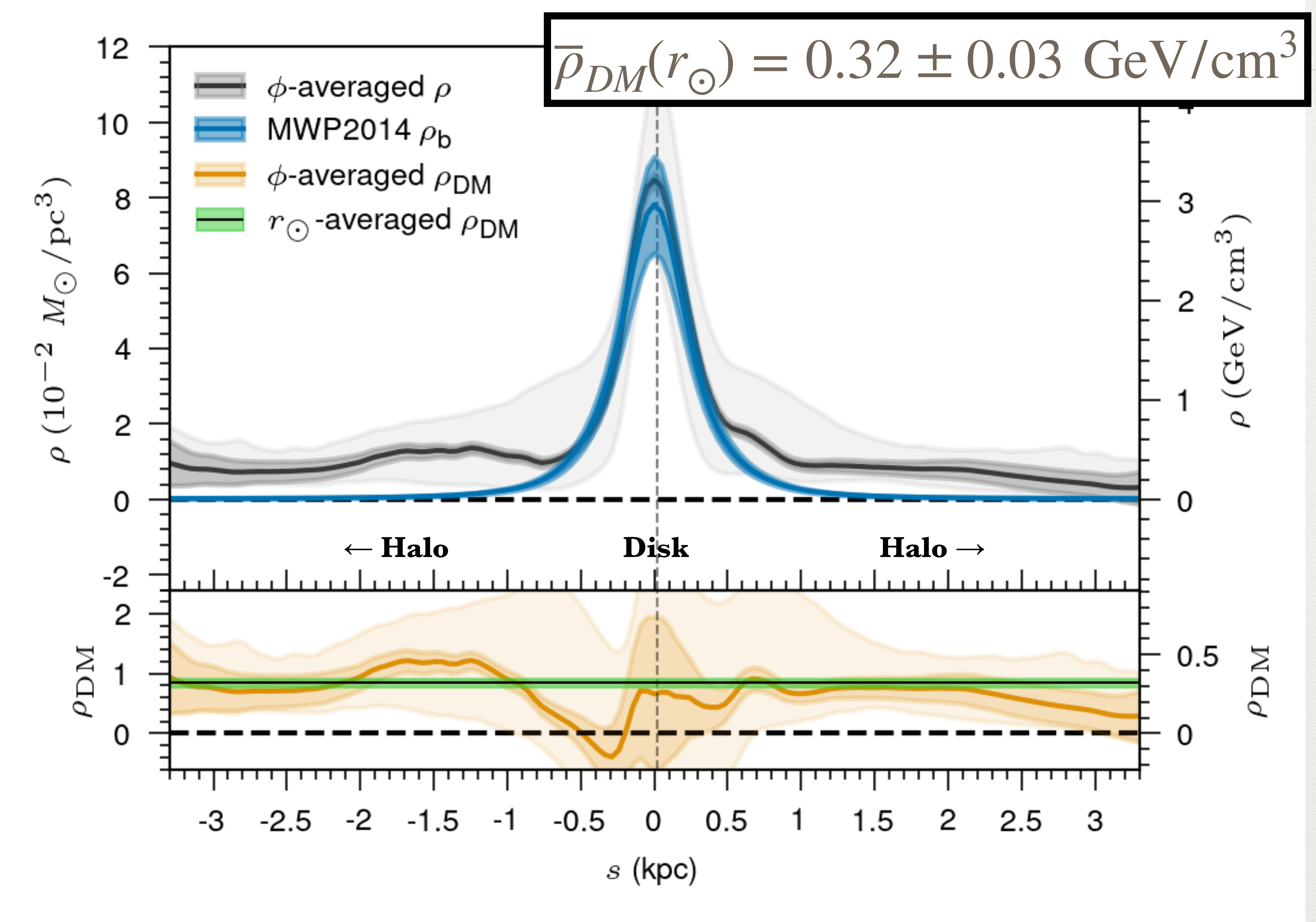
Results: accelerations

Also excellent overall agreement with recent measurements of LOS Galactic accelerations using binary pulsars

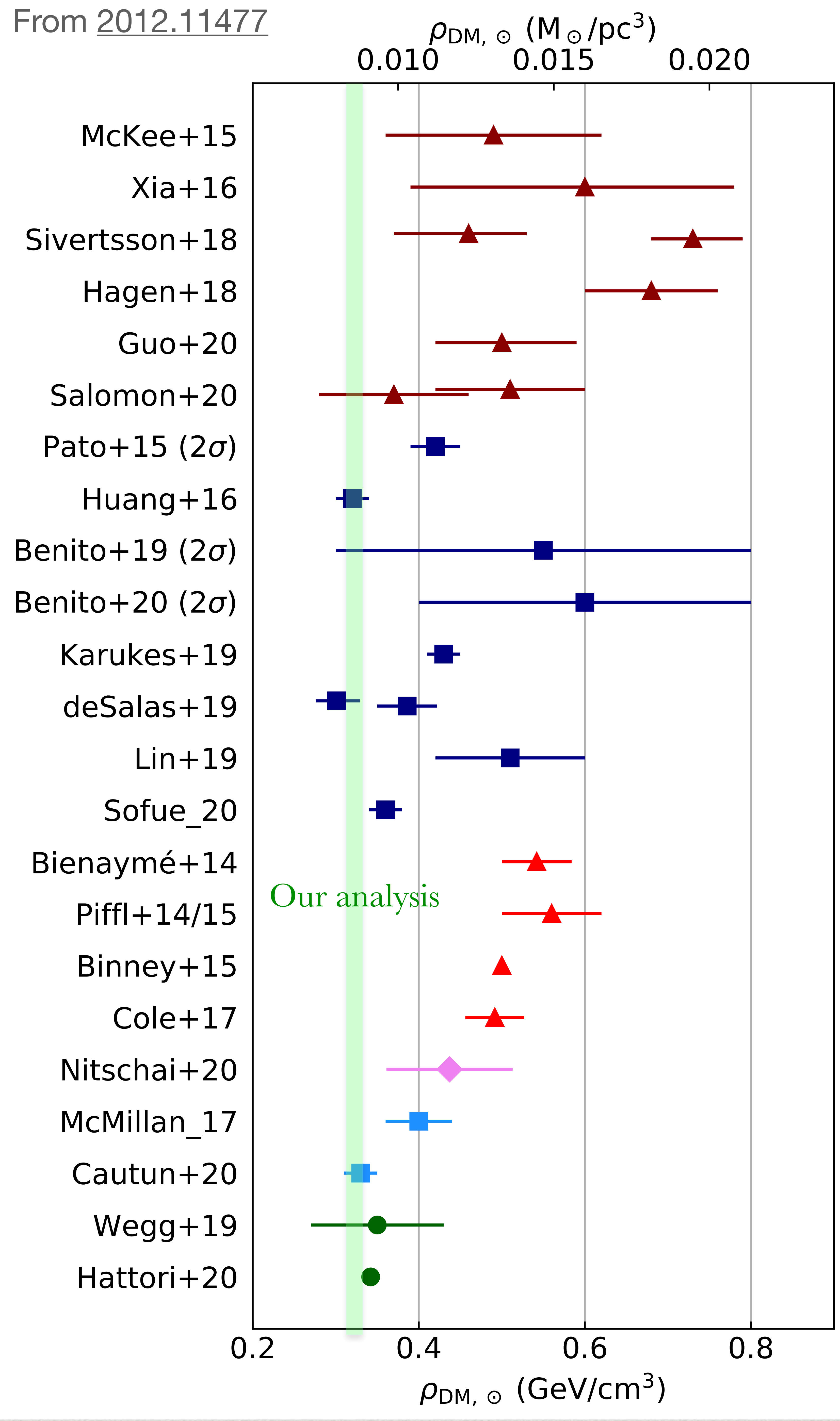
About 15% of the pulsars show significant discrepancies — could be a novel probe of disequilibrium in the Milky Way



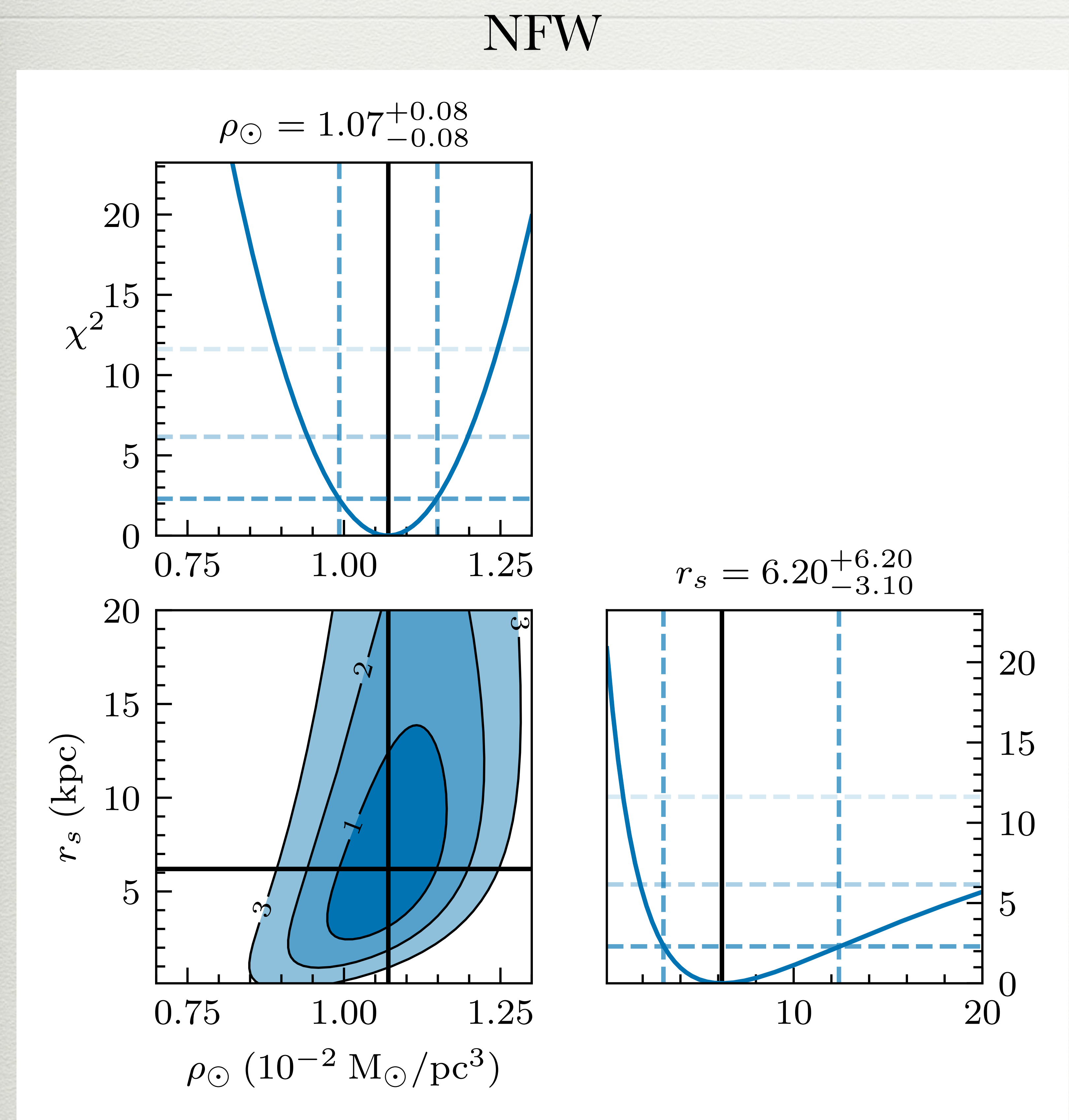
Results: mass density

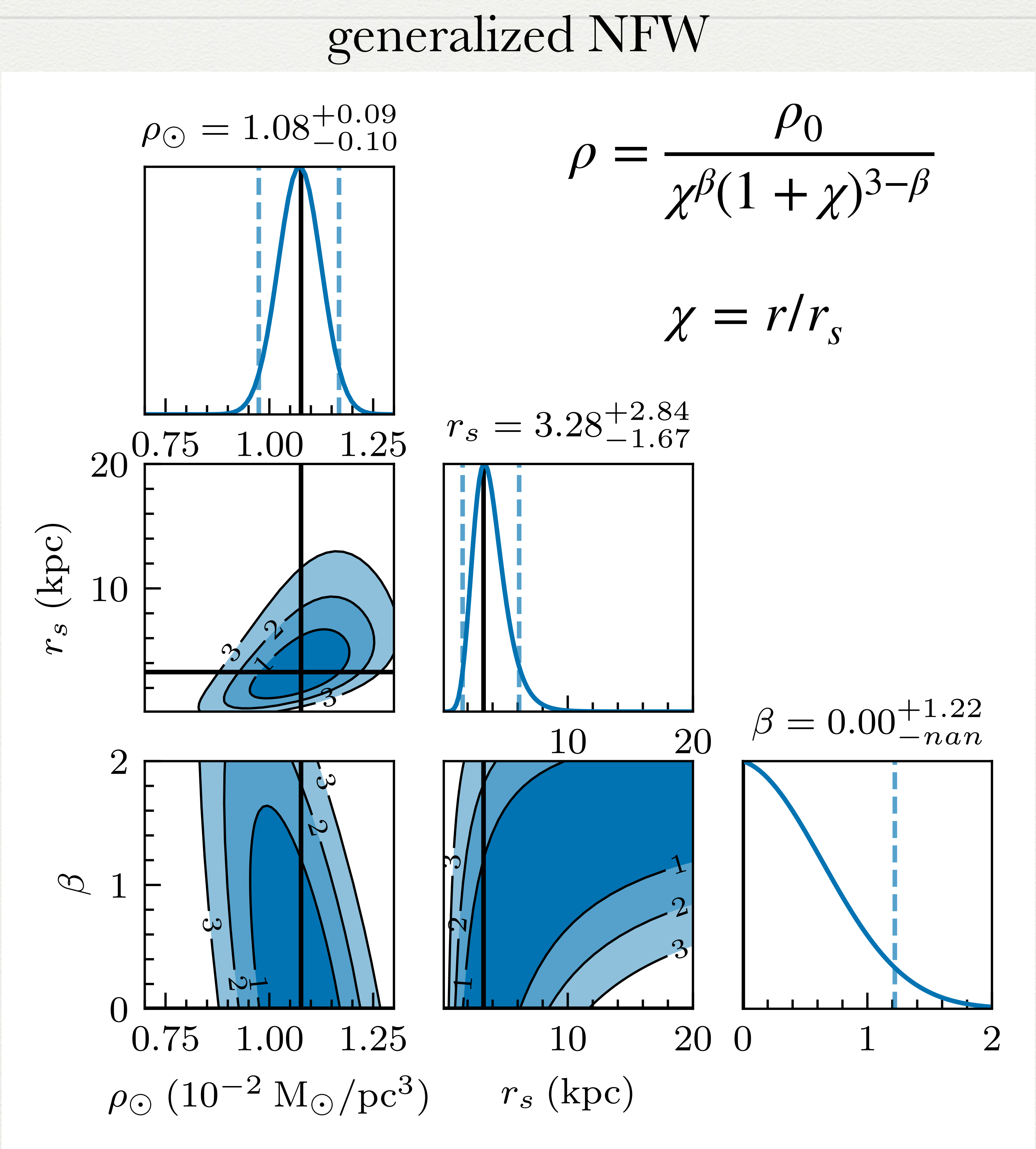


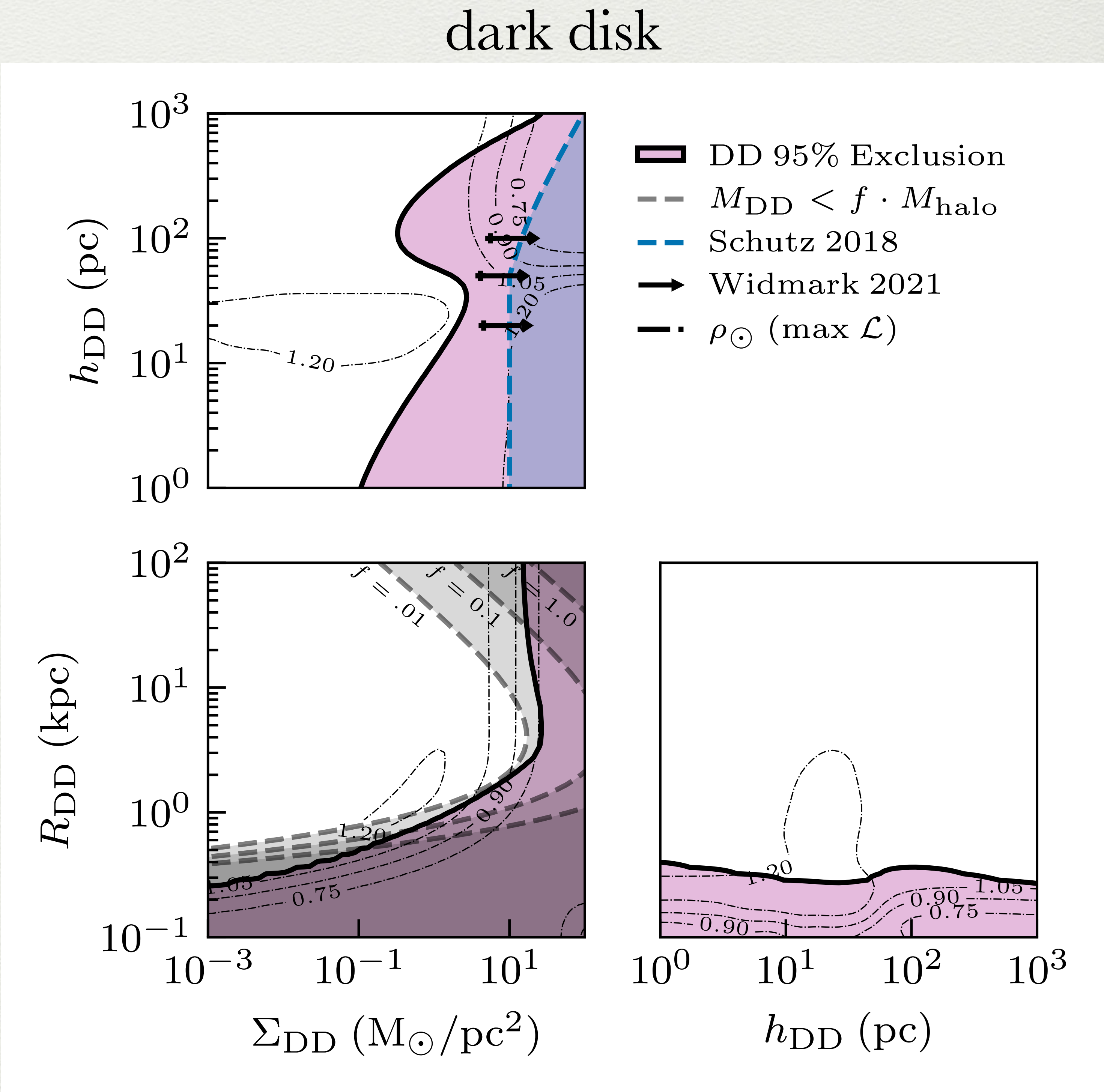
mass density vs arc height off disk, azimuthally averaged



Results: constraints on models







fits prefer smaller scale radius and highly cored profile

strongest contraints to date on thin dark disk

Summary

- We have a developed a new technique to simultaneously infer
  - the Galactic gravitational potential+acceleration field+mass density field,
  - as well as a dust efficiency map
- within a 4 kpc volume around the Sun, using normalizing flows and Gaia DR3, by leveraging the equilibrium CBE.
- Unlike previous approaches, our method is fully data-driven, model-free, unbinned and does not assume any symmetries

Outlook

- Work in progress: use our approach to set limits on dark disks, MOND…
    - **… what else?**
- Our approach should be sensitive to disequilibrium through spread of $\partial f / \partial t$ values — but how exactly?
- Gaia DR4 expected sometime in 2026 — expect further major improvements to our method!
- Large astro datasets + modern ML => huge potential for fundamental physics!

Thank you!

Credit — DALL-E + Eric Putney: “A visually striking image that represents the phrase “dark matter to dark matter, dust to dust” symbolizing the removal of interstellar dust to reveal a spiral galaxy”

Backup: uncertainties

- We attempt to quantify the following sources of uncertainty:
  - finite training statistics
  - training variance
  - Gaia measurement errors