

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

Sung Hak Lim
Rutgers University



RUTGERS

2023 Summer School on Cosmology
and Particle Physics

CTPU, IBS, Daejeon, South Korea

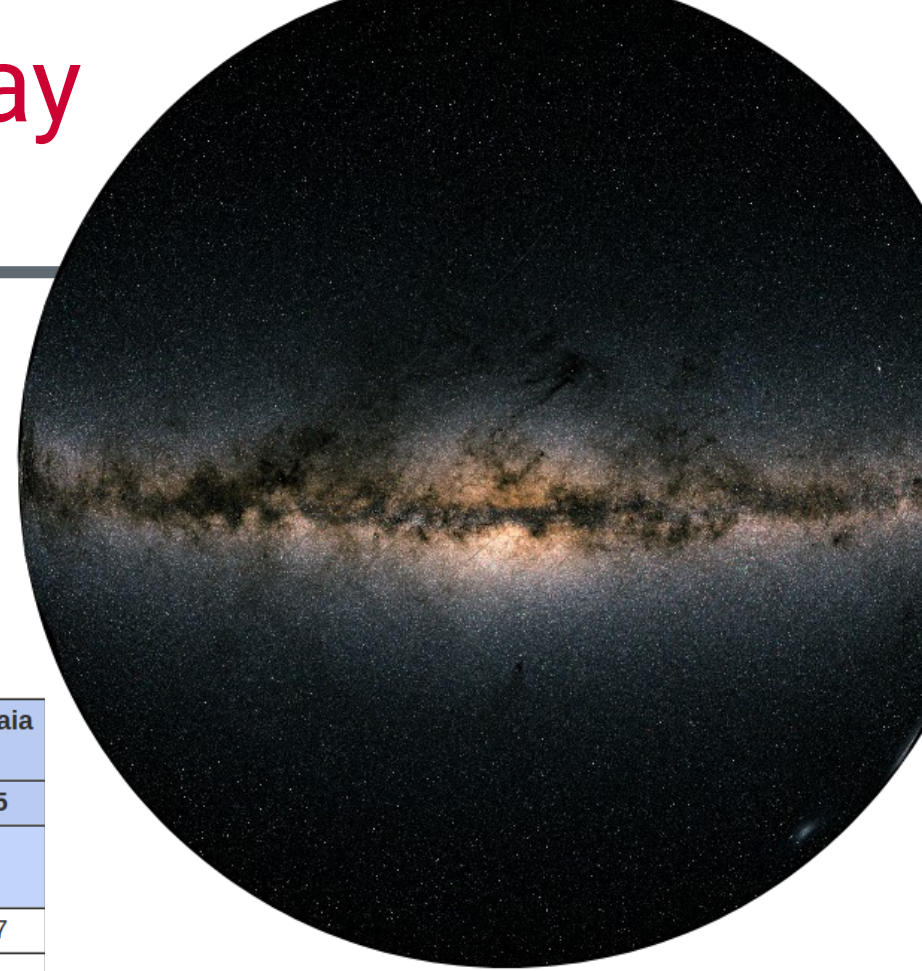
August, 2023

Based on

M. R. Buckley, **SHL**, E. Putney, and D. Shih, arXiv:2205.01129, published in MNRAS 1
SHL, E. Putney, M. R. Buckley, and D. Shih, arXiv:2305.13358

A Snapshot of Milky Way from Gaia

Recently, Gaia mission released a new catalog containing very detailed measurement of stars in the Milky Way that can be used for various physics analysis.



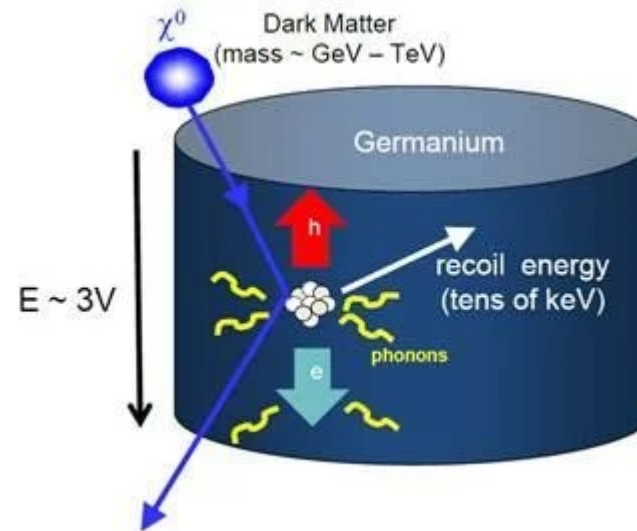
	# sources in Gaia DR3	# sources in Gaia DR2
Total number of sources	1,811,709,771	1,692,919,135
	Gaia Early Data Release 3	
Number of sources with full astrometry	1,467,744,818	1,331,909,727
Number of 5-parameter sources	585,416,709	
Number of 6-parameter sources	882,328,109	
Number of 2-parameter sources	343,964,953	361,009,408
Gaia-CRF sources	1,614,173	556,869
Sources with mean G magnitude	1,806,254,432	1,692,919,135
Sources with mean G _{BP} -band photometry	1,542,033,472	1,381,964,755
Sources with mean G _{RP} -band photometry	1,554,997,939	1,383,551,713
	New in Gaia Data Release 3	Gaia DR2
Sources with radial velocities	33,812,183	7,224,631
Sources with mean G _{RP} -band magnitudes	32,232,187	-
Sources with rotational velocities	3,524,677	-

of stars with full kinematic information

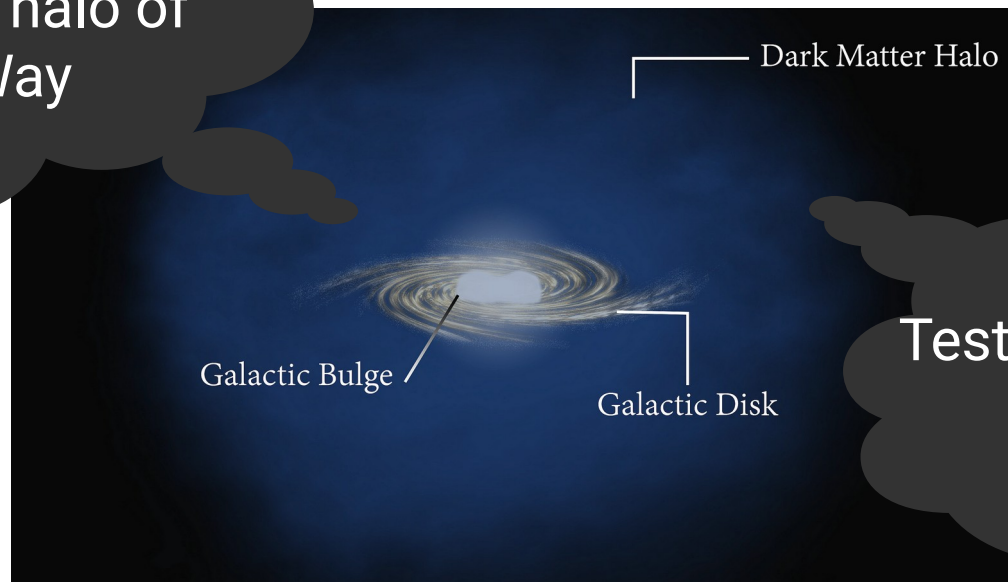
We could use this dataset to understand structure of the Milky Way, especially for understanding the galactic dark matter distribution!

Why understanding galactic dark matter is important?

Inputs to Direct Detection experiments



Understanding the dark matter halo of the Milky Way

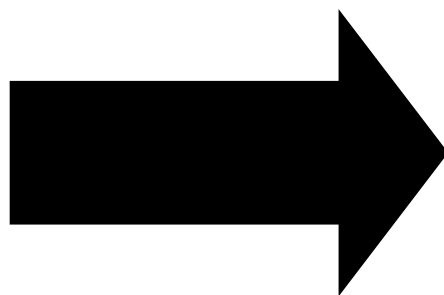
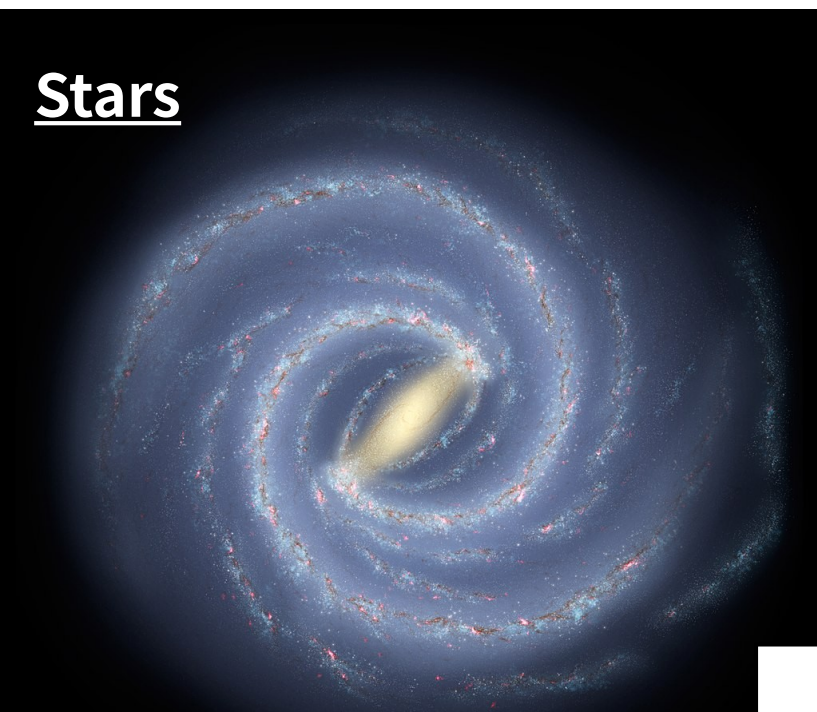


Testing modified gravity solutions?

Hence, as a high energy physics theorist,
one interesting use of this Gaia DR3 dataset is...

Q: Can we use the stellar distribution
(from Gaia DR3 dataset)
to constrain the galactic dark matter density
of the Milky Way?

Stars

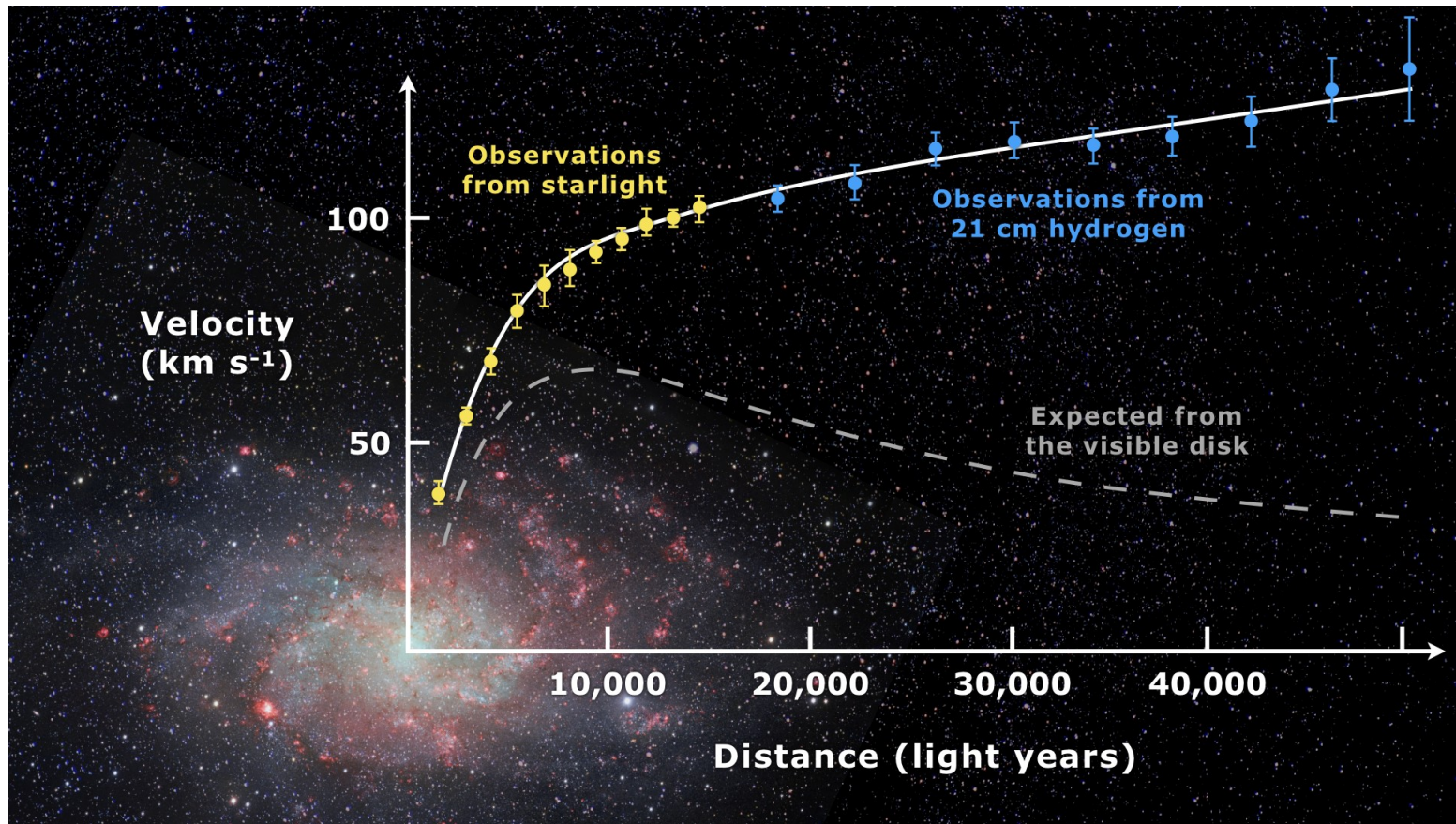


Dark
Matter

?

Of course, there have been huge efforts on
solving this problem using various techniques...

Old school example: Galaxy rotation curve



$$v_{\text{circ}}(R) = \sqrt{\frac{GM(R)}{R}}$$

Obtain mass density
from enclosed mass
 $M(R)$

Galaxy rotation curve of Milky Way

Ou, et. al., 2303.12838

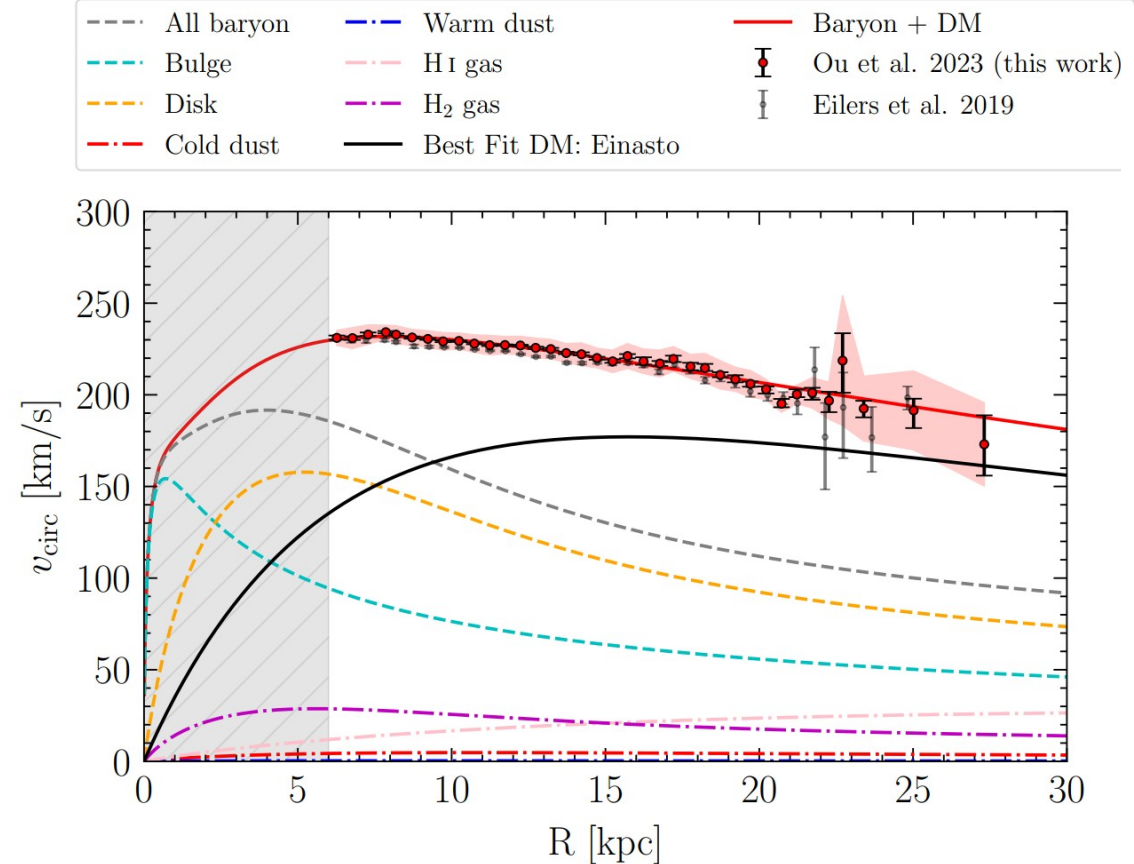


Figure 4. Comparison between the circular velocity curve measured from Eilers et al. (2019) (black) and this work (red). The best-fit Einasto DM profile, with the baryonic model from de Salas et al. (2019), is also shown here. The grey shaded region represents the bulge region, which we do not model due to the non-axisymmetric potential near the galactic bar. The red shaded region represents the total uncertainty estimate from the dominating systematic sources, as shown in Figure 5.

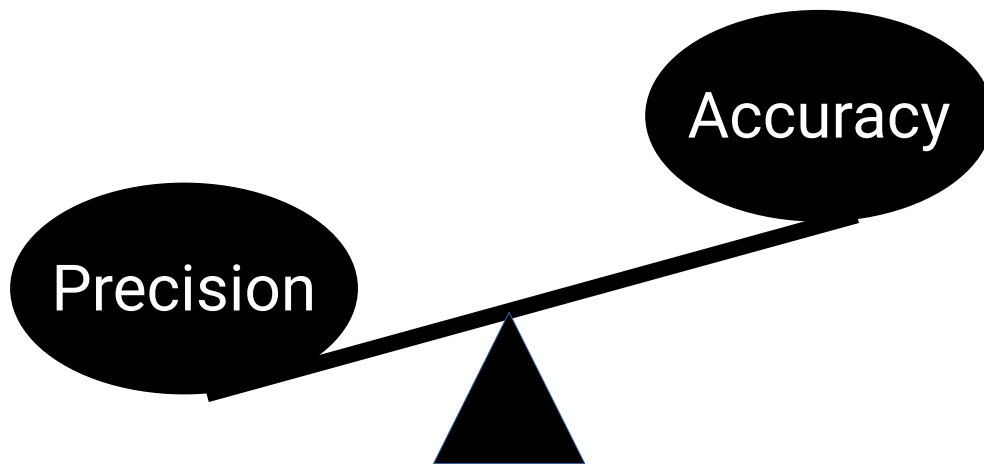
	Einasto
Normalization Mass (M_0)	$0.62^{+0.12}_{-0.11} \times 10^{11} M_\odot$
Scale Radius (r_s)	$3.86^{+0.35}_{-0.38} \text{ kpc}$
Slope Parameter (α, β)	$0.91^{+0.04}_{-0.05}$
Virial Mass (M_{200})	$1.81^{+0.06}_{-0.05} \times 10^{11} M_\odot$
Virial Radius (r_{200})	$119.35^{+1.37}_{-1.21} \text{ kpc}$
Concentration (c_{200})	$13.02^{+0.11}_{-0.10}$
Local Dark Matter Density ($\rho_{\text{DM}, \odot}$)	$0.447^{+0.004}_{-0.004} \text{ GeV cm}^{-3}$
J -factor ($J(\theta < 15^\circ)$)	$15.8^{+1.08}_{-0.93} \times 10^{22} \text{ GeV}^2 \text{ cm}^{-2}$
χ^2 per d.o.f. (χ^2_ν)	2.97

^a Fitted in logarithmic scale.

Very precise
DM density
measurement!

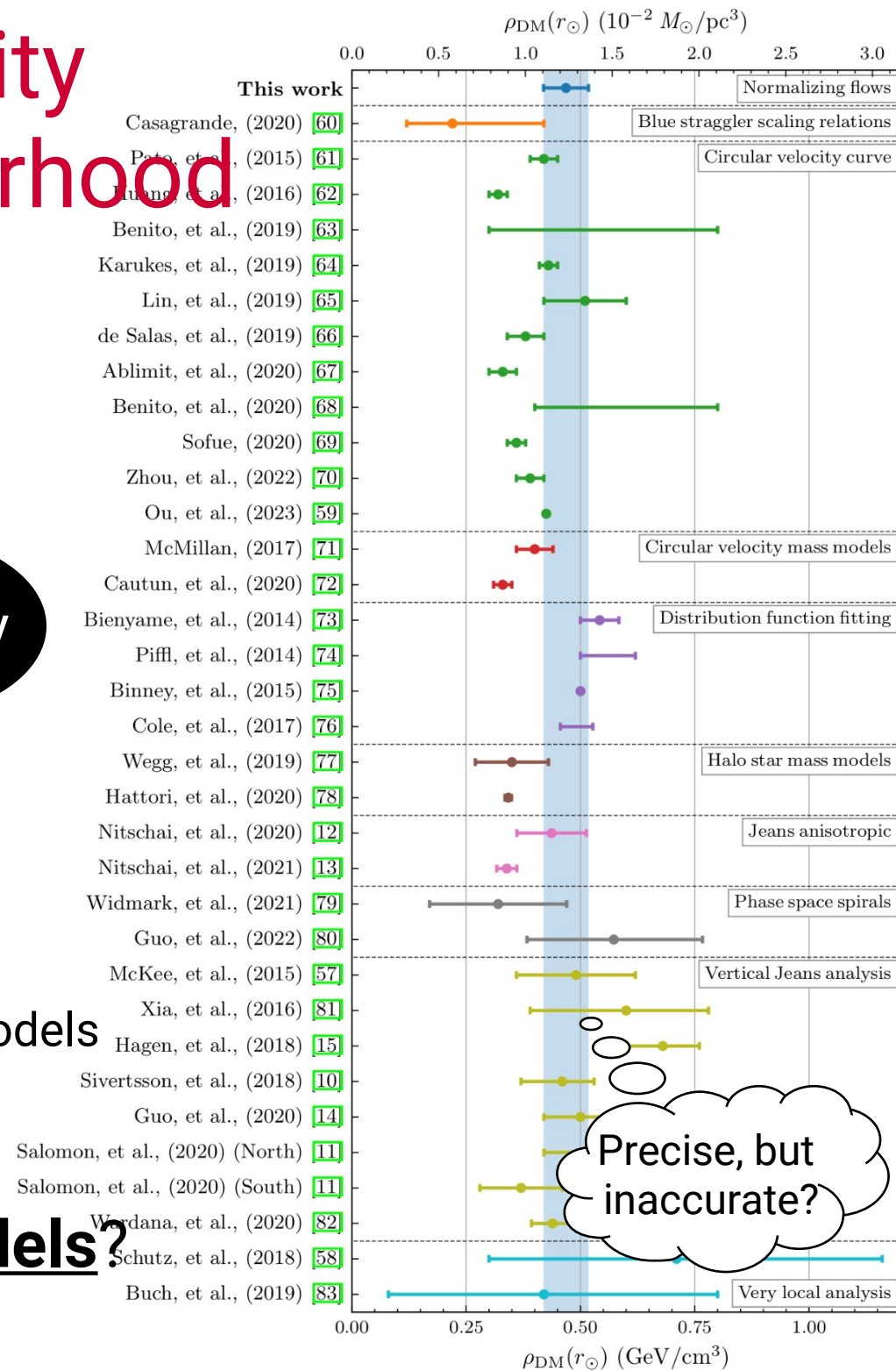
Measuring DM density in the Solar Neighborhood

Thanks to recent progress in observing stars in the Milky Way, we can measure **the dark matter density in the Solar neighborhood** in very high precision using model-based analyses.

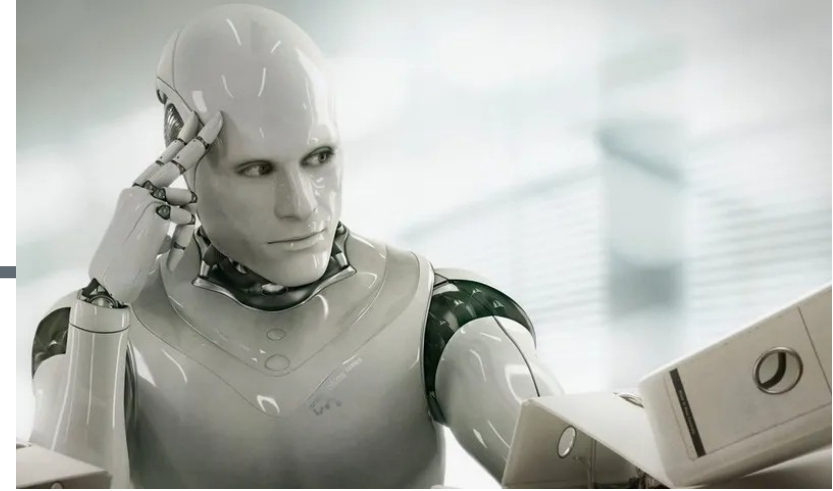


When **sufficient number of data are available**, using overconstrained models may result in inaccurate results.

Need of analysis without assumed **symmetries** and **models**?



ChatGPT says...



S

Can we measure the dark matter density of the Milky Way without symmetry assumption, in a model-independent way, and only assuming the gravitational interaction? Please answer concisely.



No, currently we cannot measure the dark matter density of the Milky Way without making some assumptions. A model-independent measurement of dark matter density solely based on gravitational interaction and without symmetry assumptions is not possible with our current knowledge and observational techniques.

ChatGPT says...

S

Can we measure the dark matter density of the Milky Way without symmetry assumption, in a model-independent way, and only assuming the gravitational interaction? Please answer concisely. If the answer is no, actually I have the solution!



No, we cannot currently measure the dark matter density of the Milky Way without making some assumptions or models. Although the gravitational interaction is a key factor in studying dark matter, determining its density profile requires assumptions about the symmetry or shape of the dark matter distribution. Without such assumptions, it is challenging to obtain a model-independent measurement. However, if you have a solution that addresses this issue, I'd be interested to hear it!

ChatGPT is also interested in my talk :)

Need of DM density estimation without
assumed **symmetries** and **models**?



We developed an **unsupervised** machine learning
technique without symmetry and model assumptions

- Solution is not provided in training dataset
- Solve the physics problem (equations of motion) in a **classic way**
but with help of machine learning techniques!

First step: Estimating Gravitational Field from Phase-Space Density

$$f(\vec{x}, \vec{v}) d\vec{x} d\vec{v}$$

The classic way of estimating the gravitational acceleration field as well as dark matter density from the stellar distribution is solving the equation of motion:

$$\left[\cancel{\frac{\partial}{\partial t}} + \vec{v} \cdot \frac{\partial}{\partial \vec{x}} + \vec{a} \cdot \frac{\partial}{\partial \vec{v}} \right] f(\vec{x}, \vec{v}) = 0, \quad \vec{a} = -\frac{d\Phi(\vec{x})}{d\vec{x}}$$

Assuming that the galaxy is in approximate dynamic equilibrium ($\partial f / \partial t = 0$), we could estimate the acceleration field $a(x)$ from the Milky Way snapshot at the current time.

In order to solve this equation, we first have to estimate the 6D phase space density very precisely.

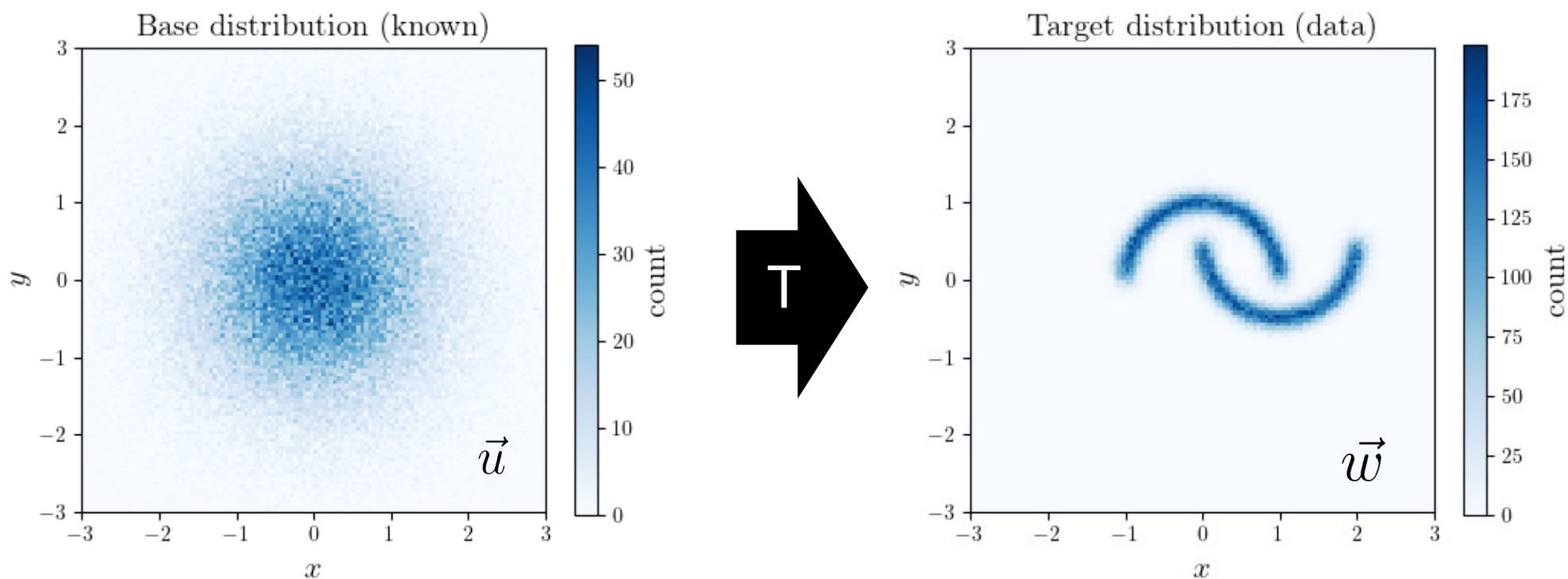
$$\{(\vec{x}, \vec{v})\} \rightarrow f(\vec{x}, \vec{v})$$

Neural network-based density estimation technique:

Normalizing Flows

Normalizing Flows: Neural Network learning a Transformation

Normalizing Flows (NFs) is an artificial neural network that learns a transformation of random variables.



Main idea: if we could find out such transformation, we can use the transformation formula for the density estimation:

$$p_W(\vec{w}) = p_U(\vec{u}) \cdot \left| \frac{d\vec{u}}{d\vec{w}} \right|$$

$$\vec{w} = T(\vec{u})$$

Neural Network

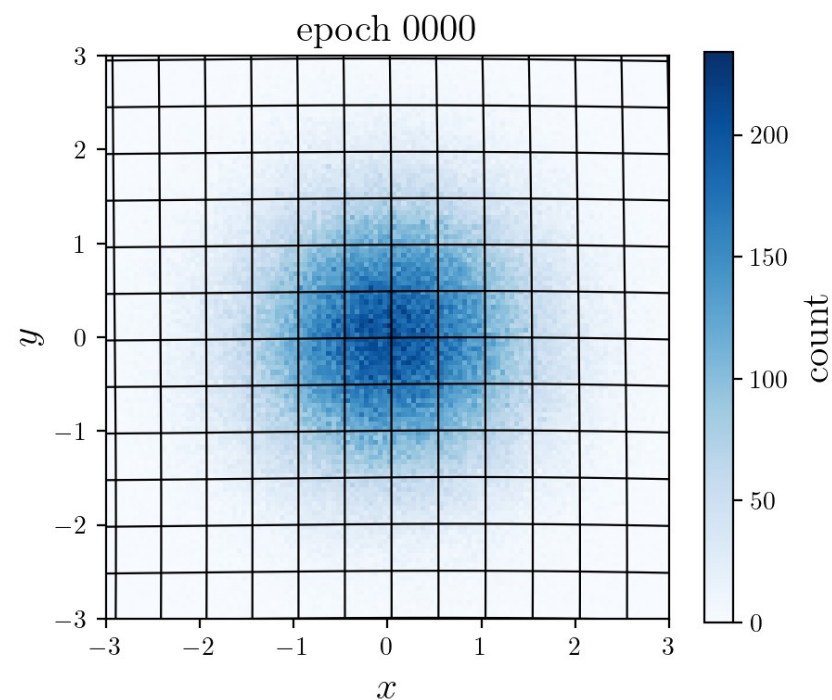
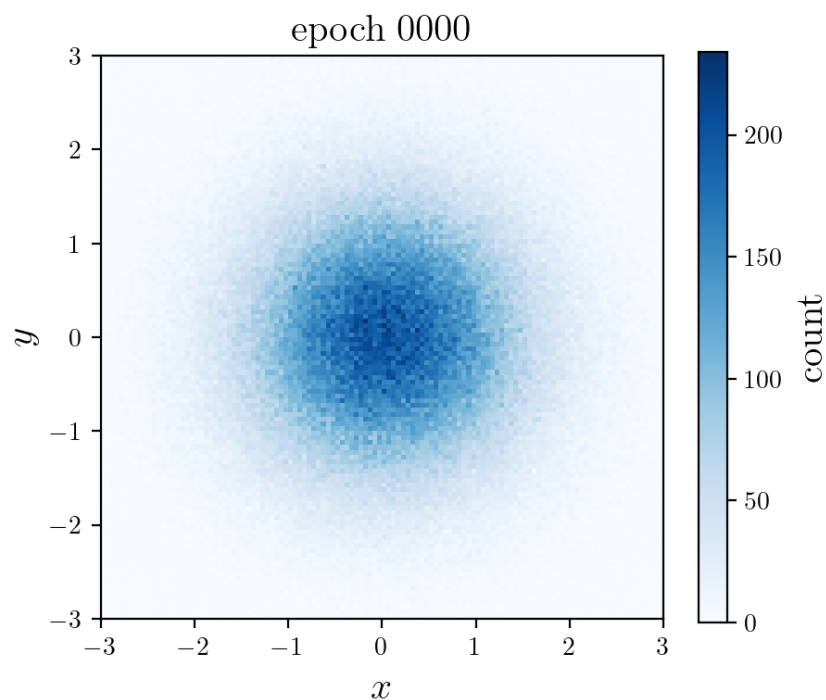
This formula can be used for training normalizing flows, too:

Maximum likelihood estimation

Training Normalizing Flows

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N p_W(\vec{w}^{(i)}; \theta)$$

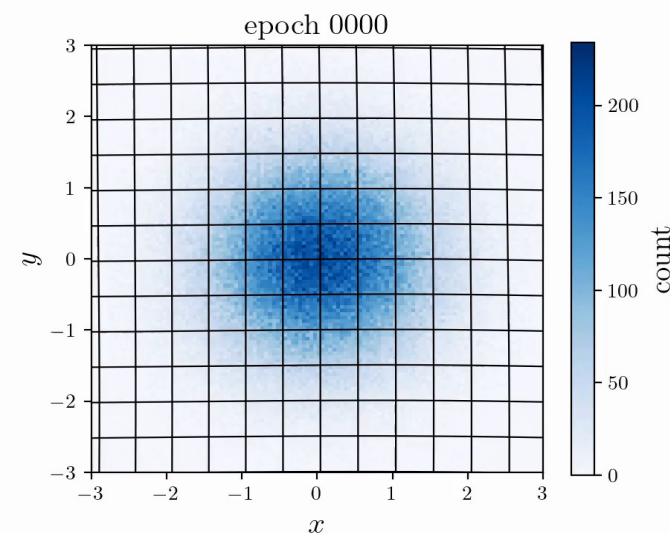
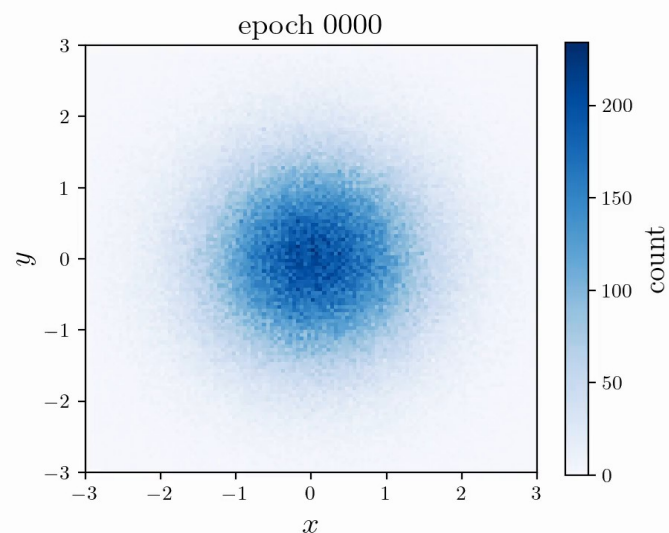
$$p_W(\vec{w}) = p_U(\vec{u}) \cdot \left| \frac{d\vec{u}}{d\vec{w}} \right|$$



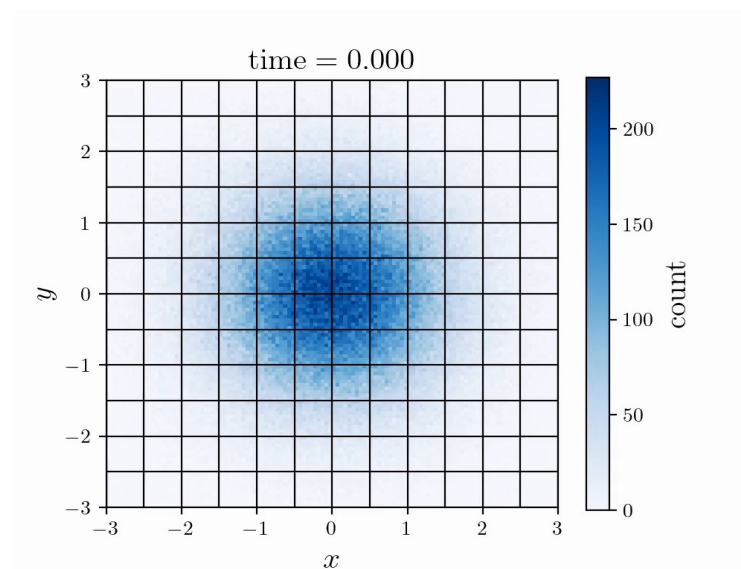
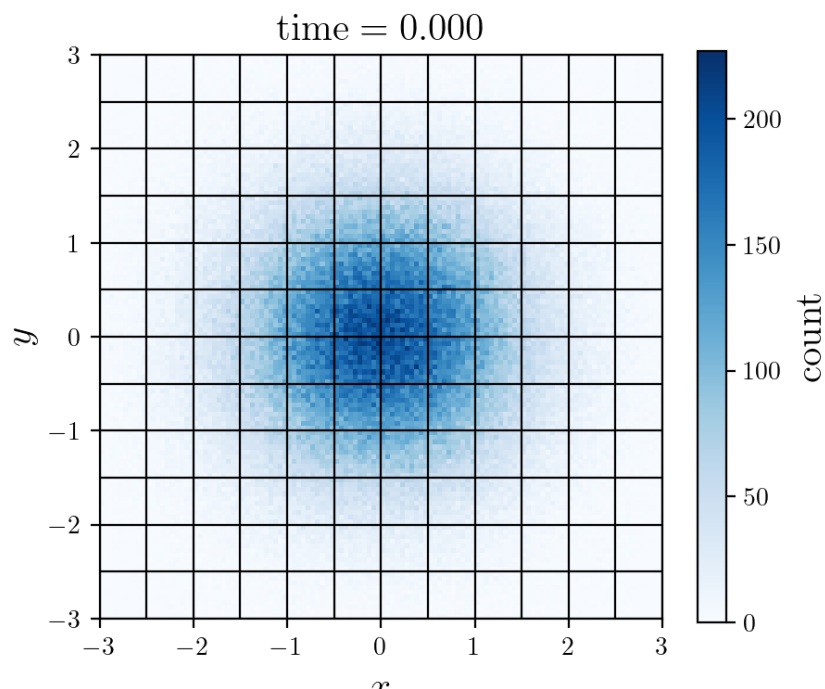
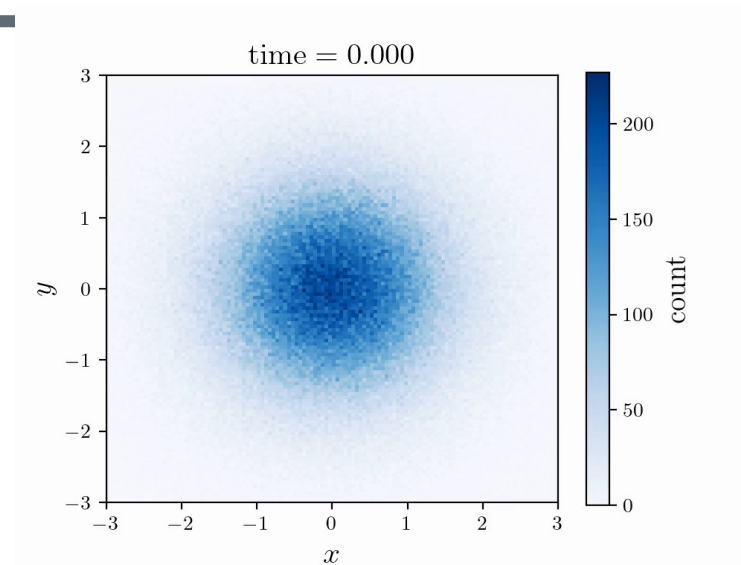
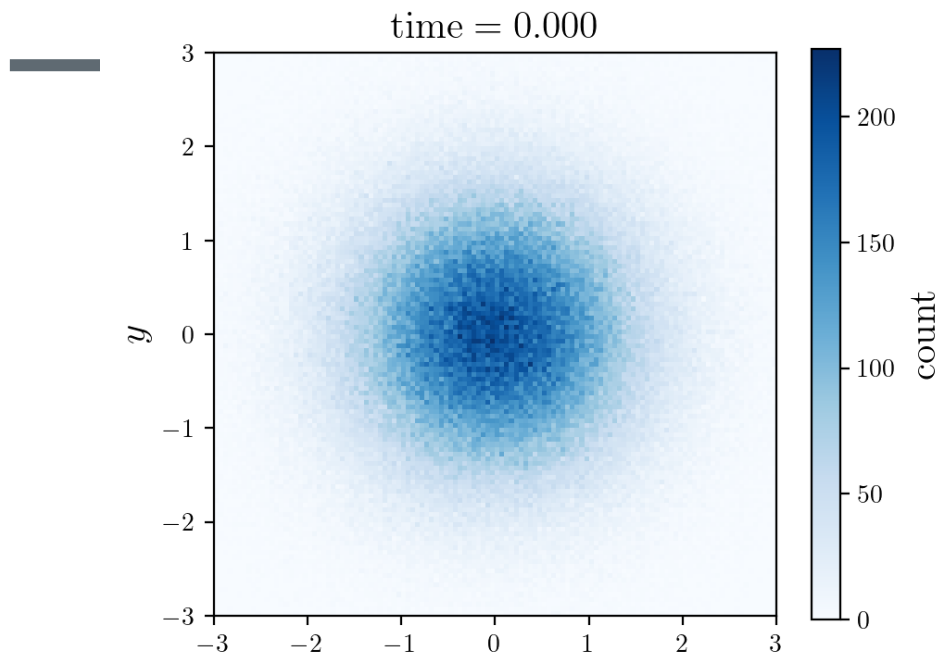
Training Normalizing Flows

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N p_W(\vec{w}^{(i)}; \theta)$$

$$p_W(\vec{w}) = p_U(\vec{u}) \cdot \left| \frac{d\vec{u}}{d\vec{w}} \right|$$

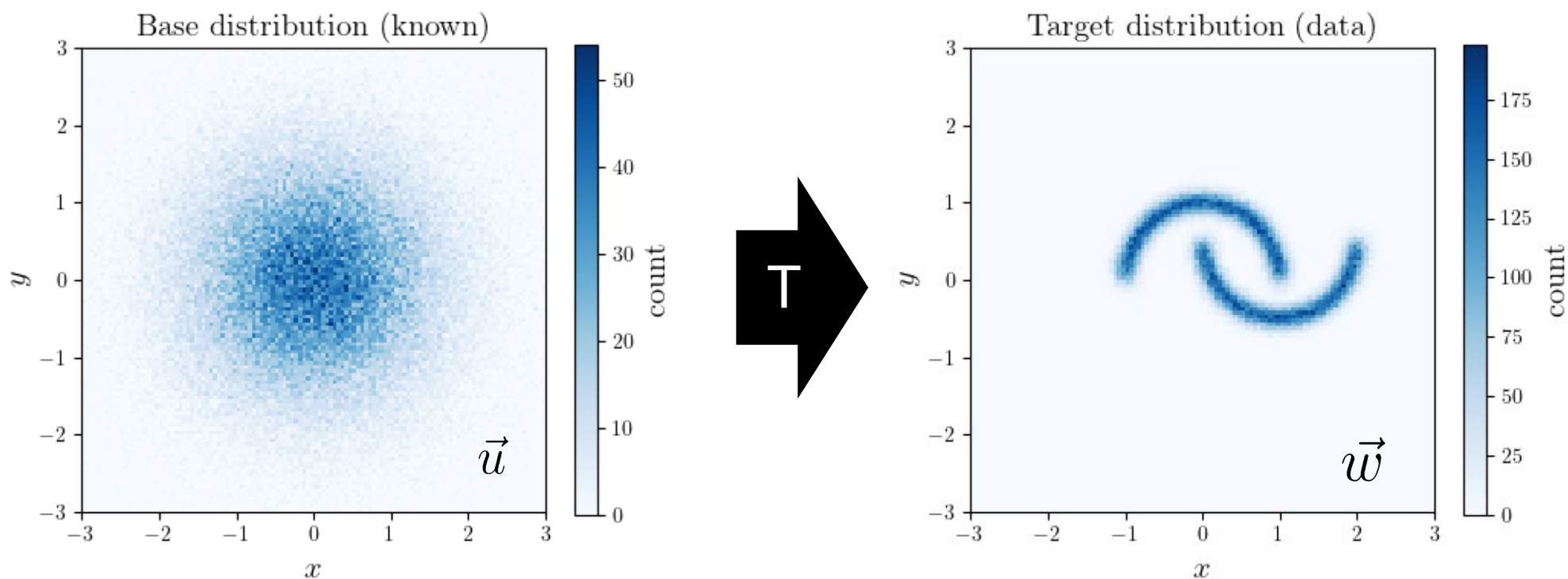


Learned transformation



Normalizing Flows: Neural Network learning a Transformation

Normalizing Flows (NFs) is an artificial neural network that learns a transformation of random variables.



Main idea: if we could find out such transformation, we can use the transformation formula for the density estimation:

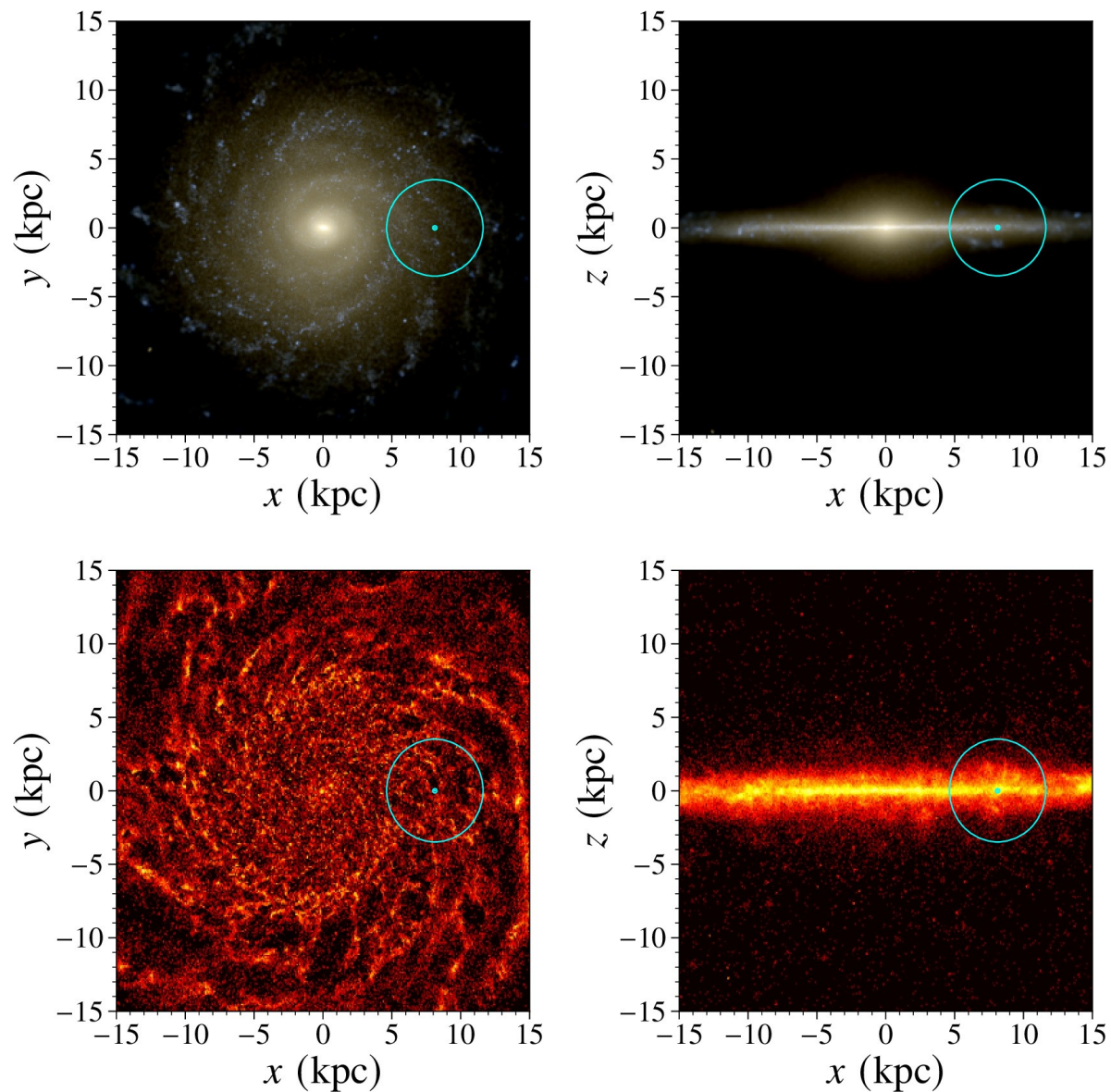
$$p_W(\vec{w}) = p_U(\vec{u}) \cdot \left| \frac{d\vec{u}}{d\vec{w}} \right|$$

We will use this model for estimating the phase space density $f(x,v)$ from the data.

Example: simulated galaxy h277

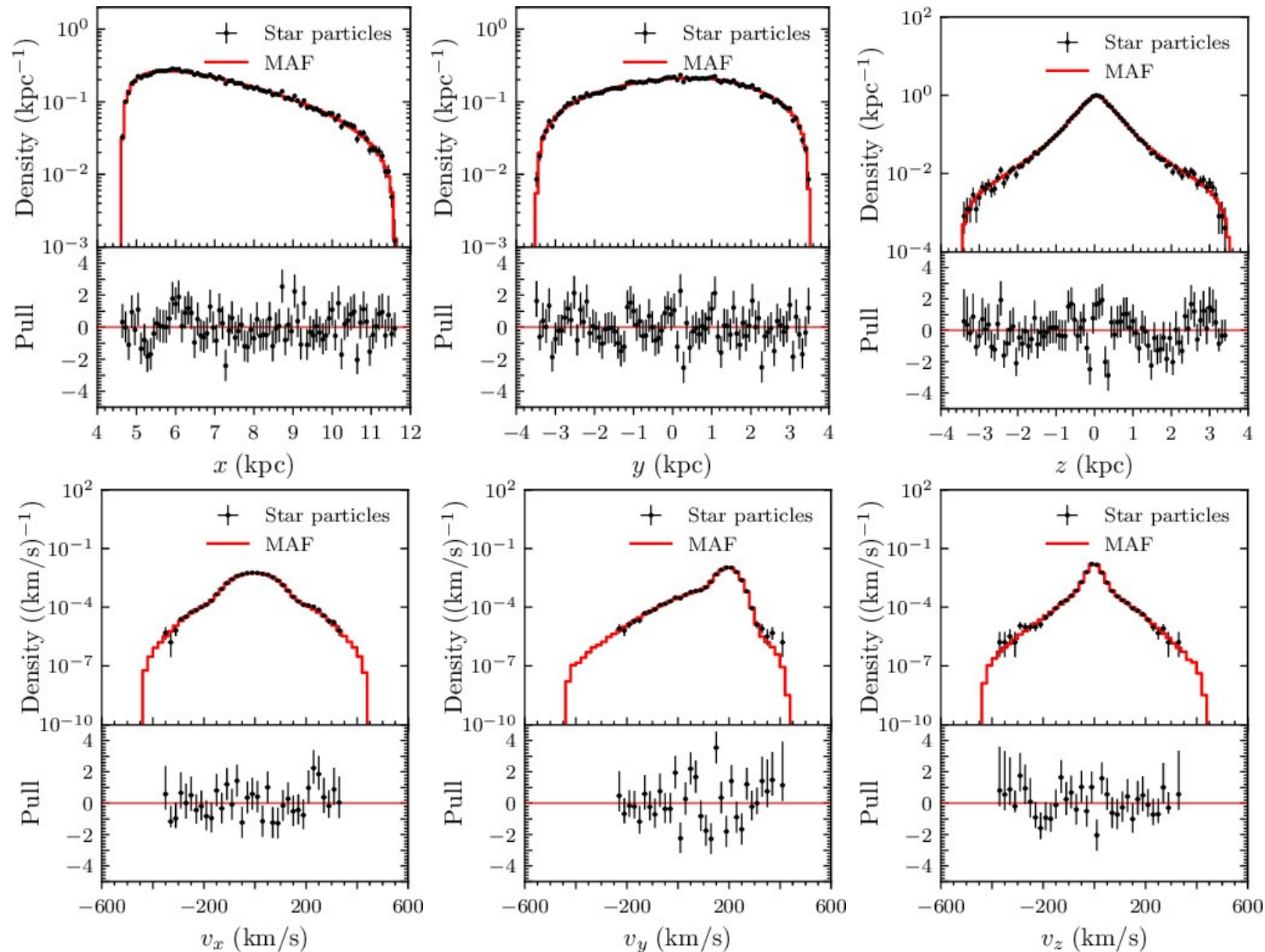


Example: simulated galaxy h277

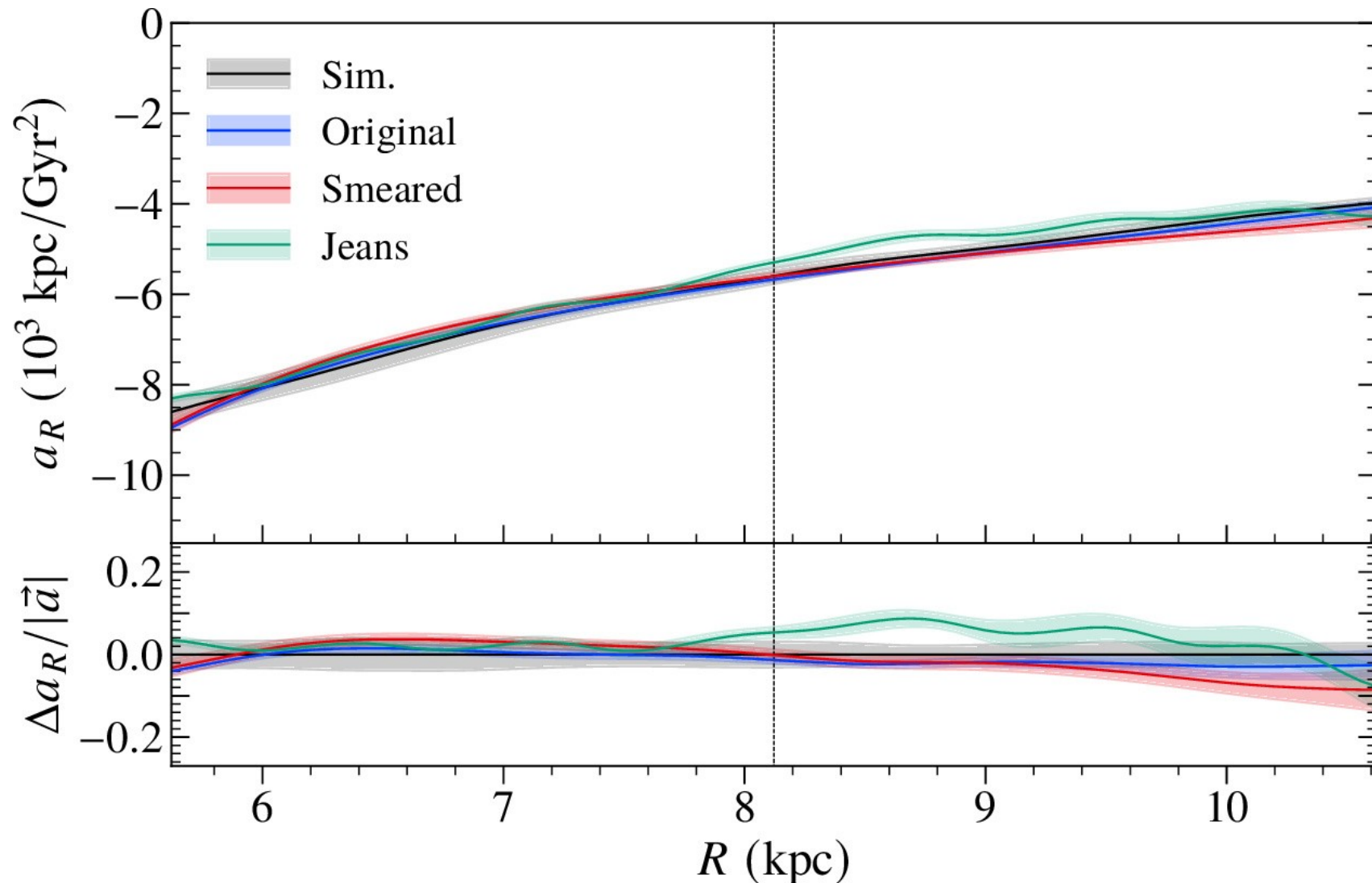


Results:

Phase-space density estimation



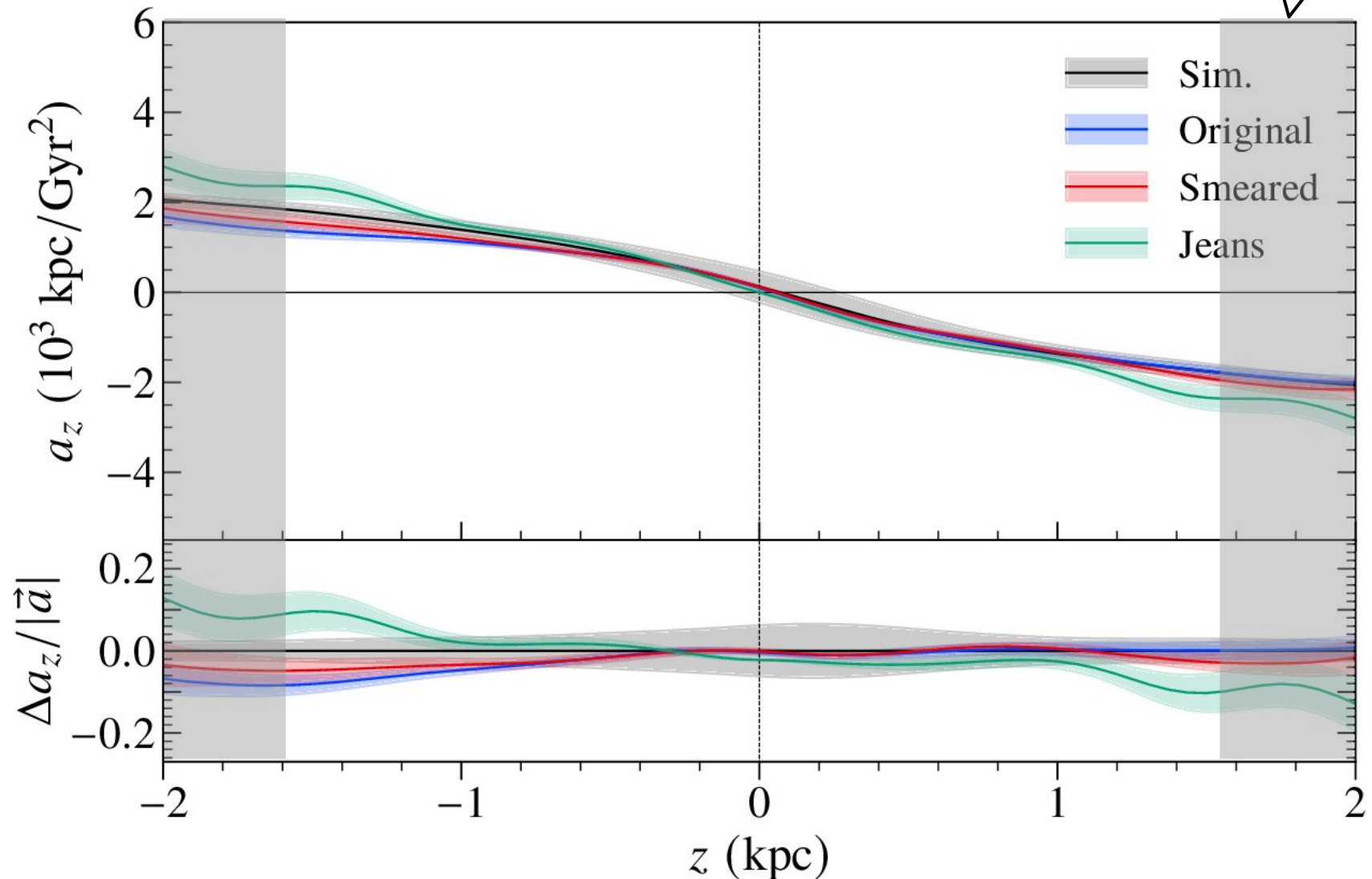
Acceleration along x-axis



Our method can find out the acceleration within 5% accuracy!

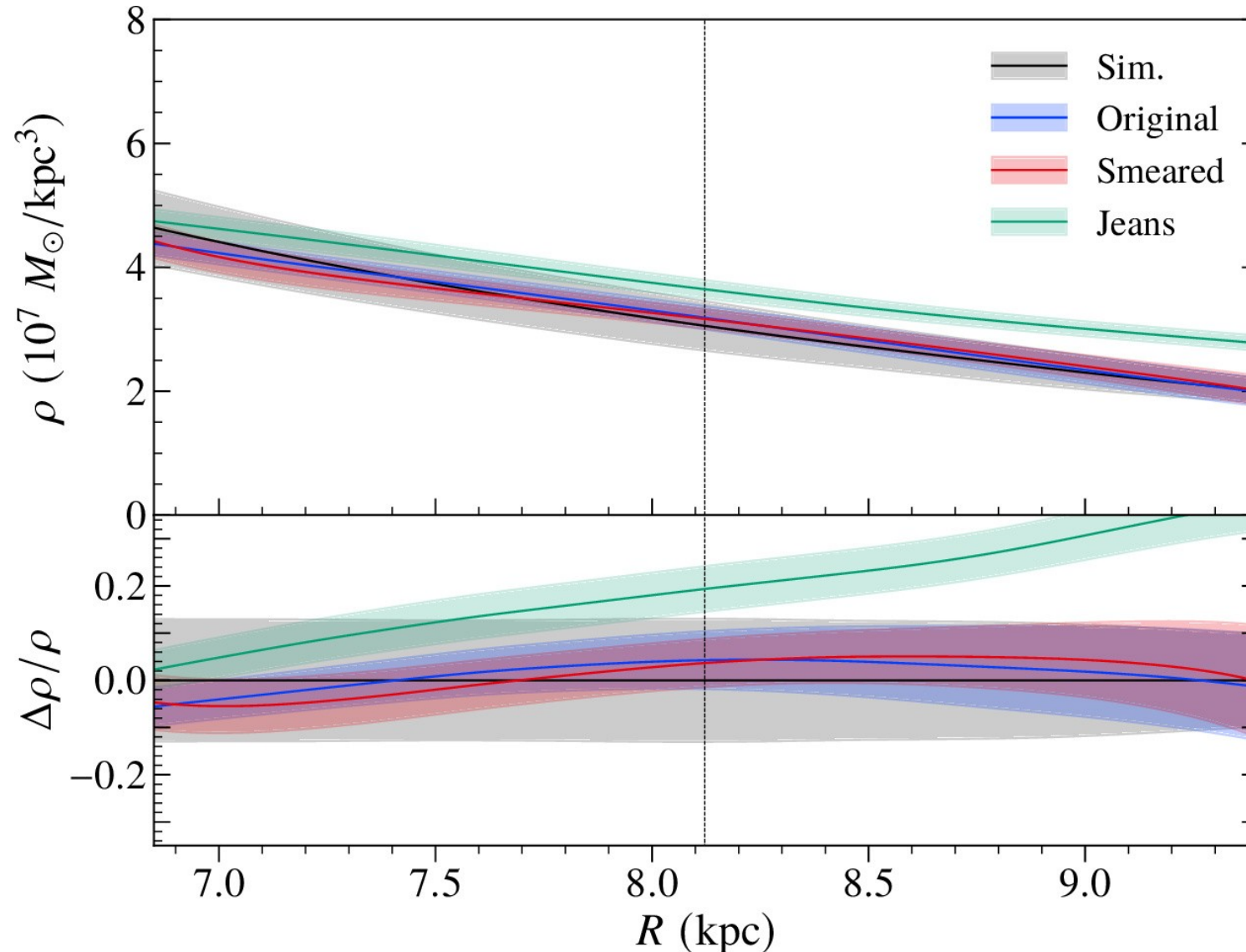
Acceleration along z-axis

Stat is low
1000~1500 / kpc³



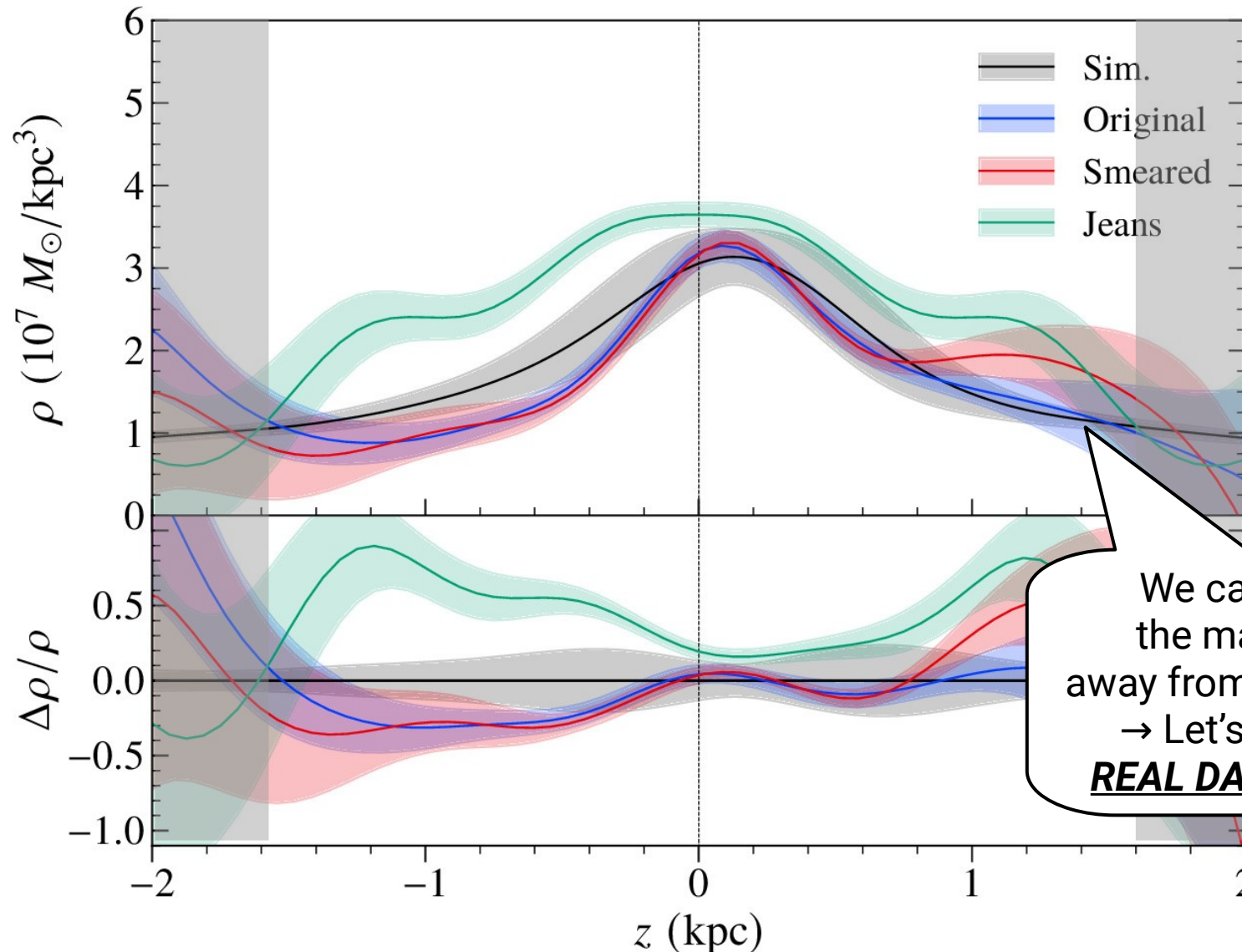
Our method can find out the acceleration within 5% accuracy!

Mass Density along x-axis



Our method can find out the mass density within 10~20% accuracy!

Mass Density along z-axis



We can measure
the mass density
away from the disk plane!
→ Let's move to the
REAL DATA ANALYSIS!

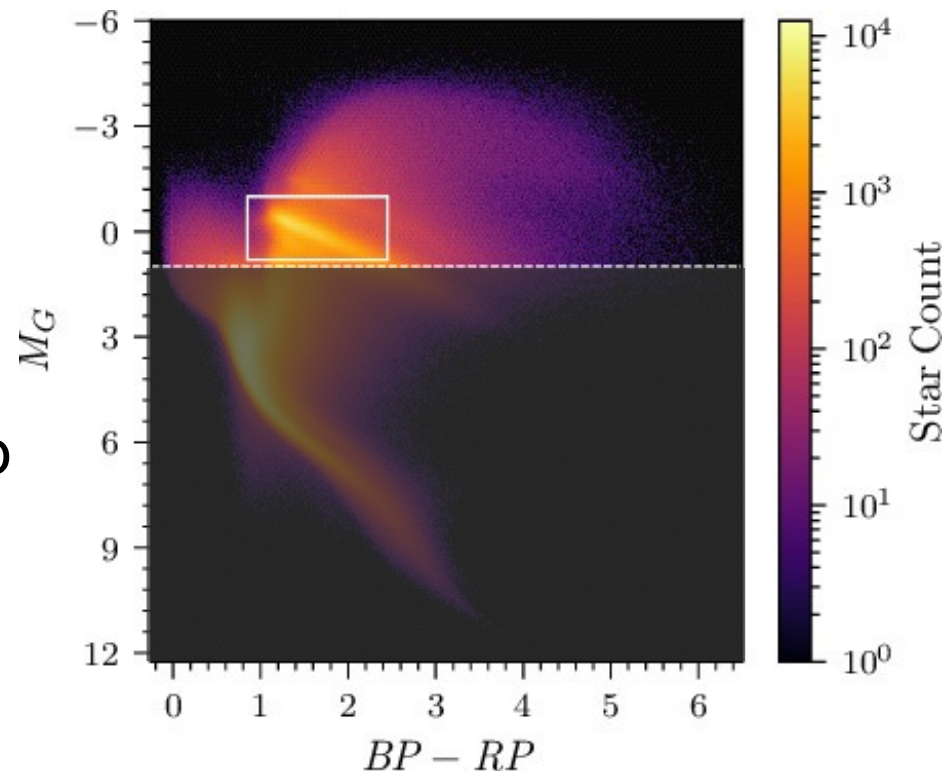
Our method can find out the mass density within 10~20% accuracy!

Tracer Population Selection

We will ignore the time derivative term for the acceleration estimation, requiring a population of stars sufficiently in equilibrium. (such as old stars, red giants)

$$\left[\cancel{\frac{\partial}{\partial t}} + \vec{v} \cdot \frac{\partial}{\partial \vec{x}} + \vec{a} \cdot \frac{\partial}{\partial \vec{v}} \right] f(\vec{x}, \vec{v}) = 0, \quad \vec{a} = -\frac{d\Phi(\vec{x})}{d\vec{x}}$$

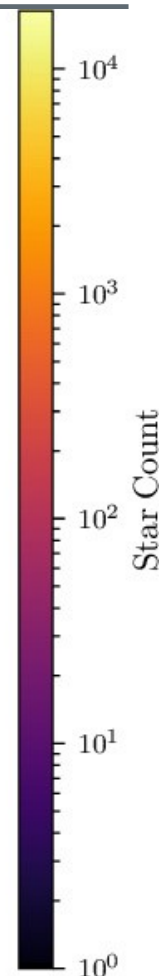
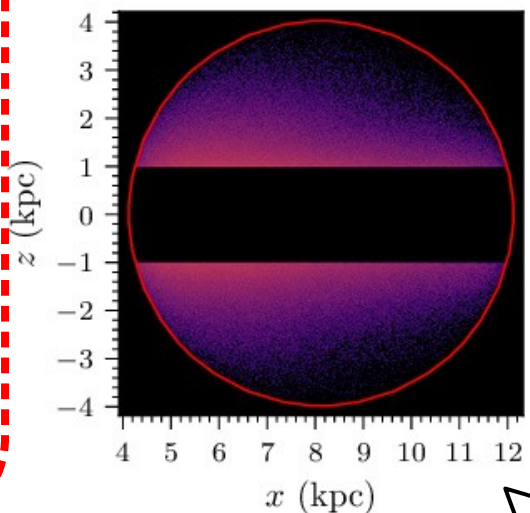
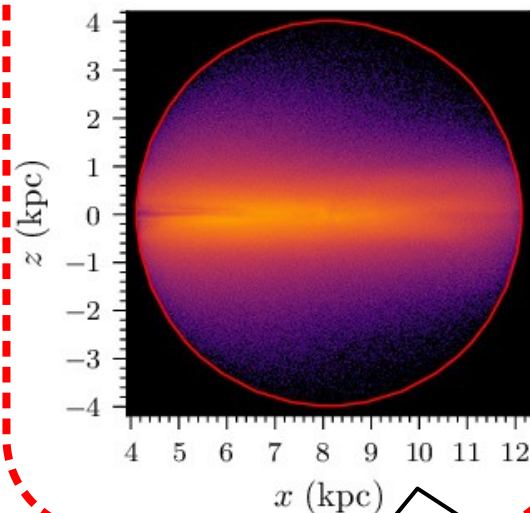
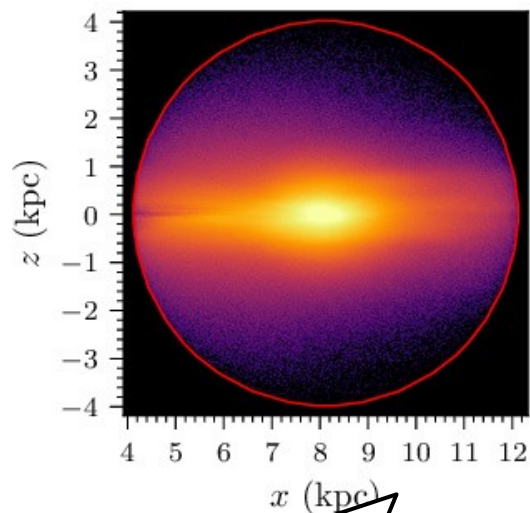
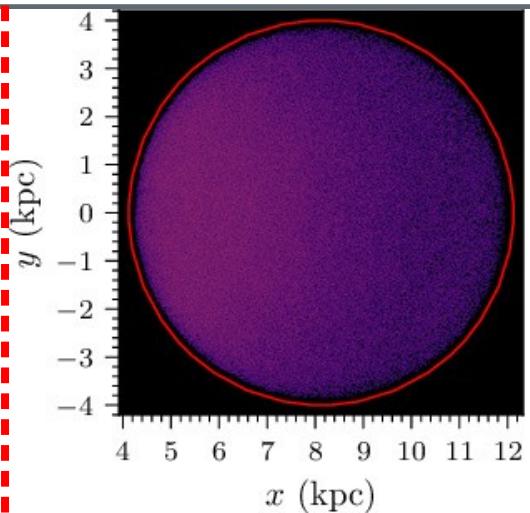
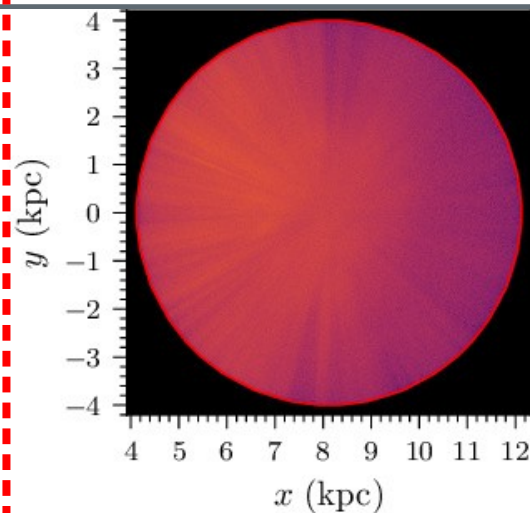
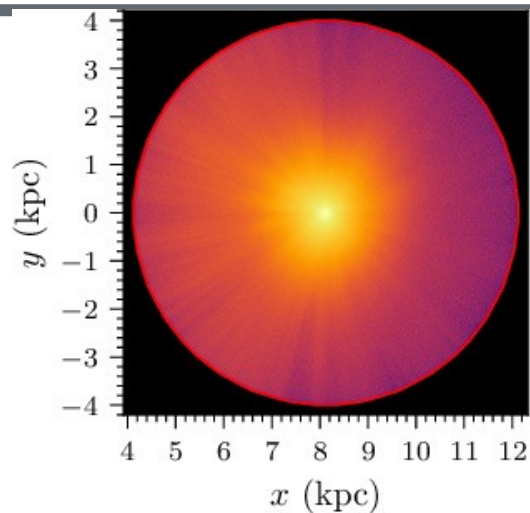
- Select stars with full **6D position and velocity** information
- **(Analysis volume)**
Distance from the Sun < 4.0 kpc
- **(Completeness)** absolute brightness must be large enough so that the stars must be visible anywhere within our analysis volume.
- **(Removing poorly measured stars)**
parallax / parallax error < 3



Stars with 6D info
+ distance < 4 kpc

Completeness
cut

Remove
disk stars

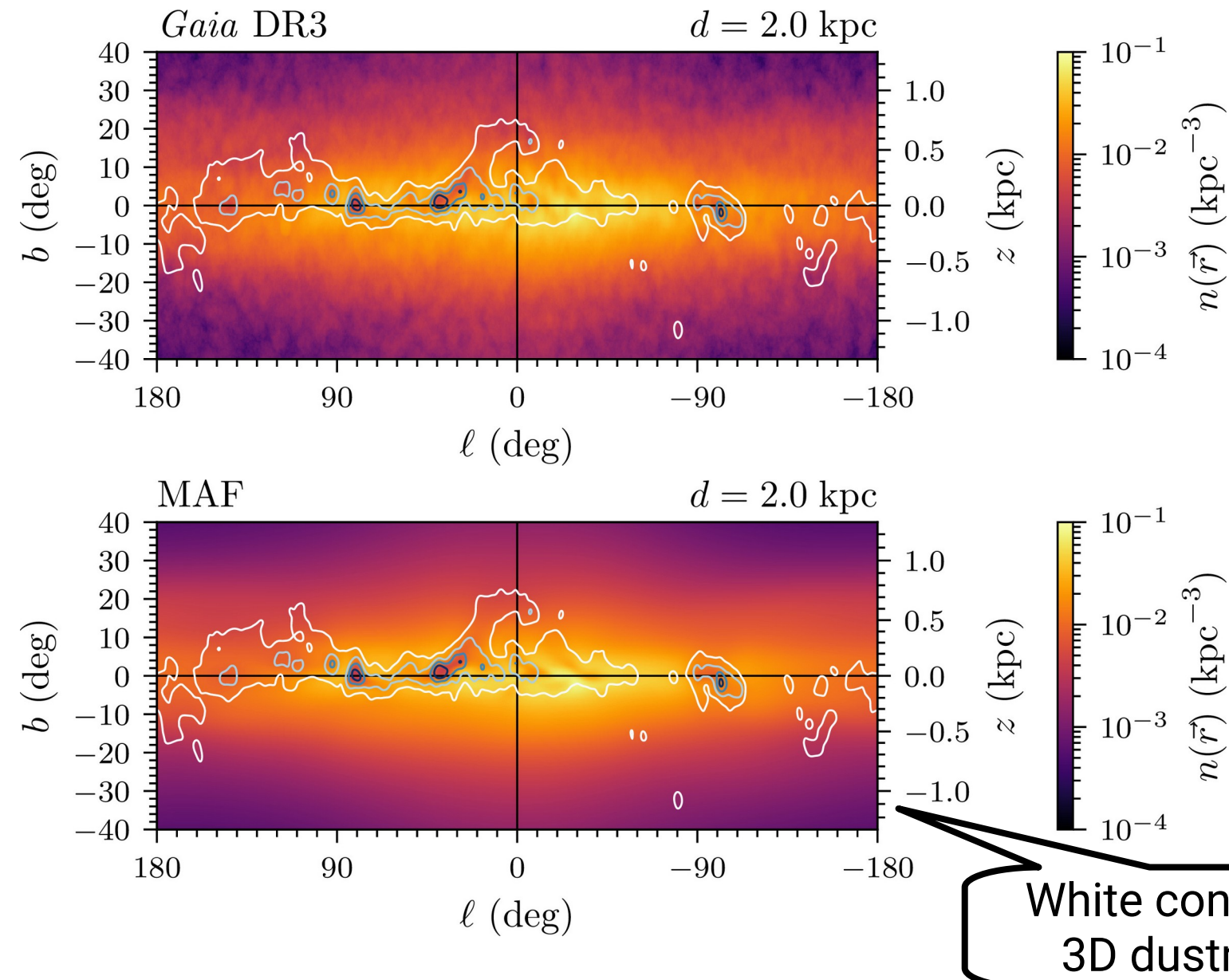


Dim and far stars
are not visible →
not a complete dataset

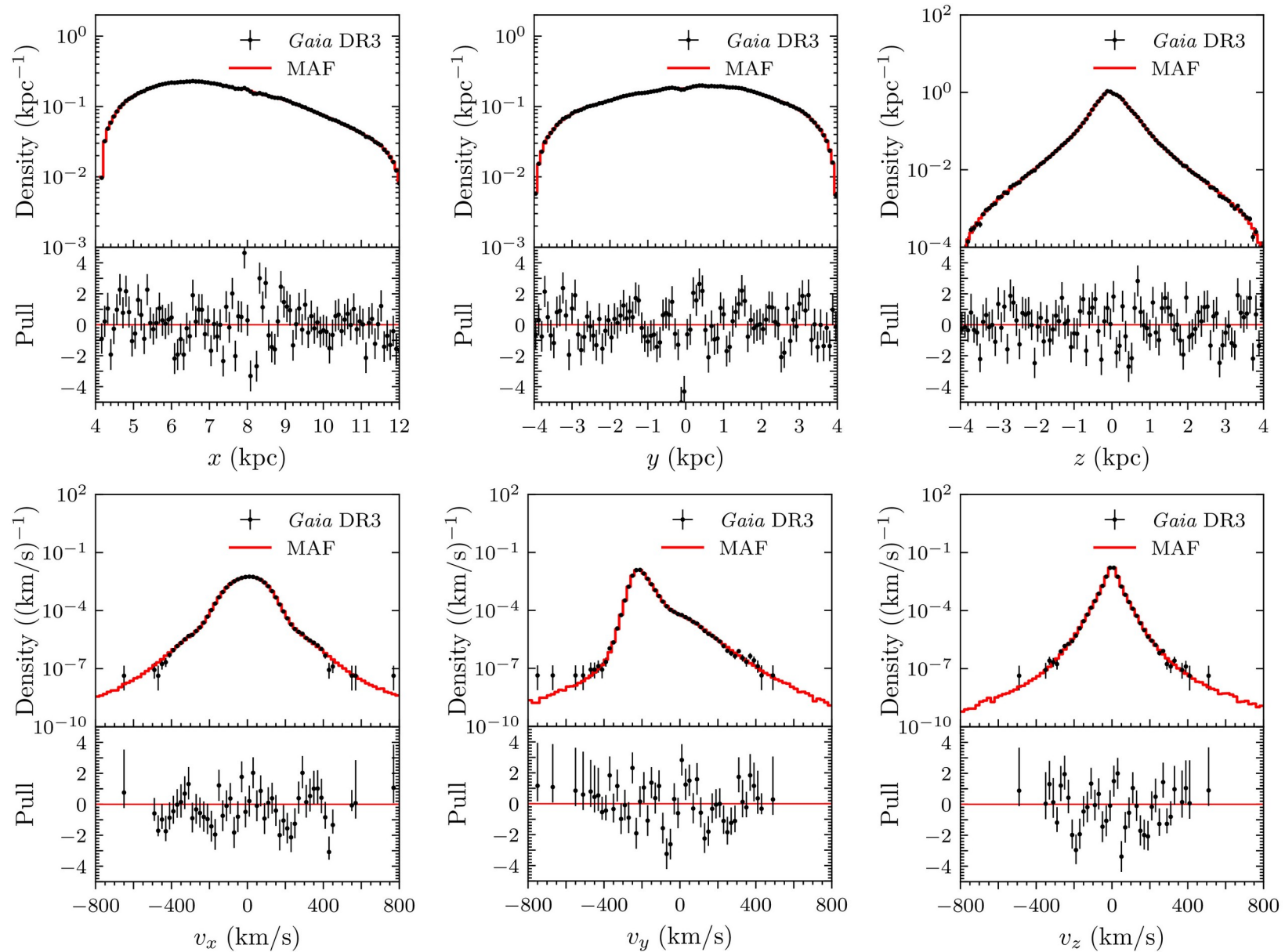
Approximately complete,
dust lanes :(

Halo region is dust-free.
Stellar distribution is smooth.

Number density estimation:



1D histograms



Acceleration Estimation: solving EOM by minimizing mean squared error

Now we have the estimated phase-space density estimation on our hand. Let's try to solve the Boltzmann equation.

$$\left[\vec{v} \cdot \frac{\partial}{\partial \vec{x}} + \vec{a} \cdot \frac{\partial}{\partial \vec{v}} \right] f(\vec{x}, \vec{v}) = 0, \quad \vec{a} = -\frac{d\Phi(\vec{x})}{d\vec{x}}$$

- Underdetermined in a point of view at each star position.
- Overdetermined in a point of view of phase-space density.

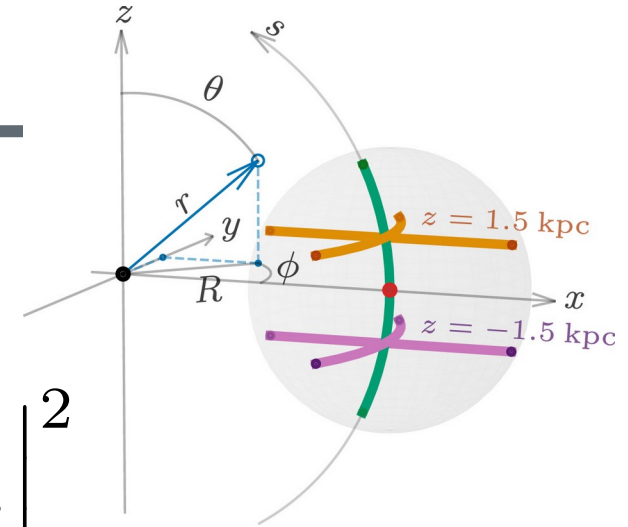
Given the fact that we could resample velocities at given position multiple times, we can solve the overdetermined system using least square minimization.

$$\mathcal{L}(\vec{x}) = \frac{1}{N} \sum_{\alpha=1}^N \left| \left[\vec{v}^{\alpha} \cdot \frac{\partial}{\partial \vec{x}} + \vec{a} \cdot \frac{\partial}{\partial \vec{v}} \right] f(\vec{x}, \vec{v}^{\alpha}) \right|^2$$

We draw 10,000 samples per position to reduce the QMC integration error below the statistical and measurement errors.

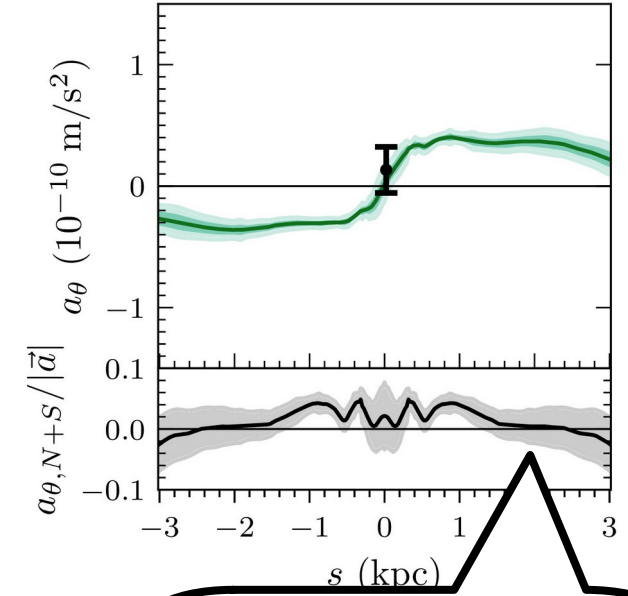
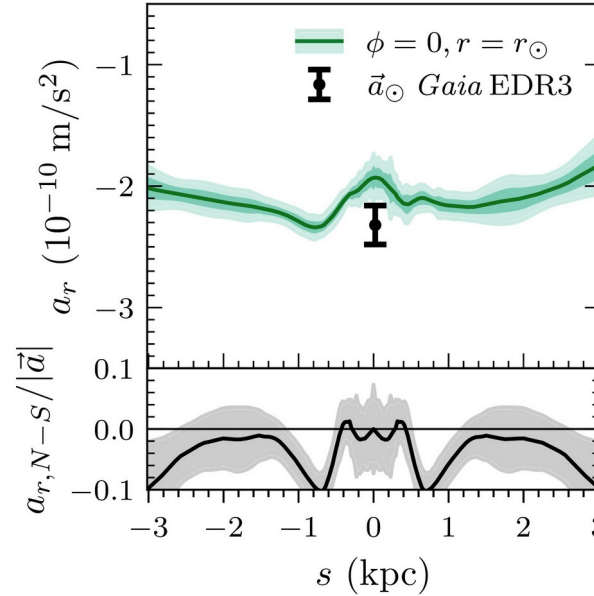
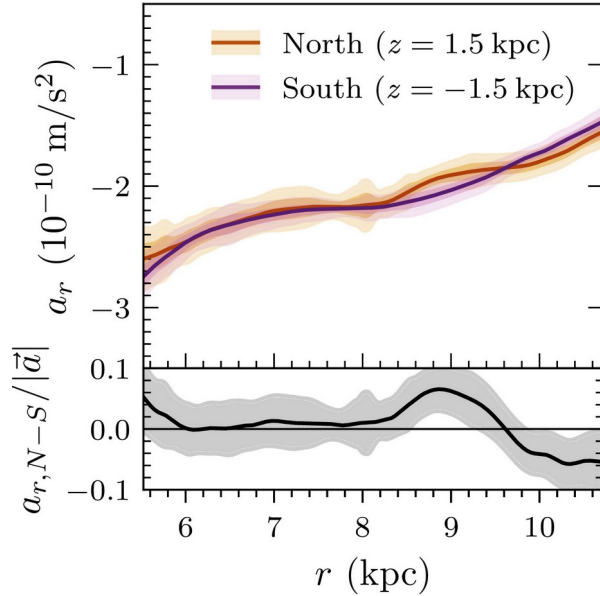
Acceleration Estimation and North-South asymmetry

Coordinate guide



$$\frac{\partial f}{\partial t} = -v_i \frac{\partial f}{\partial x_i} - a_i(\vec{x}) \frac{\partial f}{\partial v_i} = 0$$

$$\mathcal{L}_a = \int d^3\vec{v} p(\vec{v}|\vec{x}) \left| v_i \frac{\partial f}{\partial x_i} + a_i(\vec{x}) \frac{\partial f}{\partial v_i} \right|^2$$



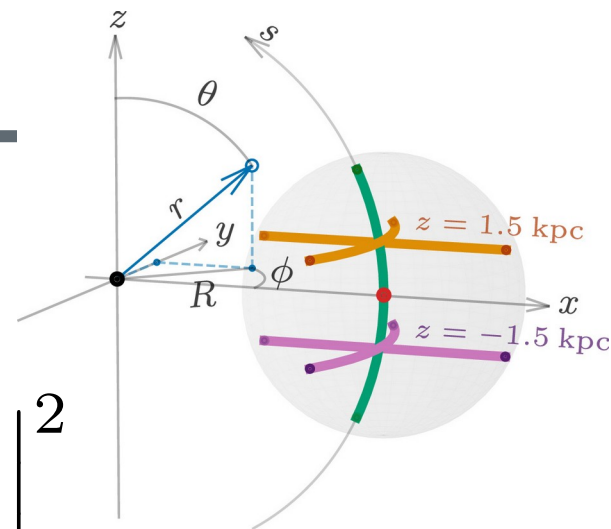
$|a| / (10^{-10} \text{ m/s}^2)$ at the Sun:

our result: **1.94±0.22**. Gaia EDR3: **2.32±0.16**

North-south
asymmetry is
less than 10%

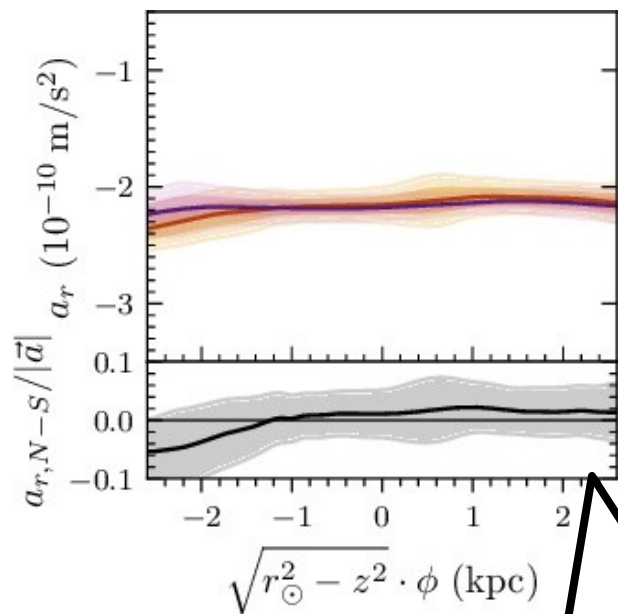
Acceleration Estimation and Axisymmetry

Coordinate guide

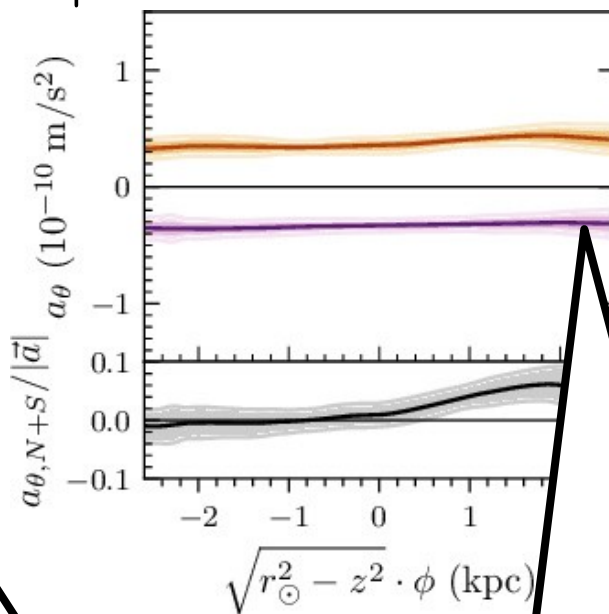


$$\frac{\partial f}{\partial t} = -v_i \frac{\partial f}{\partial x_i} - a_i(\vec{x}) \frac{\partial f}{\partial v_i} = 0$$

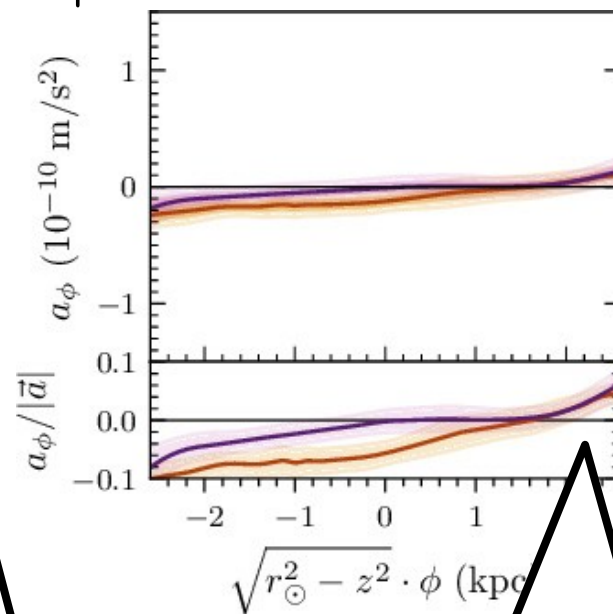
$$\mathcal{L}_a = \int d^3\vec{v} p(\vec{v}|\vec{x}) \left| v_i \frac{\partial f}{\partial x_i} + a_i(\vec{x}) \frac{\partial f}{\partial v_i} \right|^2$$



North-south
asymmetry is
less than 10%



Constant acceleration:
approximately
axisymmetric



Azimuthal
acceleration $\sim 10\%$

Solving Gauss's Equation

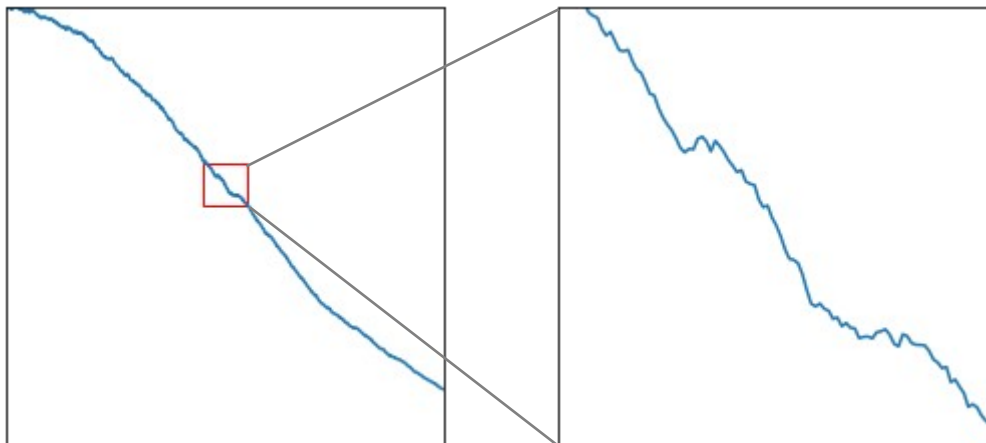
Since we have a smooth acceleration estimator, we may directly take another derivative to estimate the local mass density.

$$\text{Gauss's Law: } -4\pi G\rho = \nabla \cdot \vec{a}$$

Mass density function:
genuine point-wise feature

Compatible?

Estimated accelerations:
finite resolution
finite training samples



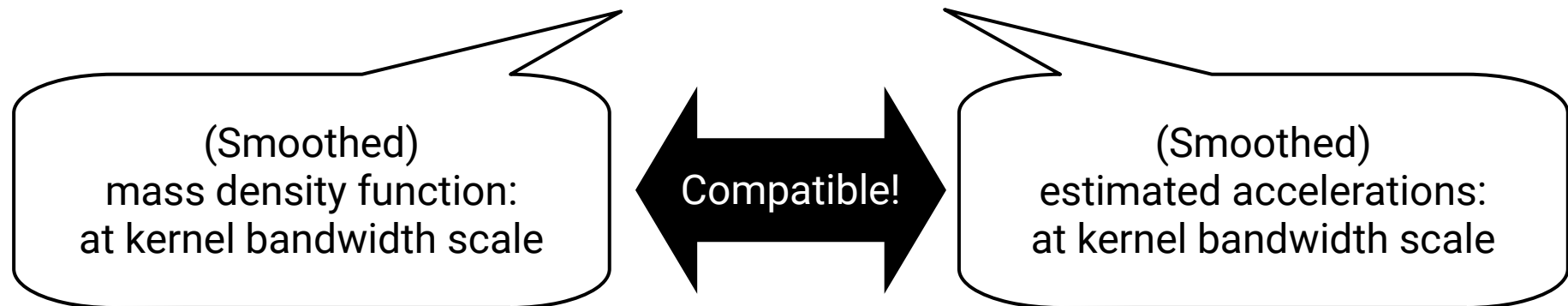
Note that differentiation is essentially an error-amplifying process.

We may end up with losing precision because of noise or inductive bias of interpolation!

Smoothed Mass Density Estimation

Instead, we will estimate kernel smoothed mass density:

$$-4\pi G\rho * K_h = (\nabla \cdot \vec{a}) * K_h$$



We use the following bandwidths,

$$(h_x, h_y, h_z) = (0.5, 0.5, 0.2) \text{ kpc}$$

No need of evaluating 2nd order derivative directly

Another advantage of using kernel smoothed mass density is that we do not need to evaluate the 2nd order derivative of the network.

$$\begin{aligned} -4\pi G\rho * K_h &= (\nabla \cdot \vec{a}) * K_h = \int d^3\vec{x}' (\nabla \cdot \vec{a})(\vec{x}') K_h(\vec{x} - \vec{x}') \\ &= \oint d^2\vec{x}' \hat{n} \cdot \vec{a}(\vec{x}') K_h(\vec{x} - \vec{x}') + \int d^3\vec{x}' \vec{a}(\vec{x}') \cdot \nabla K_h(\vec{x} - \vec{x}') \end{aligned}$$

To estimate the smoothed mass density at given position, we do the following:

1. Draw samples from kernel boundary and kernel itself.
2. Evaluate accelerations at perturbed positions.
3. Solve the above Gauss's equation

Warning: this step is very time consuming!

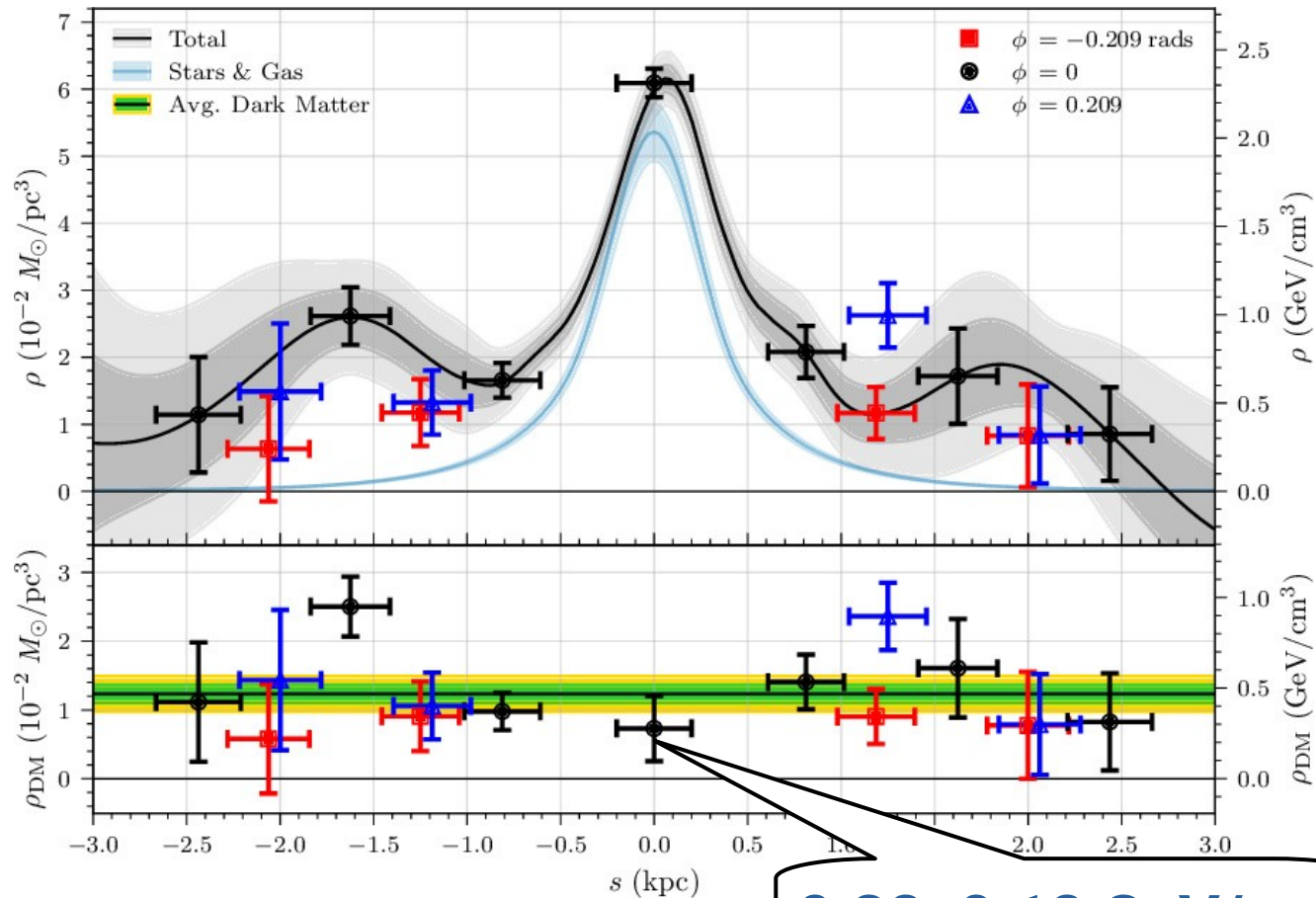
10000 x 3200

~ 30M network evaluations per point!



Mass Density Estimation

$$\nabla \cdot \vec{a} * K = -4\pi G \rho * K$$

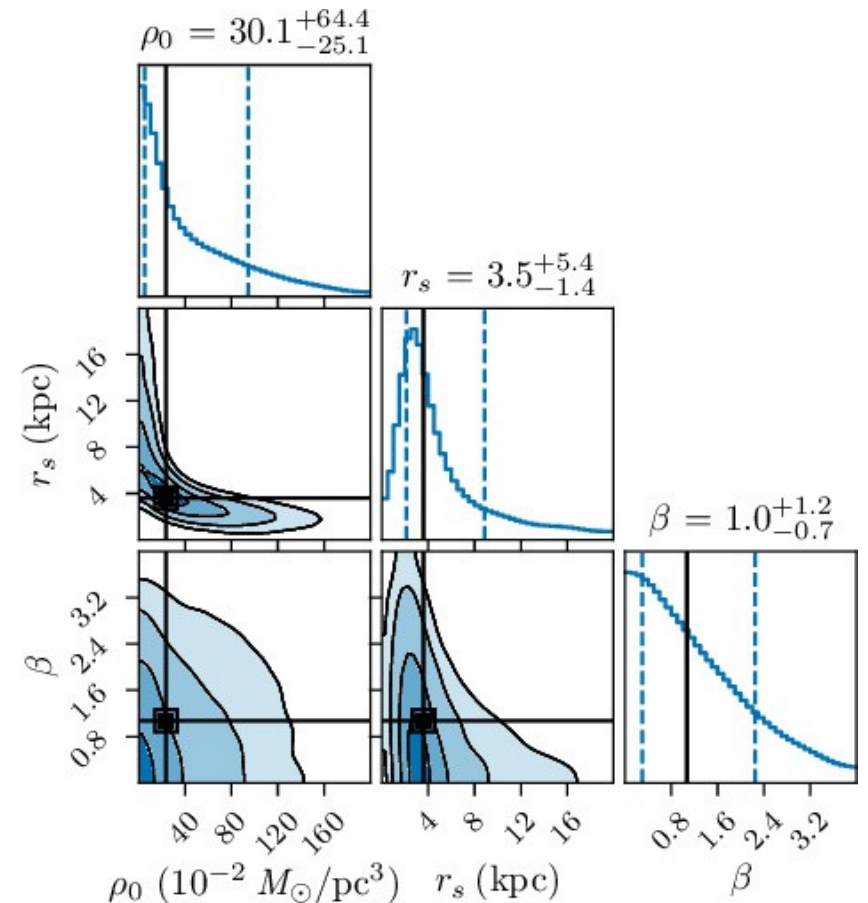
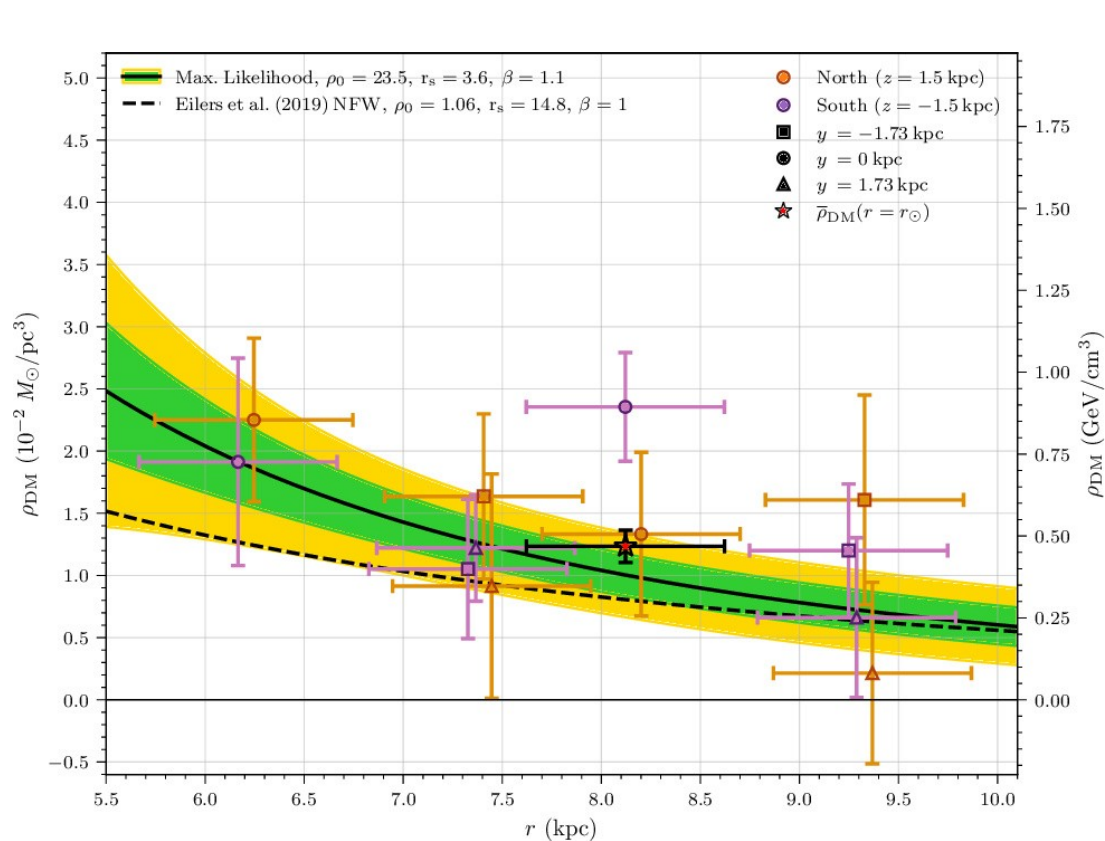


Taking the average of the DM mass density at the Solar radius, we find a local dark matter density: $0.47 \pm 0.05 \text{ GeV}/\text{cm}^3$

Compatibility to NFW profile

Our result is consistent with NFW profile:

$$\rho(\vec{r}) = \rho_0 \left(\frac{r}{r_s} \right)^\beta \left(1 + \frac{r}{r_s} \right)^{3-\beta}$$



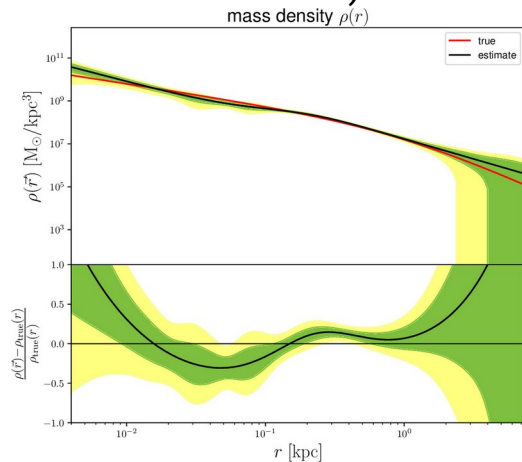
Conclusion

- We have developed unsupervised machine learning technique for measuring dark matter density of the Milky Way.
- **For the first time**, We successfully measured the dark matter density in a local volume around the Sun, without assuming the functional form of density and symmetry.
- We find **the gravitational acceleration** at the Sun: $(1.94 \pm 0.22) \times 10^{-10} \text{ m/s}^2$, which is compatible with the acceleration measurement in Gaia EDR3 using quasar by 2sigma.
- The north south asymmetry of acceleration within our analysis volume is less than 10%, indicating that the local disequilibrium effect is relatively small.
- We find **the dark matter density** at the Sun: $(0.47 \pm 0.05) \text{ GeV/cm}^3$. This result is agreeing with the results from more constrained analysis.

Other ongoing projects!

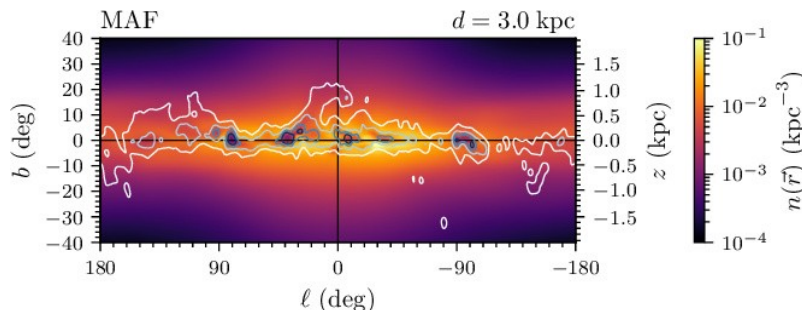
Unsupervised DM density estimation of simulated dwarf spheroidal galaxies

with K. Hayashi (Ichinoseki U.), M. N. Nojiri (KEK)
(to be appeared on arXiv soon!)

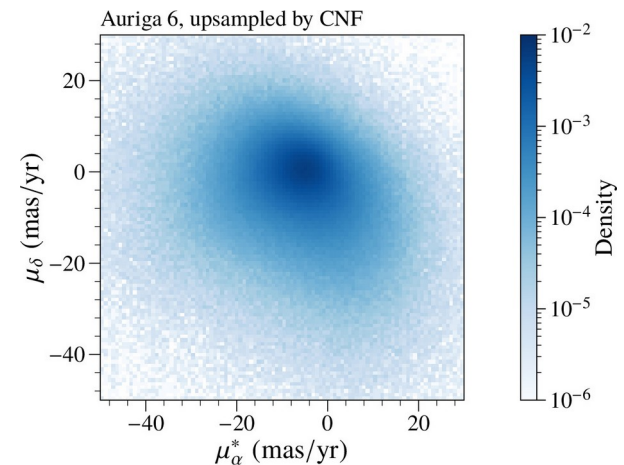
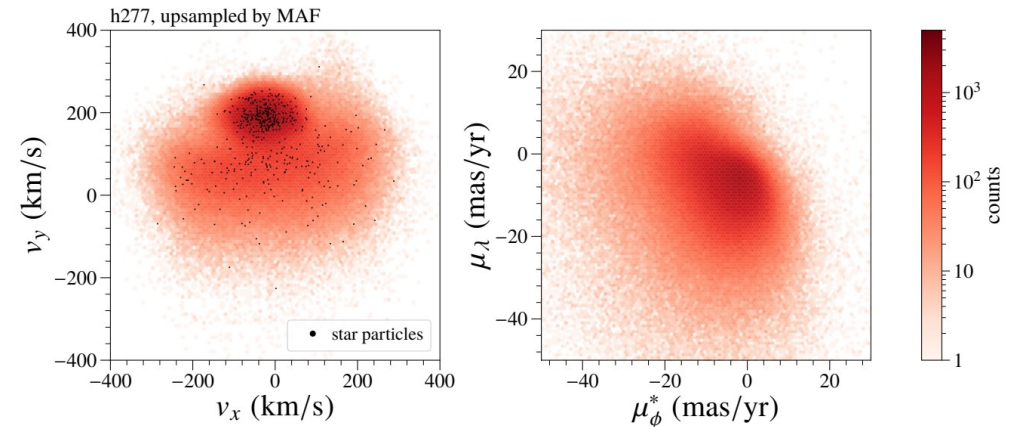


Others...

- publishing full 3D dark matter density map package
- improving density estimation performance
- measurement bias corrections and deconvolution
- playing with interstellar dust



Upsampling hydrodynamic simulation of a galaxy
with K. Raman (Berkeley Lab), M. Buckley, D. Shih (Rutgers)
(arXiv: 2211.11765)



Outline of Strategy

