Sweeping the Dust Away

Removing extinction bias from the Milky Way's potential using Gaia DR3 and unsupervised machine learning

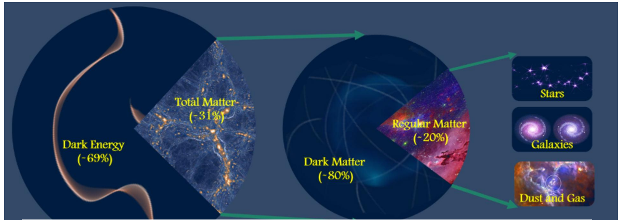
arXiv:240X.XXXXX

Eric Putney (eputney@physics.rutgers.edu)

w/ S.H. Lim, M. Buckley, D. Shih - Rutgers University

The Milky Way is a dark matter laboratory

Primordial dark matter seeded formation of all visible structures.



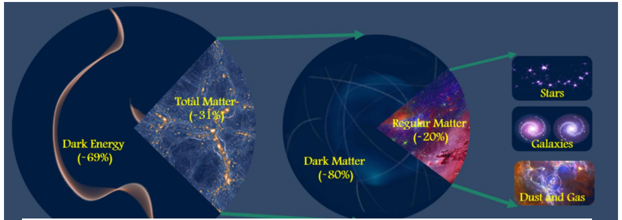
Mohamed Abdullah, University of California Riverside

The Milky Way is a dark matter laboratory

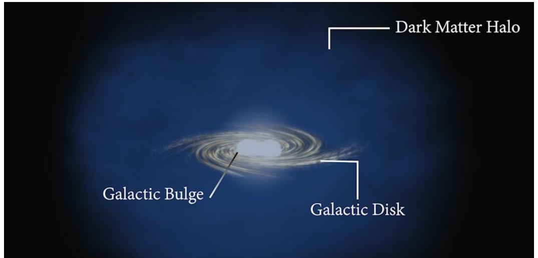
Primordial dark matter seeded formation of all visible structures.

Our galaxy sits inside of a massive halo of dark matter:

- Local density sets direct detection rate
- Density profile probes DM particle physics



Mohamed Abdullah, University of California Riverside



Credit: L. Jaramillo & O. Macias/Virginia Tech.

The Milky Way is a dark matter laboratory

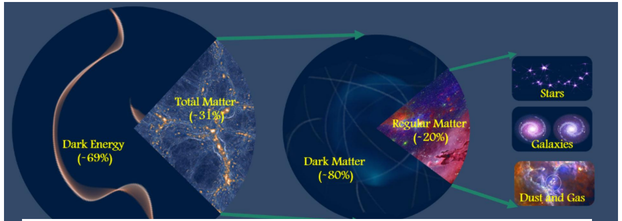
Primordial dark matter seeded formation of all visible structures.

Our galaxy sits inside of a massive halo of dark matter:

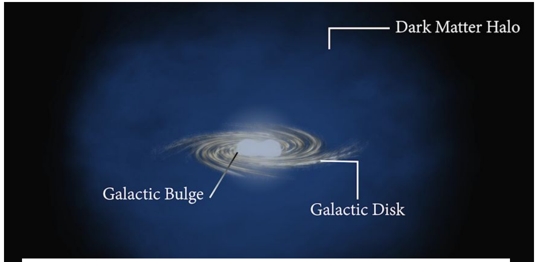
- Local density sets direct detection rate
- Density profile probes DM particle physics

Nearby stars tell us much about our host halo:

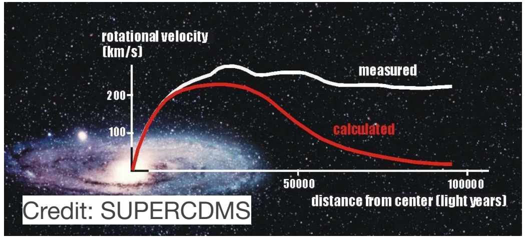
- Stellar streams sensitive to subhalos
- Moment approaches: Rotation curves, Jeans analyses, etc…



Mohamed Abdullah, University of California Riverside



Credit: L. Jaramillo & O. Macias/Virginia Tech.



Credit: SUPERCDMS

The Milky Way is a dark matter laboratory

Primordial dark matter seeded formation of all visible structures.

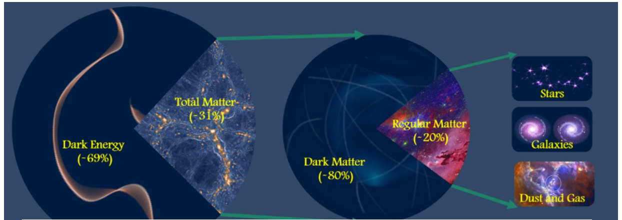
Our galaxy sits inside of a massive halo of dark matter:

- Local density sets direct detection rate
- Density profile probes DM particle physics

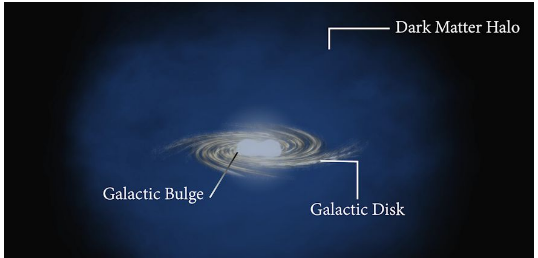
Nearby stars tell us much about our host halo:

- Stellar streams sensitive to subhalos
- Moment approaches: Rotation curves, Jeans analyses, etc…

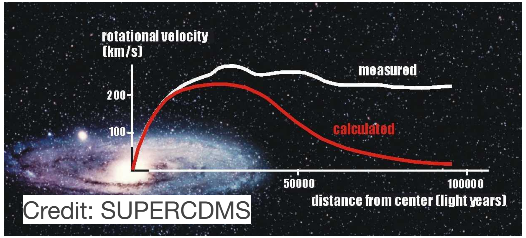
Potential encoded in stellar phase space **f(x,v)**



Mohamed Abdullah, University of California Riverside



Credit: L. Jaramillo & O. Macias/Virginia Tech.



Credit: SUPERCDMS

Phase Space of the Milky Way

Local stellar **f(x,v)** encodes the local **gravitational potential $\boldsymbol{\Phi}$** via the equilibrium collisionless Boltzmann equation (CBE):

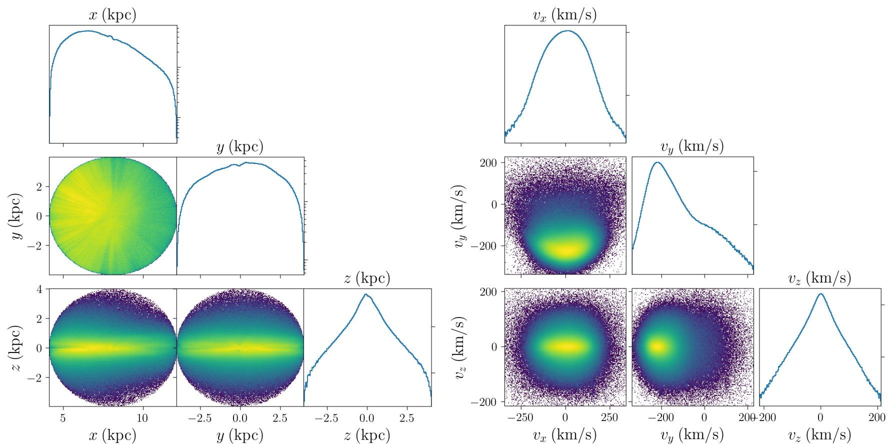
$$\frac{d}{dt}\ln f = \left[\cancel{\frac{\partial}{\partial t}} + v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right]\ln f = 0$$

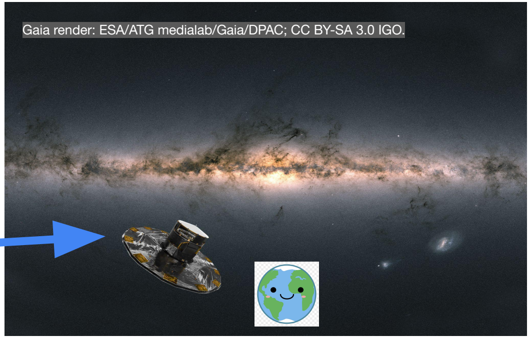
Phase Space of the Milky Way

Local stellar **f(x,v)** encodes the local **gravitational potential Φ** via the equilibrium collisionless Boltzmann equation (CBE):

$$\frac{d}{dt}\ln f = \left[\cancel{\frac{\partial}{\partial t}} + v_i\frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right]\ln f = 0$$

Gaia DR3 provides the 6D kinematics of millions of nearby stars. We use a "complete" sample of ~6 million Red Giant Branch stars w/in 4 kpc.





Phase Space of the Milky Way

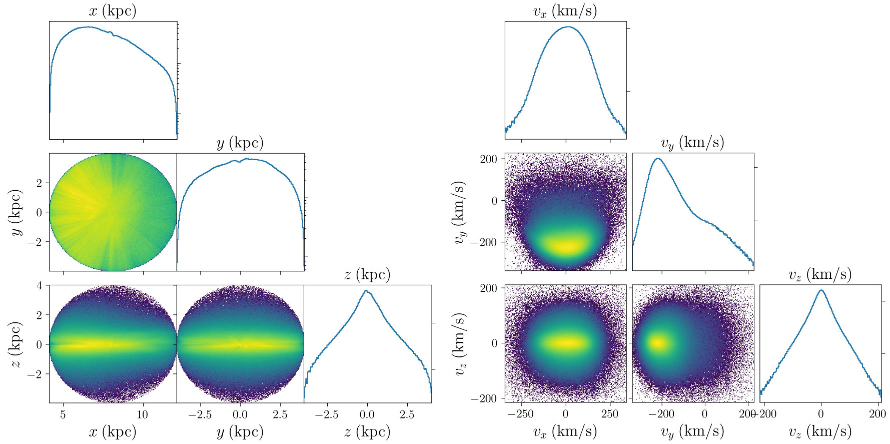
Local stellar **f(x,v)** encodes the local **gravitational potential $\boldsymbol{\Phi}$** via the equilibrium collisionless Boltzmann equation (CBE):

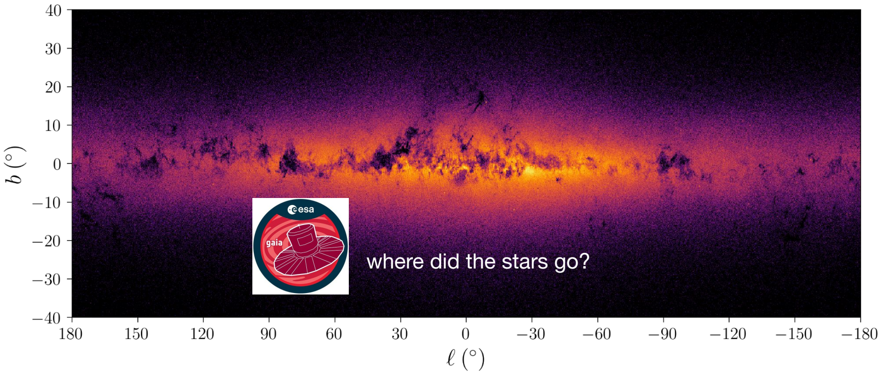
$$\frac{d}{dt}\ln f = \left[\cancel{\frac{\partial}{\partial t}} + v_i\frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right]\ln f = 0$$

Gaia DR3 provides the 6D kinematics of millions of nearby stars. We use a "complete" sample of ~6 million Red Giant Branch stars w/in 4 kpc.

Interstellar dust absorbs or scatters starlight, which dims stars below *Gaia*'s sensitivity limit.

Suppression of **f(x,v)** biases our model of the MW's gravitational potential.





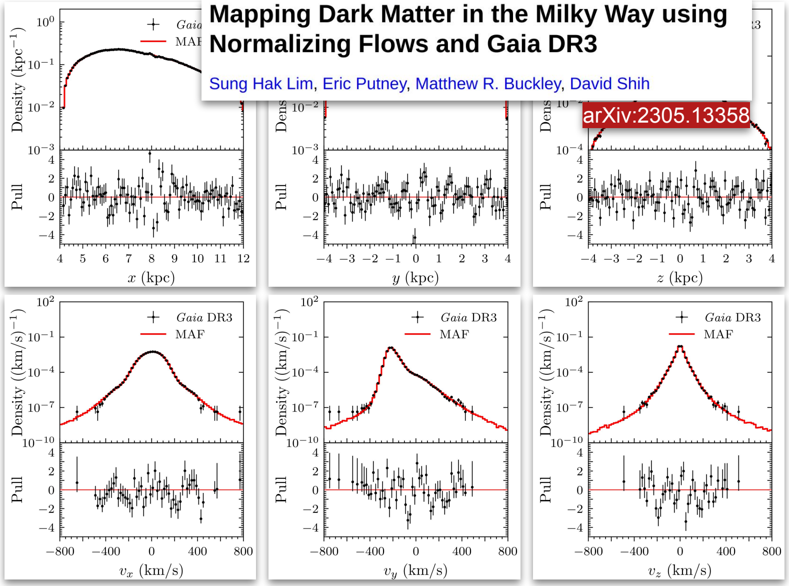
How far can we go with dusty data?

We learned **f(x,v)** from dusty stellar data using masked autoregressive flows (MAFs), generative ML density estimators.

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

Sung Hak Lim, Eric Putney, Matthew R. Buckley, David Shih

arXiv:2305.13358



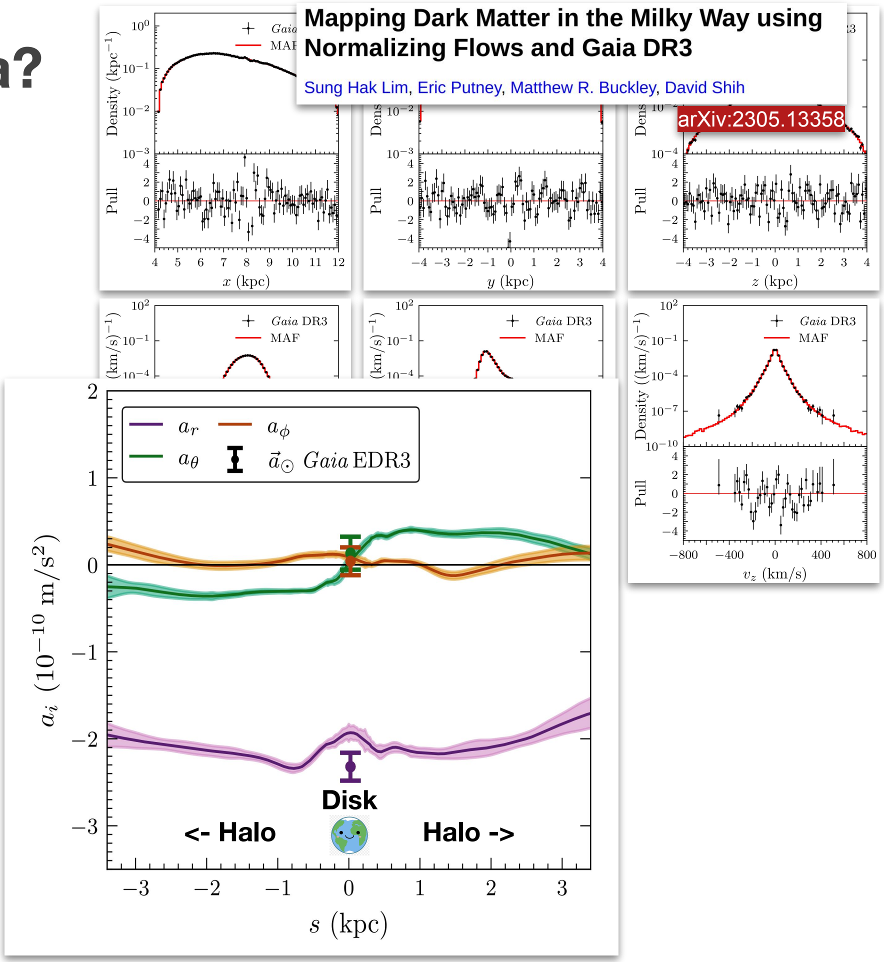
How far can we go with dusty data?

We learned **f(x,v)** from dusty stellar data using masked autoregressive flows (MAFs), generative ML density estimators.

Used f(x,v) to solve CBE for the local **acceleration a(x)** pointwise in dust-free regions:

$$\left[v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right] \ln f = 0 \qquad -\frac{\partial \Phi}{\partial x_i} \equiv a_i$$

Local **acceleration a(x)** was consistent with prior parametric studies.



How far can we go with dusty data?

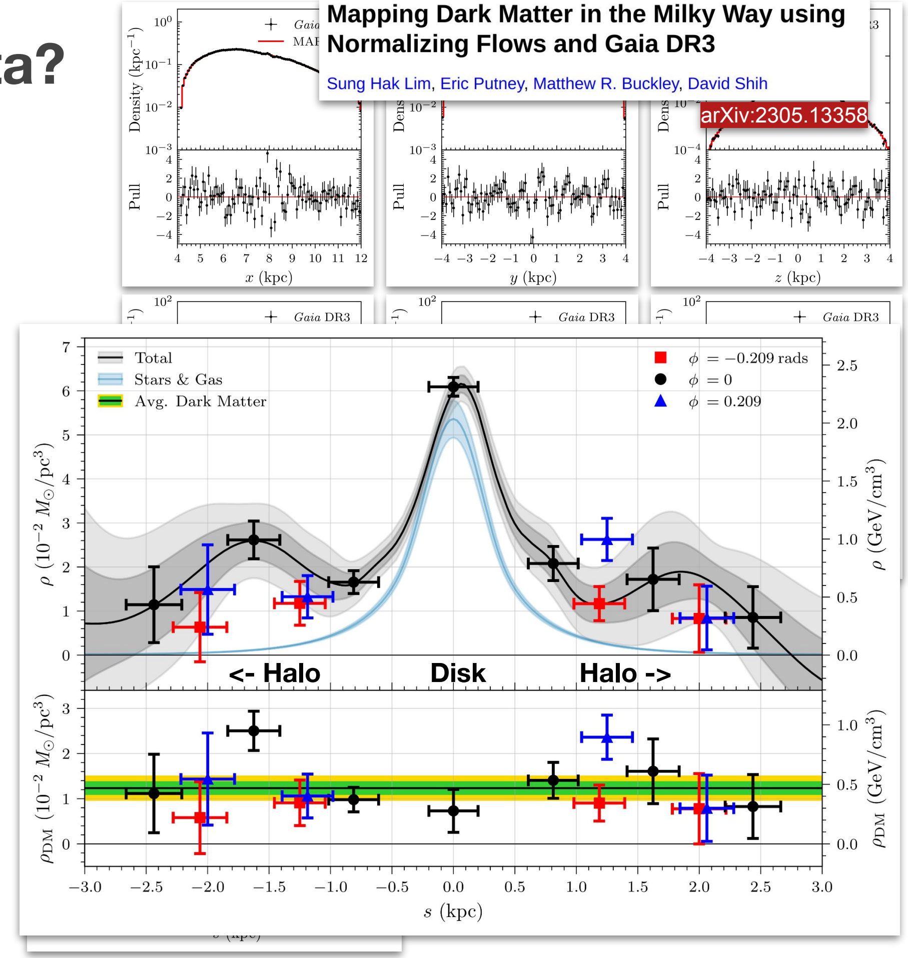
We learned **f(x,v)** from dusty stellar data using masked autoregressive flows (MAFs), generative ML density estimators.

Used f(x,v) to solve CBE for the local **acceleration a(x)** pointwise in dust-free regions:

$$\left[v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right] \ln f = 0 \qquad -\frac{\partial \Phi}{\partial x_i} \equiv a_i$$

Local **acceleration a(x)** was consistent with prior parametric studies.

Gradient of **a(x)** yields the **mass density field ρ(x)**.



How far can we go with dusty data?

We learned **f(x,v)** from dusty stellar data using masked autoregressive flows (MAFs), generative ML density estimators.

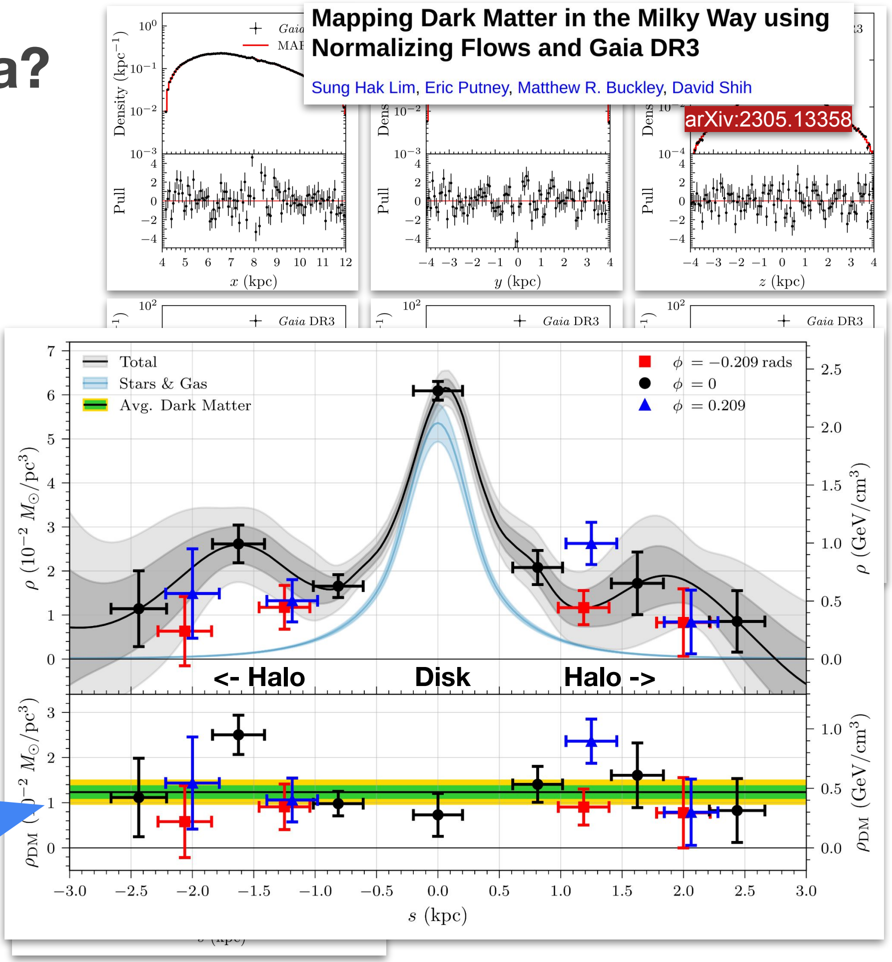
Used f(x,v) to solve CBE for the local **acceleration a(x)** pointwise in dust-free regions:

$$\left[v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right] \ln f = 0 \qquad -\frac{\partial \Phi}{\partial x_i} \equiv a_i$$

Local **acceleration a(x)** was consistent with prior parametric studies.

Gradient of **a(x)** yields the **mass density field ρ(x)**.

Measured the **dark matter** halo density around the Earth and the halo density profile!



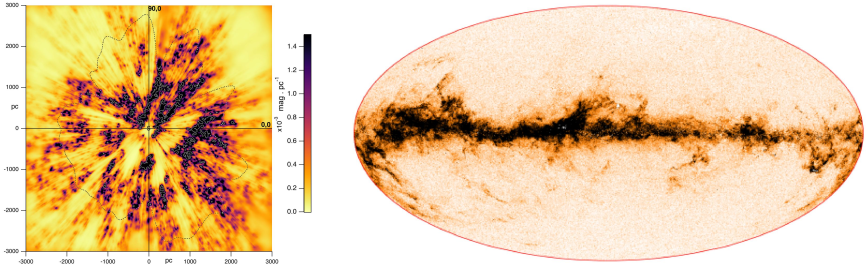
Can we *undo* the effect of dust?

Astronomers craft “dust maps” based on knowledge of the local ISM, extinction laws, etc…

Dust maps undo the dimming of observed stars, but they don’t give insight into missing stars.

We note three key insights:

1. Only **f(x,v)** satisfies the CBE
2. Dust bias factorizable into **ϵ(x)**, a position-dependent “dust efficiency”
3. Velocity dependence of terms breaks degeneracy between **Φ(x)** and **ϵ(x)**



Credit: Lallement et al. 2019 A&A 625, A135

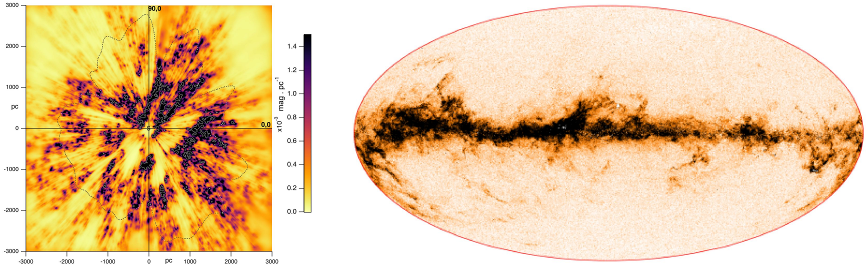
Can we *undo* the effect of dust?

Astronomers craft “dust maps” based on knowledge of the local ISM, extinction laws, etc...

Dust maps undo the dimming of observed stars, but they don’t give insight into missing stars.

We note three key insights:

1. Only **f(x,v)** satisfies the CBE
2. Dust bias factorizable into **ϵ(x)**, a position-dependent “dust efficiency”
3. Velocity dependence of terms breaks degeneracy between **Φ(x)** and **ϵ(x)**



Credit: Lallement et al. 2019 A&A 625, A135

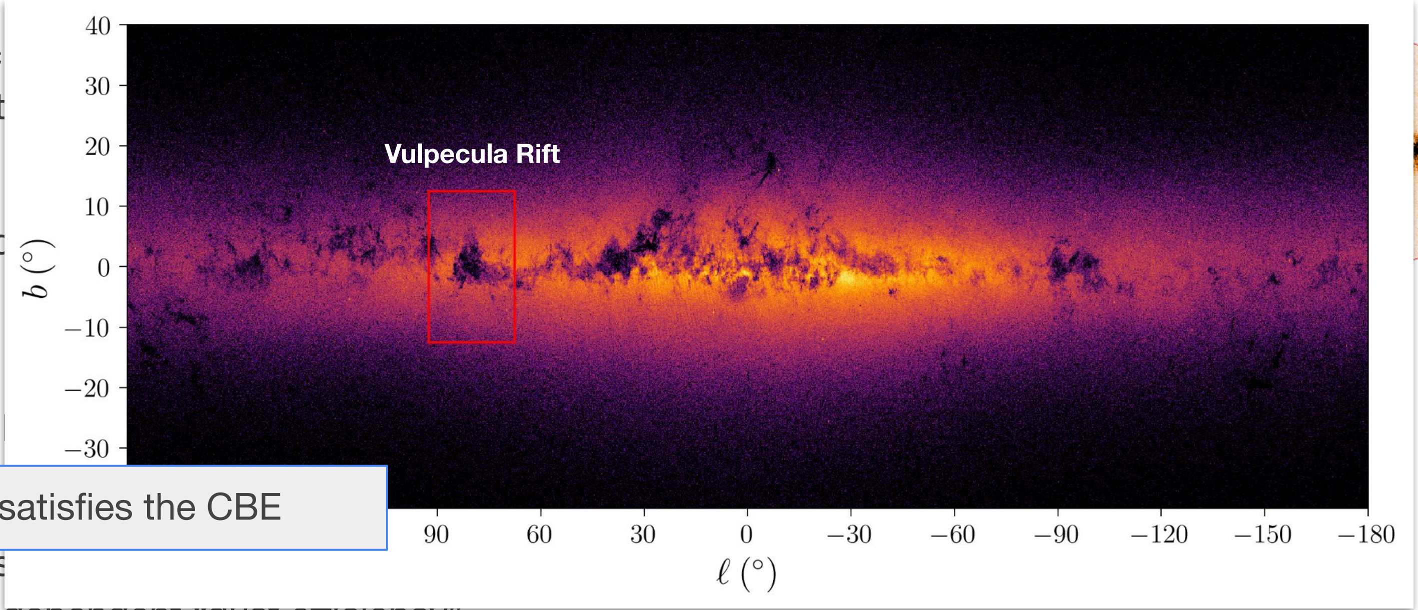
$$\left[v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right] \ln f = 0$$

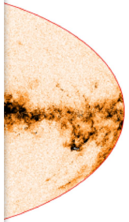
Can we *undo* the effect of dust?

Astronomers c
knowledge of t
etc...

Dust maps unc
stars, but they
stars.

We note three





1. Only **f(x,v)** satisfies the CBE
2. Dust bias
   position-dependent "dust efficiency"
3. Velocity dependence of terms breaks degeneracy between **Φ(x)** and **ϵ(x)**

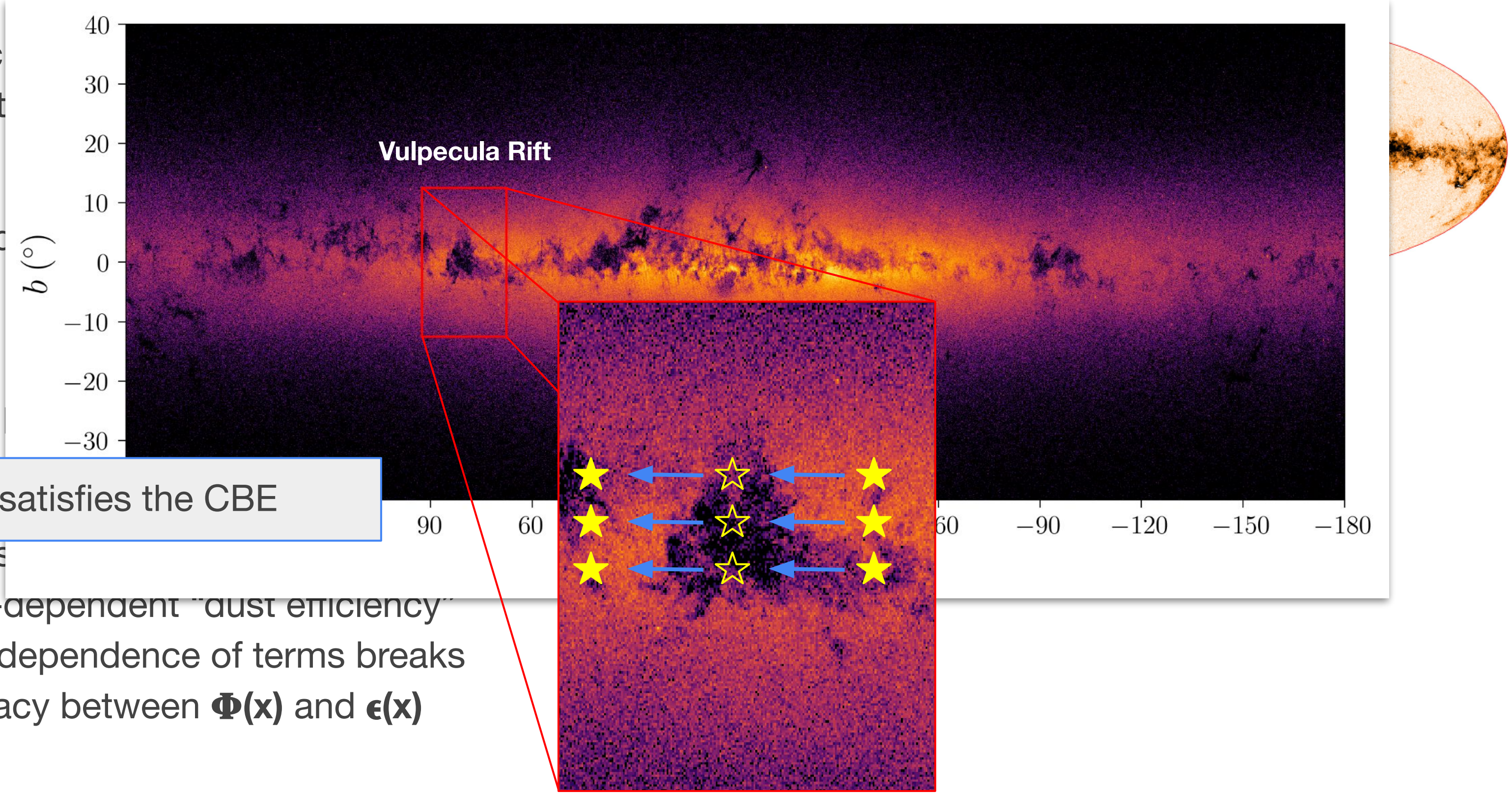
Can we *undo* the effect of dust?

Astronomers c
knowledge of t
etc...

Dust maps unc
stars, but they
stars.

We note three

1. Only **f(x,v)** satisfies the CBE
2. Dust bias
position-dependent "dust efficiency"
3. Velocity dependence of terms breaks degeneracy between **Φ(x)** and **ϵ(x)**



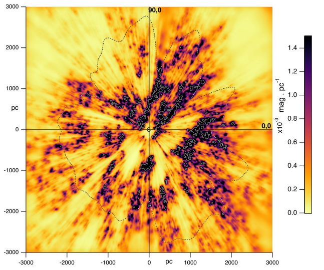
Can we *undo* the effect of dust?

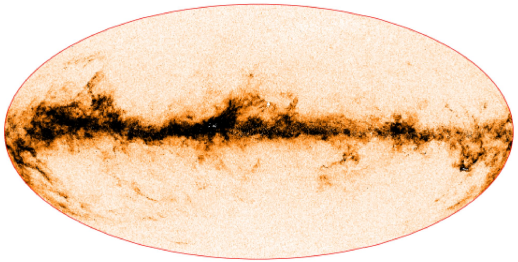
Astronomers craft “dust maps” based on knowledge of the local ISM, extinction laws, etc...

Dust maps undo the dimming of observed stars, but they don’t give insight into missing stars.

We note three key insights:

1. Only **f(x,v)** satisfies the CBE
2. Dust bias factorizable into **ϵ(x)**, a position-dependent “dust efficiency”
3. Velocity dependence of terms breaks degeneracy between **Φ(x)** and **ϵ(x)**





Credit: Lallement et al. 2019 A&A 625, A135

$$\left[v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \frac{\partial}{\partial v_i} \right] \ln f = 0$$

$$f_{\mathrm{obs}}(x, v) \propto \epsilon(x) f(x, v)$$

(up to normalization)

Can we *undo* the effect of dust?

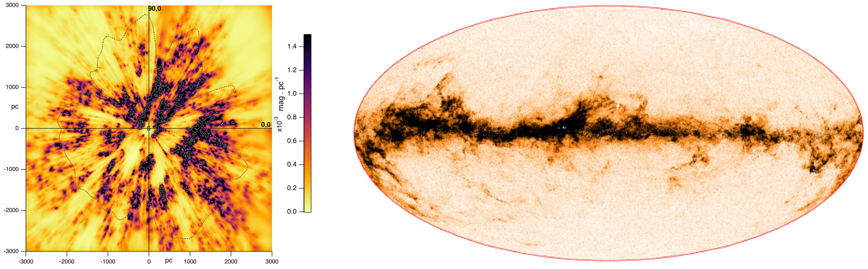
Astronomers craft “dust maps” based on knowledge of the local ISM, extinction laws, etc...

Dust maps undo the dimming of observed stars, but they don’t give insight into missing stars.

We note three key insights:

1. Only **f(x,v)** satisfies the CBE
2. Dust bias factorizable into **ϵ(x)**, a position-dependent “dust efficiency”

3. Velocity dependence of terms breaks degeneracy between **Φ(x)** and **ϵ(x)**



Credit: Lallement et al. 2019 A&A 625, A135

$$\left[v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial}{\partial v_i}\right] \ln f = 0$$

$$f_{\mathrm{obs}}(x, v) \propto \epsilon(x) f(x, v)$$

$$v_i \frac{\partial \ln f_{\mathrm{obs}}}{\partial x_i} - v_i \frac{\partial \ln \epsilon}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial \ln f_{\mathrm{obs}}}{\partial v_i} = 0$$

Solving the dust-corrected CBE

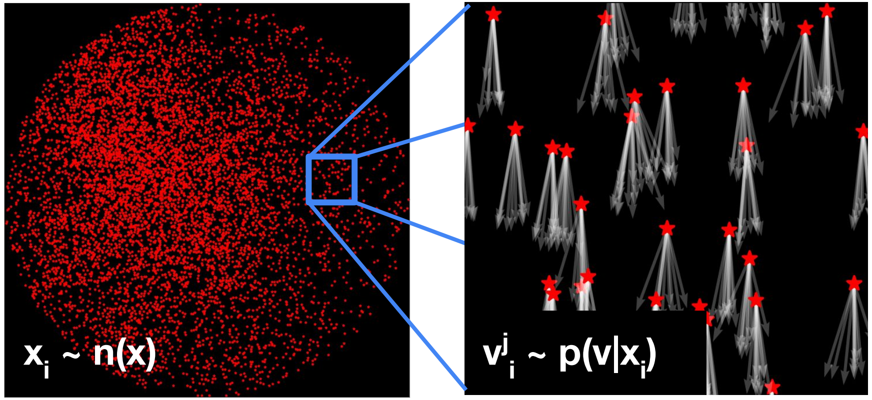
Given $\mathbf{f_{obs}(x,v)}$, we can solve the CBE for $\mathbf{\Phi(x)}$ and $\boldsymbol{\epsilon}\mathbf{(x)}$ simultaneously. We parameterize both with multi-layer perceptrons (MLPs).

$\mathbf{\Phi(x)}$, $\boldsymbol{\epsilon}\mathbf{(x)}$ training steps:

- Sample **(x,v)** pairs from $\mathbf{f_{obs}(x,v)}$
- Compute **df/dx**, **df/dv** at **(x,v)**
- Minimize the MSE of the CBE as a loss

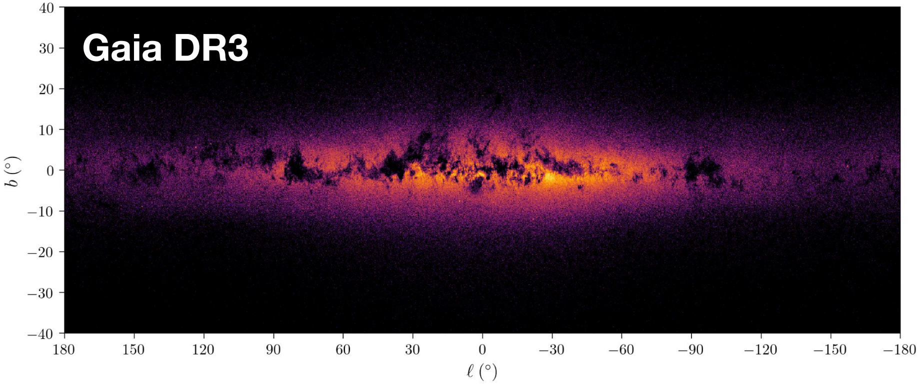
MLP architecture:

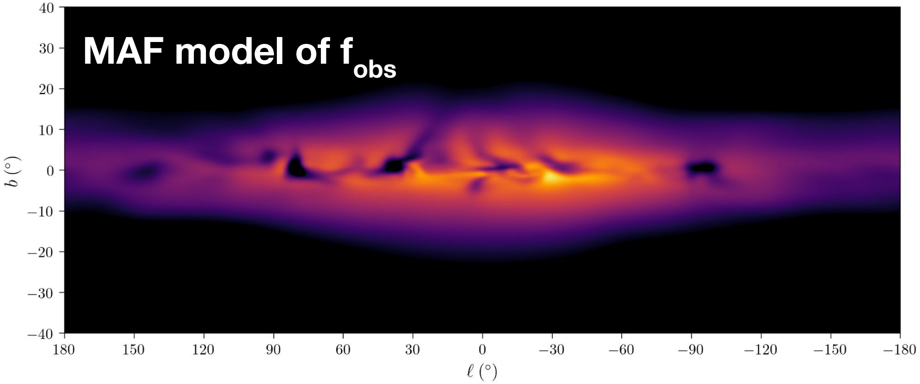
- 5x hidden linear layers, 100 wide
- GeLU activation
- Outputs of $\boldsymbol{\epsilon}\mathbf{(x)}$ are clamped within (0,1)



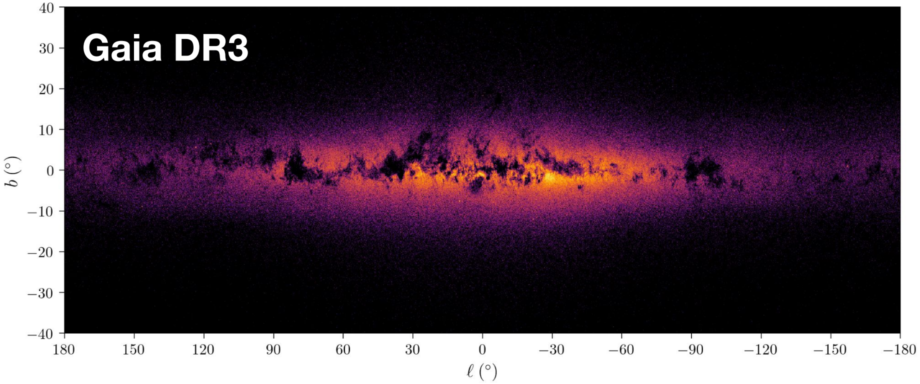
$$\mathcal{L}(\theta) = \sum_j \left| v_i \frac{\partial \ln f_{\rm obs}}{\partial x_i} - v_i \frac{\partial \ln \epsilon_\theta}{\partial x_i} - \frac{\partial \Phi_\theta}{\partial x_i} \frac{\partial \ln f_{\rm obs}}{\partial v_i} \right|^2_{x_j, v_j}$$

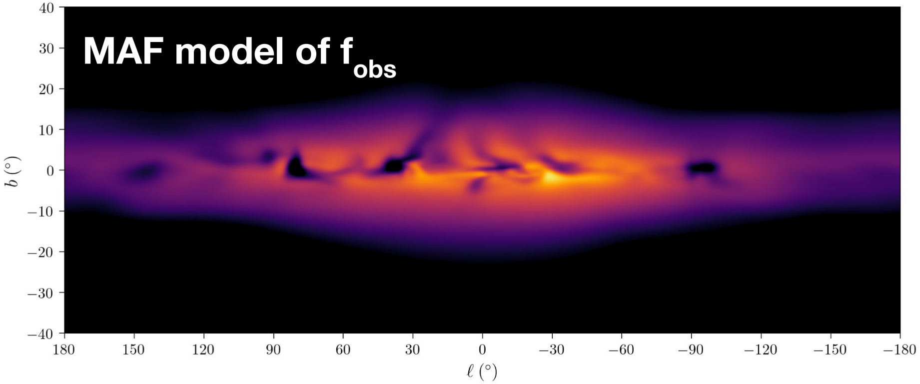
Unbiased Phase Space Density

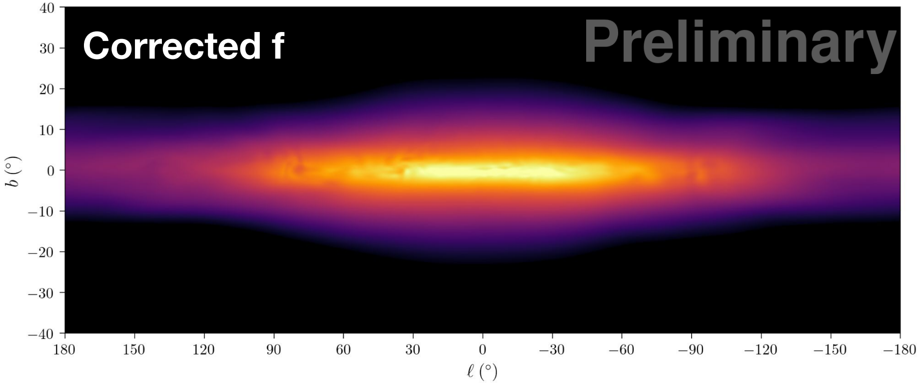




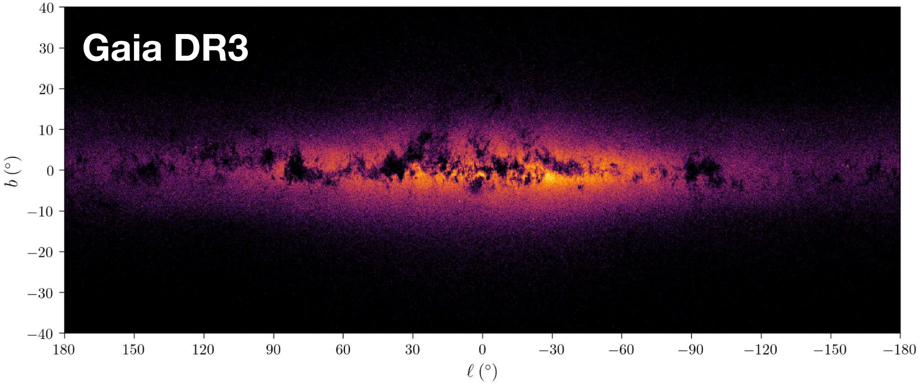
Unbiased Phase Space Density

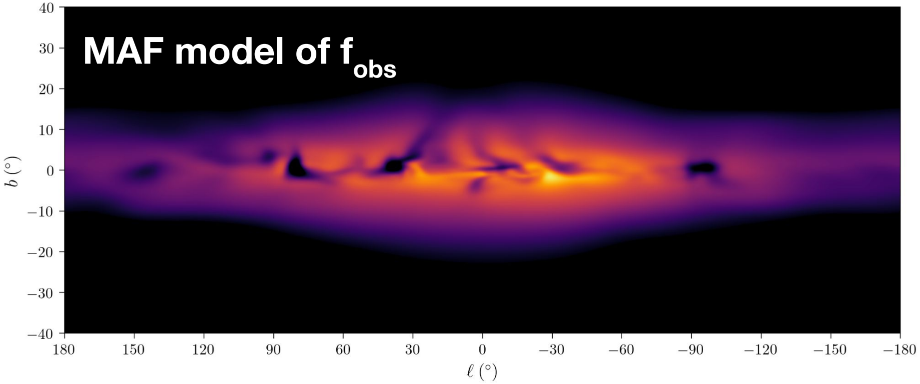


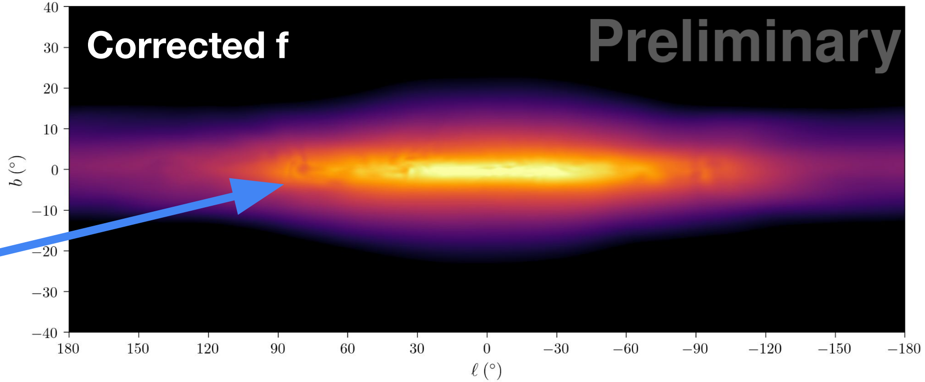




Unbiased Phase Space Density





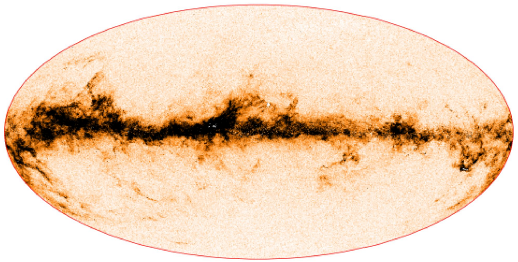


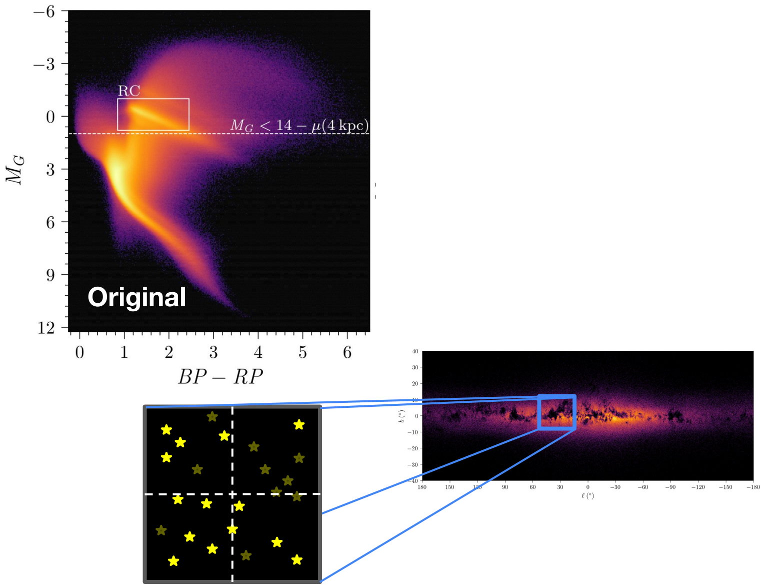
Some wrinkles visible at edges

Validating ϵ

How do we verify that ϵ is working?

Construct an effective ϵ based on a 3D dust map (Lallement 2022[1] i.e. L22).





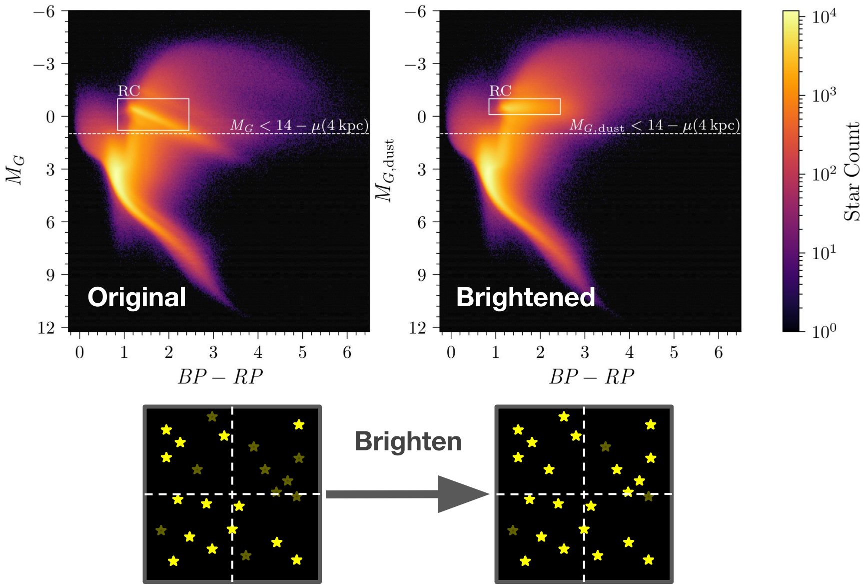
Validating ϵ

How do we verify that ϵ is working?

Construct an effective ϵ based on a 3D dust map (Lallement 2022[1] i.e. L22).

Steps:

- Brighten observed but dimmed stars



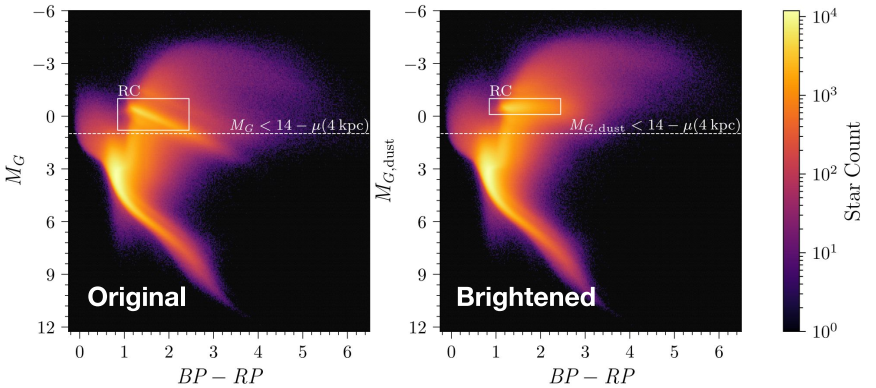
Validating ϵ

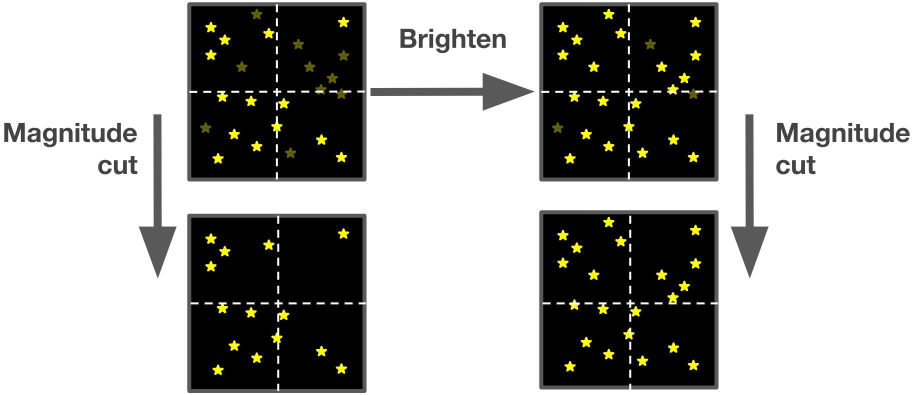
How do we verify that ϵ is working?

Construct an effective ϵ based on a 3D dust map (Lallement 2022[1] i.e. L22).

Steps:

- Brighten observed but dimmed stars
- Repeat magnitude cut for completeness





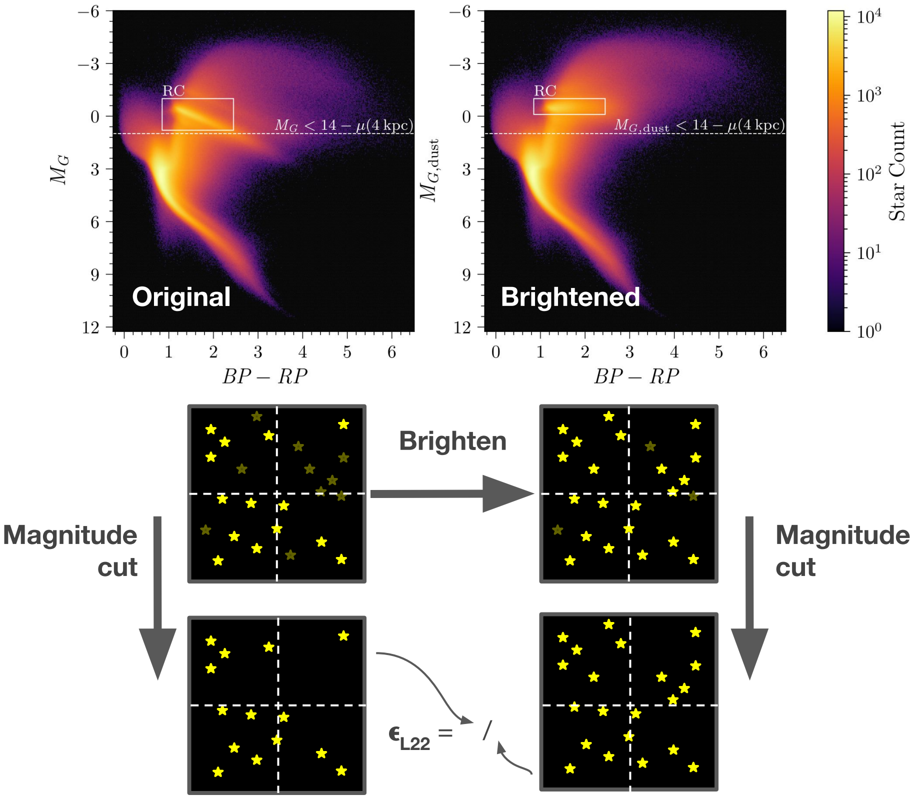
Validating ϵ

How do we verify that ϵ is working?

Construct an effective ϵ based on a 3D dust map (Lallement 2022[1] i.e. L22).

Steps:

- Brighten observed but dimmed stars
- Repeat magnitude cut for completeness
- Compute ratio of number densities: ϵ_{L22}



Validating ϵ

How do we verify that ϵ is working?

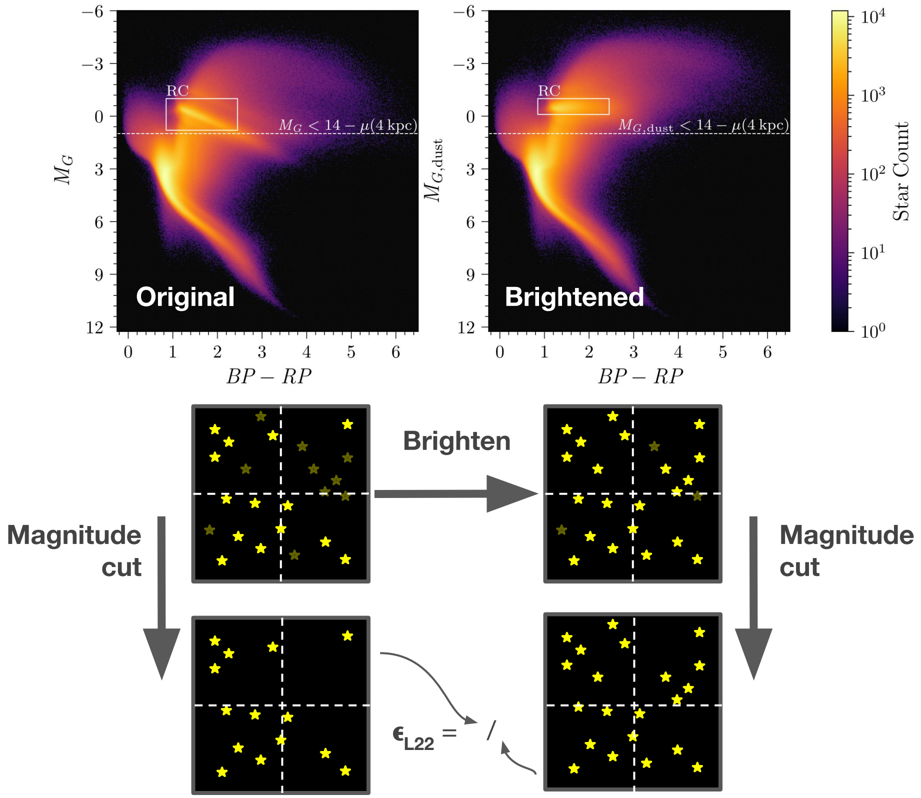
Construct an effective ϵ based on a 3D dust map (Lallement 2022[1] i.e. L22).

Steps:

- Brighten observed but dimmed stars
- Repeat magnitude cut for completeness
- Compute ratio of number densities: ϵ_{L22}

Limitations:

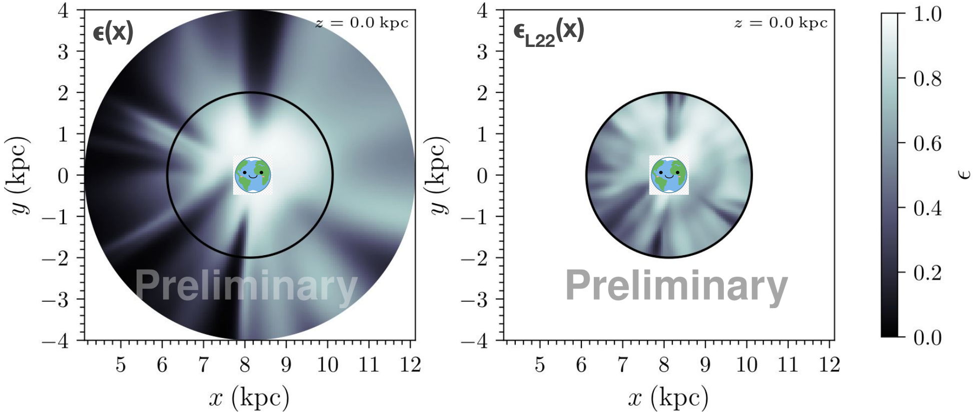
- Limited to nearby, totally complete regions of space i.e. within ~2.5 kpc
- ϵ captures dust and over-smoothing, should look like a blurred version of ϵ_{L22}



Comparing ϵ and ϵ_{L22}

ϵ and ϵ_{L22} consistent over the comparison volume!

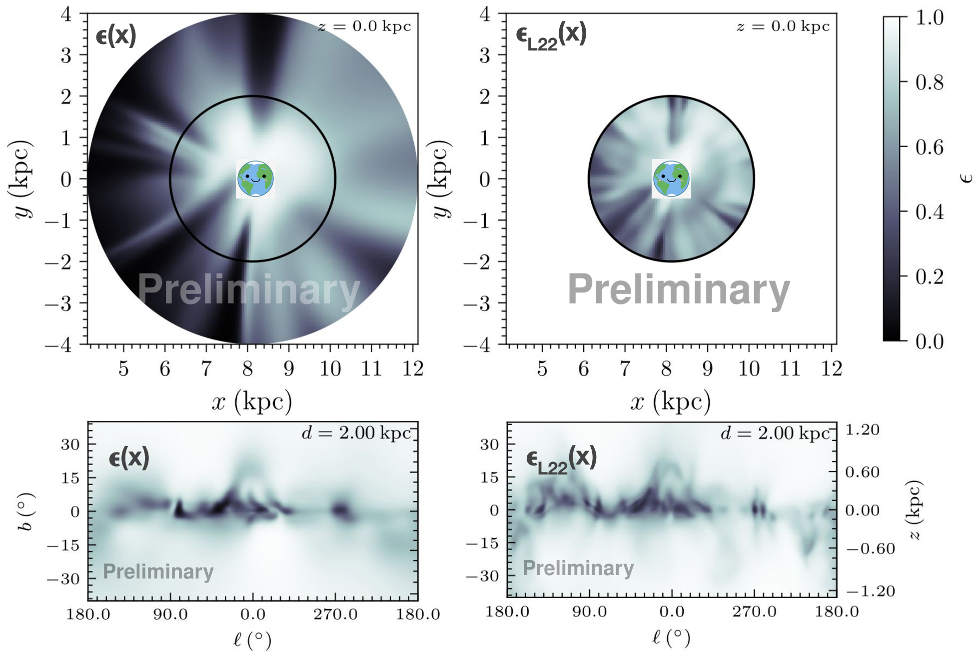
- We are estimating the most dust-dense regions accurately



Comparing ϵ and ϵ_{L22}

ϵ and ϵ_{L22} consistent over the comparison volume!

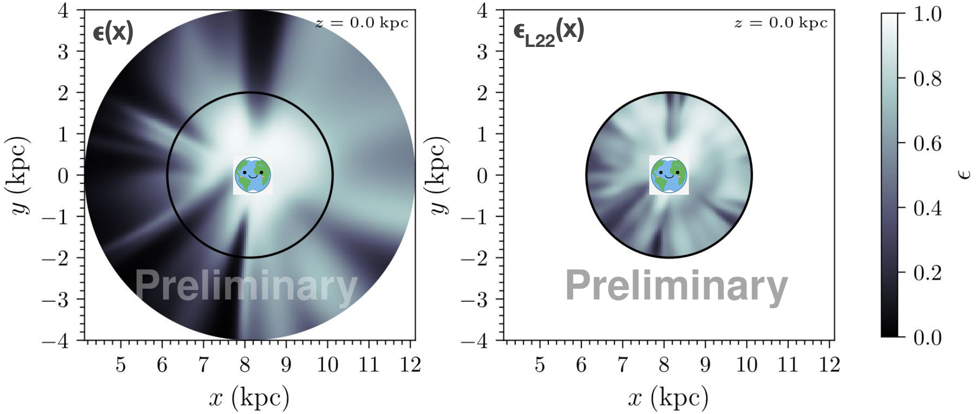
- We are estimating the most dust-dense regions accurately
- Dust cloud complexes are reproduced when projected on the sky

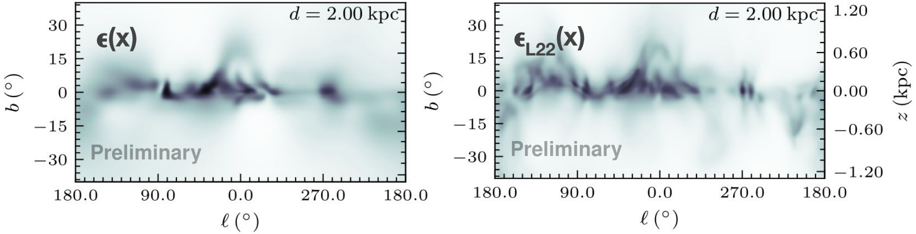


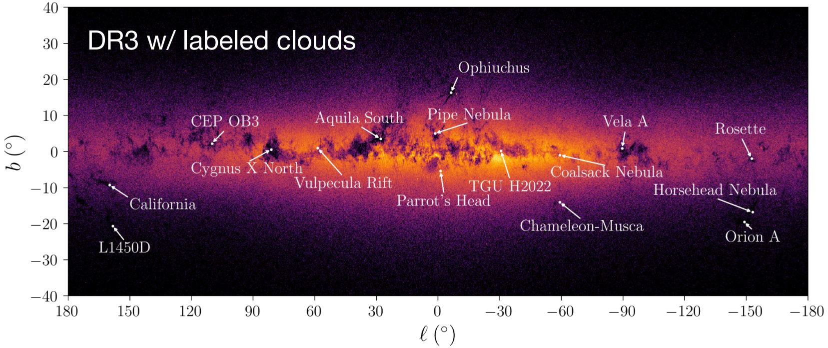
Comparing ϵ and ϵ_{L22}

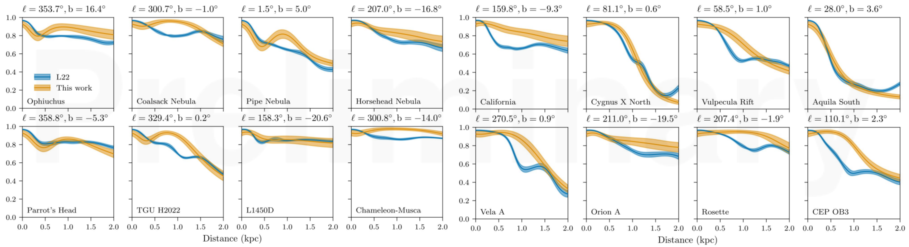
ϵ and ϵ_{L22} consistent over the comparison volume!

- We are estimating the most dust-dense regions accurately
- Dust cloud complexes are reproduced when projected on the sky
- Extinction profiles consistent along dusty lines of sight





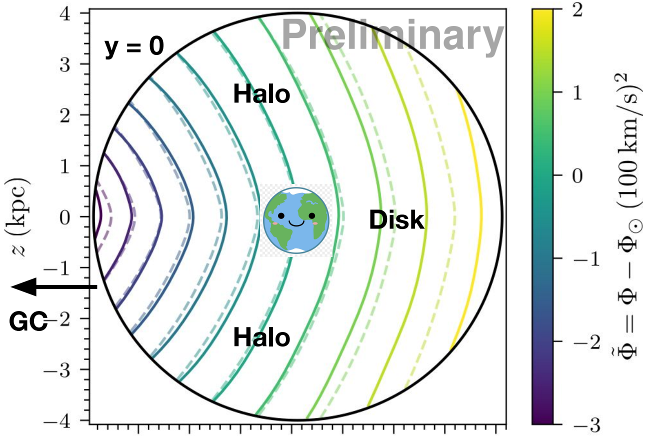




Dust-Corrected Potential

We compare **Φ(x)** to the standard benchmark Milky Way potential MWPotential2014[2].

- We recover a smoother potential in previously dust-obscured regions of the disk

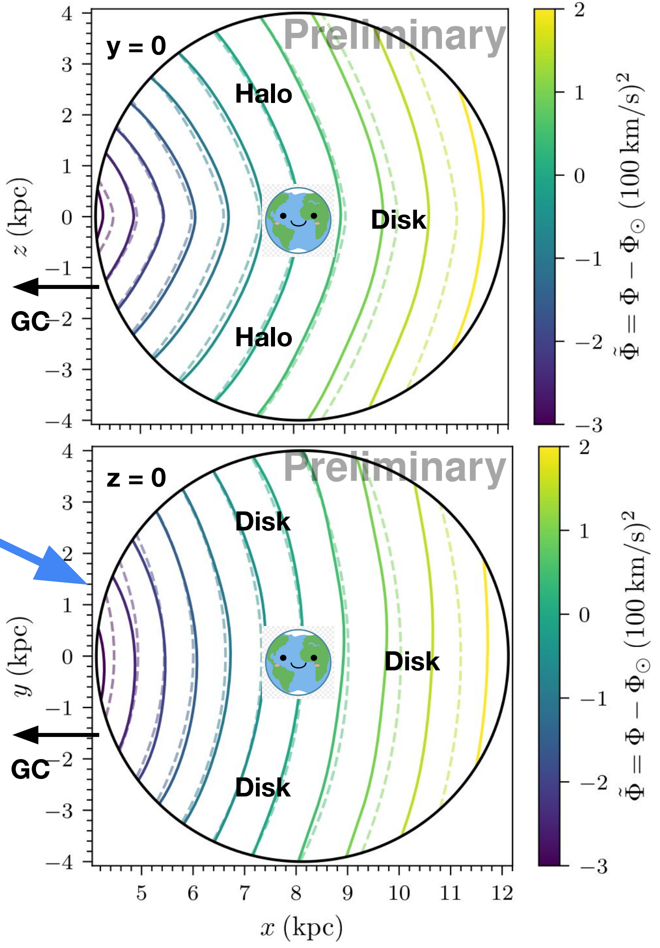


Solid: CBE-derived **Φ(x)**
Dashed: MWPotential2014 **Φ(x)**

Dust-Corrected Potential

We compare **Φ(x)** to the standard benchmark Milky Way potential MWPotential2014[2].

- We recover a smoother potential in previously dust-obscured regions of the disk
- Deviations from the standard axisymmetric picture are apparent

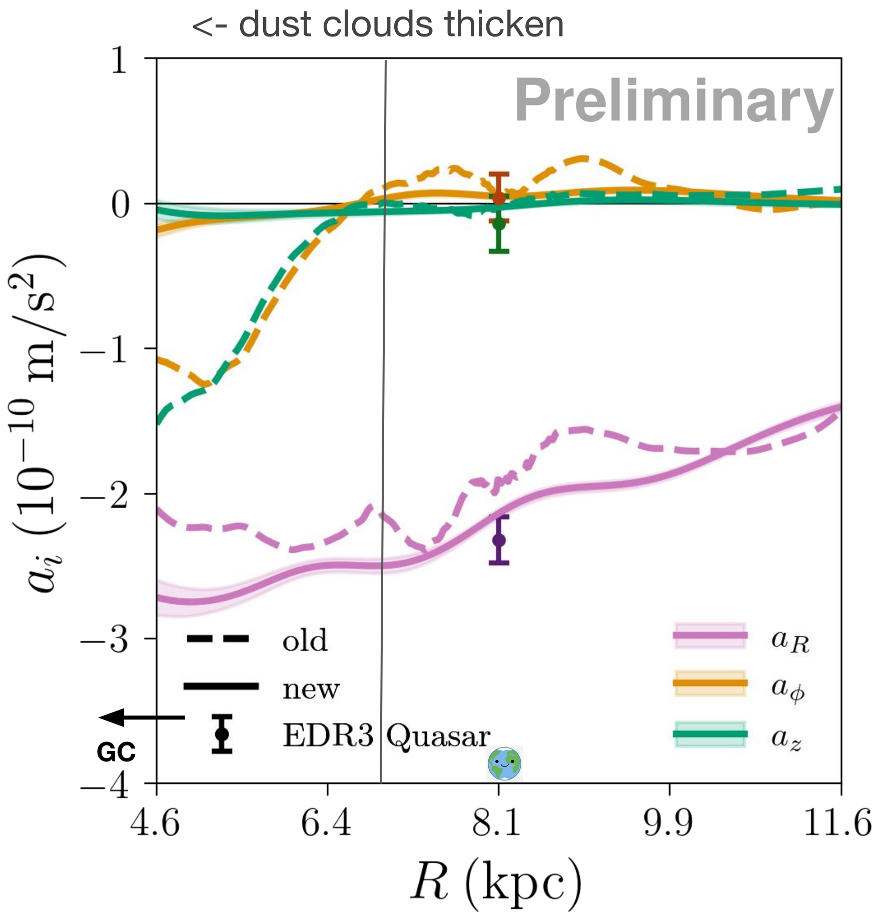


Solid: CBE-derived **Φ(x)**
Dashed: MWPotential2014 **Φ(x)**

Acceleration & Mass Density

Acceleration improvements:

- Parameterizing Φ(x) with an MLP smooths **a(x)** significantly
- Dust-correction stabilizes **a(x)** in dusty regions of the sky. Can access the disk!



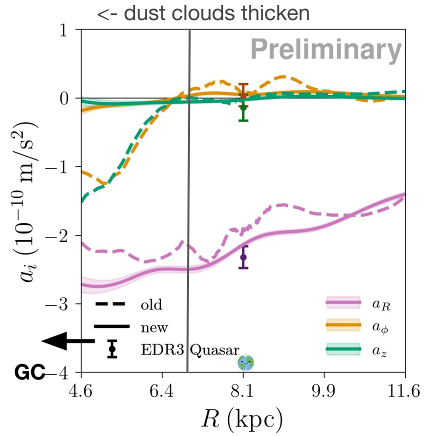
Acceleration & Mass Density

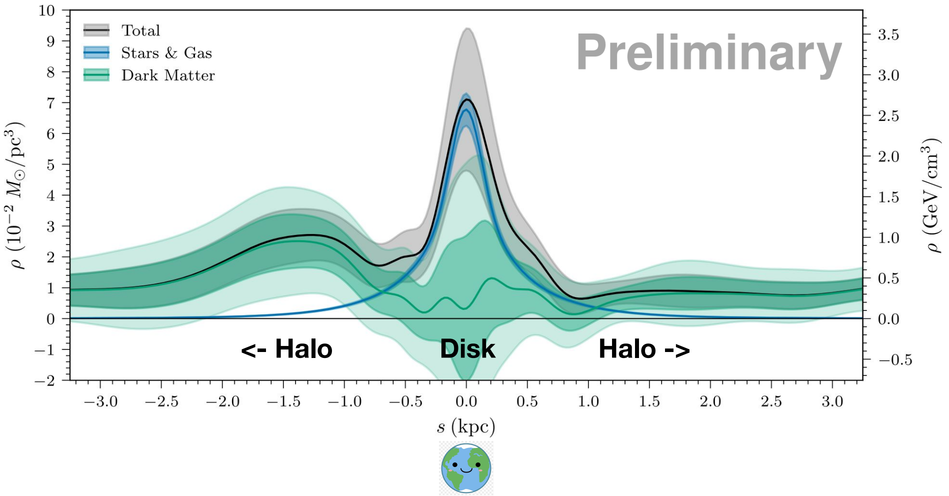
Acceleration improvements:

- Parameterizing Φ(x) with an MLP smooths **a(x)** significantly
- Dust-correction stabilizes **a(x)** in dusty regions of the sky. Can access the disk!

Finally, we estimate the local dust-corrected mass density field **ρ(x).**

- Similar smoothing benefit, estimate is stable in the disk
- Exhibits noisy oscillations in some regions

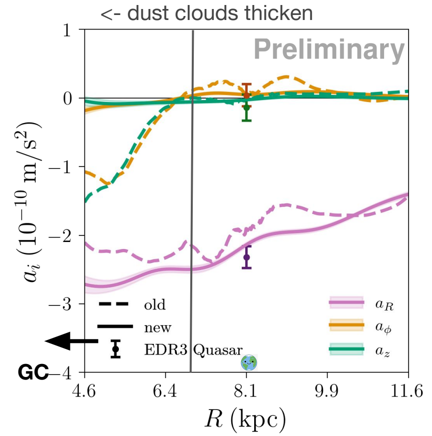




Acceleration & Mass Density

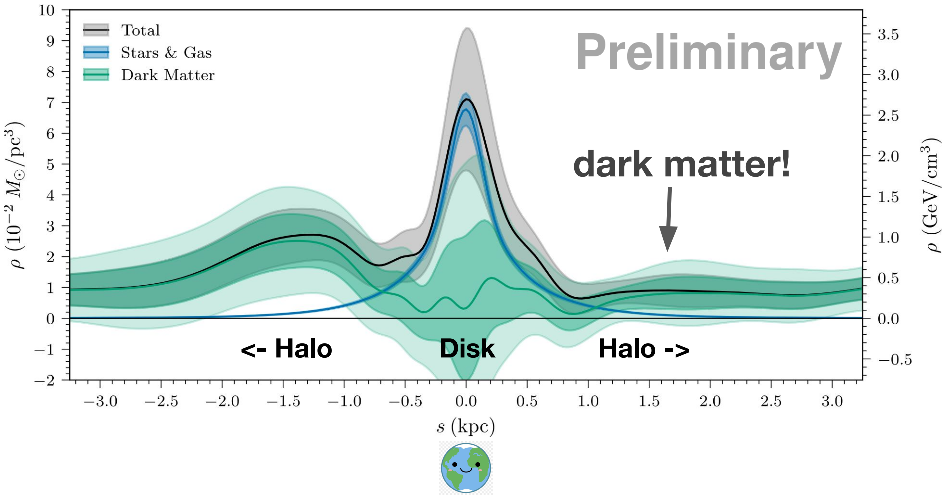
Acceleration improvements:

- Parameterizing Φ(x) with an MLP smooths **a(x)** significantly
- Dust-correction stabilizes **a(x)** in dusty regions of the sky. Can access the disk!



Finally, we estimate the local dust-corrected mass density field **ρ(x).**

- Similar smoothing benefit, estimate is stable in the disk
- Exhibits noisy oscillations in some regions
- Residual map of dark matter coming into focus…



Summary

We have developed a data-driven technique for removing dust extinction bias from kinematic phase space.

Our estimate of this bias is consistent with and builds upon existing 3D dust maps.

Used the unbiased f(x,v) to map the local potential. Closer to a fully data-driven map of dark matter in the Milky Way.

DALL-E: “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”

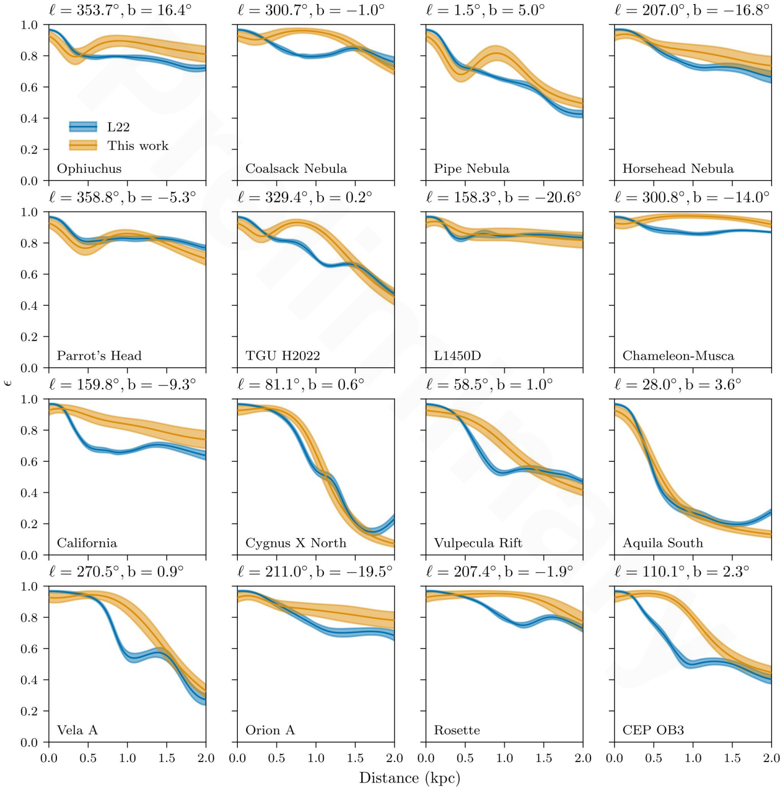
Thank you! arXiv:240X.XXXXX

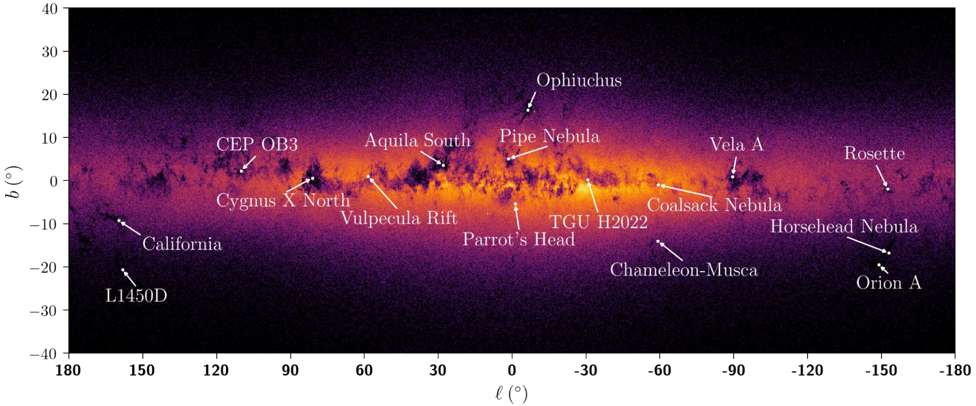
References:

1. R. Lallement, J. L. Vergely, C. Babusiaux, and N. L. J. Cox, Astronomy & Astrophysics 661, A147 (2022), ISSN 1432-0746
2. J. Bovy, 216, 29 (2015), 1412.3451

Additional Slides

Flying through the clouds (legible edition)

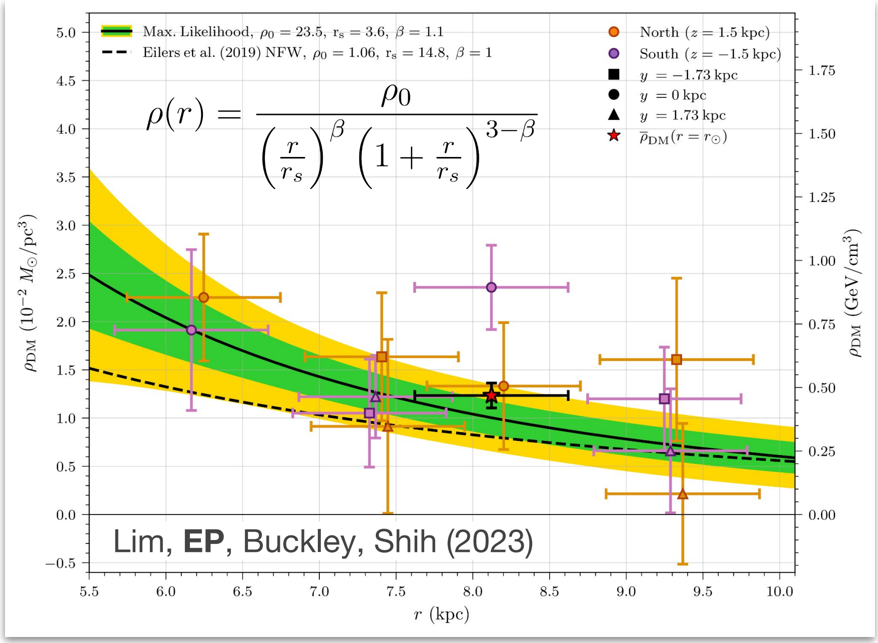


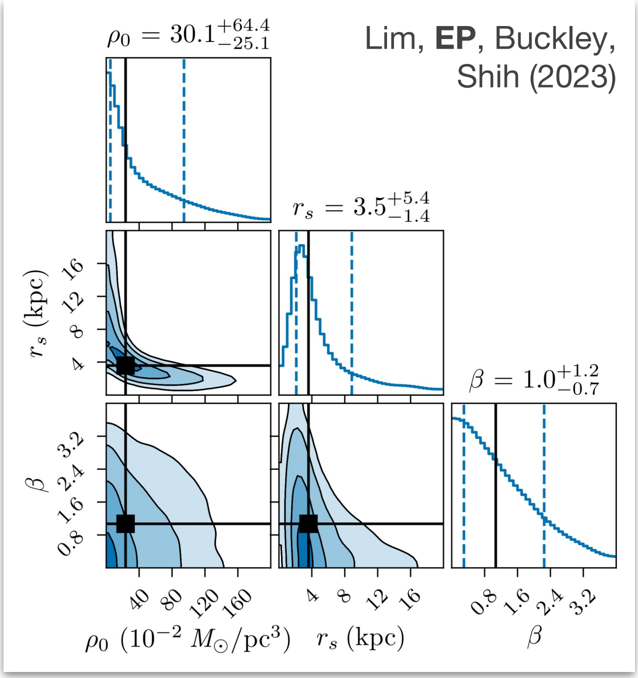


- Consistent profiles for most clouds
- Overshoot distance to some clouds
- We miss fainter clouds
- Dust map error bars are very tight!

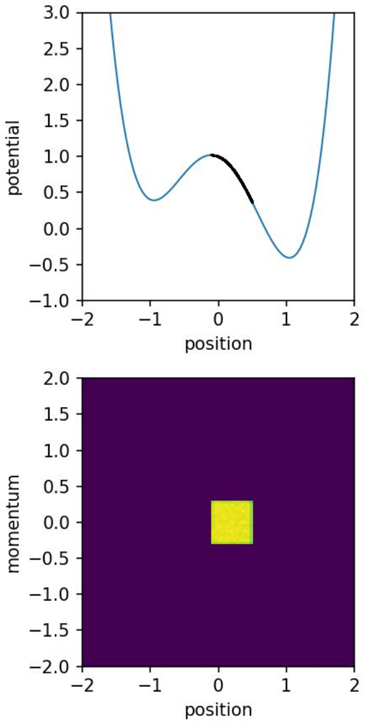
Radial Profile of the Milky Way's Dark Matter Halo

Can place loose constraints on the MW halo. Radial profile consistent with NFW (β=1) when fit to a generalized NFW (gNFW) profile.

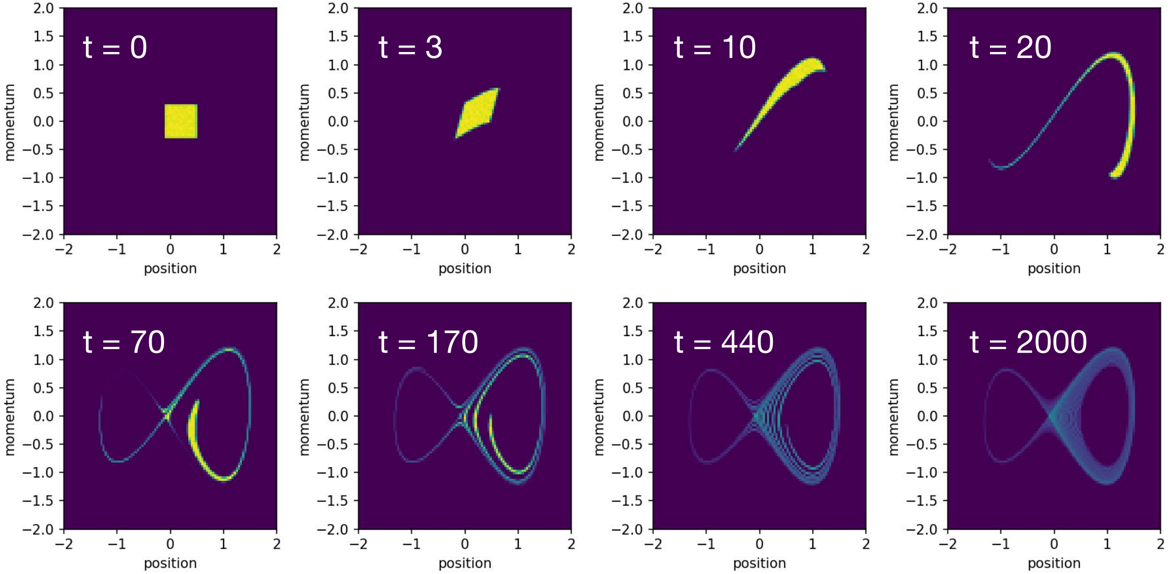




Phase mixing & equilibrium

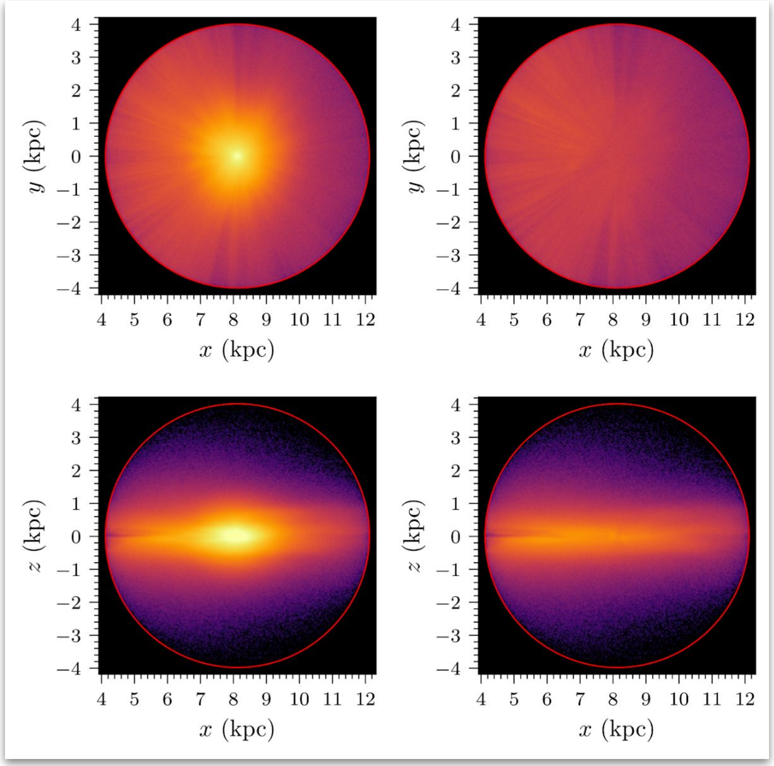


$$\frac{df}{dt} = \left[\frac{\partial}{\partial t} + v_i \frac{\partial}{\partial x_i} + a_i \frac{\partial}{\partial v_i}\right] f = 0$$

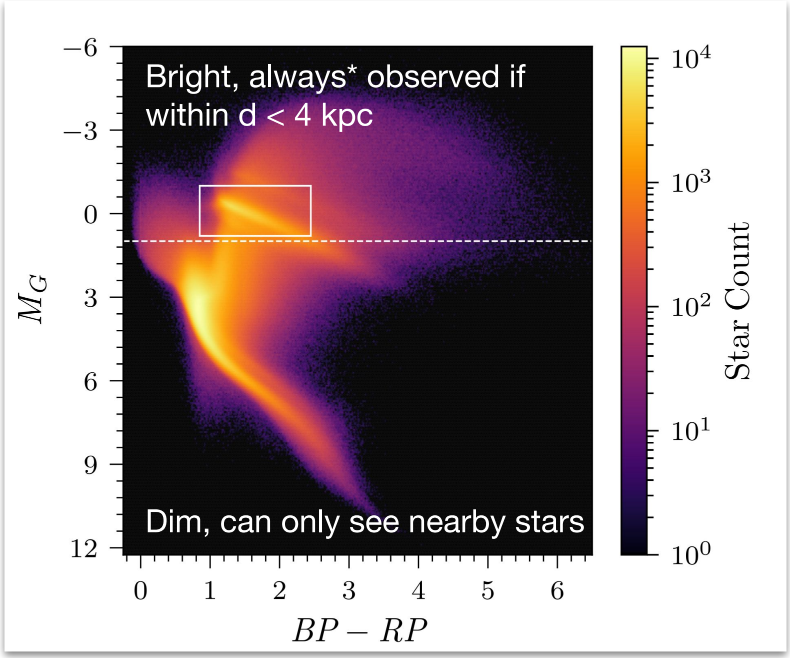


$$\frac{df}{dt} = \left[\cancel{\frac{\partial}{\partial t}} + v_i \frac{\partial}{\partial x_i} + a_i \frac{\partial}{\partial v_i}\right] f = 0$$

Gaia Dataset: The Red Clump (and friends)



Lim, **EP**, Buckley, Shih (2023)



Jeans Equations

Moments of the CBE yield the Jeans equations:

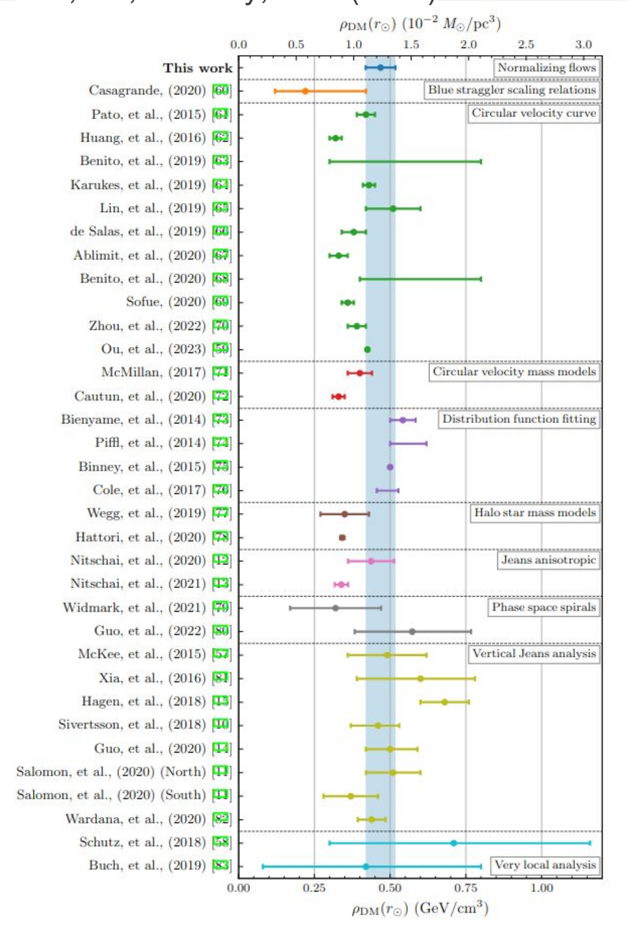
$$\int d^3p(p_z * \mathrm{CBE}) \longrightarrow \quad a_z = \frac{1}{\nu}\frac{\partial(\nu\overline{v_z^2})}{\partial z} + \frac{1}{\nu R}\frac{\partial(\nu R\overline{v_R v_z})}{\partial R}$$

$$\int d^3p(p_R * \mathrm{CBE}) \longrightarrow \quad a_R = \frac{1}{\nu}\frac{\partial(\nu\overline{v_R^2})}{\partial R} + \frac{1}{\nu}\frac{\partial(\nu\overline{v_R v_z})}{\partial z} + \frac{\overline{v_R^2} - \overline{v_\phi^2}}{R}$$

Key notes:

- Axisymmetry implicitly assumed when $a_\phi = 0$.
- Must model each velocity moment and number density
- Many terms are noisy. Calculation of ρ can involve high order derivatives of fits to noisy data.

Lim, **EP**, Buckley, Shih (2023)



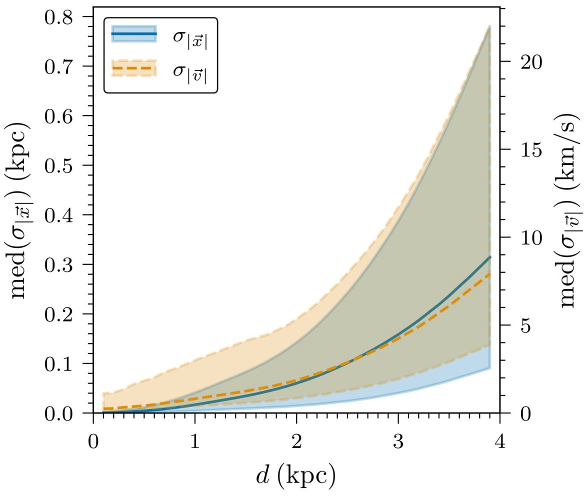
Lim, **EP**, Buckley, Shih (2023)

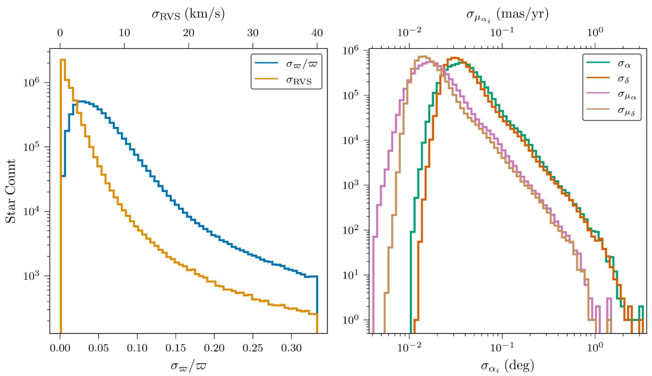
Gaia Measurement Errors

Astrometric errors and propagated from ICRS to Cartesian coordinates.

Primary uncertainties for stars are parallax and radial velocities.

- We cut relative stars w/ relative parallax errors > .33



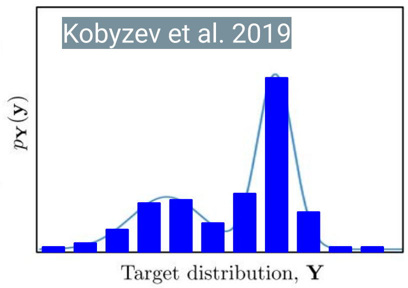


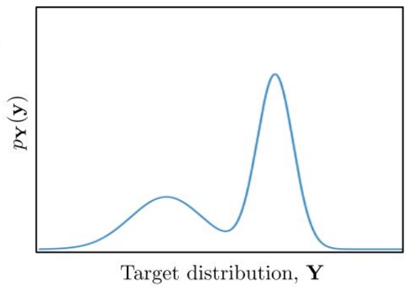
Normalizing Flows

Stats definition: A transformation **g** of a base distribution **p(z)** into a target distribution **p(y)** by a sequence of invertible and differentiable mappings.

Key Ideas:

- There are simple distributions **p(z)** that we understand.
- Our data was drawn from some unknown distribution **p(y)**.
- Training a normalizing flow -> finding a smooth transformation **g** (with inverse **f**) between **p(z)** and **p(y)**. **g** may can be a sequence of smooth and invertible transformations.





Chain rule lets us write the following:

$$p_Y(y) = p_Z(f(y))\frac{df(y)}{dy}$$

Maximize likelihood of data = minimizing negative log-likelihood of data

Masked Autoregressive Flows

Autoregressive model: transformation modeled as a product of conditionals. MAFs enforce autoregressive property by masking network connections, allowing for complex non-linear transformations.

$$p(x_1, ..., x_n) = p(x_1)p(x_2|x_1)p(x_3|x_2, x_1)...p(x_n|x_{n-1}, ..., x_1)$$

Main idea: The order you specify dimensions should not matter. (x,y,z vs z,y,x)

If there are correlations in the data, it should find them on its own!