

Precision Galactic Archaeology : Revealing the Milky Way's
Engines through the *Statistical Alignments* of Stars

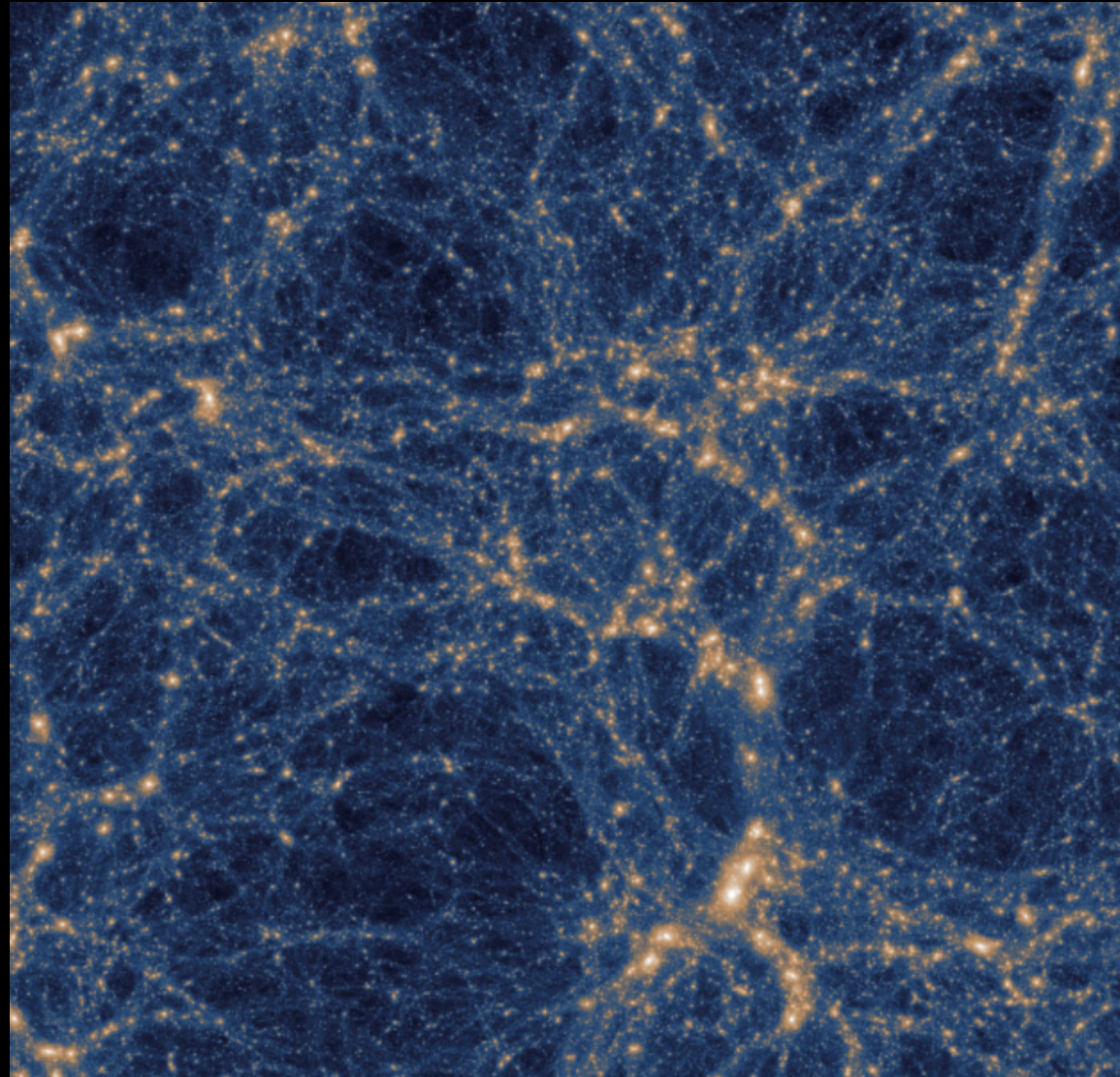
Yuan-Sen Ting

Australian National University

@TingAstro

Can we study the formation of star clusters *without* looking
for *clustering* of stars?

Studying the statistical alignments of objects is *not new*



Cosmology

Large-scale structure

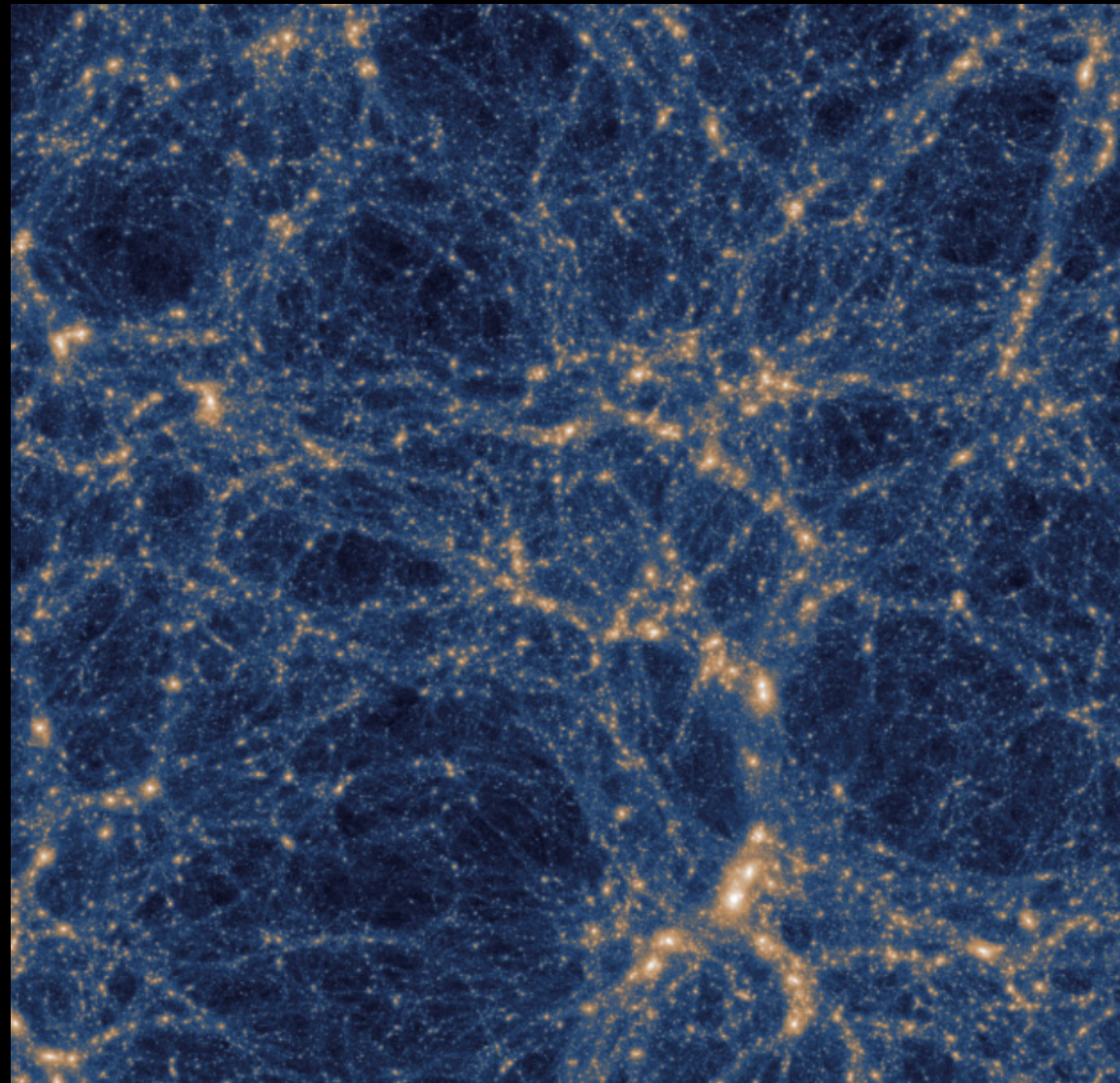
TNG Simulation

Studying the statistical alignments of objects is *not new*



Cosmology

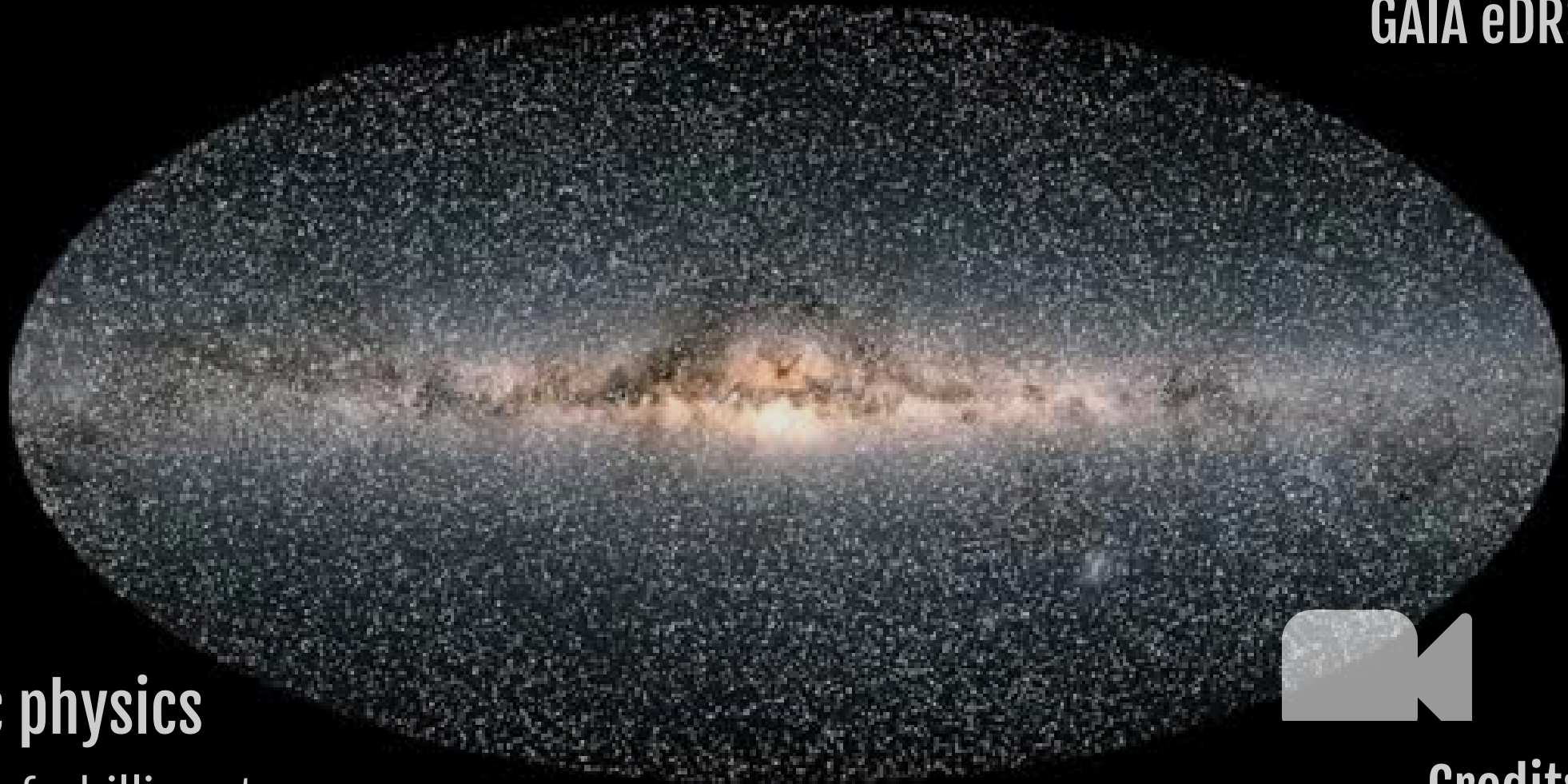
Large-scale structure



**TNG
Simulation**

Recently, Gaia has moved "*statistical astrophysics*" closer to home !

GAIA eDR3



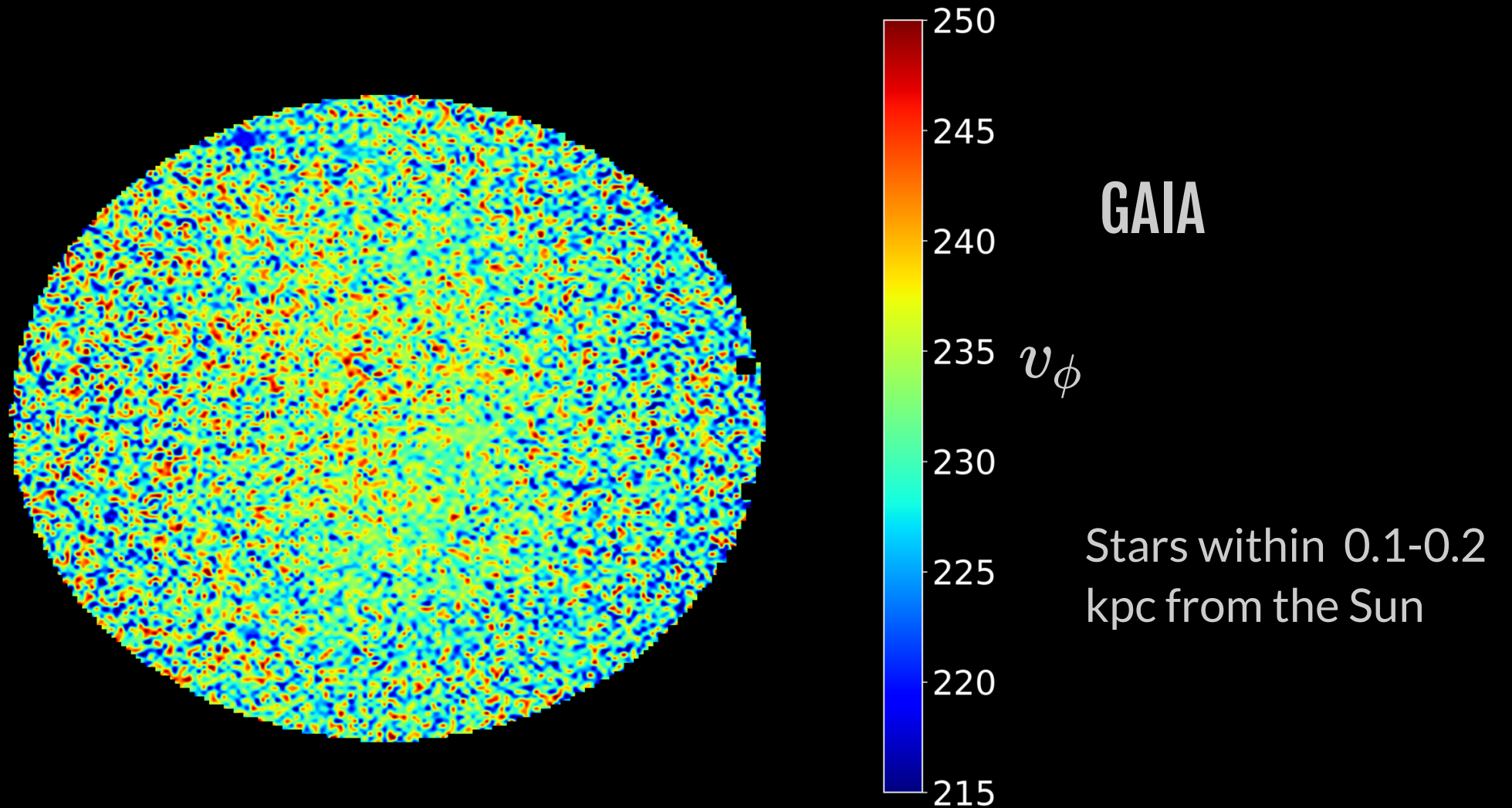
Galactic physics

Motions of a billion stars



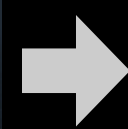
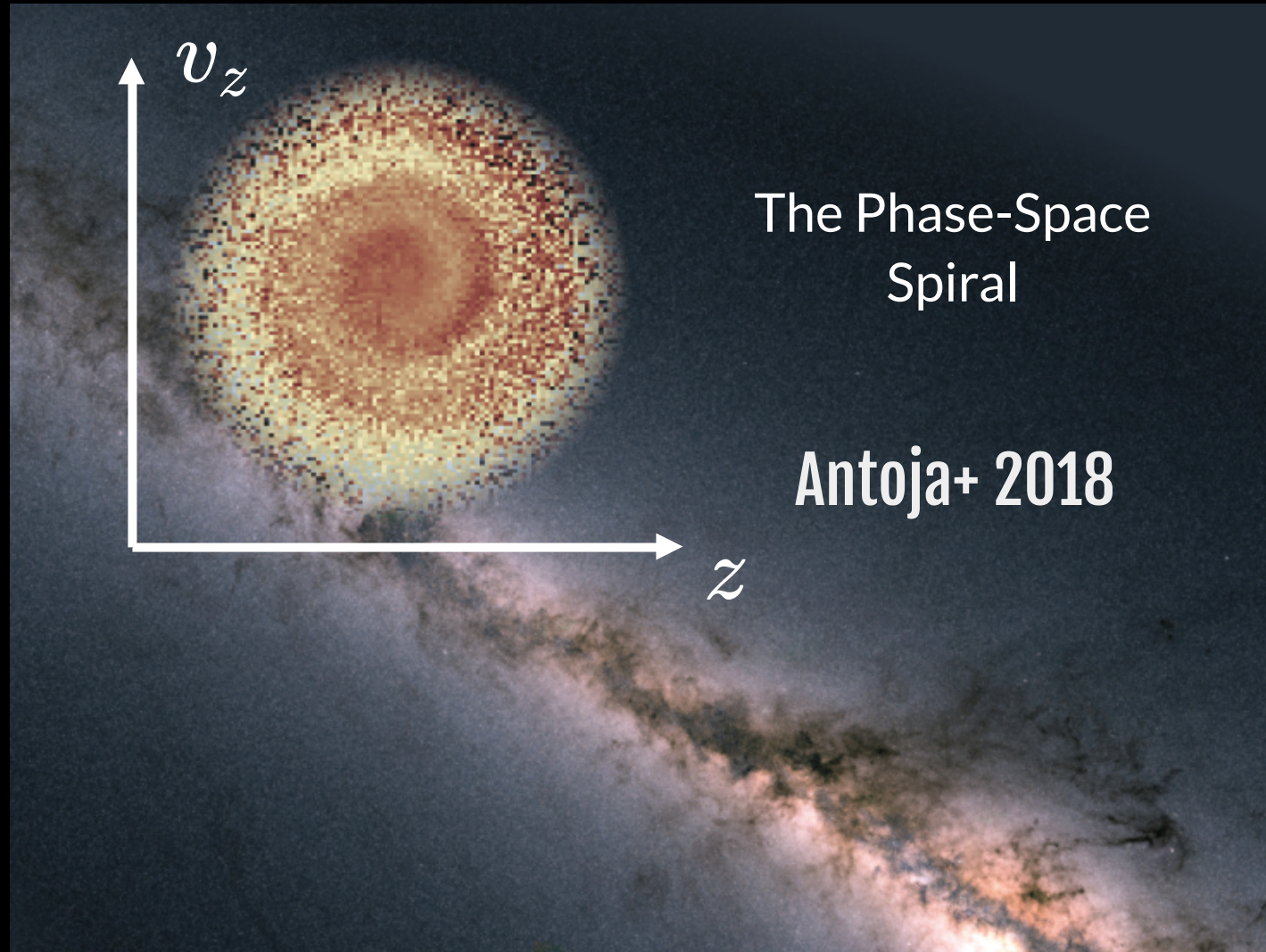
Credit: ESA

Gaia reveals that the Milky Way is *not in equilibrium*



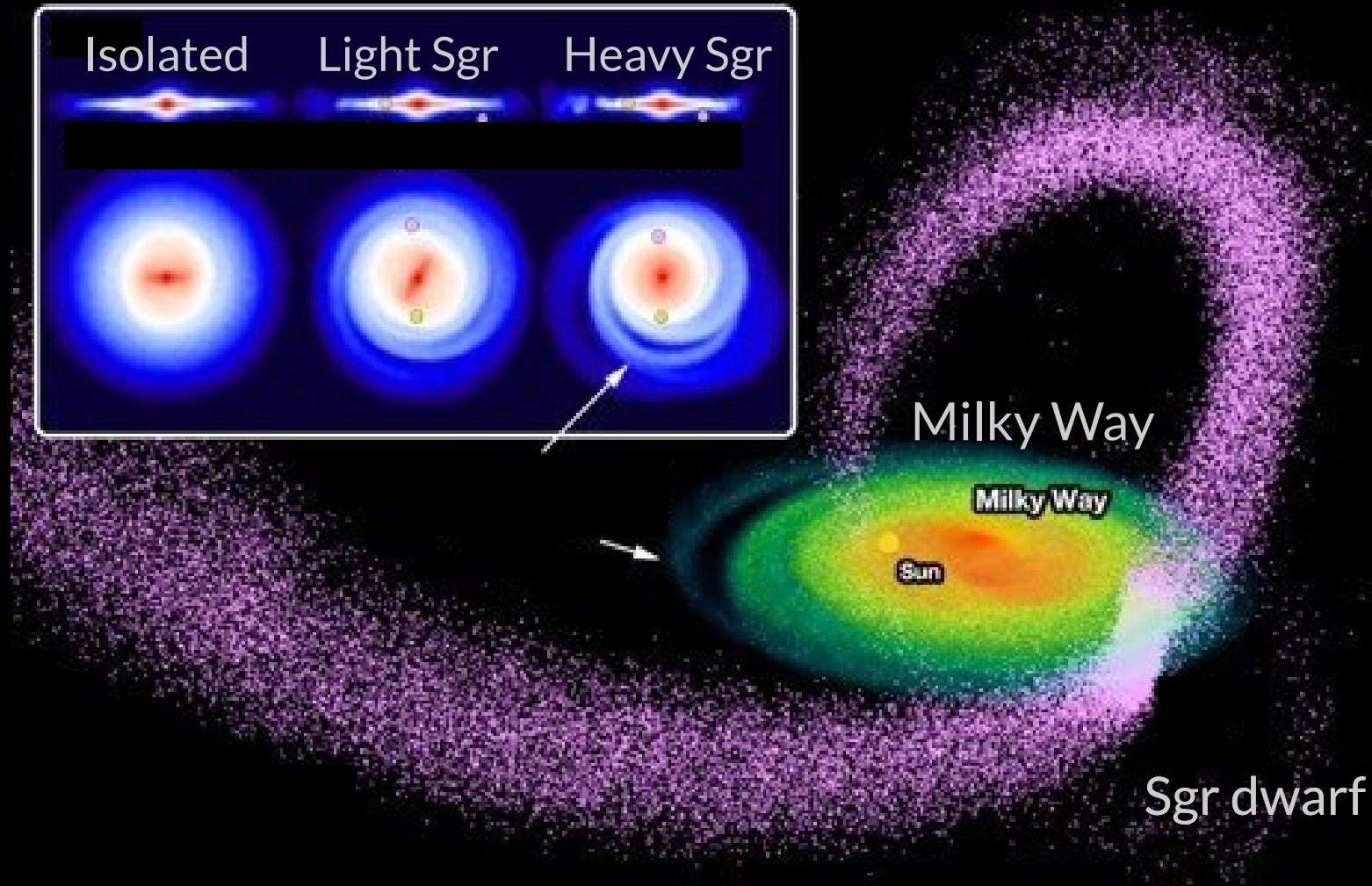
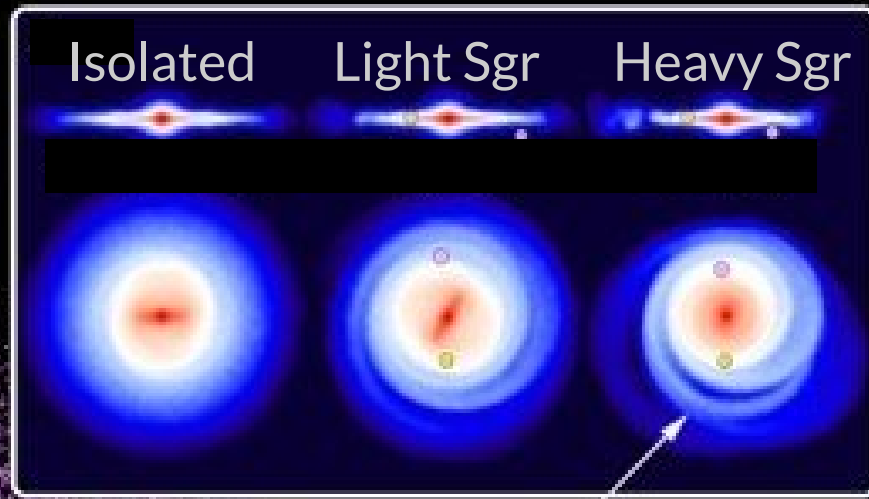


Quantifying the *"ripples"* of the Milky Way in action !

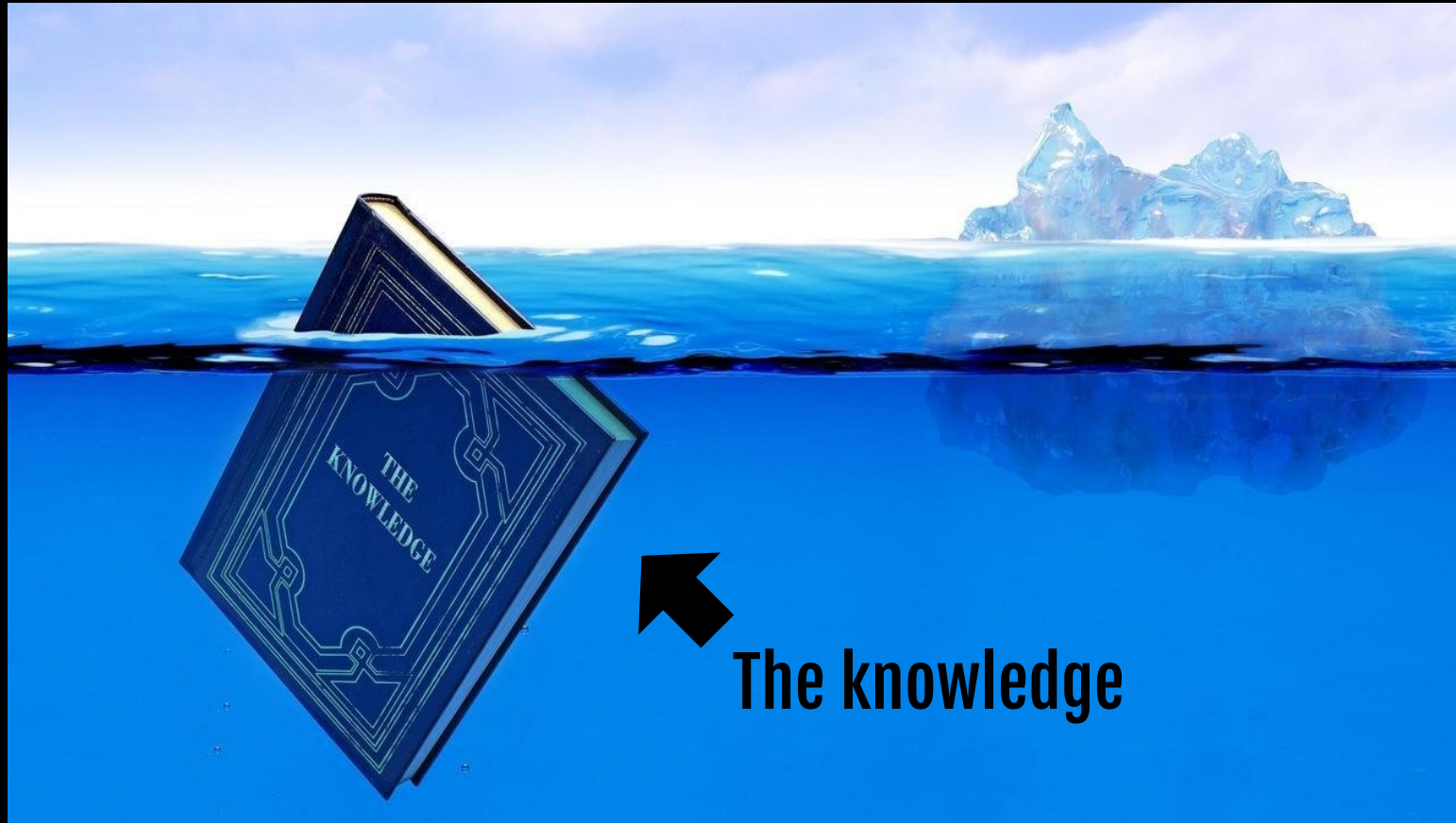


Perturbation from the Sagittarius dwarf galaxy, ~5 Gyr ago

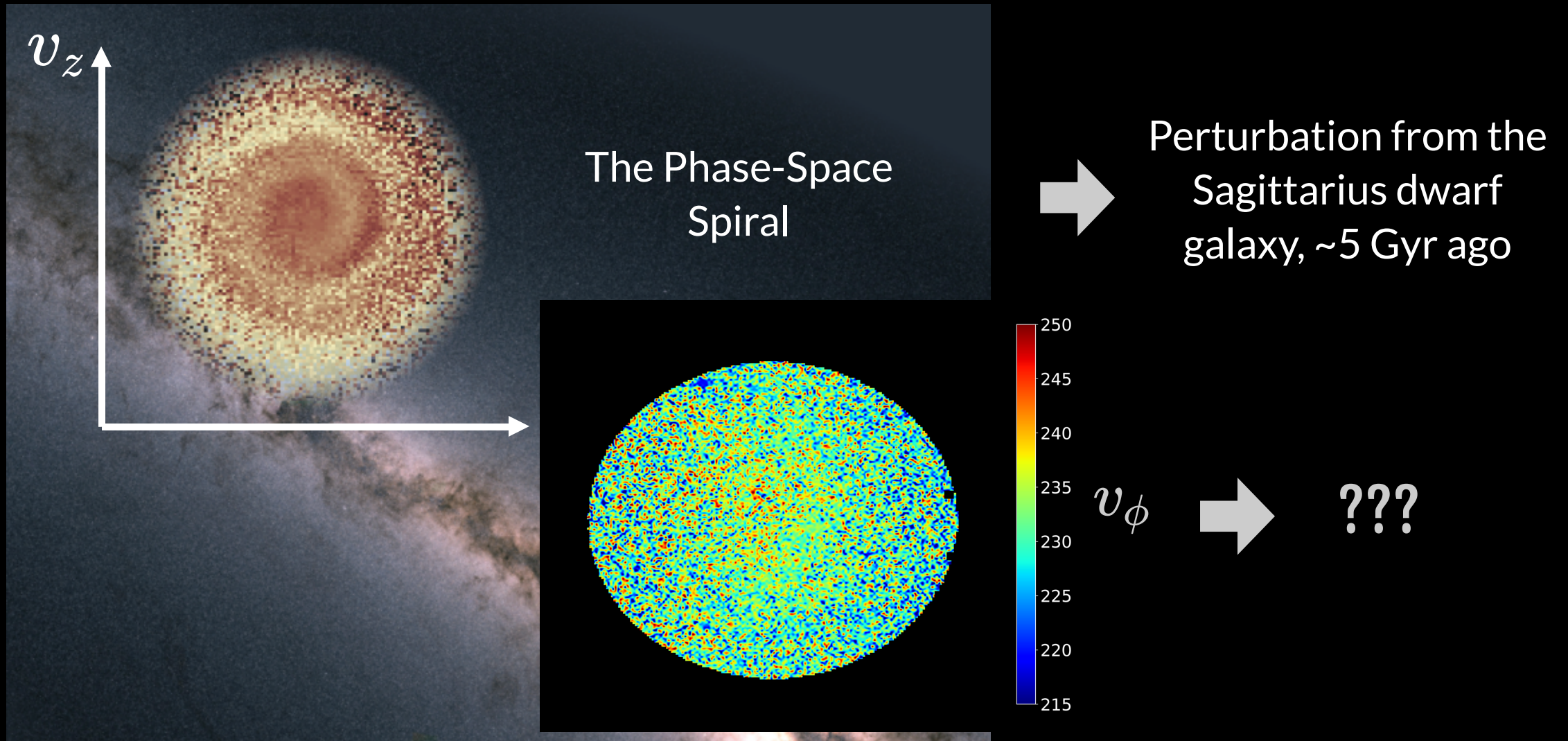
Quantifying the *"ripples"* of the Milky Way in action !



But, we have focused mostly on the *"blockbuster" events*



Non-equilibrium at *different scales* can reveal *different physics*

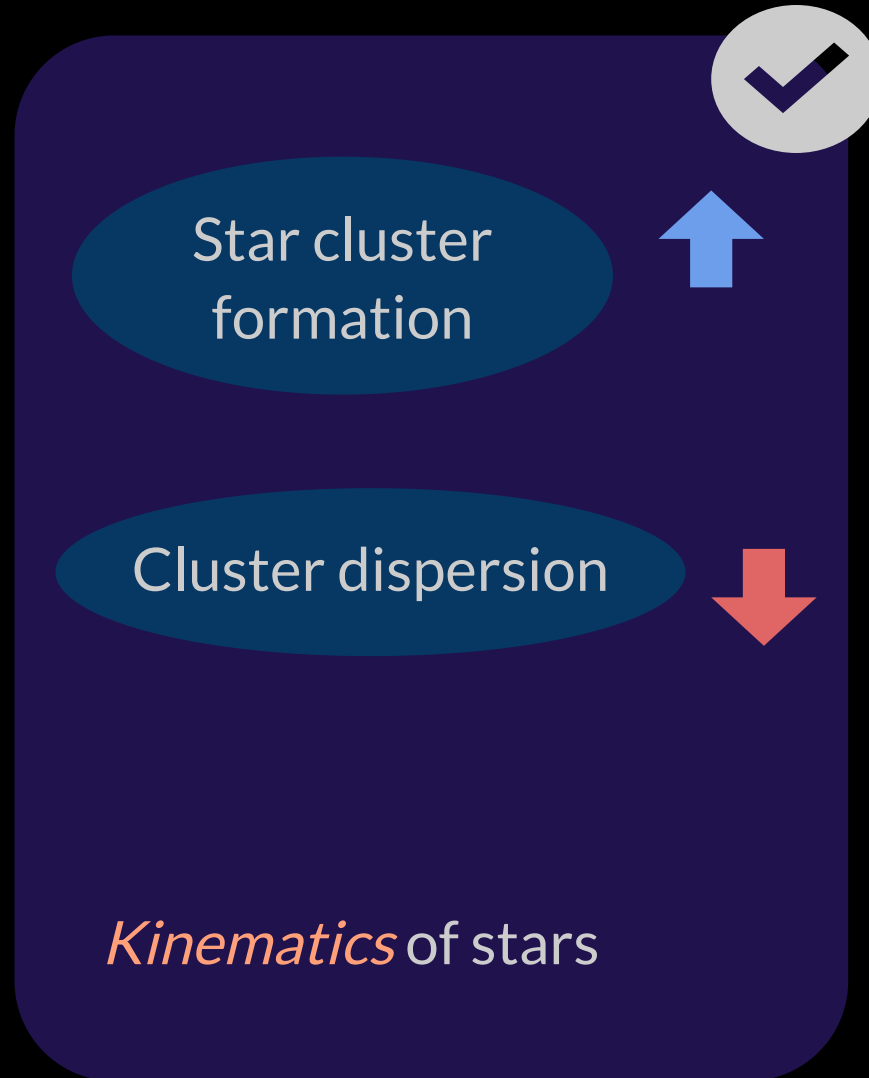


Statistical alignments of stars can constrain various *subgrid physics*



The Milky Way is never in equilibrium because of *"subgrid" physics*

Entropy level



Quantifying statistical alignments with *power spectrum*

Number of pairs in the *non-equilibrium* case

Number of pairs in the *smooth* case ("random-random")



Two-point correlation function

$$\xi(\Delta r) \equiv$$

$$\frac{DD(\Delta r) - RR(\Delta r)}{RR(\Delta r)}$$



The separation of two stars



Normalization

Fourier transform of "power spectrum"

see also, Bovy+2015, Khanna+ 2019, Lancaster+2019, ...

On how to estimate "*random-random*" / the smooth Milky Way

Two-point correlation function

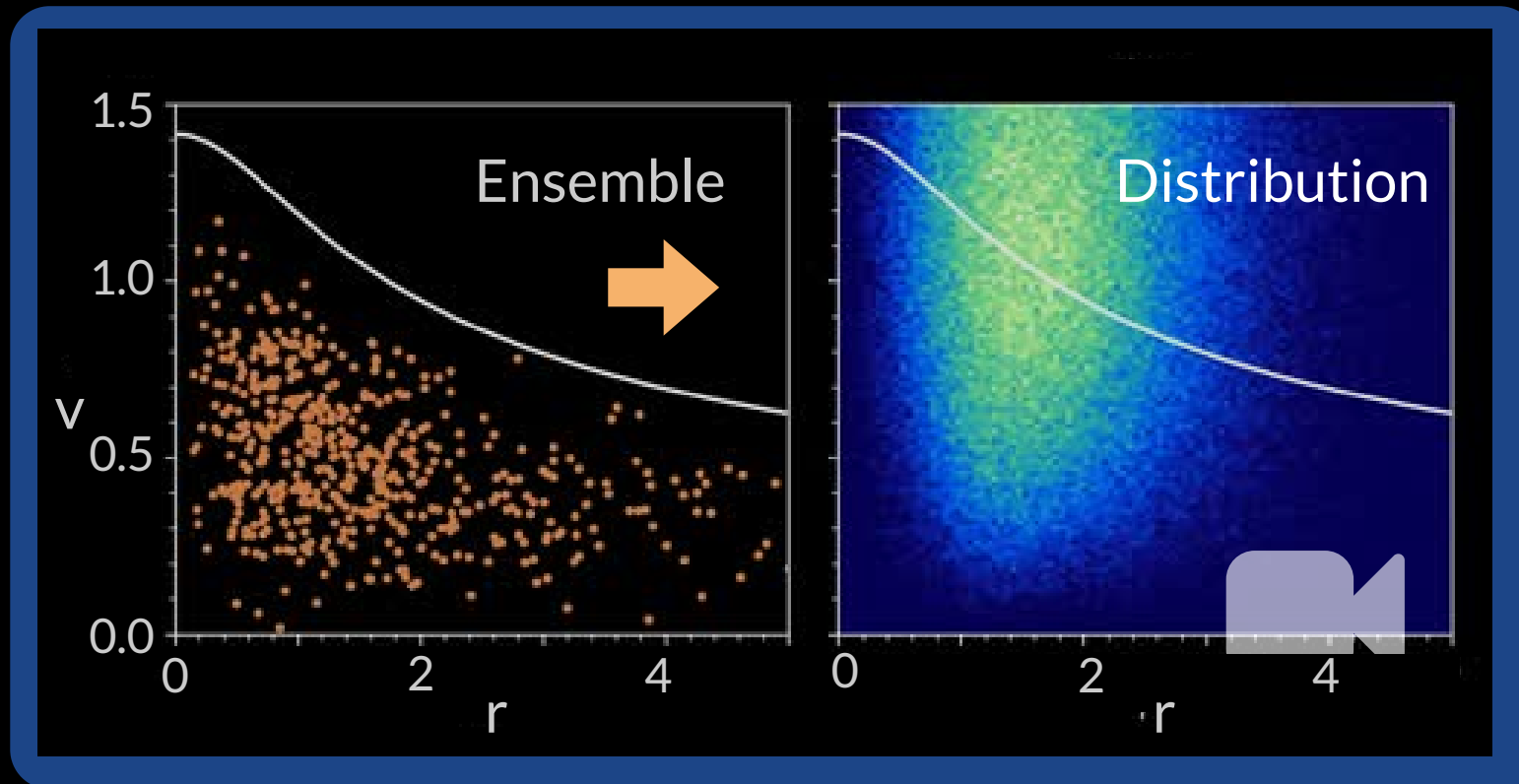
$$\xi(\Delta r) \equiv \frac{DD(\Delta r) - RR(\Delta r)}{RR(\Delta r)}$$

On describing arbitrary distributions through *normalizing flows*

$$\{x_i, v_i\}_{Gaia} \quad \rightarrow \quad p(x, v)$$

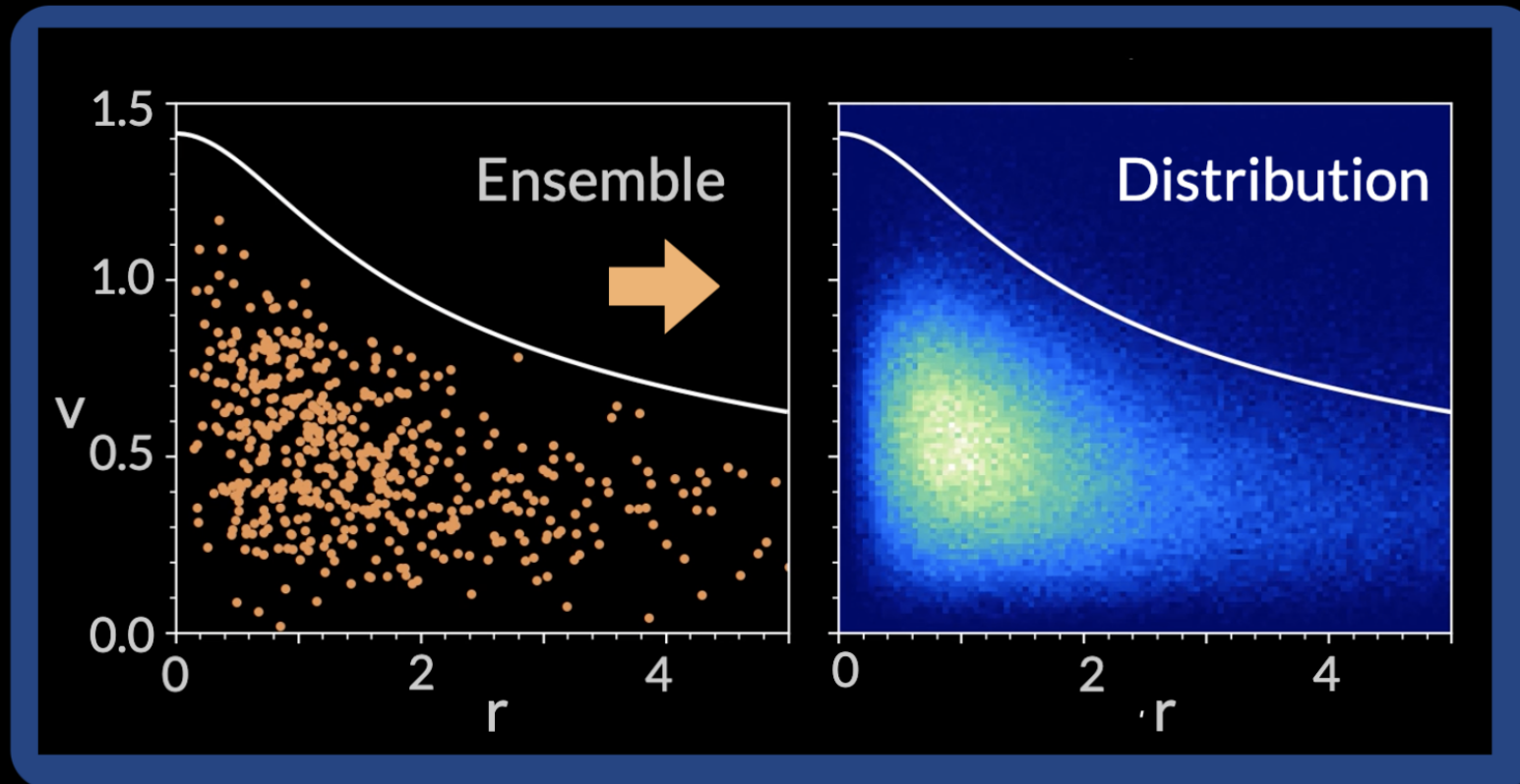
On describing arbitrary distributions through *normalizing flows*

$$\{x_i, v_i\}_{Gaia} \quad \rightarrow \quad p(x, v)$$

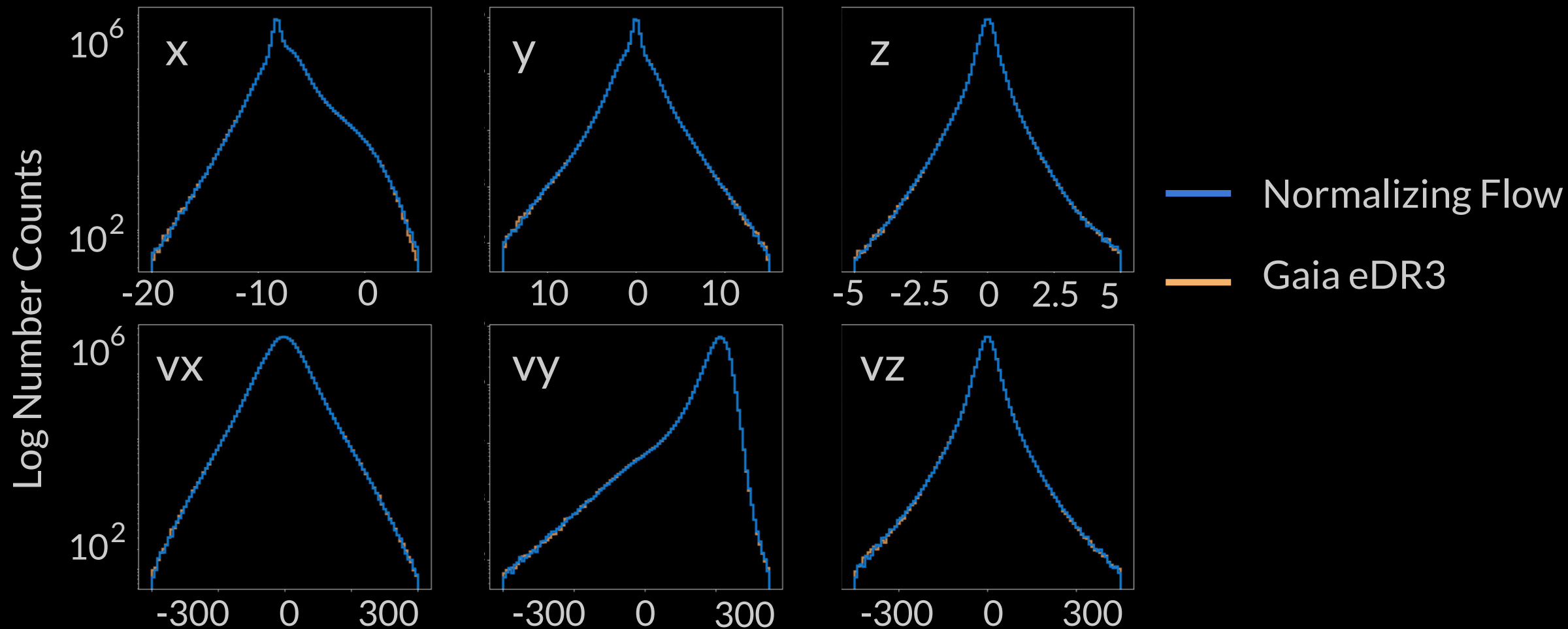


On describing arbitrary distributions through *normalizing flows*

$$\{x_i, v_i\}_{Gaia} \quad \rightarrow \quad p(x, v)$$

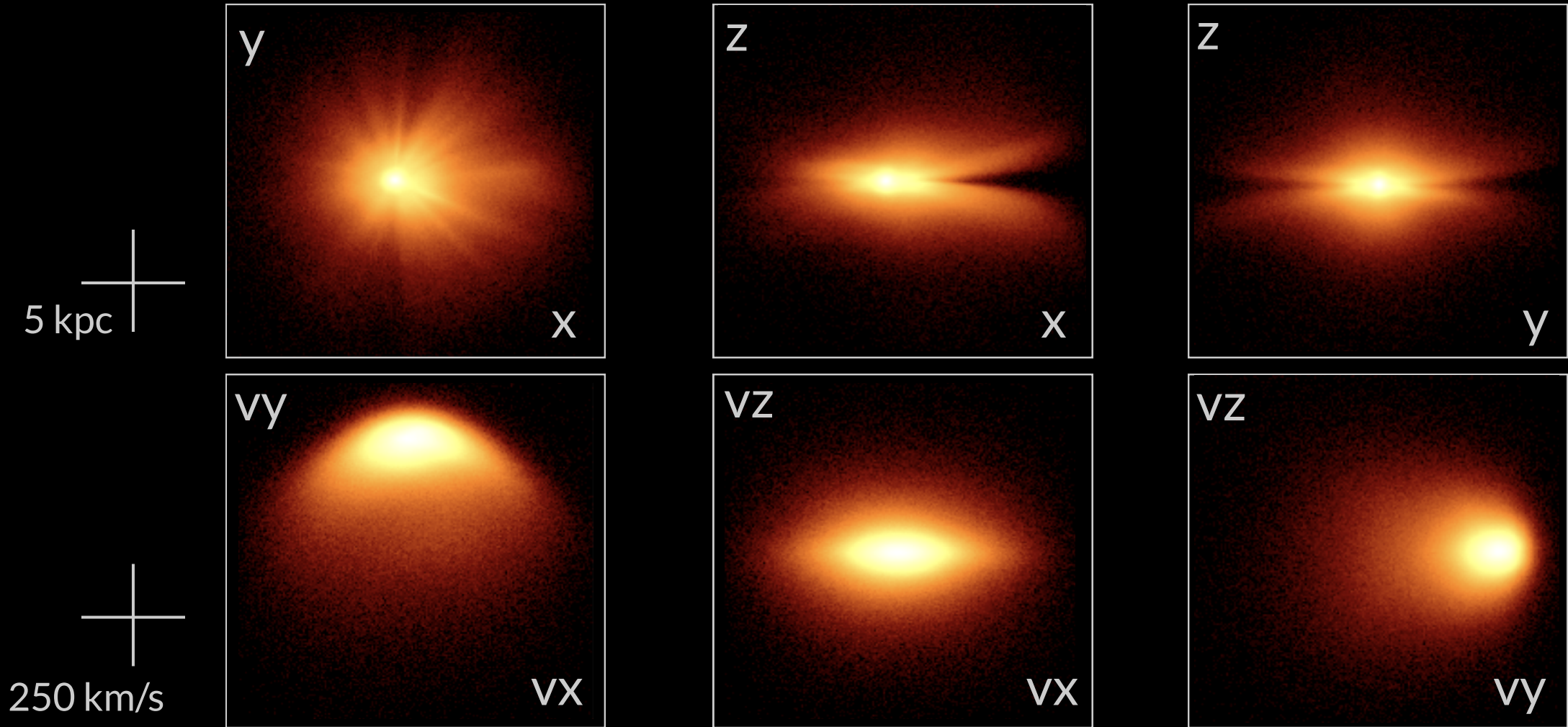


Constructing a robust *Gaia smooth catalog* with normalizing flow



Green & YST, 2020; YST, Kamdar, Koppelman, in prep.

Constructing a robust *Gaia smooth catalog* with normalizing flow



Green & YST, 2020; YST, Kamdar, Koppelman, in prep.

Quantifying statistical alignments with *power spectrum*

Number of pairs in the *non-equilibrium* case

Number of pairs in the *smooth* case ("random-random")

Two-point correlation function $\xi(\Delta r) \equiv \frac{DD(\Delta r) - RR(\Delta r)}{RR(\Delta r)}$

The separation of two stars

Normalization

Diagram description: The equation $\xi(\Delta r) \equiv \frac{DD(\Delta r) - RR(\Delta r)}{RR(\Delta r)}$ is shown. A green arrow points from the text 'The separation of two stars' to Δr . A red arrow points from the text 'Normalization' to the denominator $RR(\Delta r)$. An orange arrow points from the text 'Number of pairs in the non-equilibrium case' to $DD(\Delta r)$. A blue arrow points from the text 'Number of pairs in the smooth case ("random-random")' to $RR(\Delta r)$. Three question marks '???' are placed above the minus sign in the numerator.

Fourier transform of "power spectrum"

see also, Bovy+2015, Khanna+ 2019, Lancaster+2019, ...

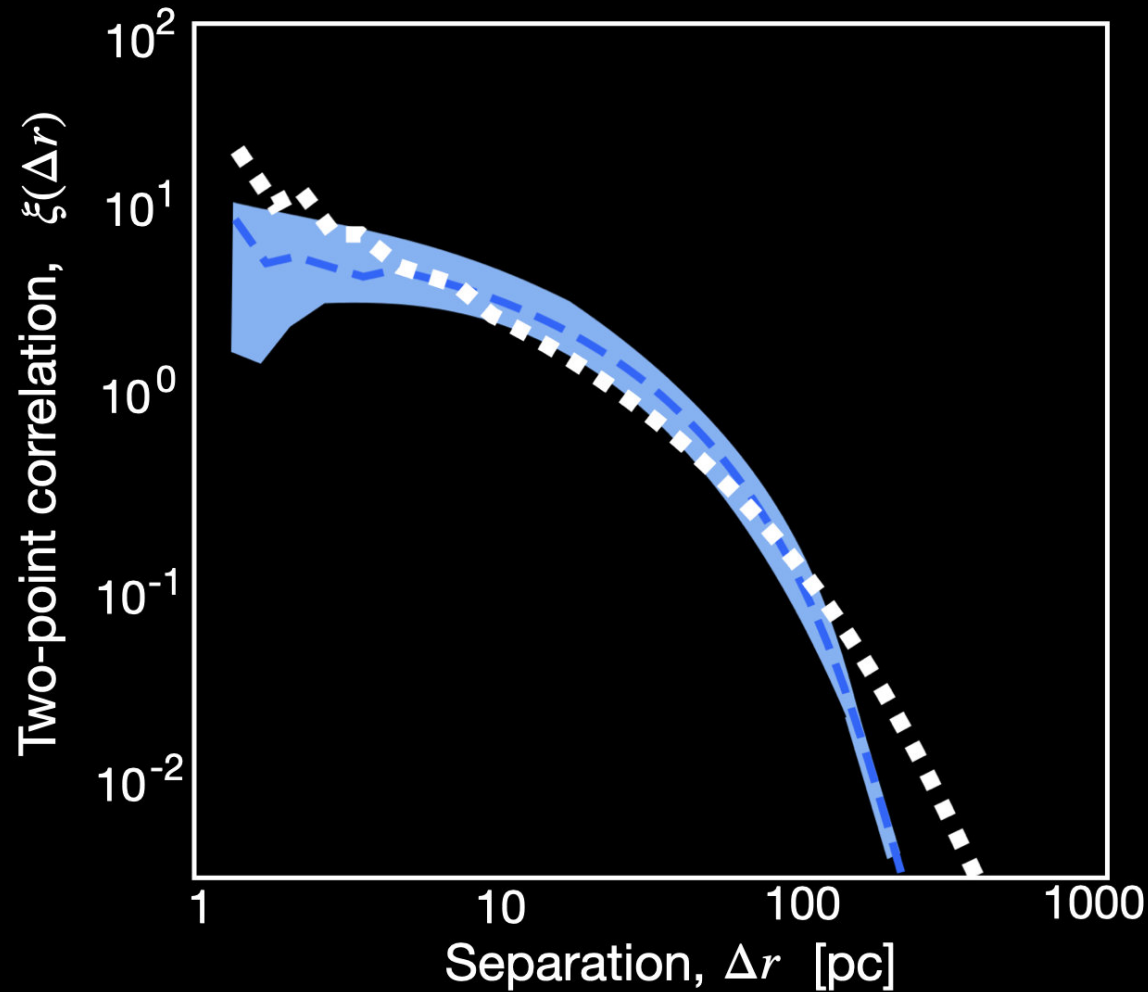
Simulating orbits of billions of stars

- Probing the statistical alignments of *individual stars* requires simulations with *solar-mass resolution*
- Orbit integration with *time varying potentials*, including the giant molecular clouds, bar and spiral arms

Kamdar, Conroy, YST+ 2019a,b



Two-point correlations of stars in the Solar Neighbourhood

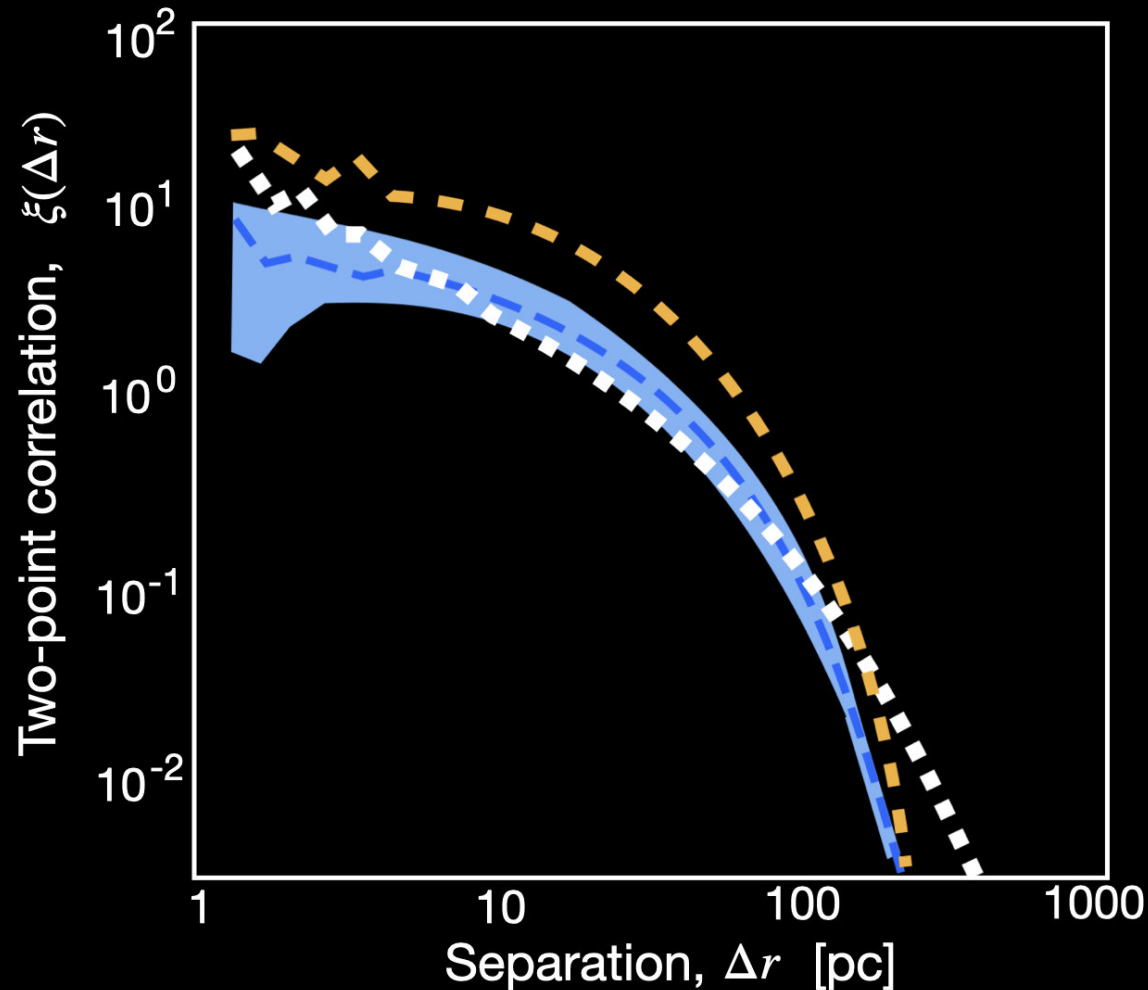


Velocity difference of the pair < 2 km/s

- Gaia data
- - - Our fiducial simulation

Kamdar, Conroy, YST+ 2020

Two-point correlations of stars in the Solar Neighbourhood



Velocity difference of the pair < 2 km/s

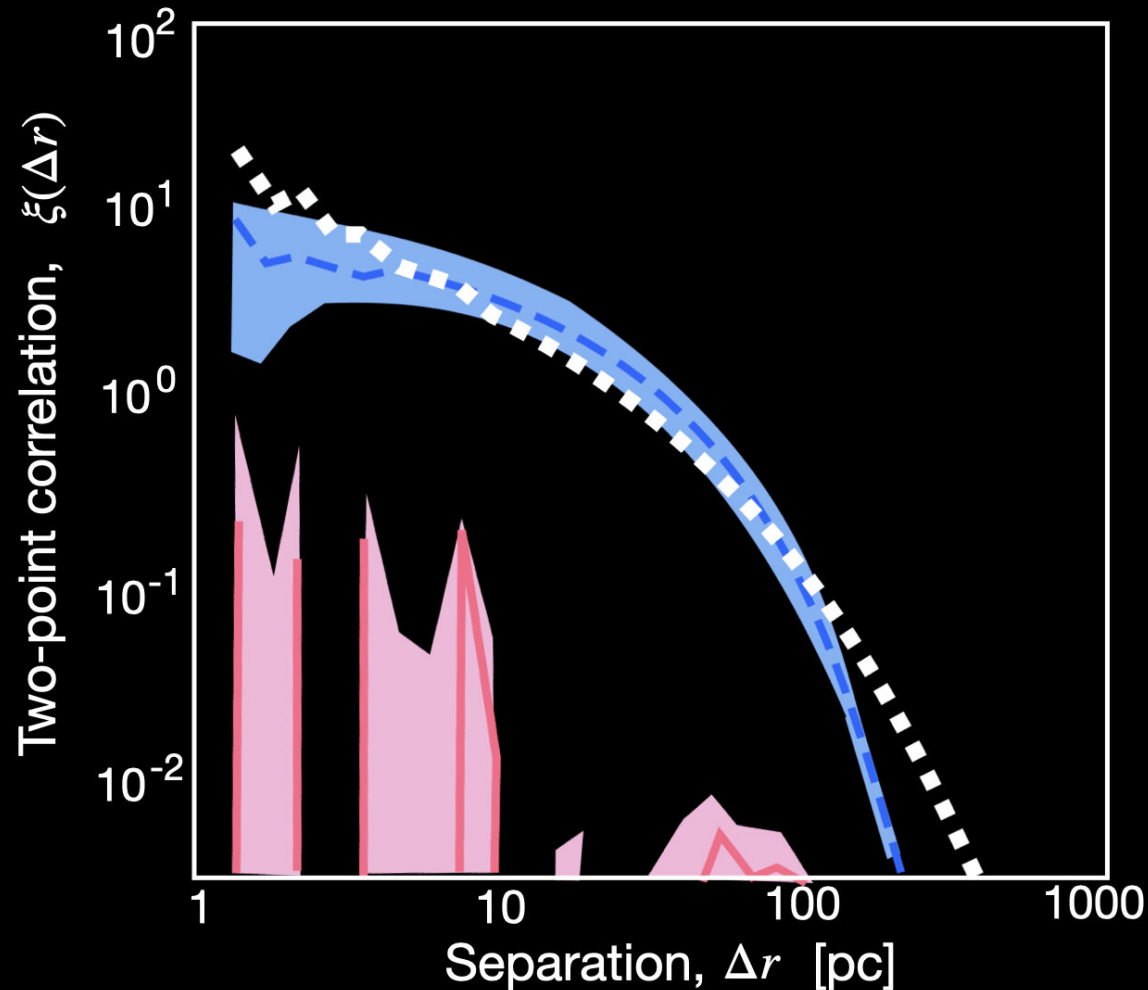
--- Simulation *without* perturbations from giant molecular clouds

•••• Gaia data

--- Our fiducial simulation

Kamdar, Conroy, YST+ 2020

Two-point correlations of stars in the Solar Neighbourhood

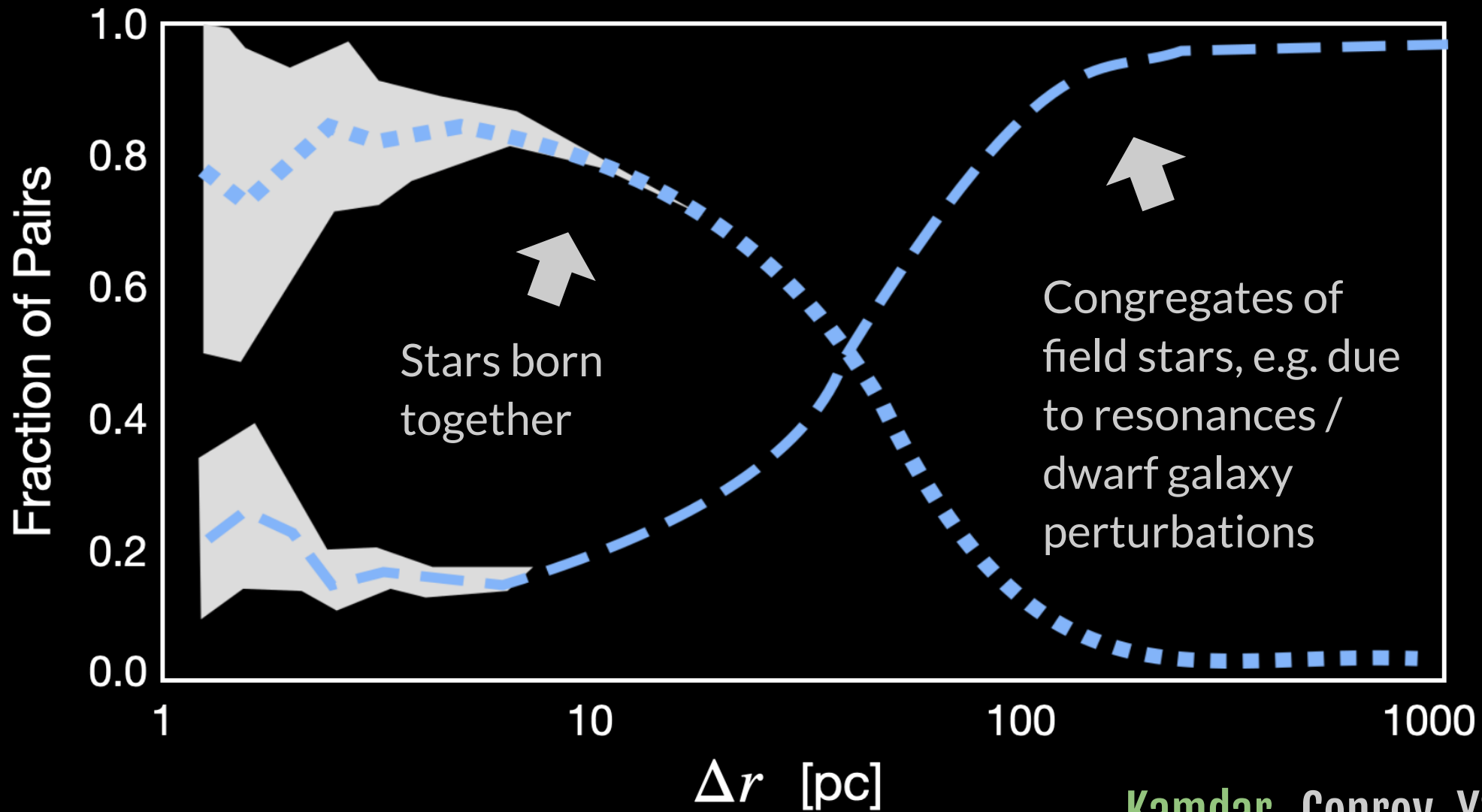


Velocity difference of the pair < 2 km/s

- Gaia data
- - - Our fiducial simulation
- Without* clustered star formation

Kamdar, Conroy, YST+ 2020

Contributions from both star formation & Galactic perturbation



Kamdar, Conroy, YST+ 2019b

Summary :

- The *unprecedented amount* of Gaia data has propelled Galactic Archaeology to a phase *akin to cosmology* a few decades ago
- The statistical alignments of stars can constrain many *subgrid physics* that are otherwise unattainable by other astronomical probes
- The *power spectrum* of stellar phase space reveals how stars formed and dispersed and the density of giant molecular clouds
- Modern days *machine learning techniques* play a crucial role in characterizing these subtle fluctuations / patterns

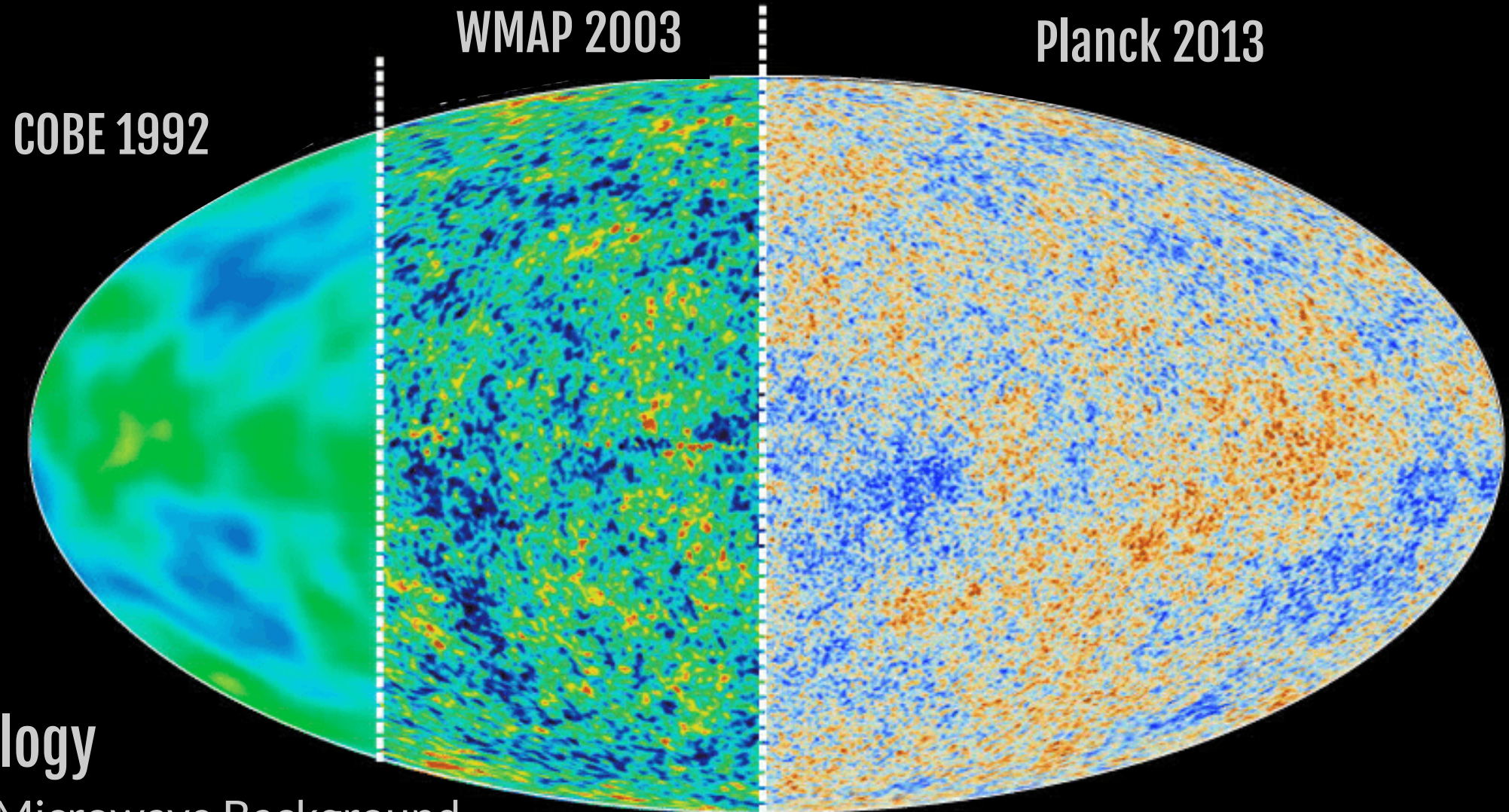
Extra Slides

The classical inferences *assume* that all stars are *statistically i.i.d*

- $\text{star}_i \sim_{i.i.d.} p(\vec{x}, \vec{v})$

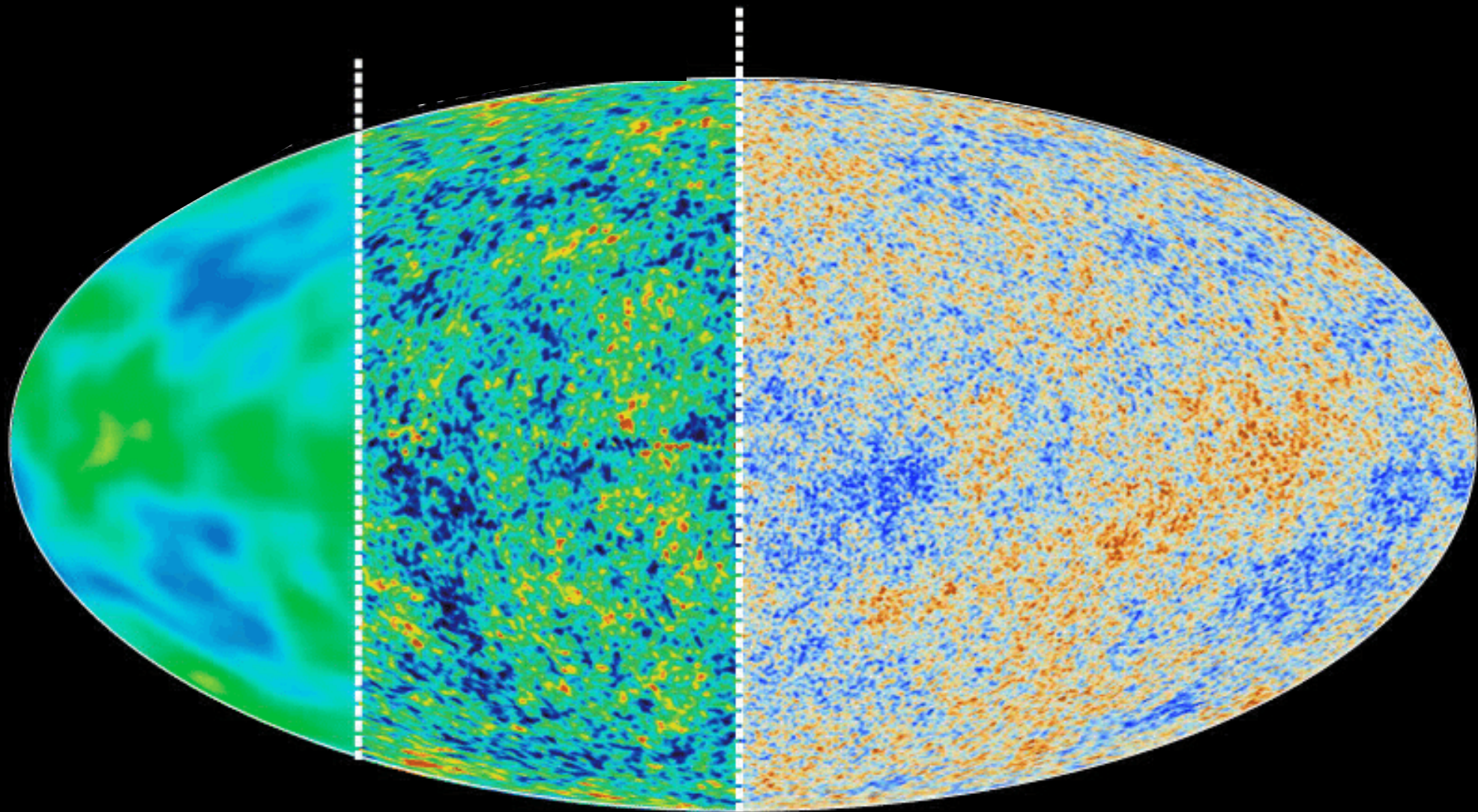
- $p(\vec{x}_i, \vec{v}_i, \vec{x}_j, \vec{v}_j) = p(\vec{x}_i, \vec{v}_i) \cdot p(\vec{x}_j, \vec{v}_j)$

We are entering a new era of *"high-definition"* astronomy

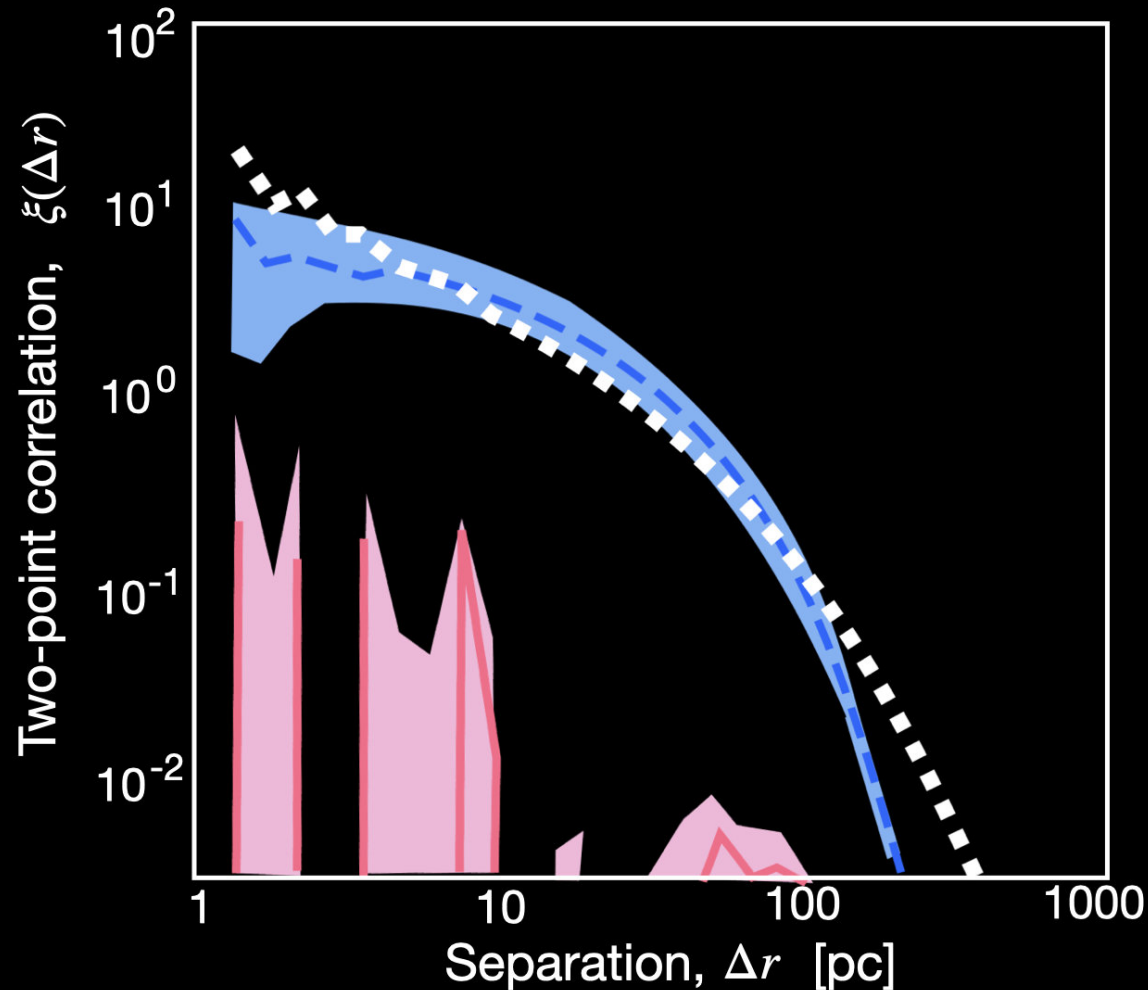


Cosmology

Cosmic Microwave Background



Two-point correlations of stars in the Solar Neighbourhood

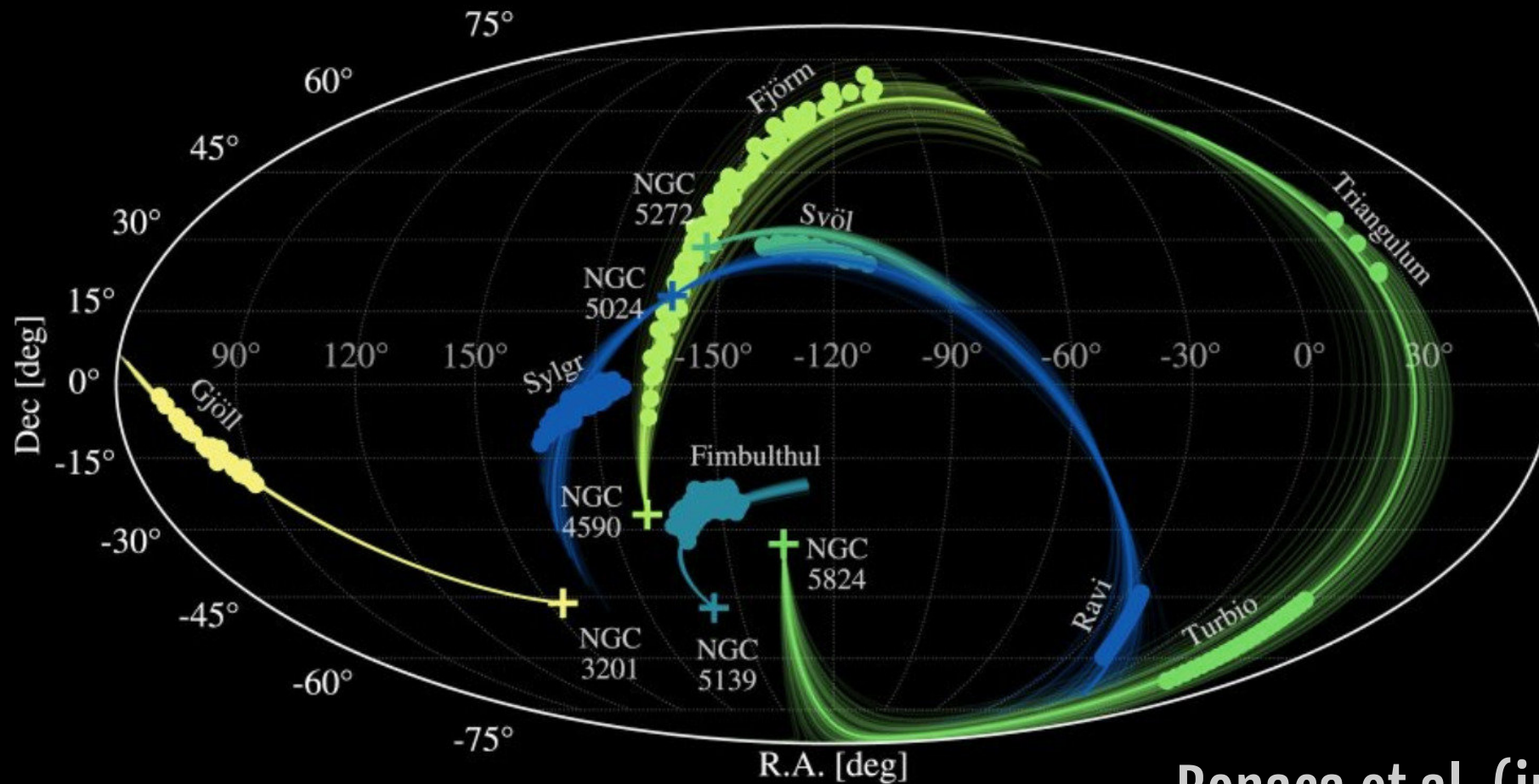


Velocity difference of the pair < 2 km/s

- Gaia data
- - - Our fiducial simulation
- Without* clustered star formation

Kamdar, Conroy, YST+ 2020

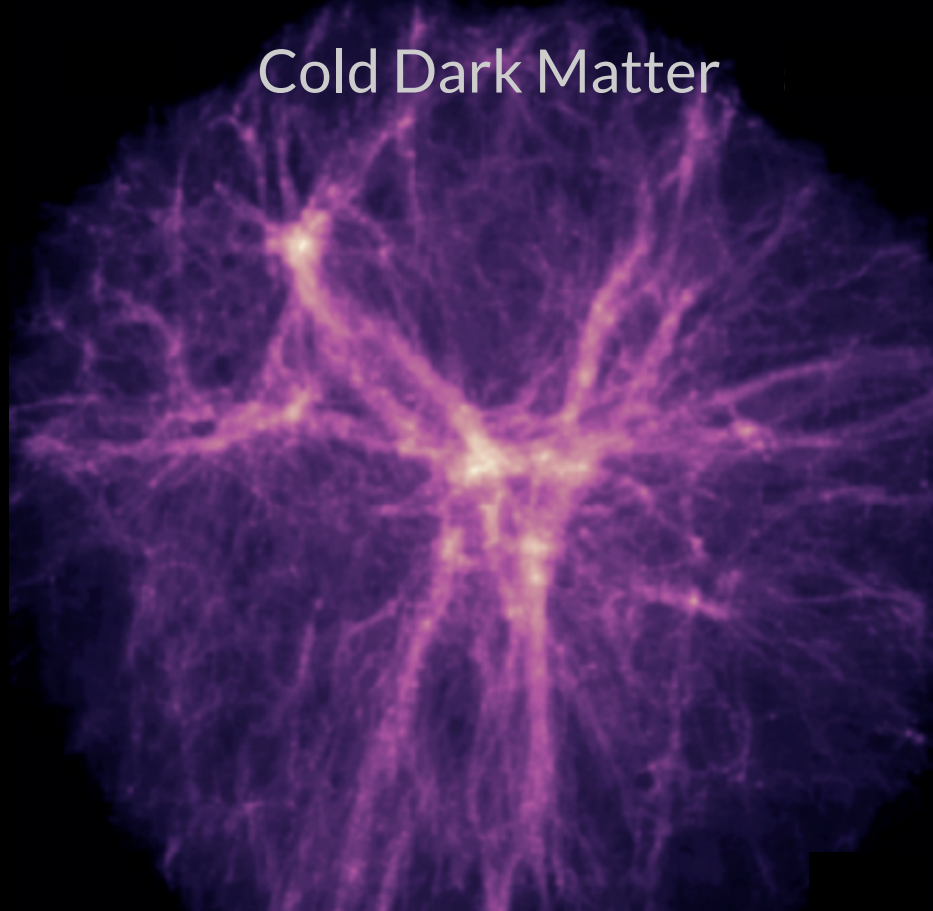
We focused mostly on the *"blockbuster" events*



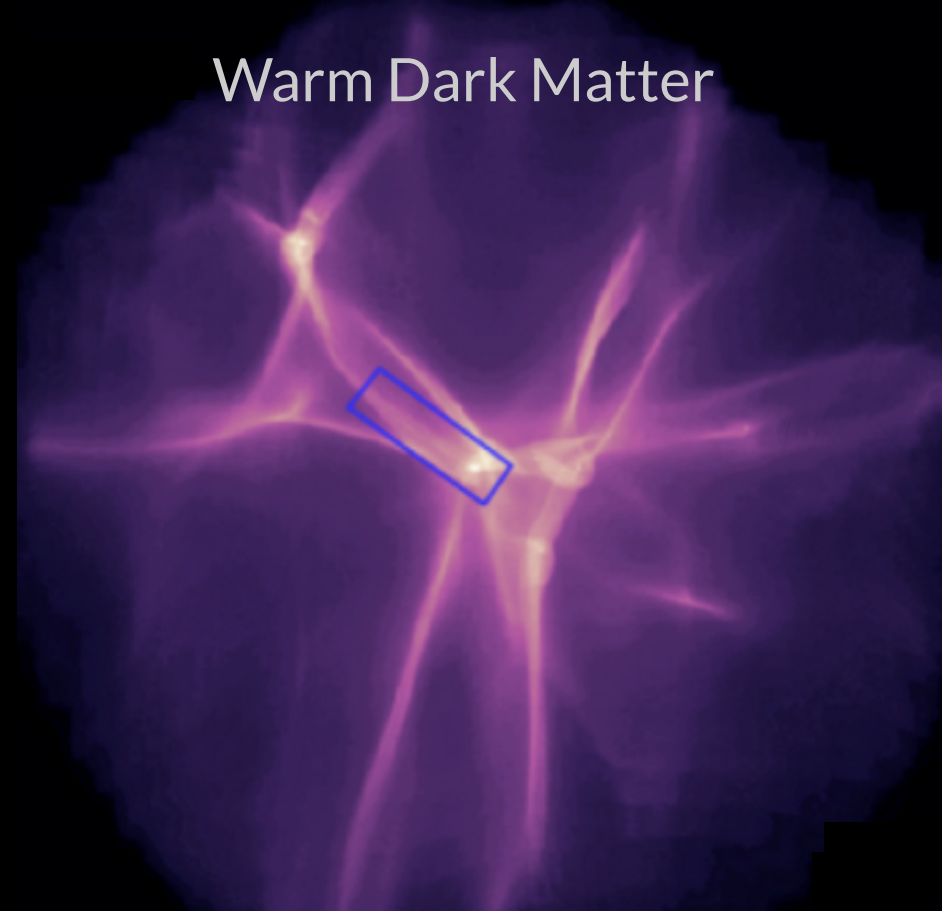
Bonaca et al. (incl. YST), 2021

Studying the statistical alignments of objects is *not new*

Cold Dark Matter



Warm Dark Matter



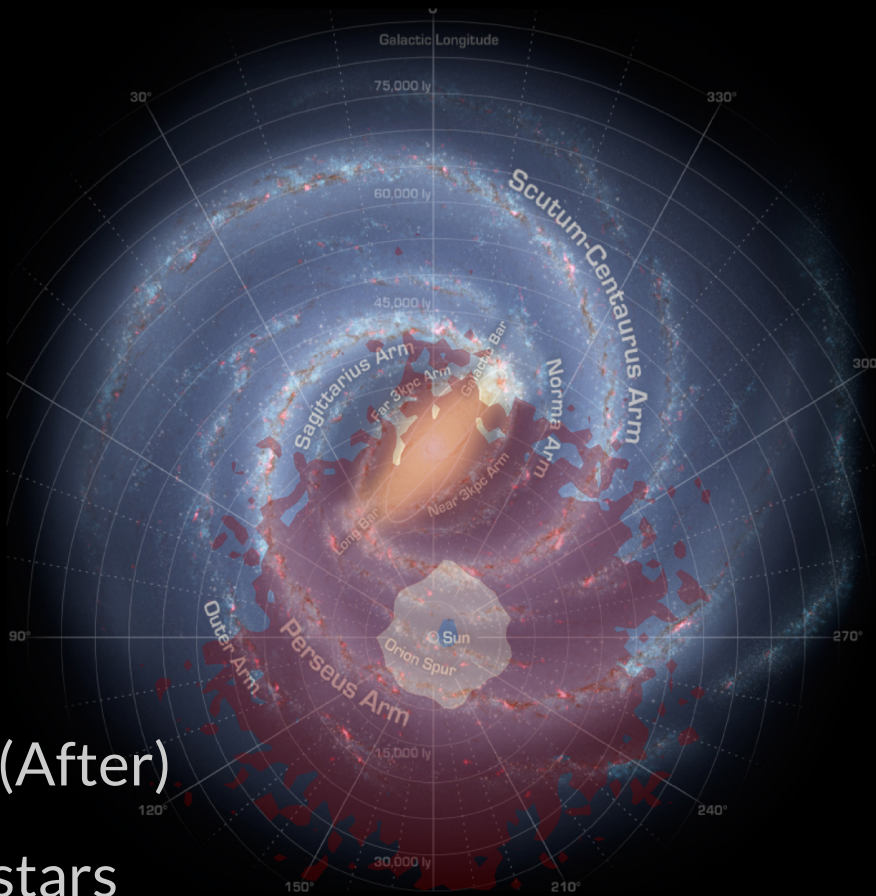
Cosmology

Large-scale structure

Gao & Theuns 2007

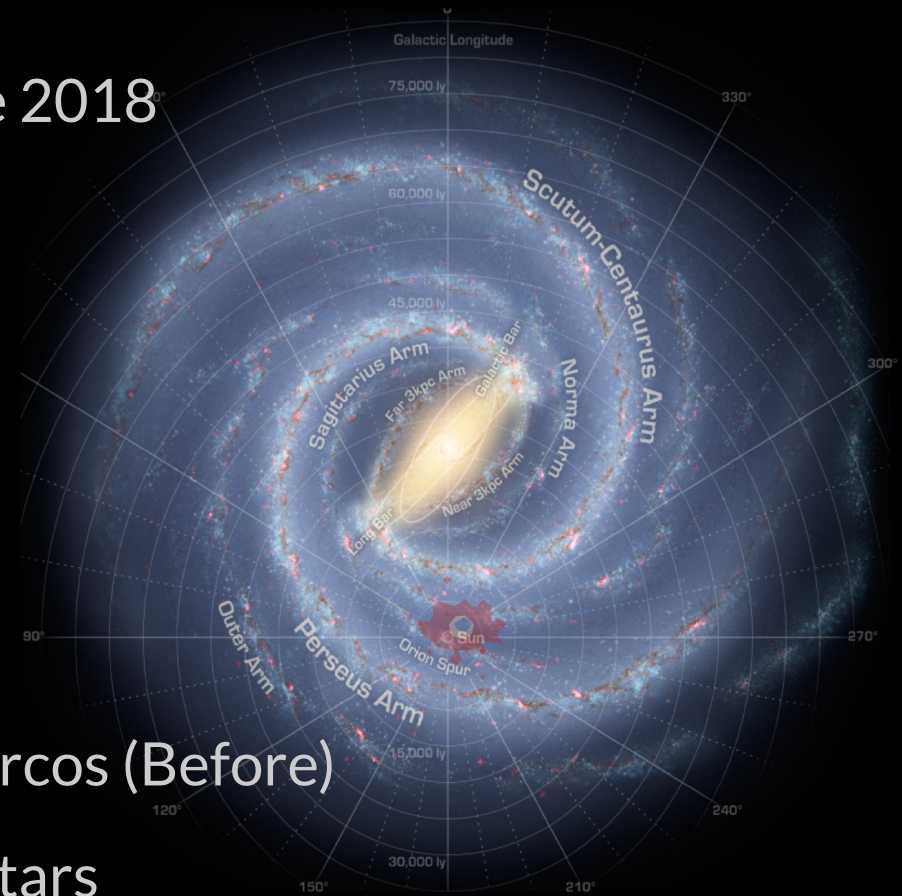
There is an *explosion of data* about stars in the Milky Way

Now



Gaia (After)
 10^9 stars

Before 2018

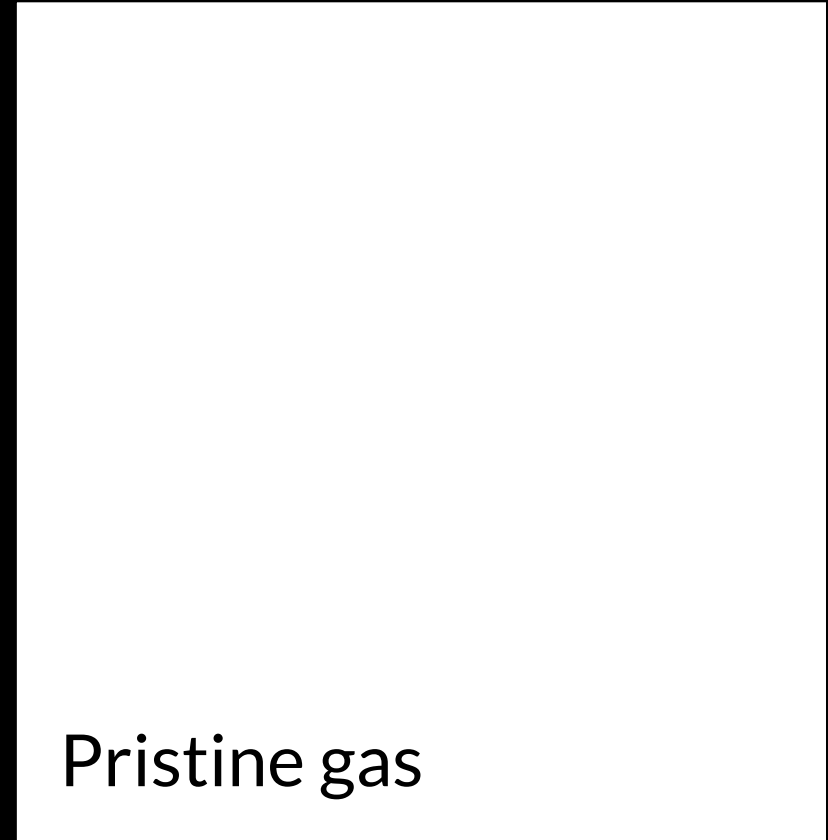
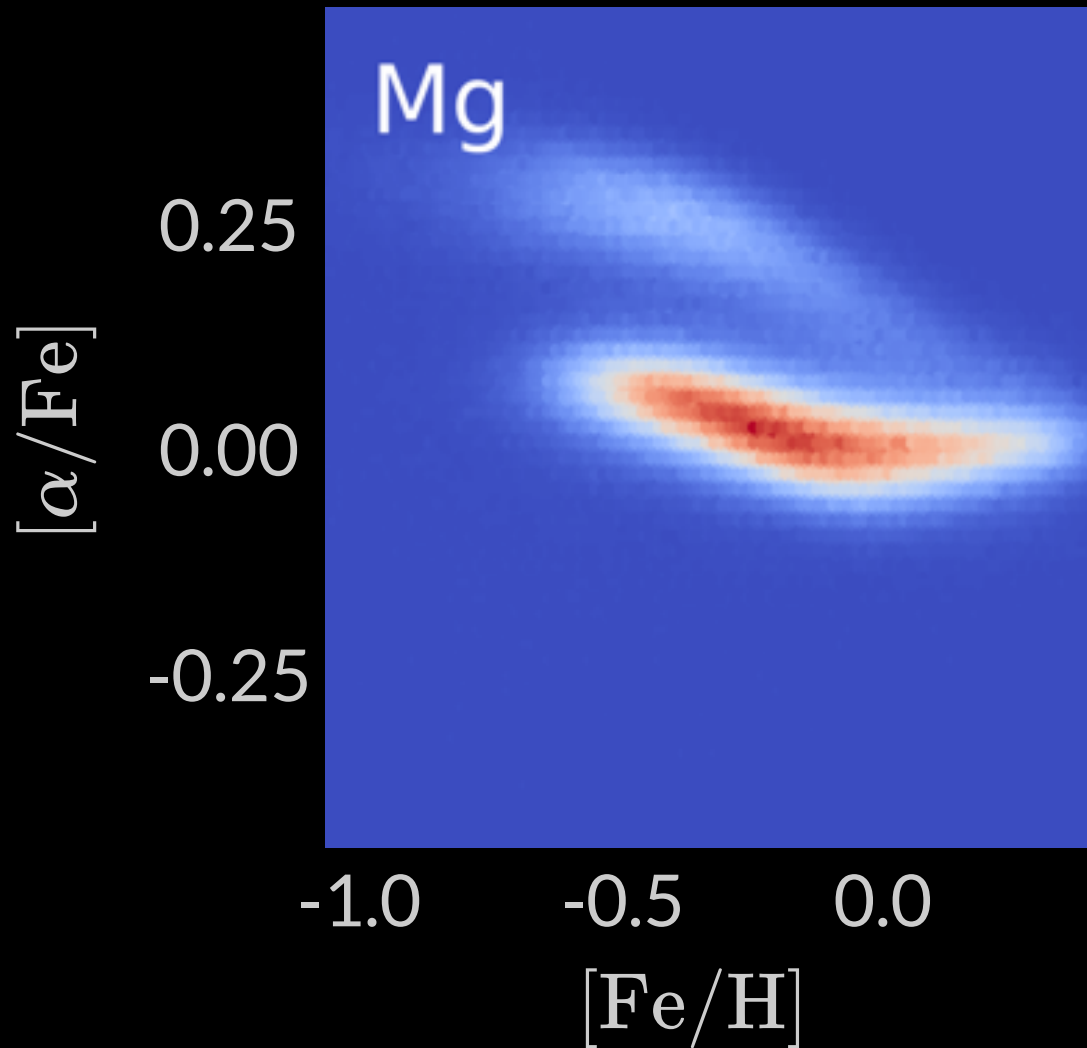


Hipparcos (Before)
 10^5 stars

Stars with *kinematic* information

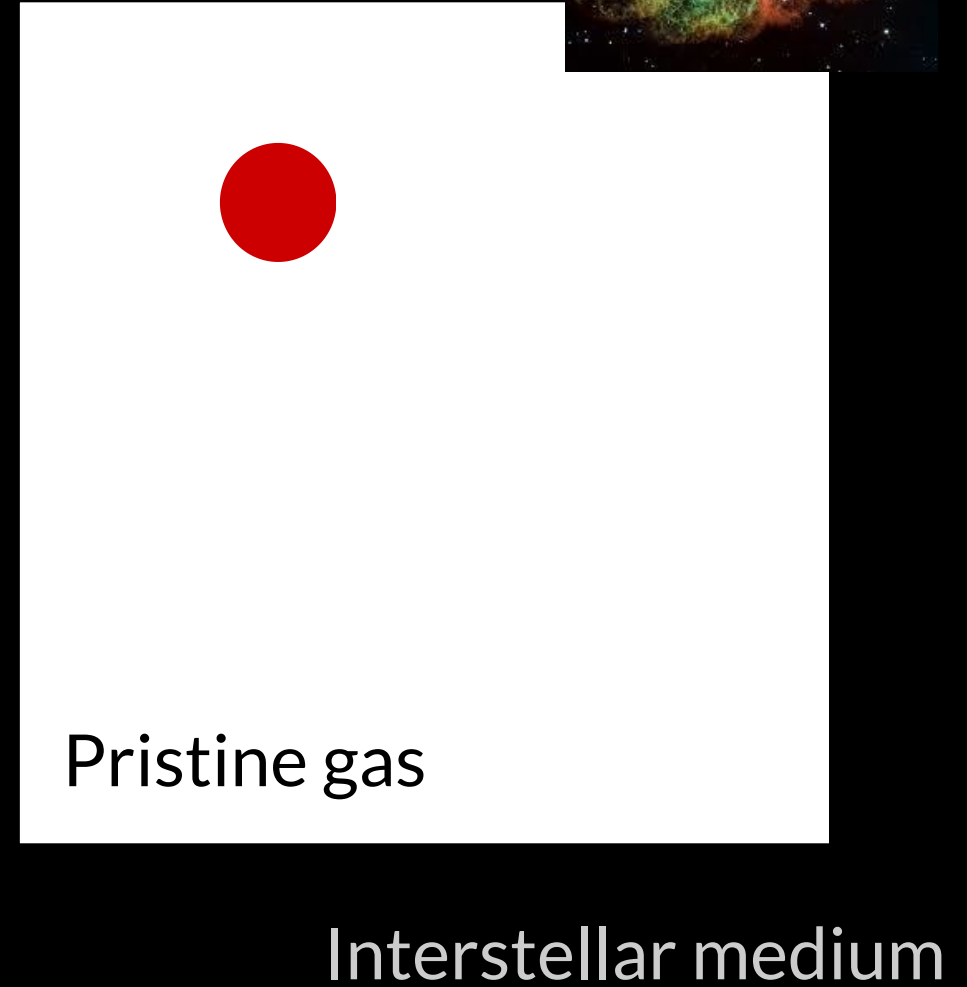
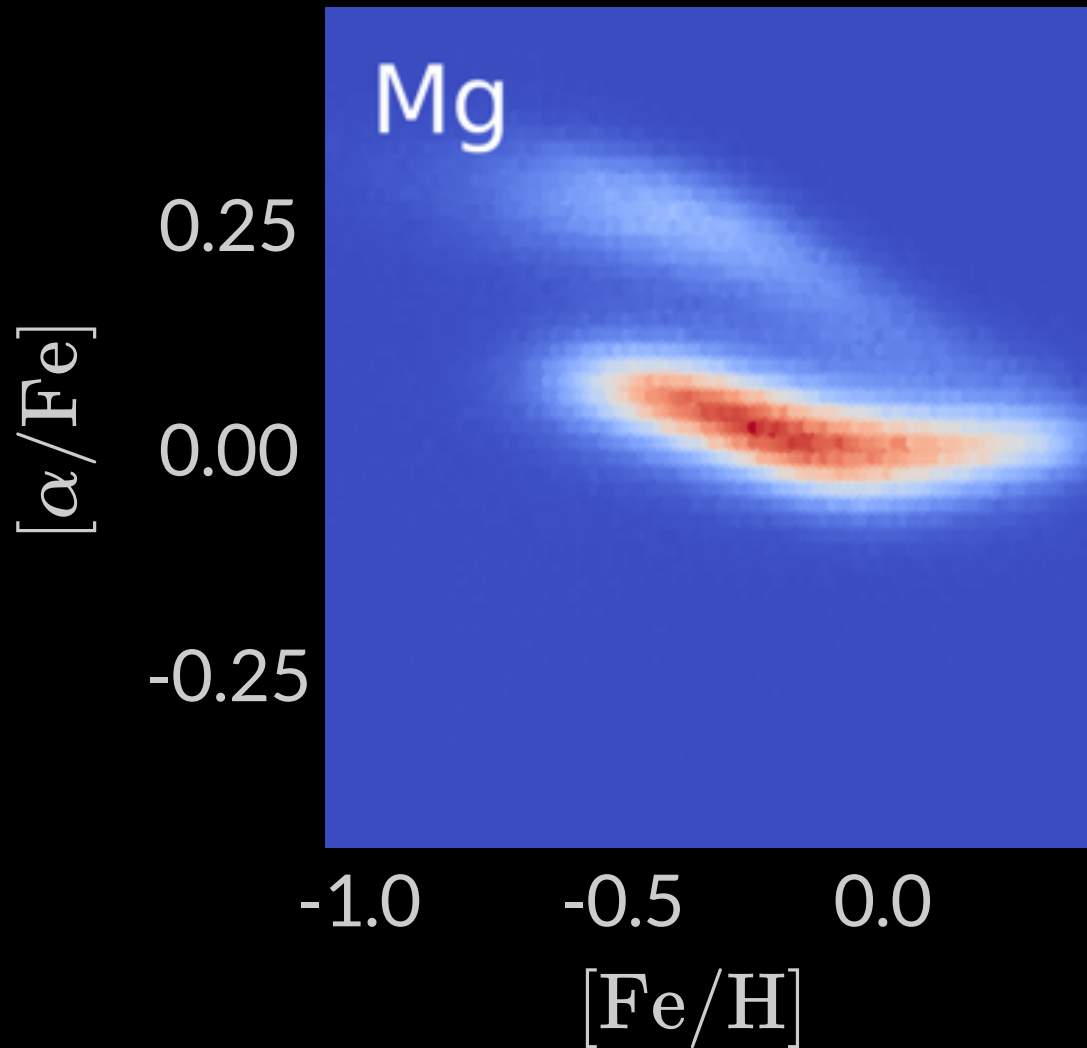
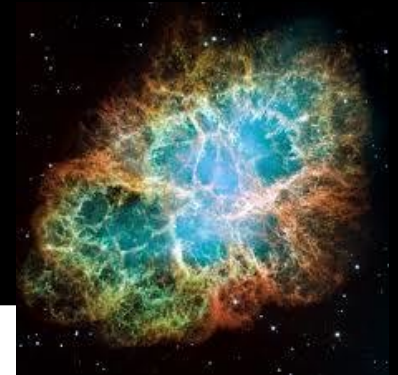
How about the fluctuations in the
chemical space?

A simplistic view of chemical evolution

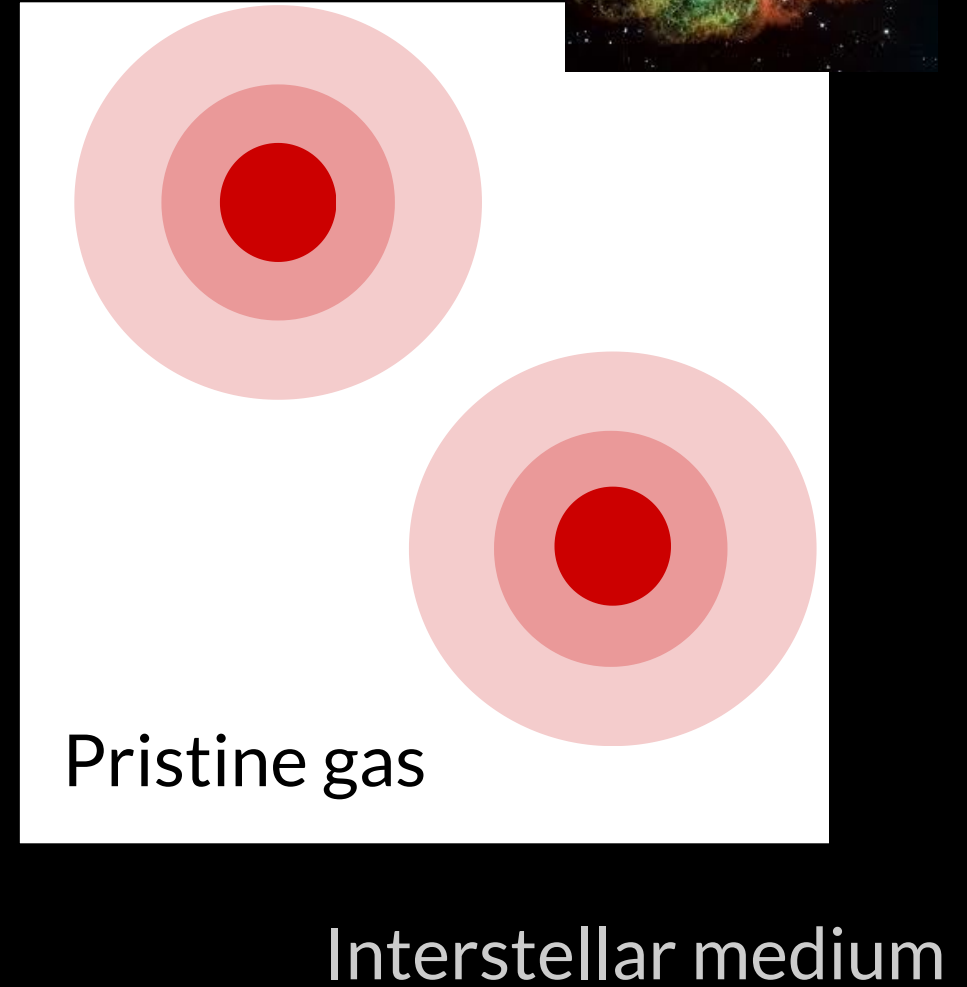
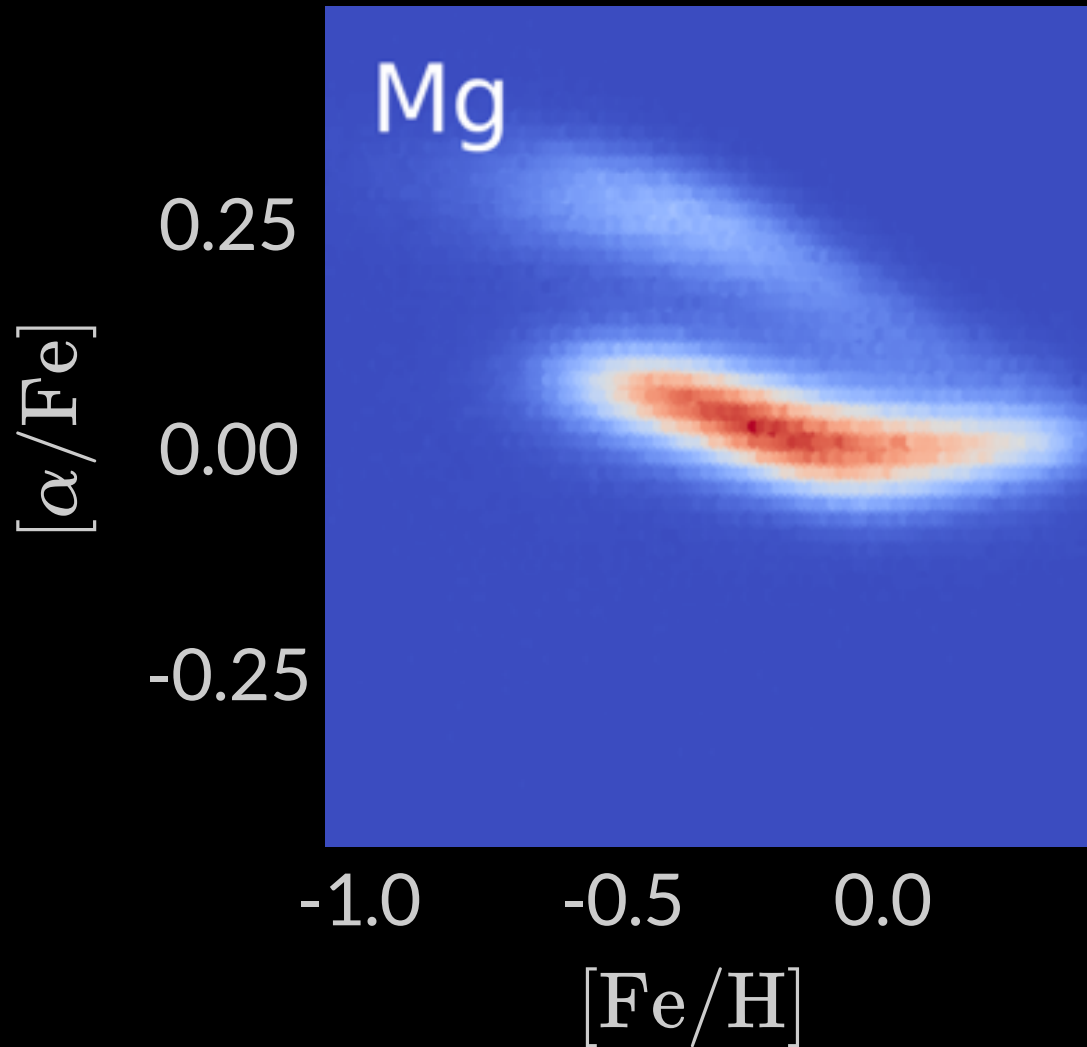
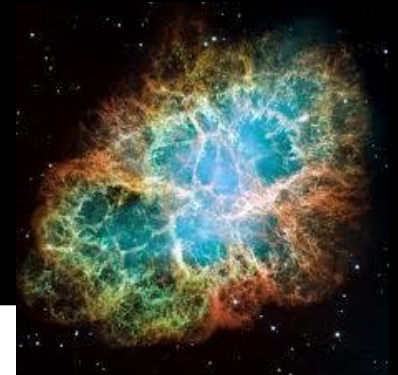


Interstellar medium

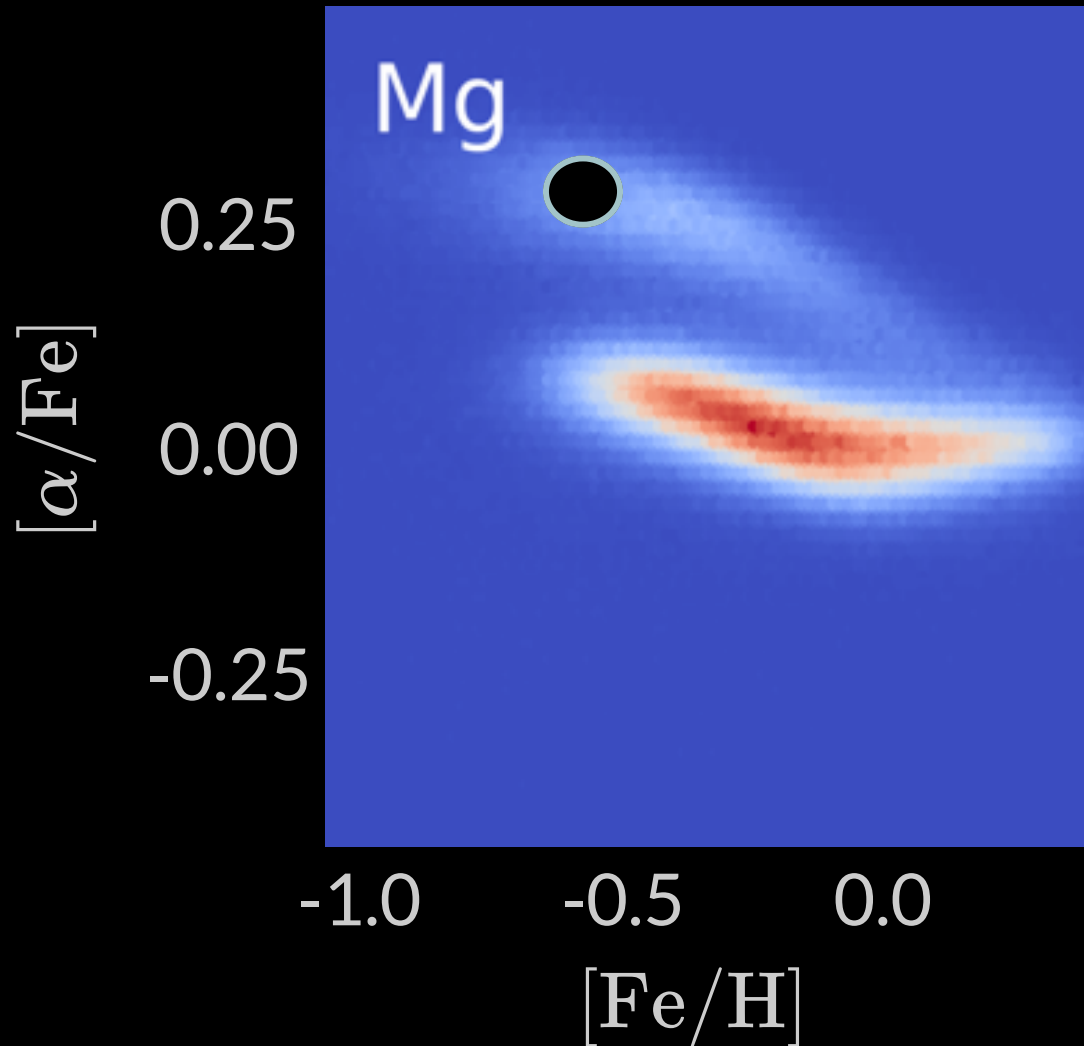
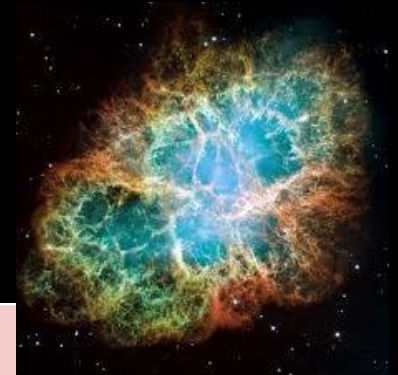
A simplistic view of chemical evolution



A simplistic view of chemical evolution



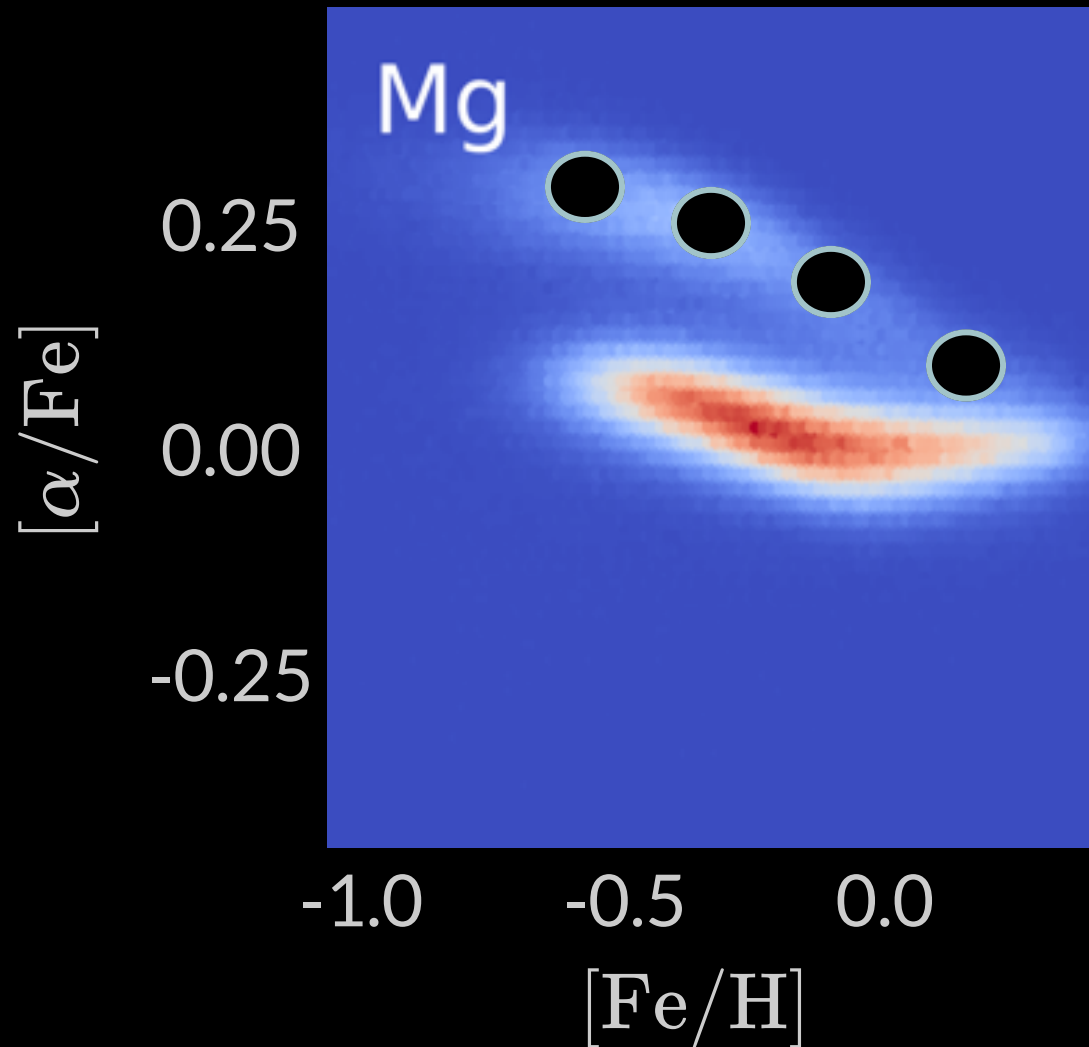
A simplistic view of chemical evolution



If ISM mixing is
instantaneous

Interstellar medium

A simplistic view of chemical evolution

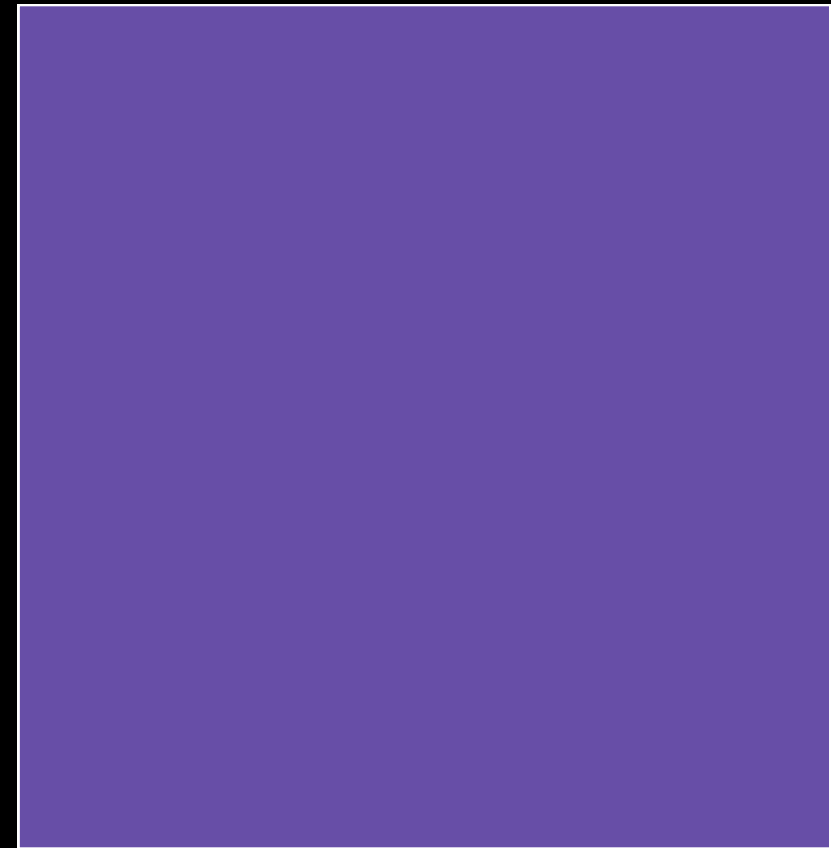
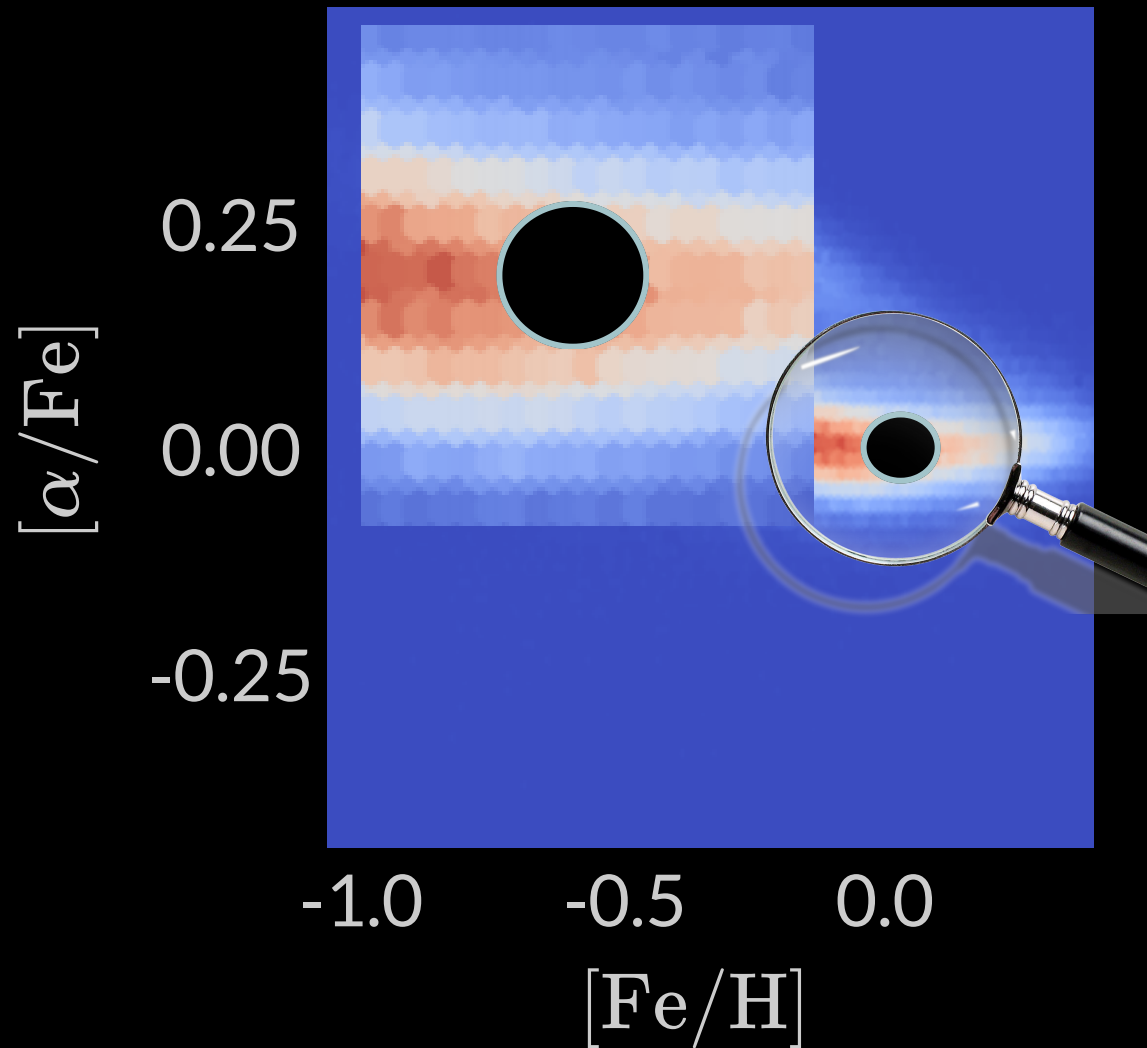


If ISM mixing is instantaneous

Interstellar medium

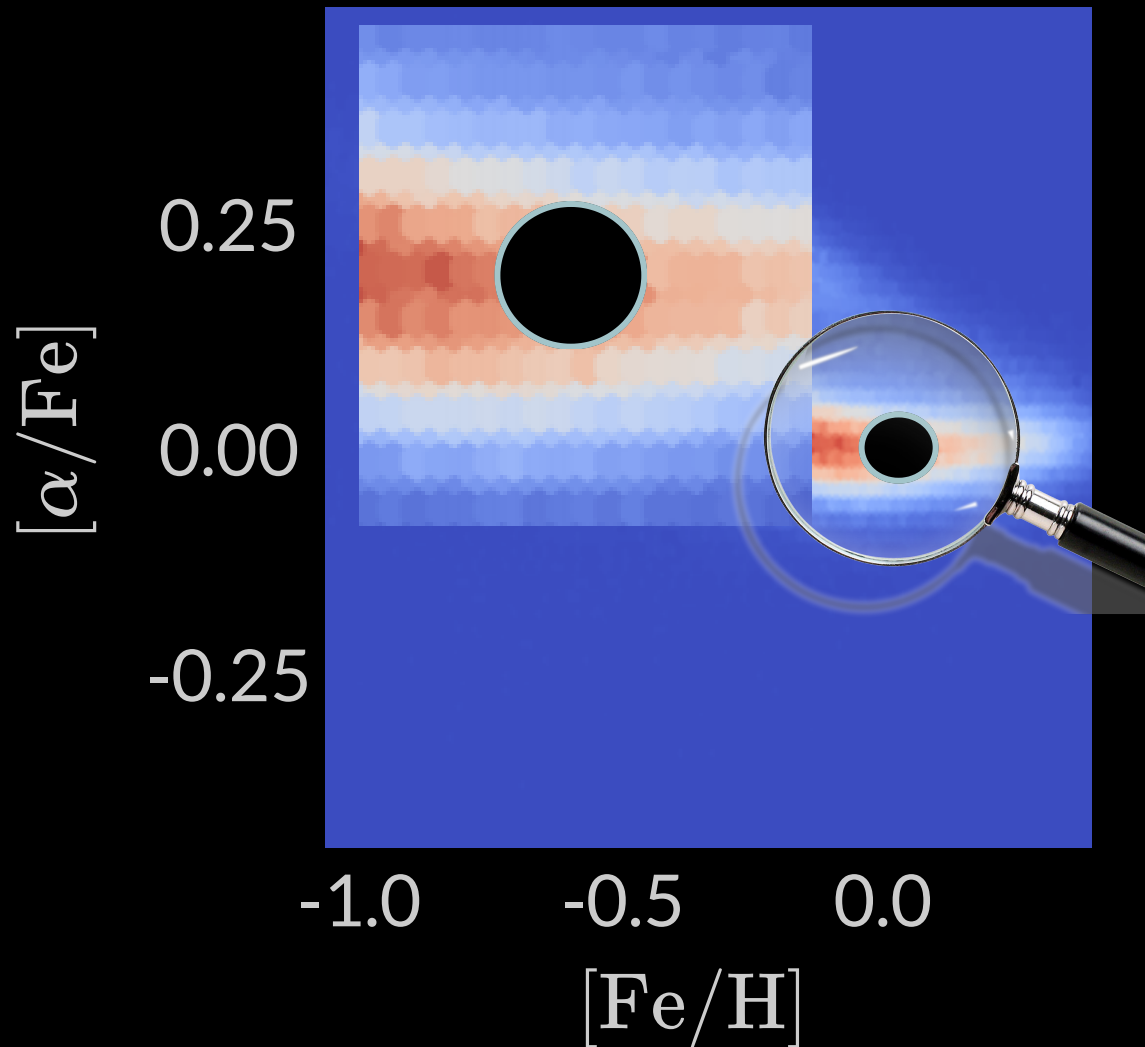
Is there any information *beyond* the mean chemical track ?

Information beyond the chemical track



Interstellar medium

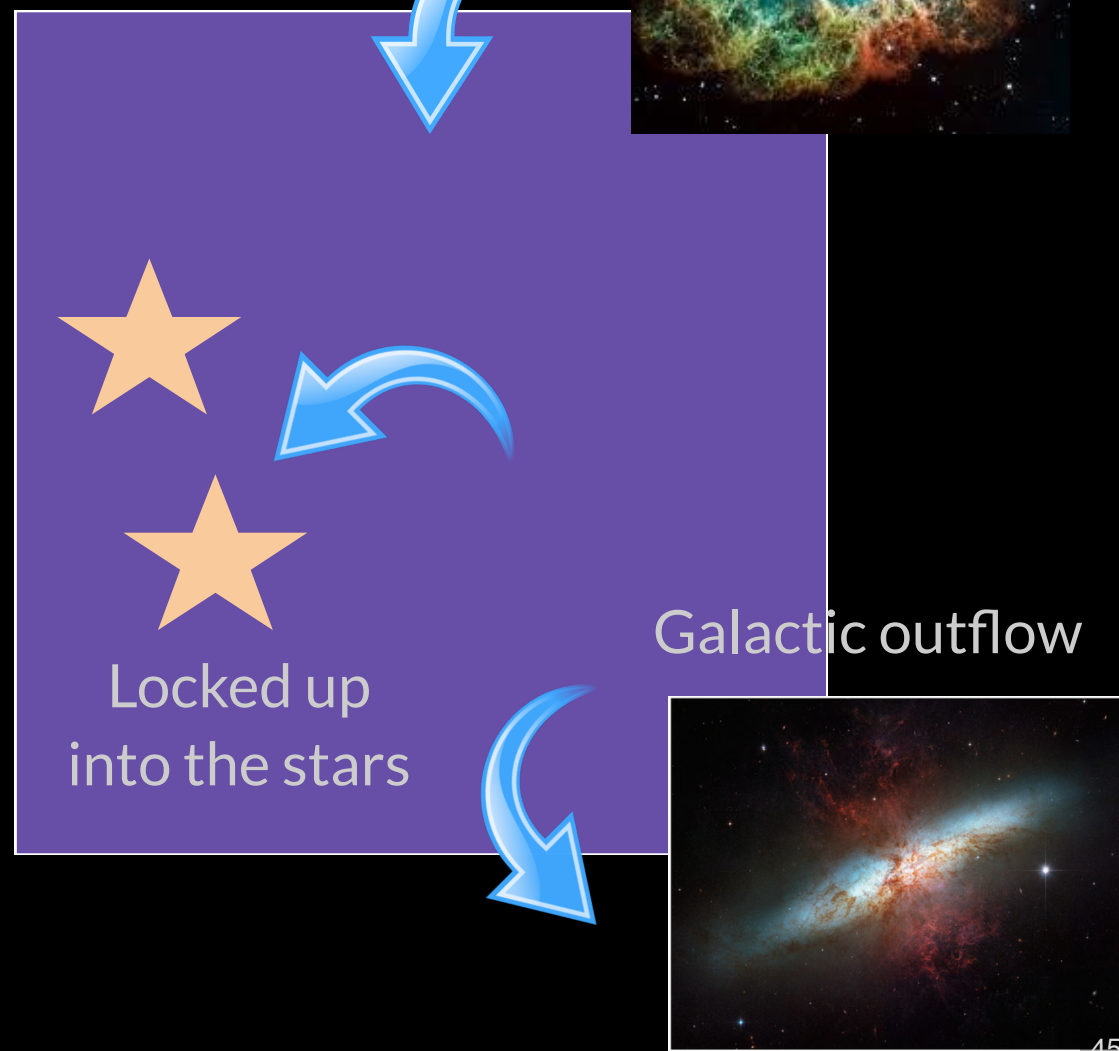
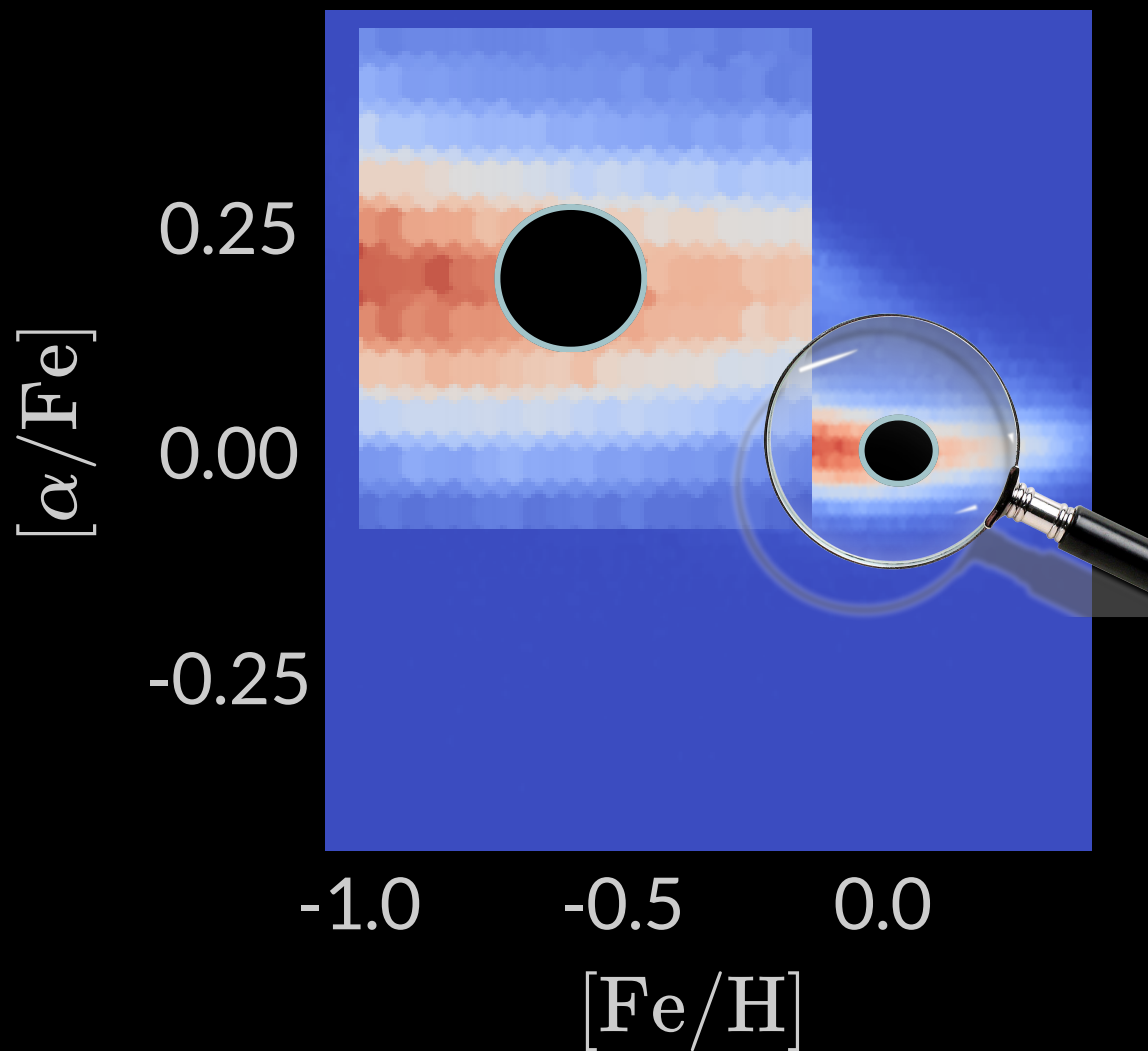
Information beyond the chemical track



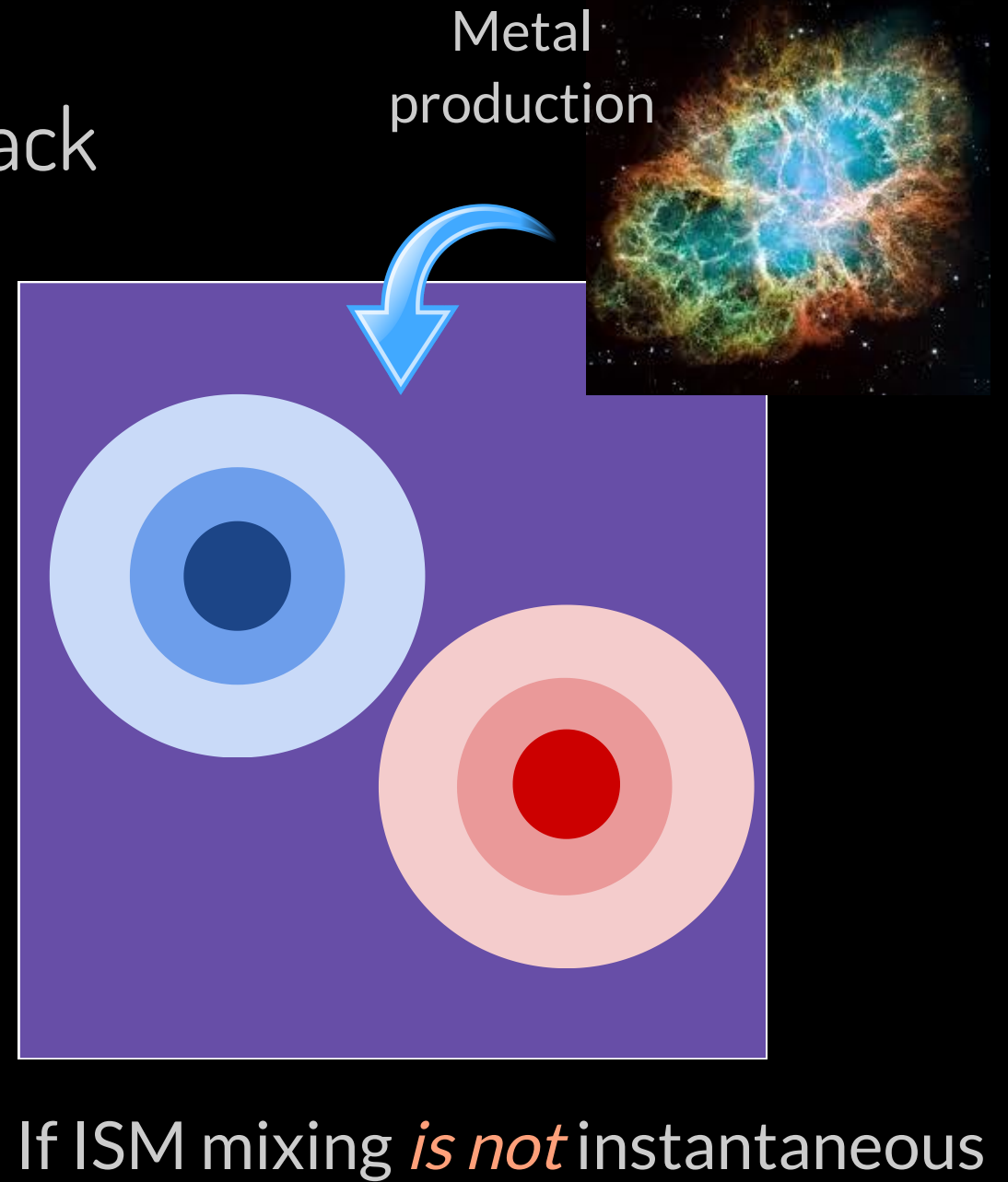
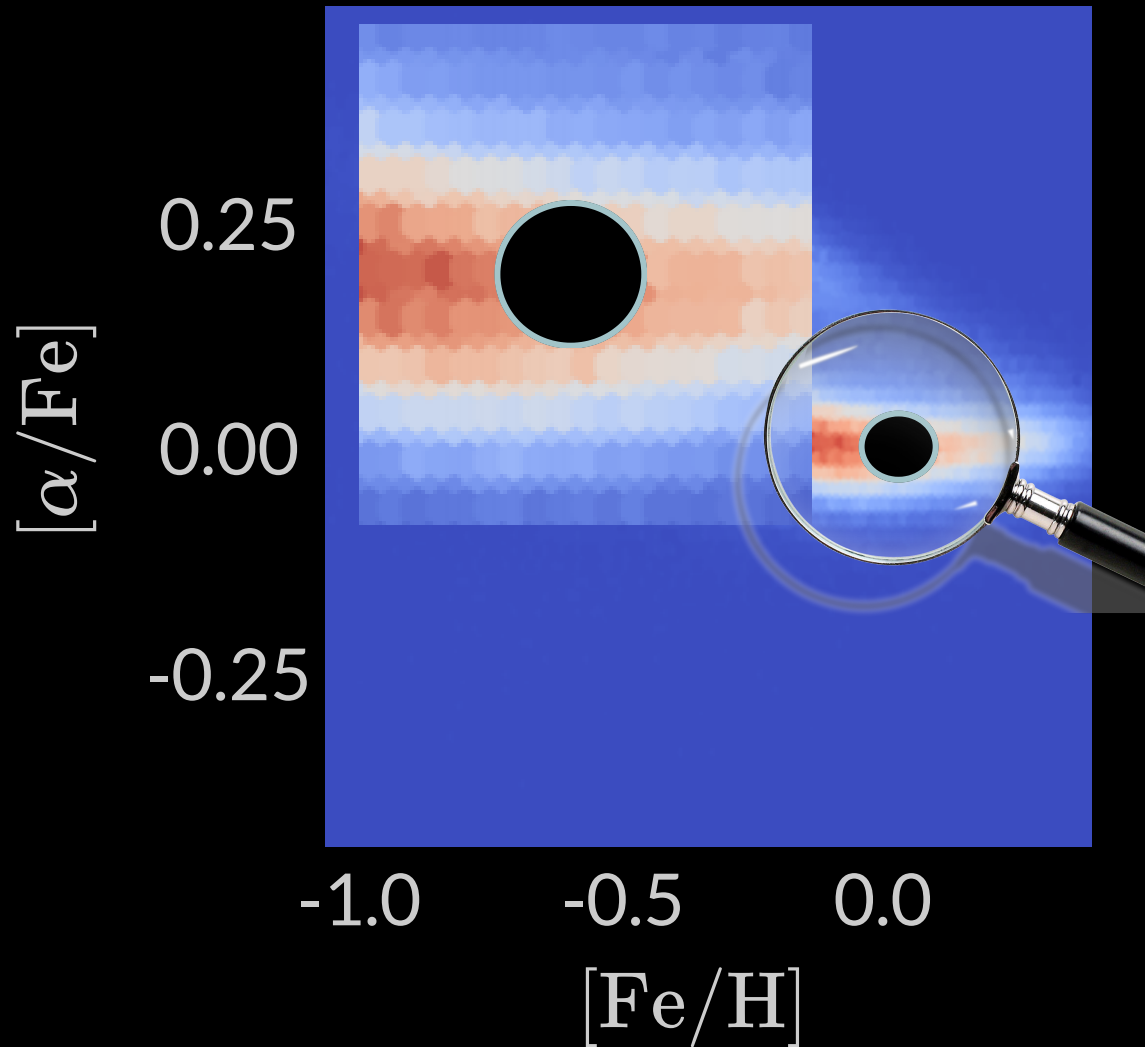
If ISM mixing is instantaneous

Interstellar medium

Information beyond the chemical track



Information beyond the chemical track



Can we quantify the "*ripples*" in the chemical space ?

Yes !

YST & Weinberg 2021

APOGEE data reveals *subtle but non-trivial* chemical fluctuation

APOGEE abundance correlation

