Toward the fabric of the Milky Way

The density of disk streams from a local 250^3 pc^3 volume

Sebastian Ratzenböck @ CfA

April 15, 2025

Toward the fabric of the Milky Way

The density of disk streams from a local 250^3 pc^3 volume

Joint work with:
João Alves, Emily Hunt, Núria Miret-Roig, Stefan Meingast & Torsten Möller

Sebastian Ratzenböck @ CfA

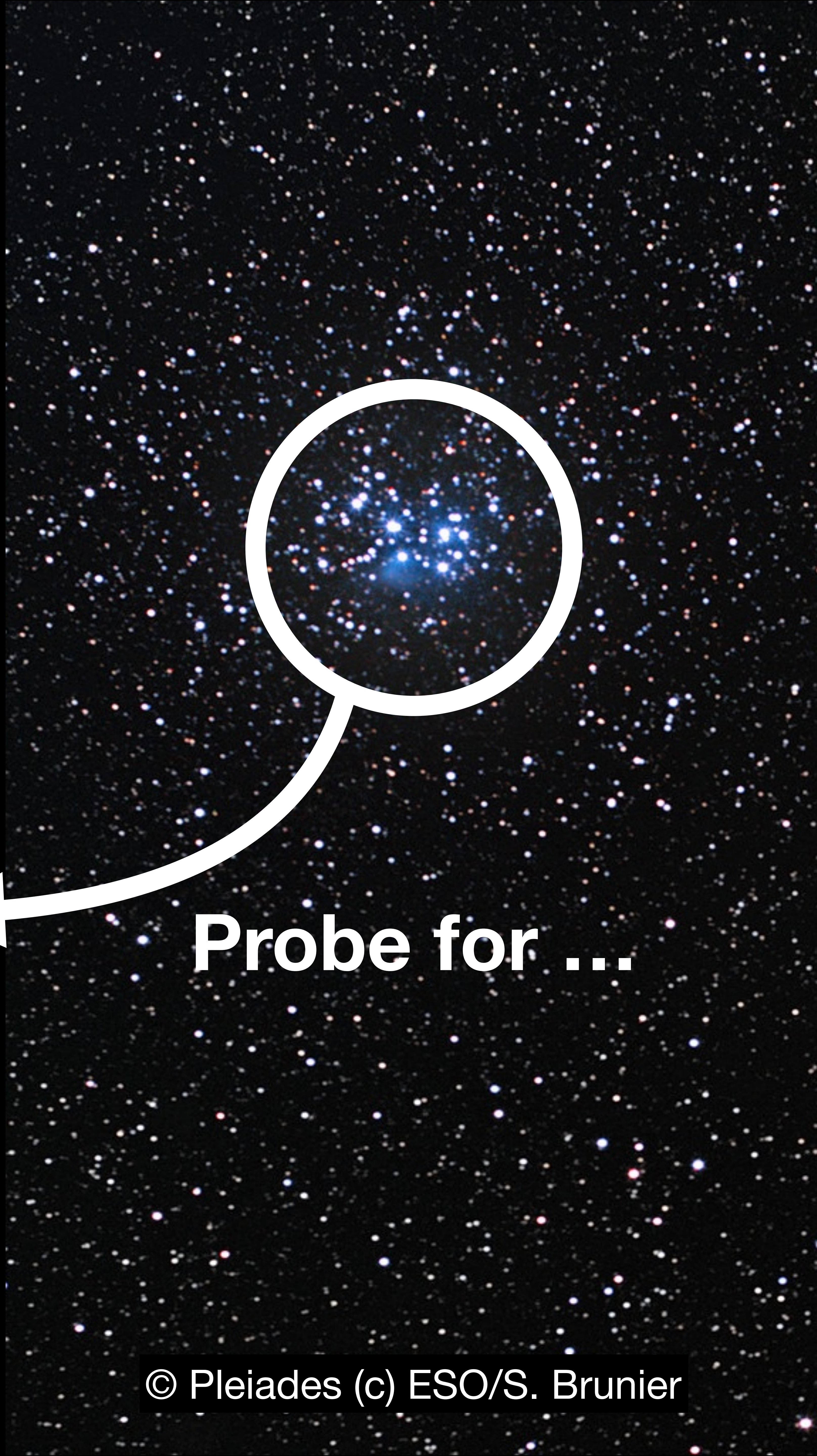
April 15, 2025

Stellar populations in the disk*

Born from same molecular cloud

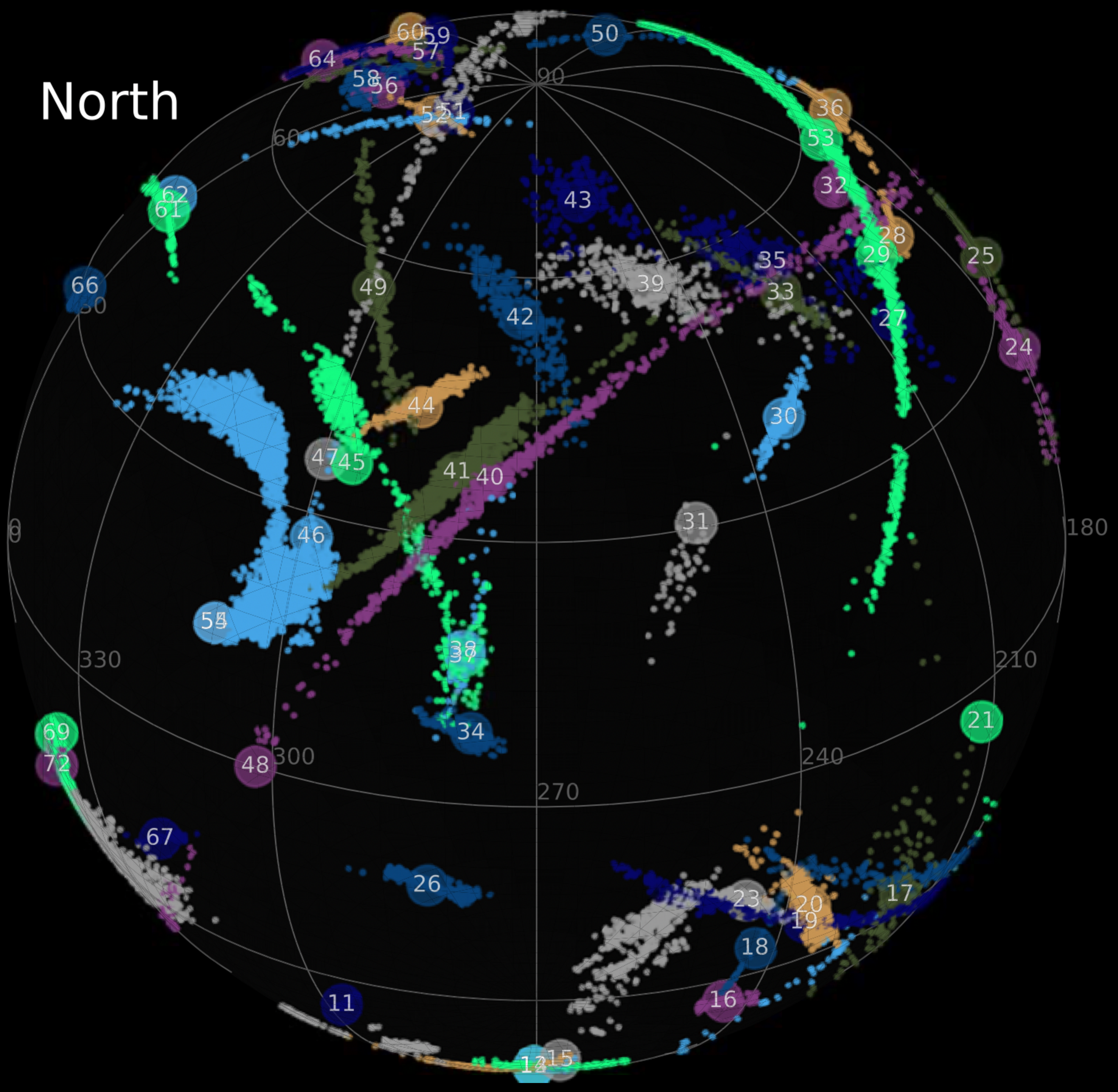
- Thought to be birthplace of most stars
  (Lada & Lada 2003; Parker & Goodwin 2007)
- Structure formation and evolution
- Chemical composition of Milky Way
- Exoplanet formation and evolution
- Stellar initial mass function

*stellar over-density over background




© S. Payne-Wardenaar / K. Malhan, MPIA


North
South
1 C-20
2 SGP-S
3 NGC288
4 ATLAS
5 Phoenix
6 Kwando
7 New-1
8 New-2
9 C-13
10 Gaia-12
11 Hydrus
12 NGC1261
13 NGC1261a
14 NGC1261b
15 Indus
16 NGC1851
17 New-3
18 New-4
19 Leiptr
20 C-12
21 New-5
22 New-6
23 NGC2298
24 C-25
25 C-11
26 NGC2808
27 New-7
28 New-8
29 New-9
30 C-10
31 New-10
32 New-11
33 Gaia-10
34 Gjoll
35 C-9
36 C-24
37 New-12
38 New-13
39 Slidr
40 New-14
41 Ylgr
42 Sylgr
43 New-15
44 Gaia-7
45 Gaia-8
46 New-16
47 Fjorm
48 Orphan
49 Gaia-1
50 C-23
51 LMS-1
52 C-22
53 GD1
54 Fimbulthul
55 Fimbulthul-S
56 New-17
57 NGC5466
58 New-18
59 Gaia-6
60 New-19
61 Pal-5
62 M5
63 Kshir
64 Svol
65 Gaia-9
66 Ophiuchus
67 NGC6101
68 M92
69 NGC6397
70 Gaia-11
71 Hrid
72 C-7
73 New-20
74 New-21
75 Phlegethon
76 NGC7089
77 NGC7099
78 New-22
79 New-23
80 New-24
81 New-25
82 New-26
83 New-27
84 NGC7492
85 C-19
86 Jhelum
87 Tuc-3



Probes for

- Galactic & DM halo potential shape
  (e.g., Dubinski+1999; Law & Majewski 2010; Malhan & Ibata 2019; Nibauer+2023)

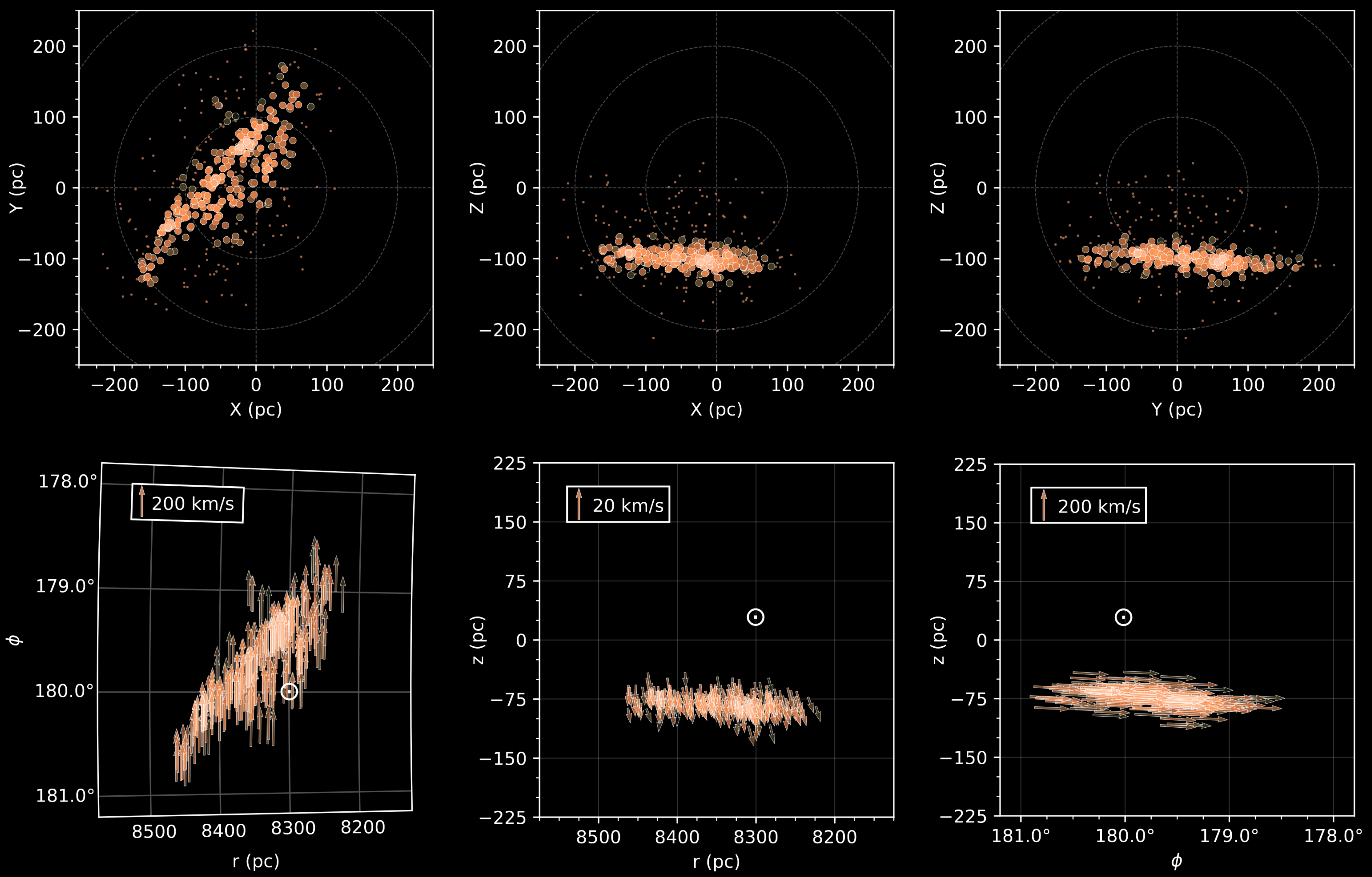
- DM subhalos
  (e.g., Johnston+2002; Ibata+2002; Siegal-Gaskins & Valluri+2008; Bonaca+2019)

- Hierarchical assembly of Galaxy
  (e.g., Ibata+1994; Helmi+1999; Myeong+2019)

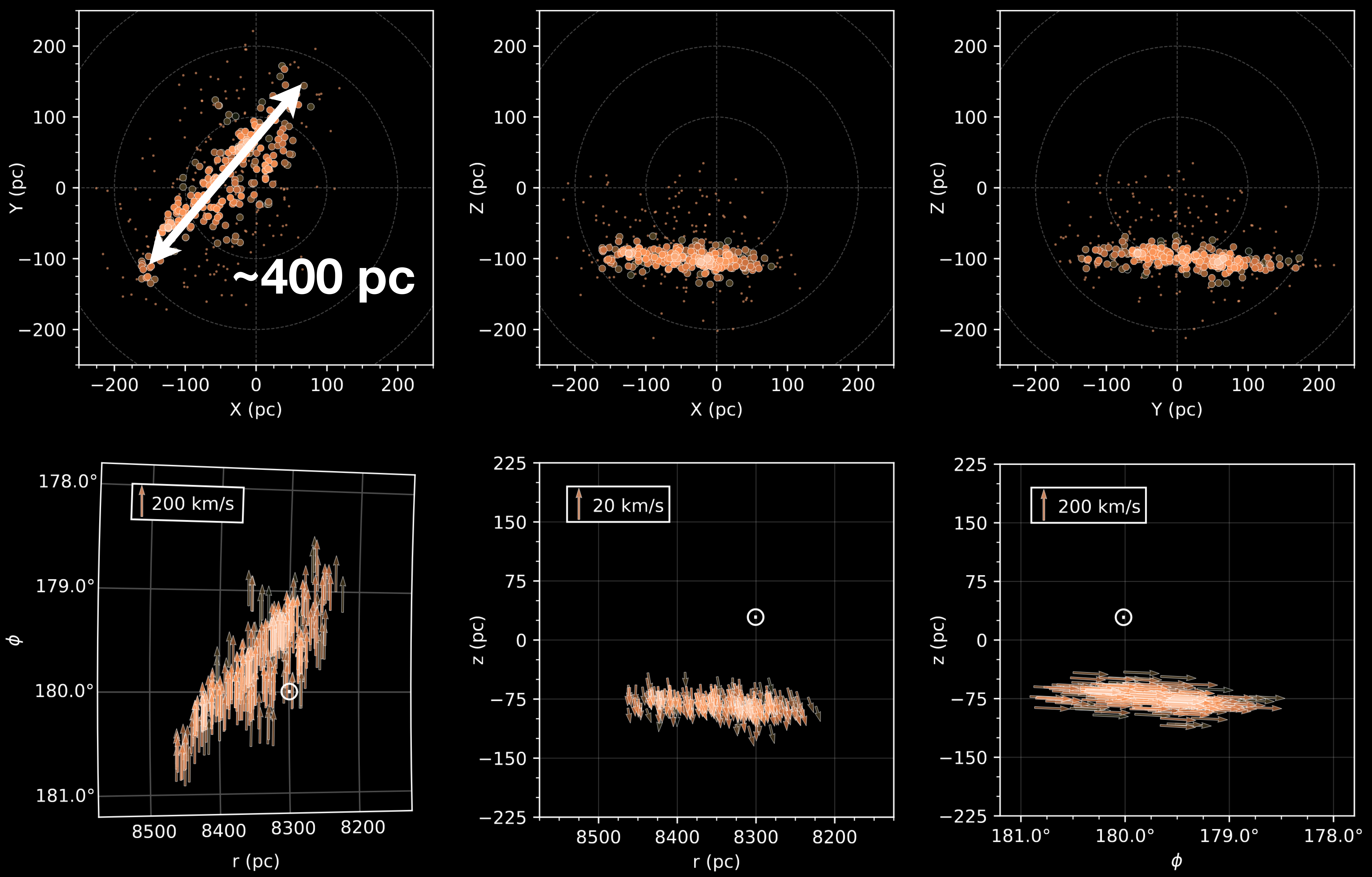
- Chemical tagging
  (e.g., Hasselquist+2019; Cunningham+2023)

What about the disk?

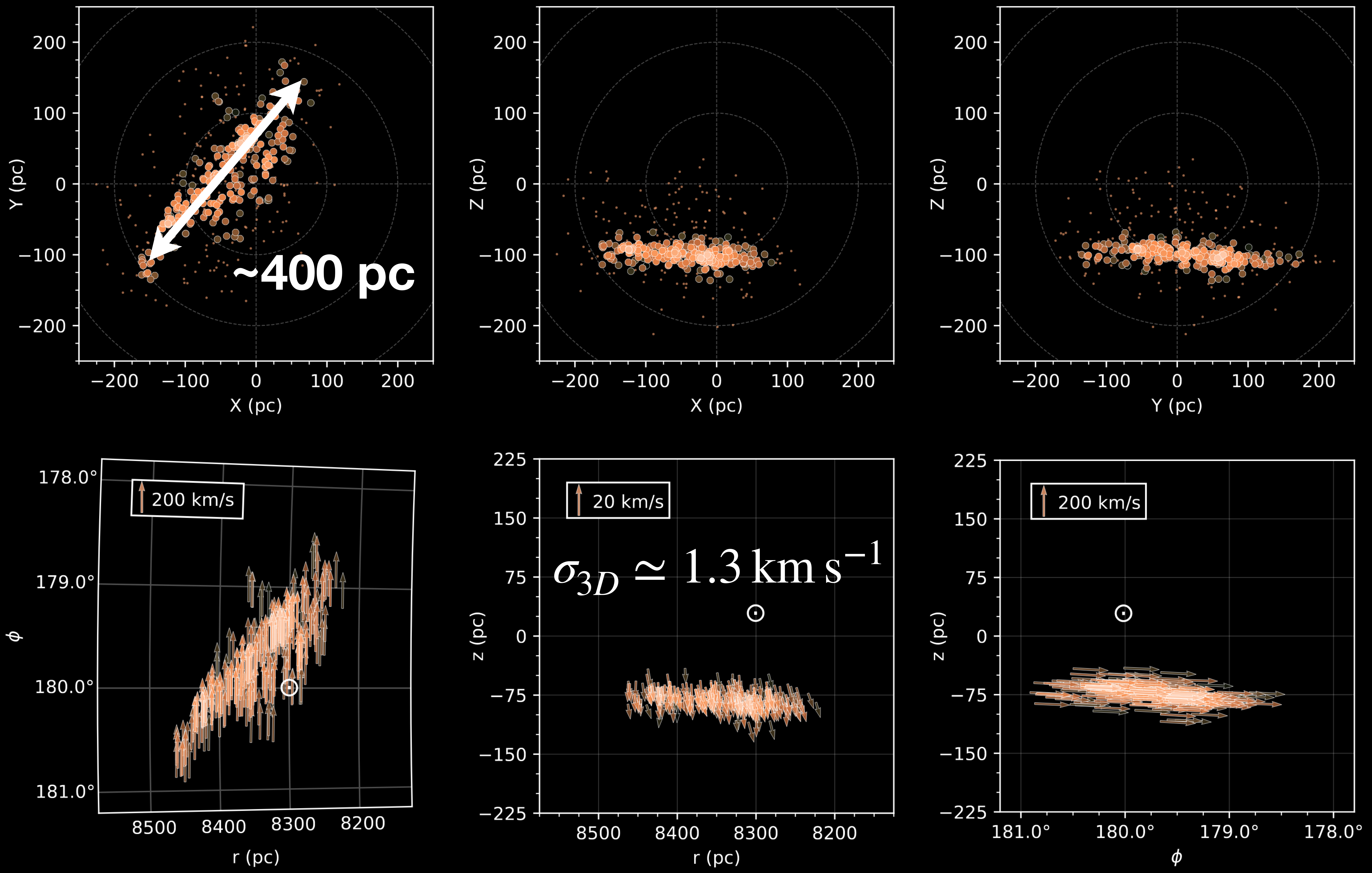
What about the disk?



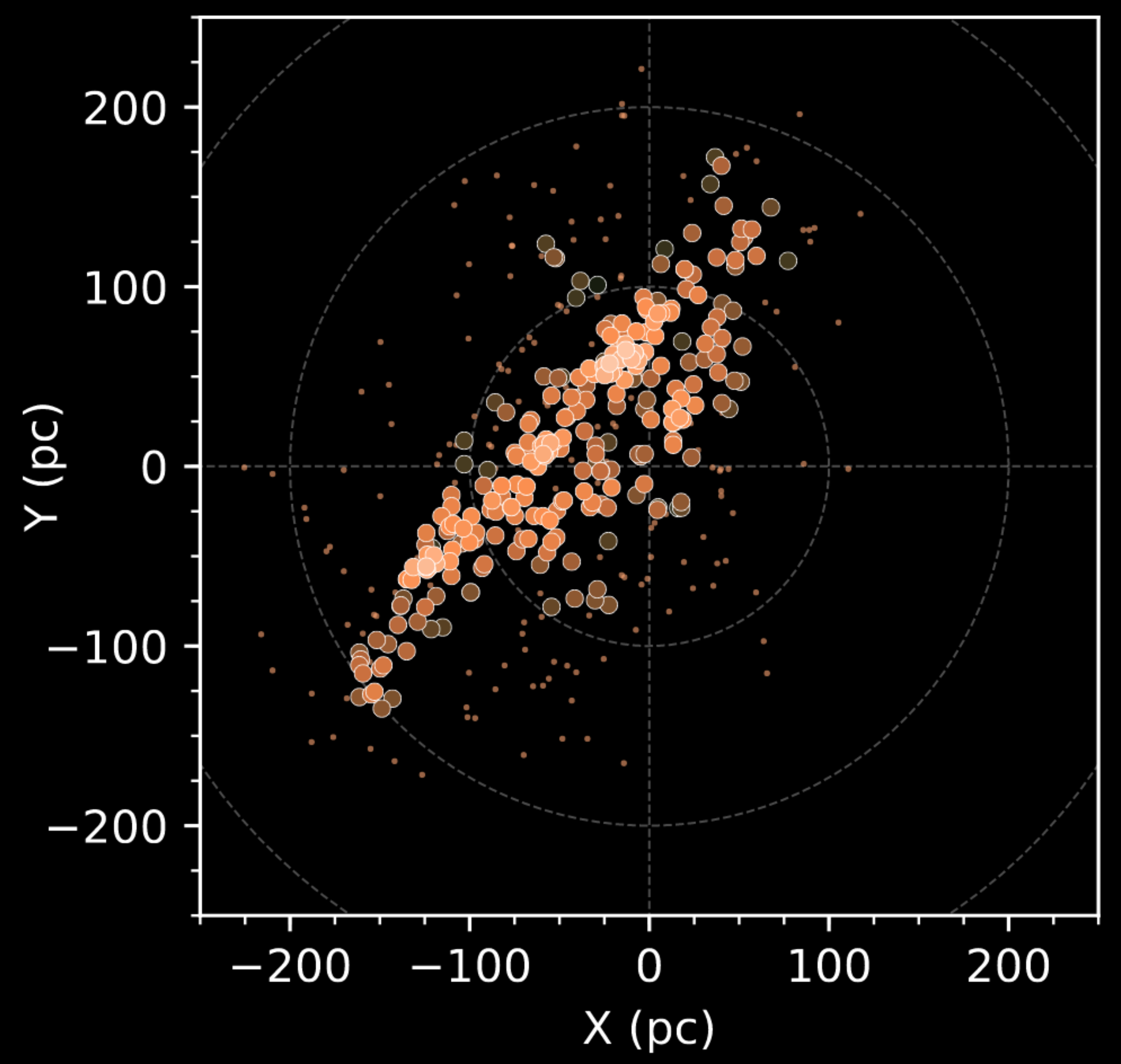
What about the disk?



What about the disk?

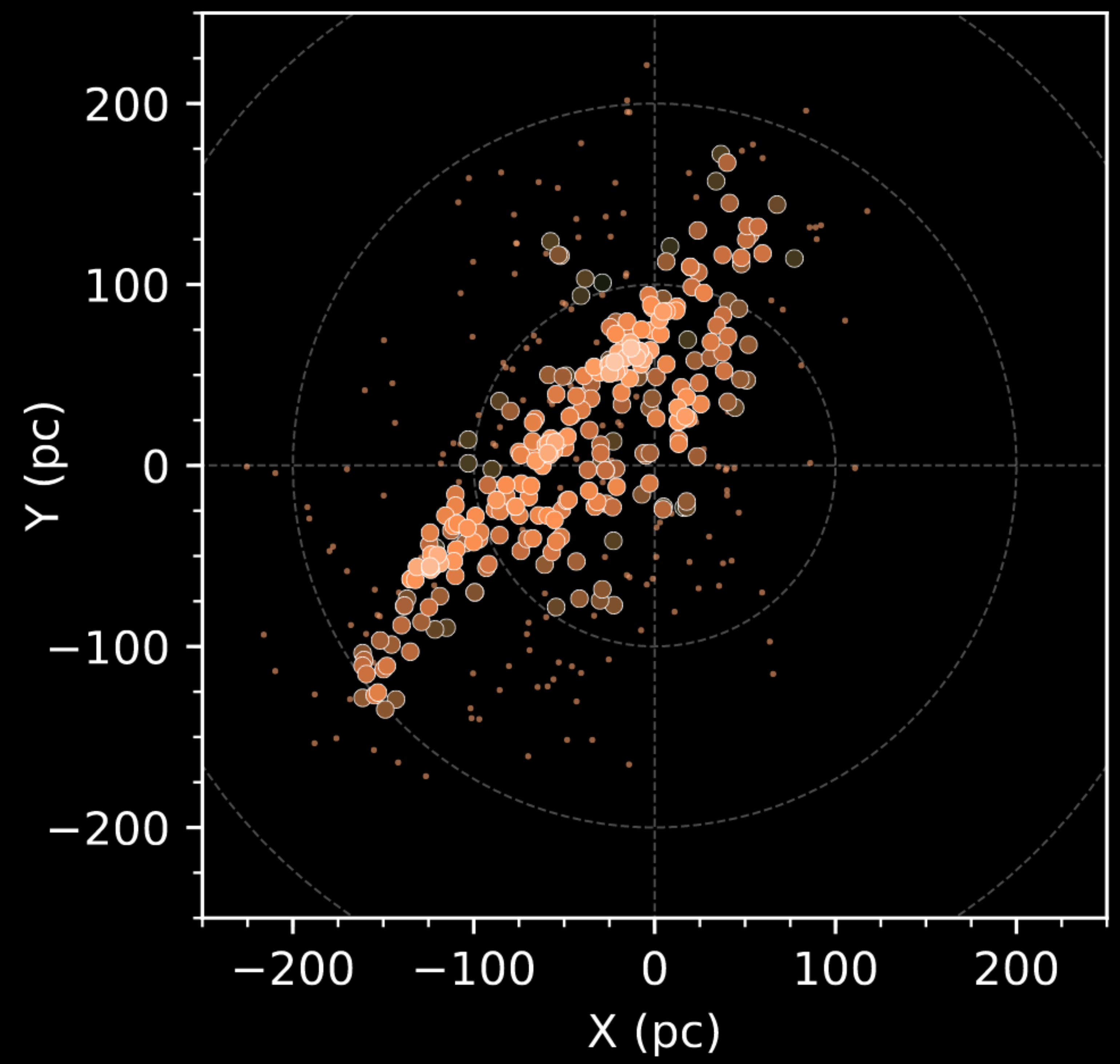


What can we learn?



Difficulties with assessing what we can learn

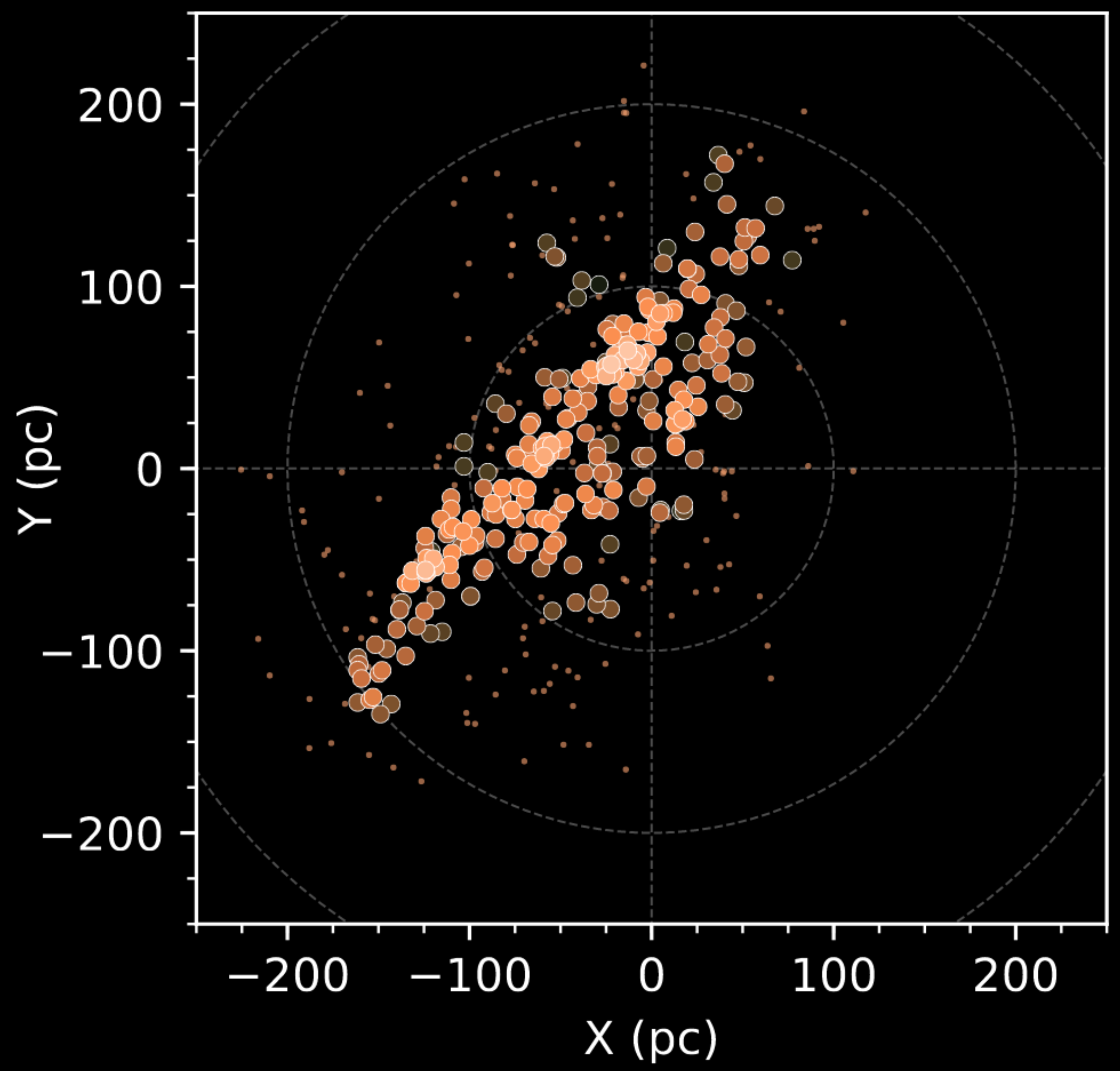
- Perturbations from spiral arms and bar more likely on disk-like orbits



© Meingast+2019b

Difficulties with assessing what we can learn

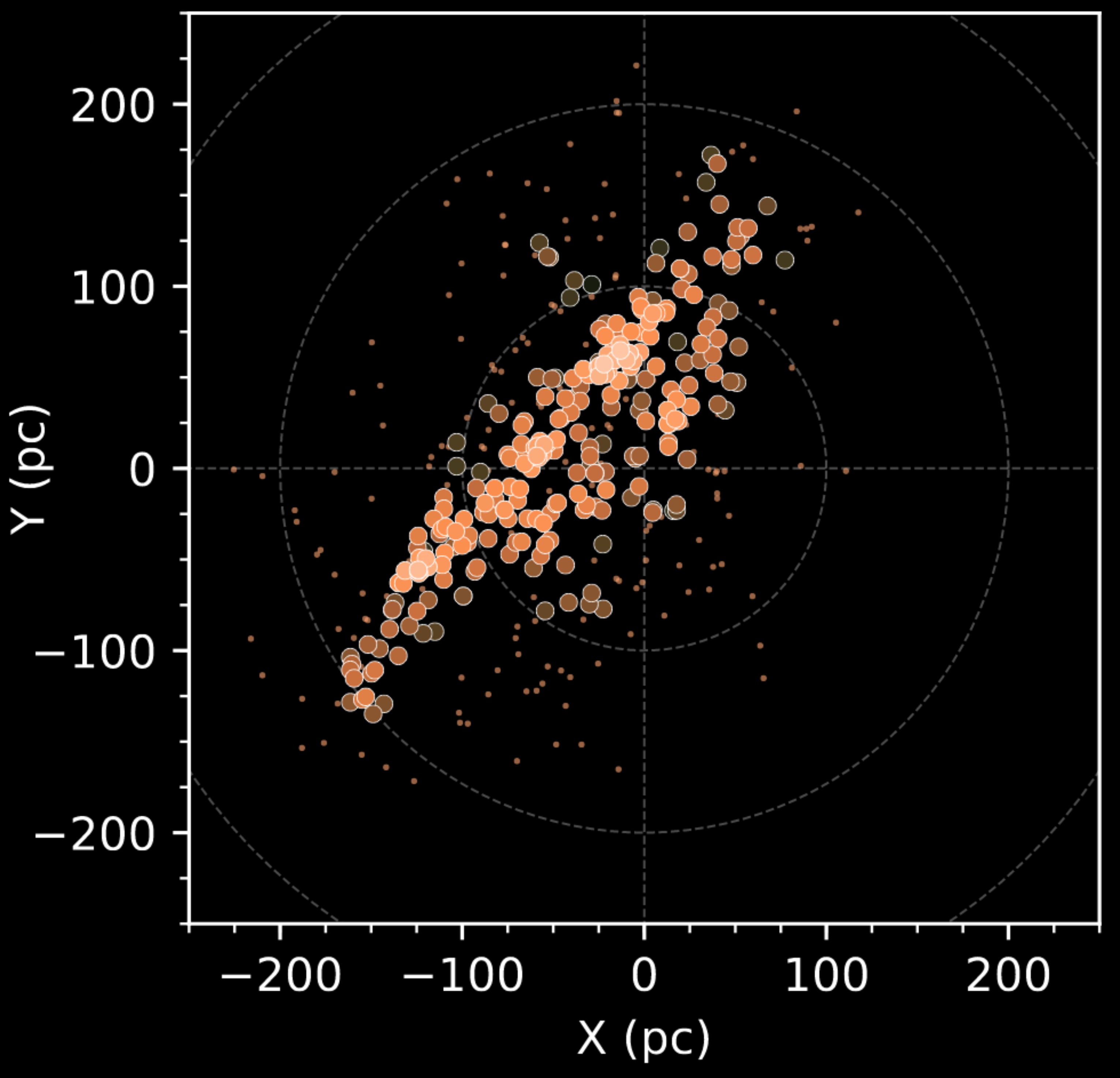
- Perturbations from spiral arms and bar more likely on disk-like orbits
- Unclear how long detectable in phase space
  ➡ much shorter dynamical time
  ➡ complex dynamical signature of stellar feedback and gas expulsion
  (e.g., Dinnbier & Kroupa 2020)



© Meingast+2019b

What can we learn?

- Cluster dissolution into the field
  (e.g., Almeida+2025)
- Tidal tails likely sensitive to shape of Galactic potential & presence of GMCs
  (e.g., Jerabkova+2021)
- Number density, age, and rewinding accuracy
  → GMC number density, mass function, lifetime, & velocity dispersion
  (e.g., Kamdar+2021)
- Rewinding accuracy can probe bar mass
  (e.g., Kamdar+2021)



© Meingast+2019b

What can we learn?

- Cluster dissolution into the field
  (e.g., Almeida+2025)

- Tidal tails likely sensitive to shape of Galactic potential & presence of GMCs
  (e.g., Jerabkova+2021)

- Number density, age, and rewinding accuracy
  → GMC number density, mass function, lifetime, & velocity dispersion
  (e.g., Kamdar+2021)

- Rewinding accuracy can probe bar mass
  (e.g., Kamdar+2021)

Prediction

~**5–80** streams per kpc^3

Goal: provide first estimate of *disk stream number density* from control volume

Search volume

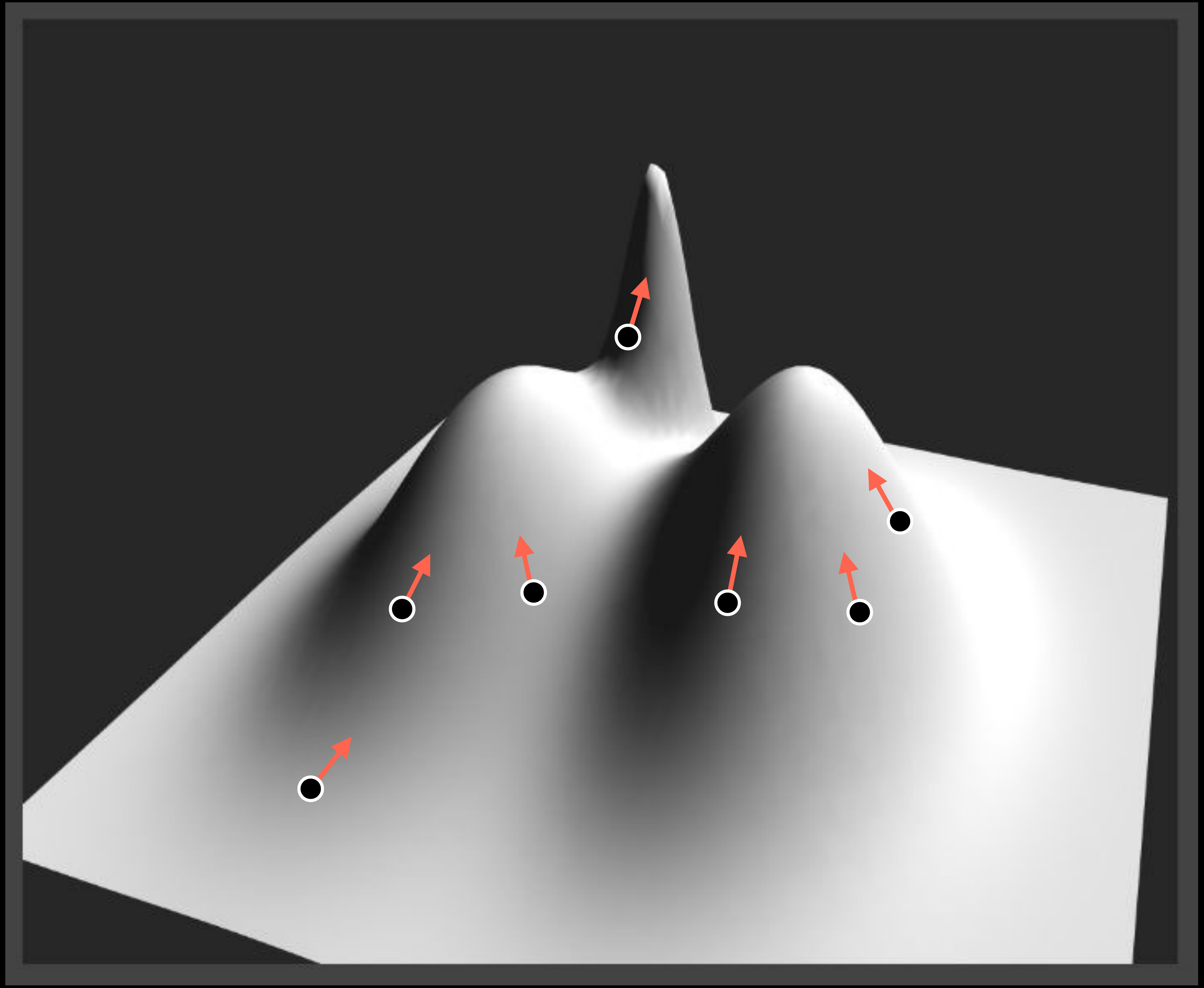
Search volume

Apply *SigMA*

Recap: Density based clustering

Nonparametric, density-based clustering
Problem definition

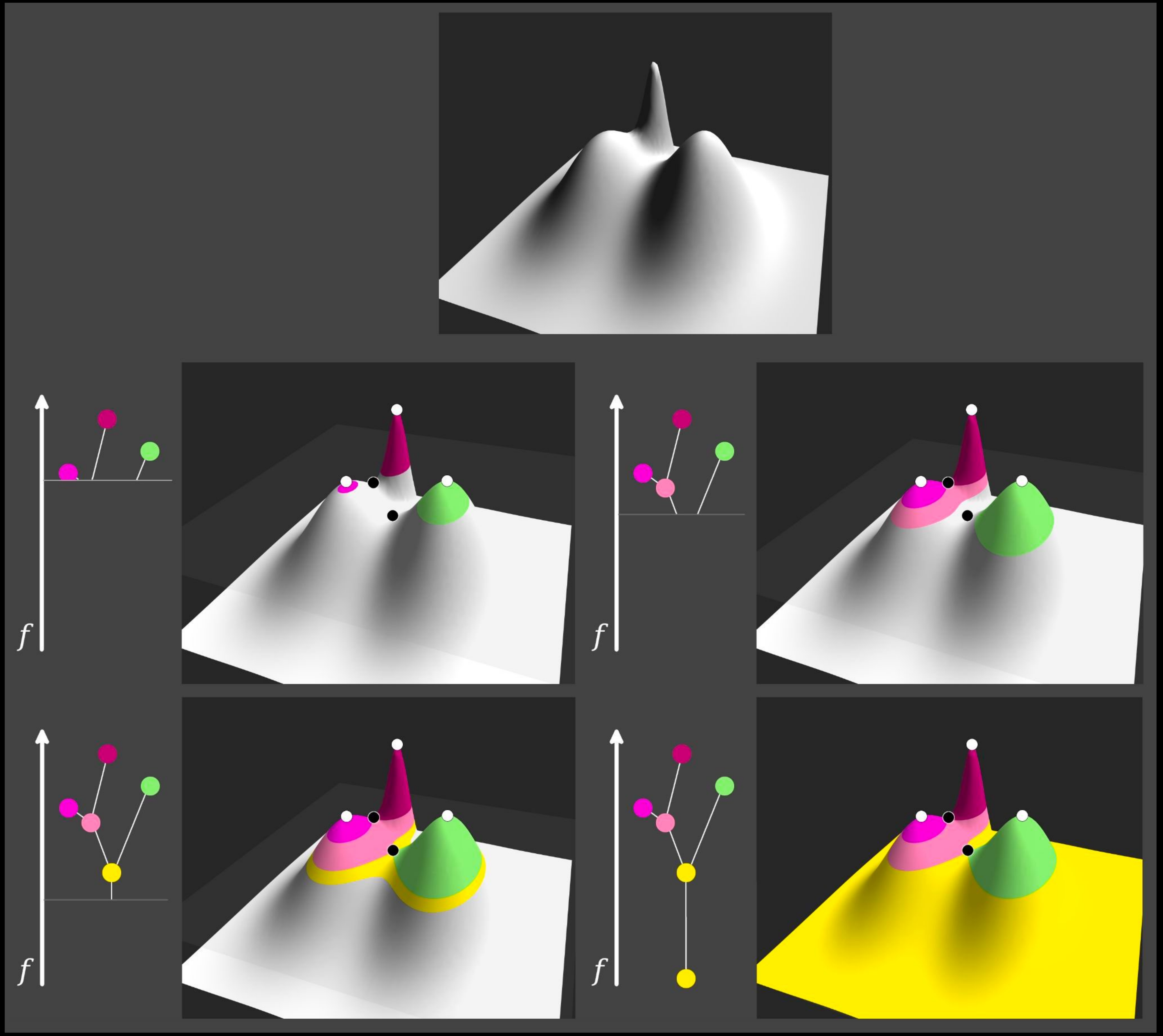
- **Wishart (1969) cluster definition**
  - $\mathbf{x}_i$ associated with modes of f
  - Propagate $\mathbf{x}_i$ along ∇f



Nonparametric, density-based clustering

Problem definition

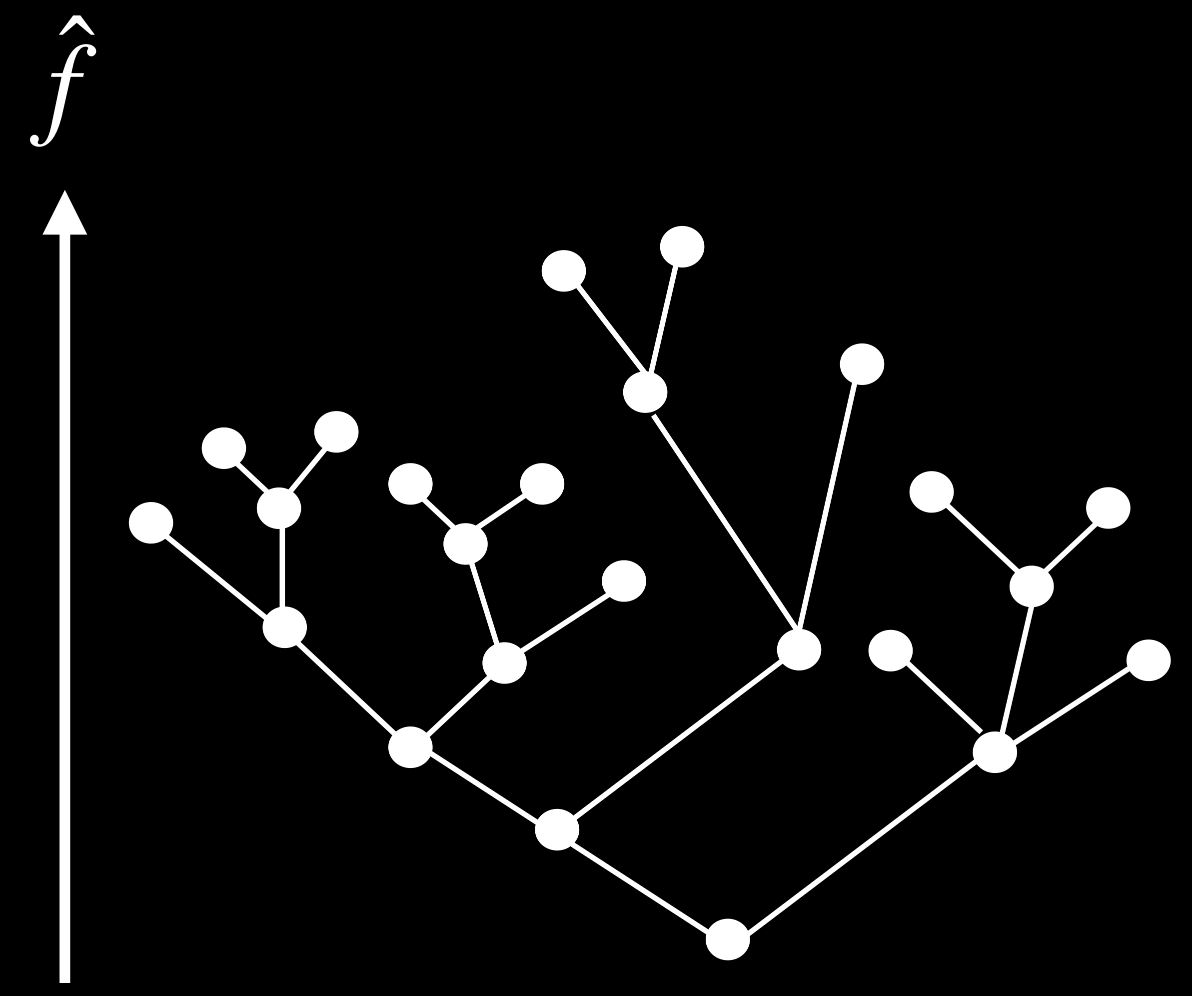
- Level set: $L(\lambda) = \{f(\mathbf{x}) \geq \lambda\}$
- **Hartigan (1975) cluster definition**
  - Connected components of $L(\lambda)$
  - Cluster tree: vary λ: $\infty \rightarrow -\infty$

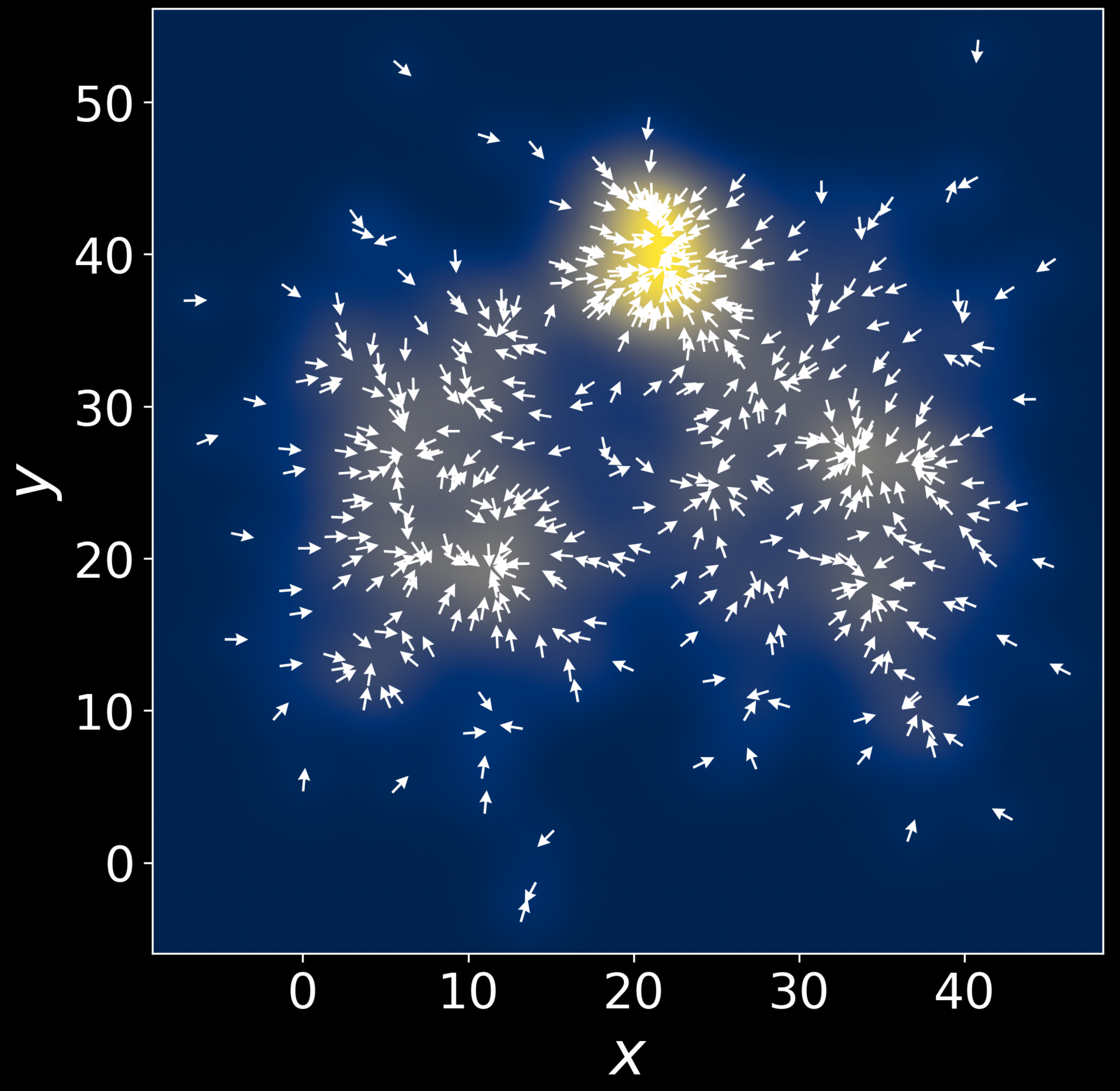


Reality:
Estimate density from data

Clustering pipeline: SigMA

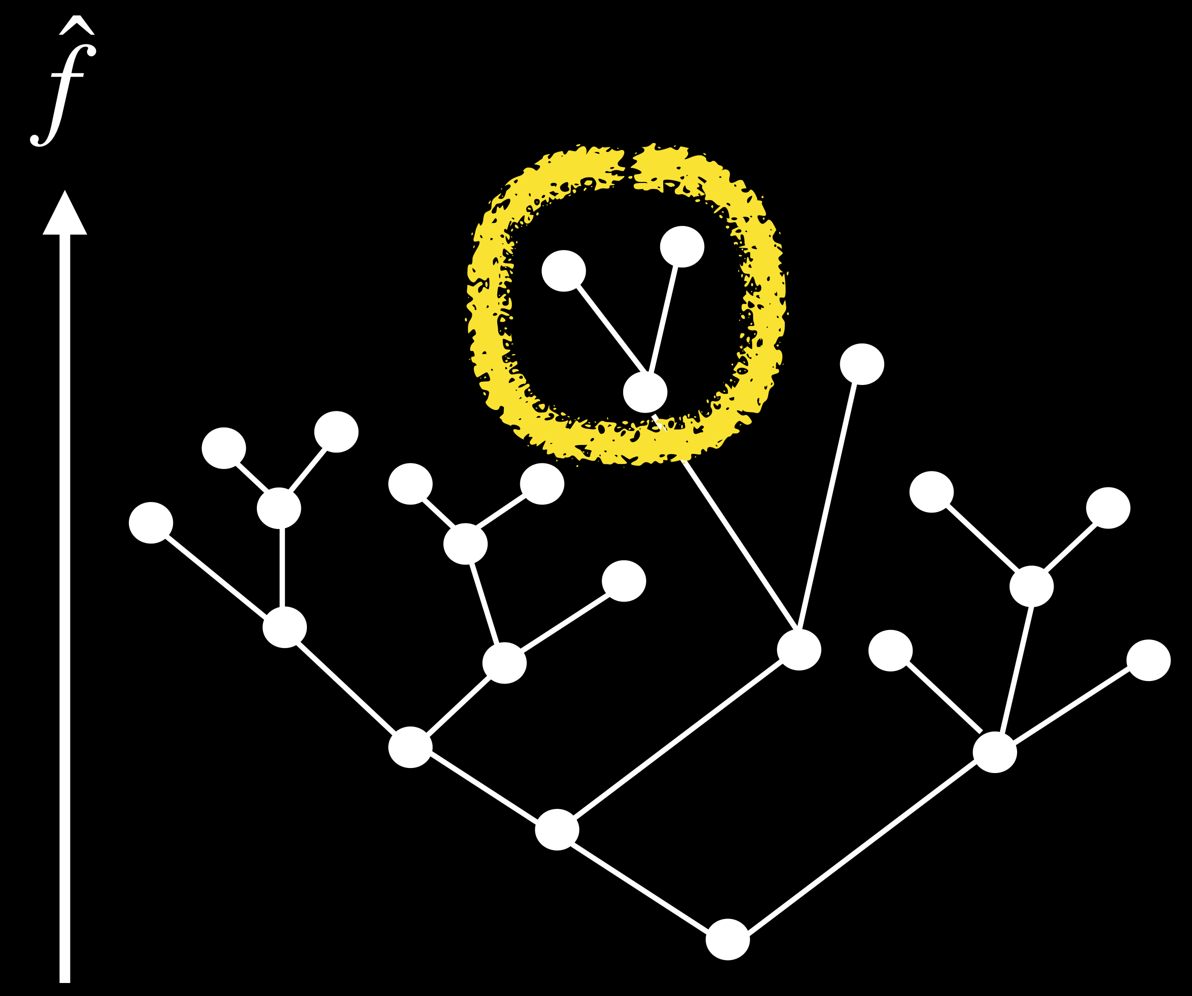
1. Gradient ascent step — cluster tree

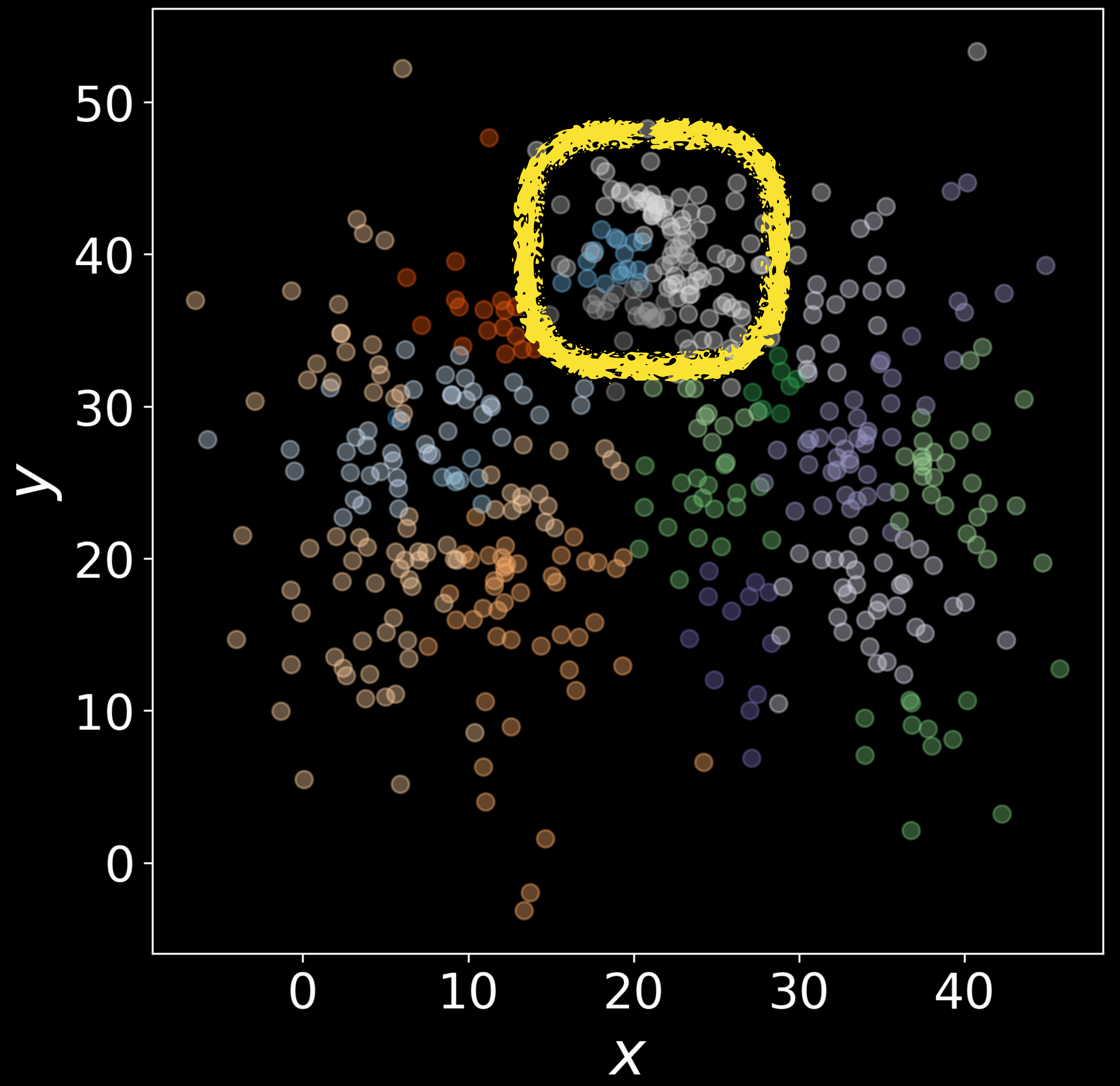




Clustering pipeline: SigMA

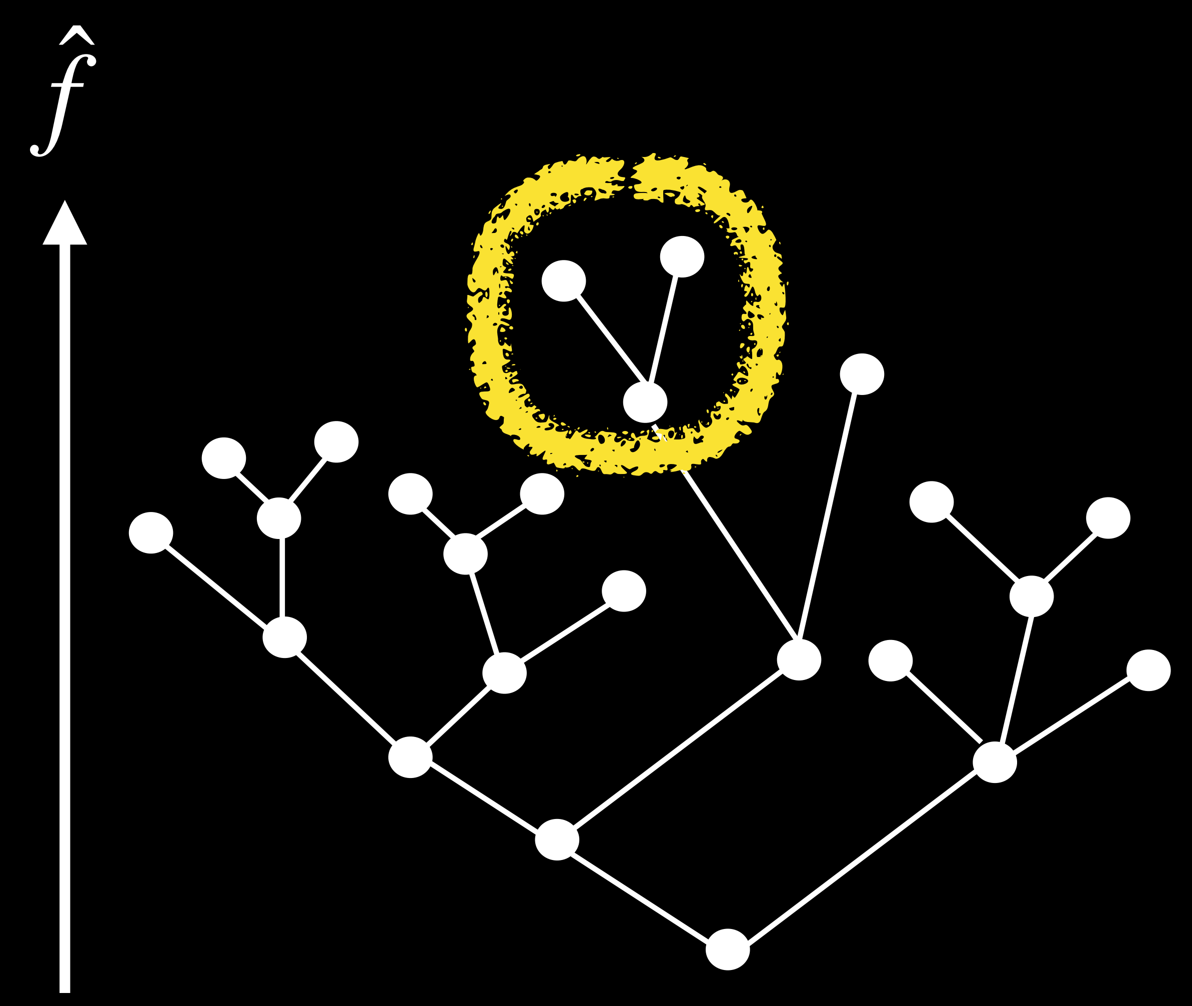
1. Gradient ascent step
2. Scan saddle points: $\max \hat{f} \rightarrow \min \hat{f}$

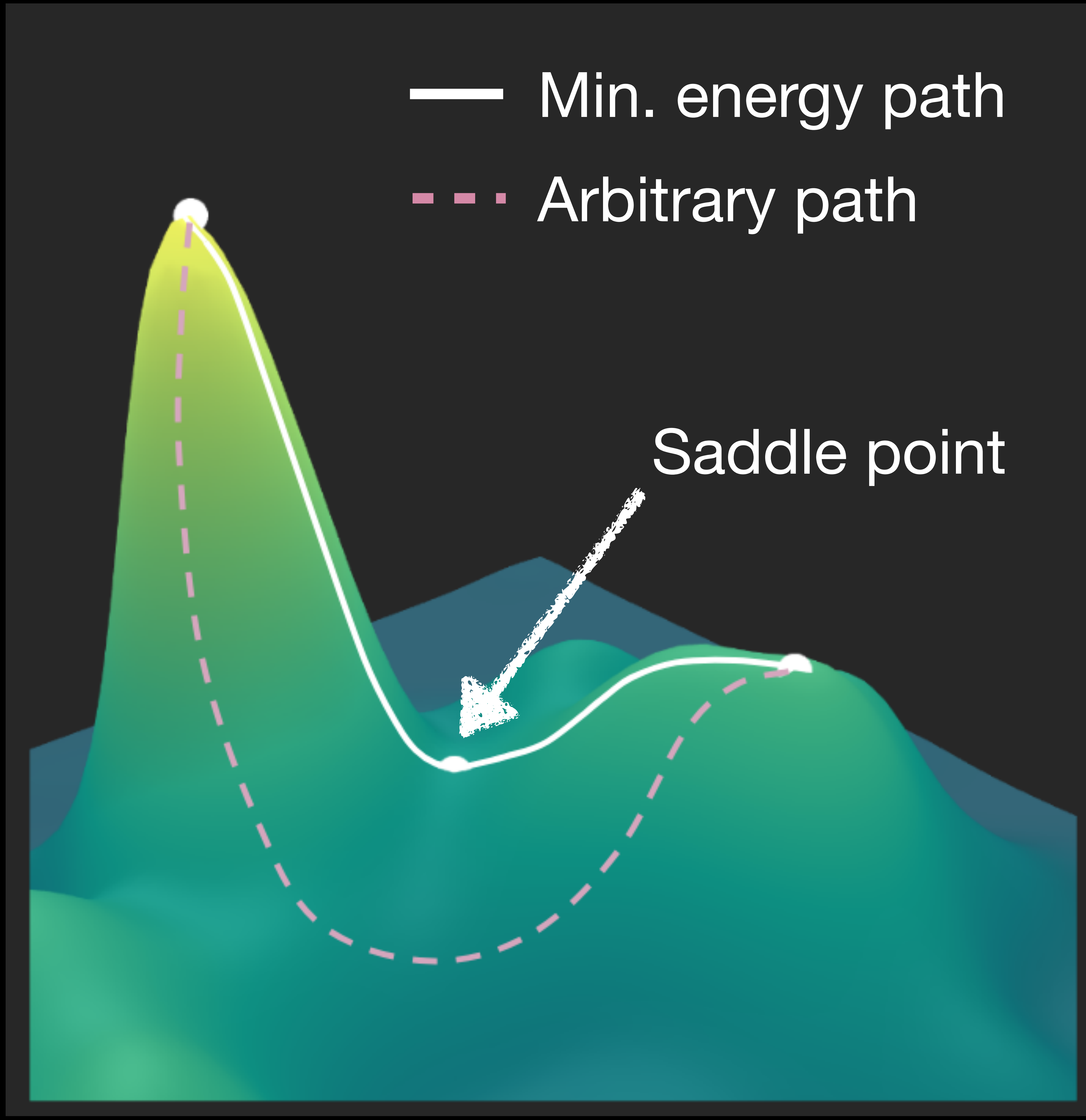




Clustering pipeline: SigMA

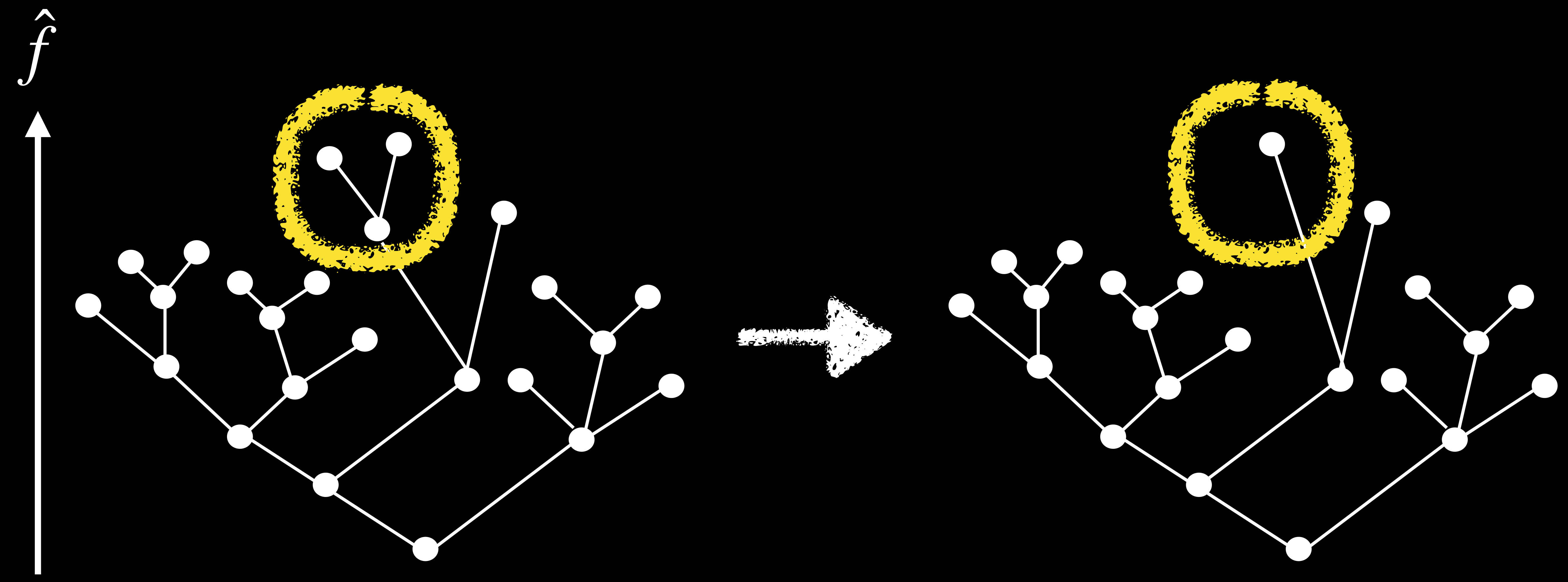
1. Gradient ascent step
2. Scan saddle points: max $\hat{f} \rightarrow$ min $\hat{f}$
   A. Test modality between modes

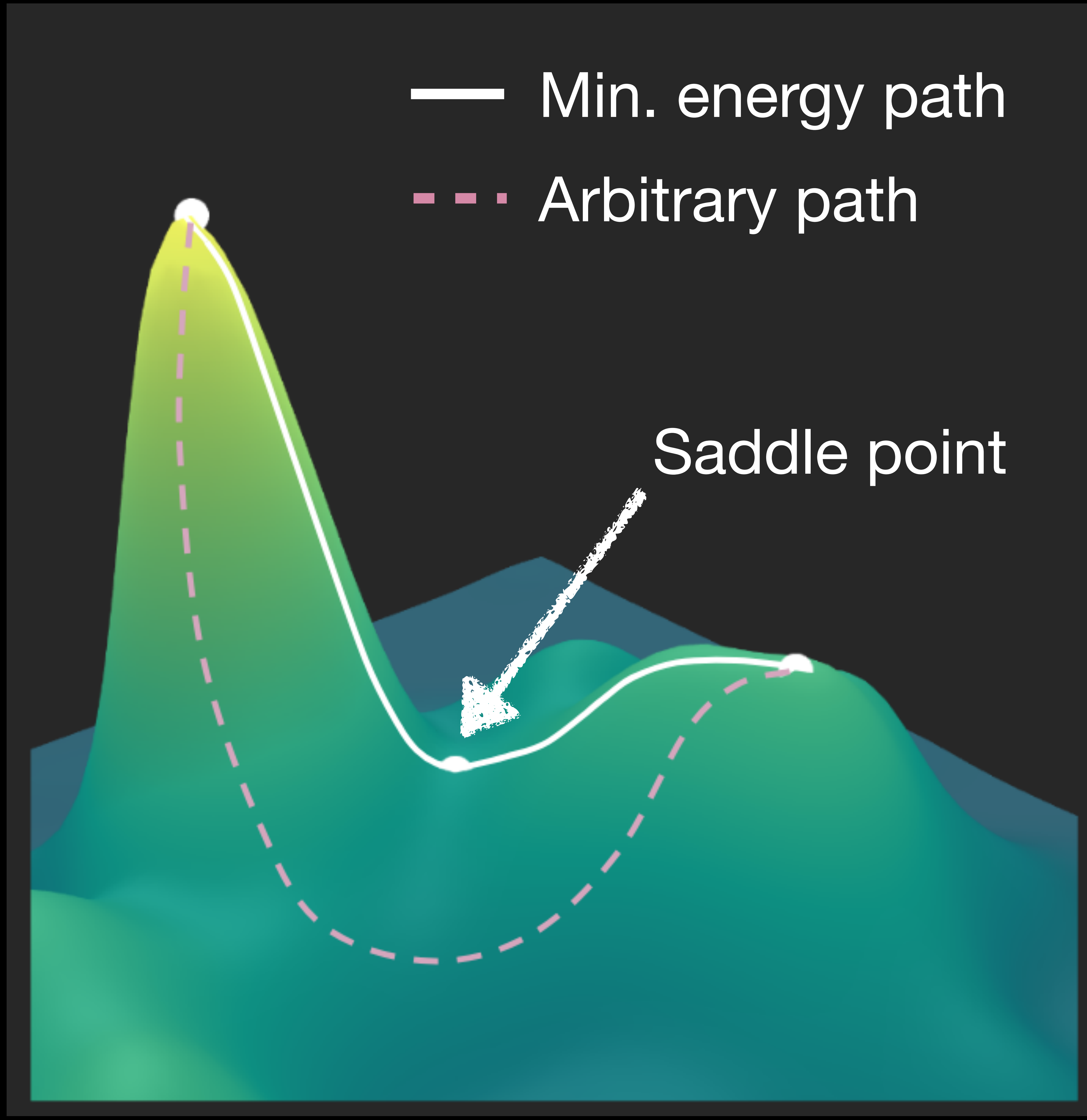




Clustering pipeline: SigMA

1. Gradient ascent step
2. Scan saddle points: $\max \hat{f} \rightarrow \min \hat{f}$
   A. Test modality between modes
   B. If H_0 cannot be rejected — merge

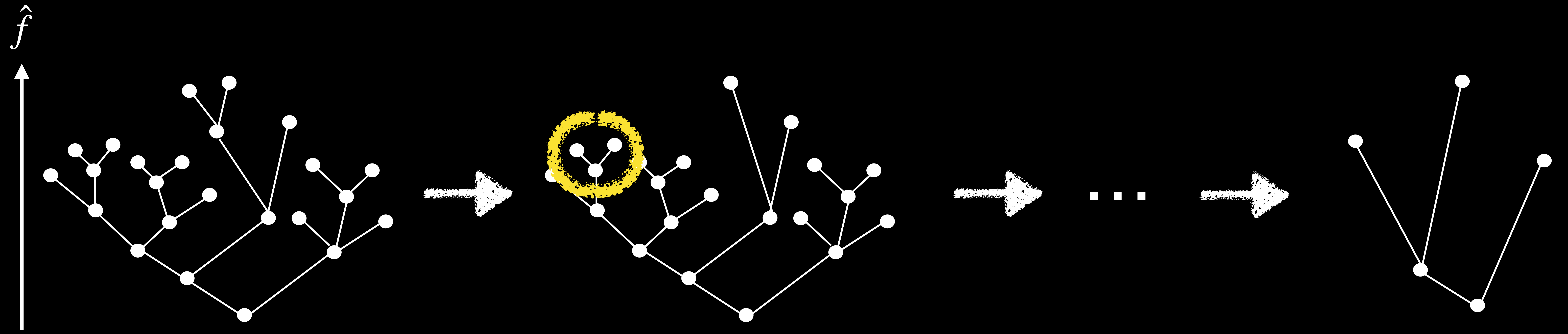




Clustering pipeline: SigMA

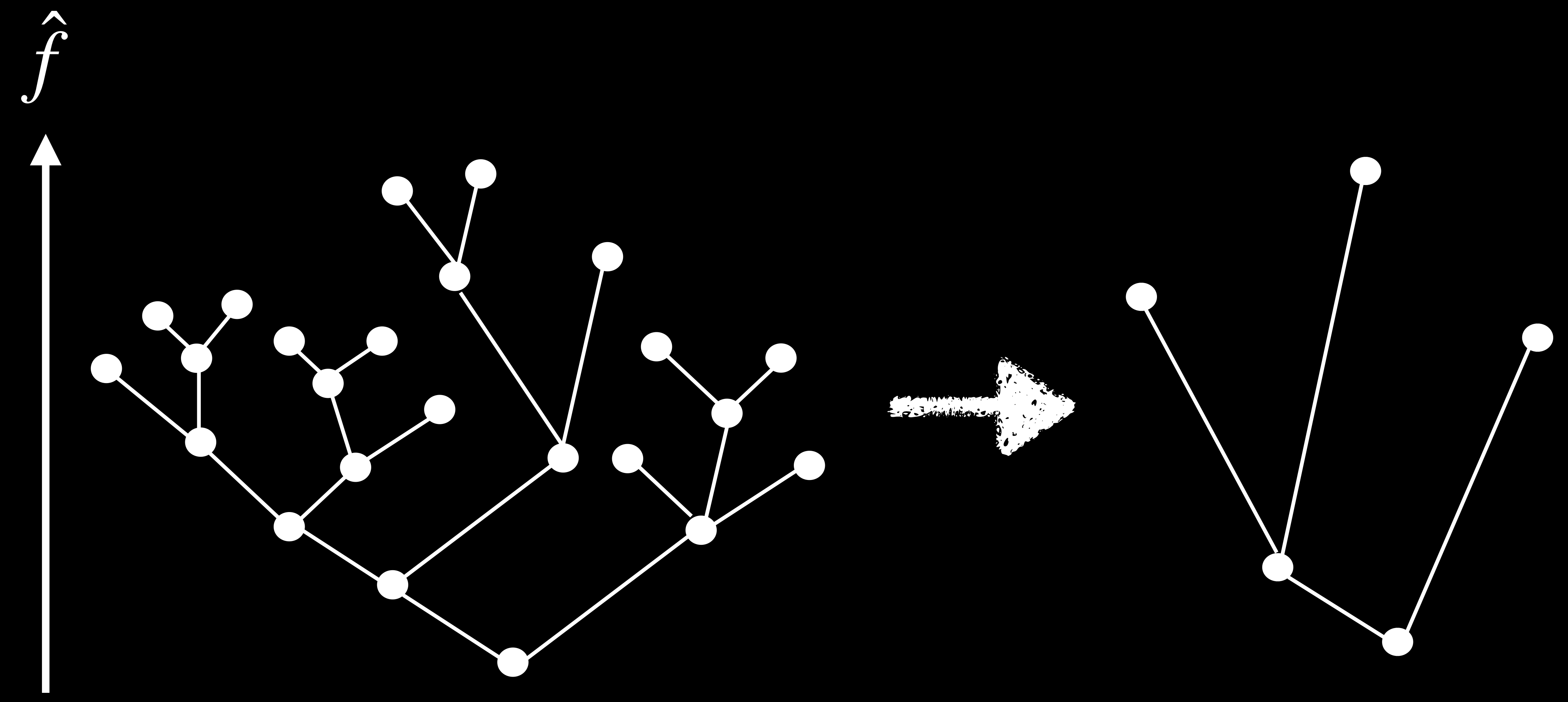
1. Gradient ascent step
2. Scan saddle points: max $\hat{f} \rightarrow$ min $\hat{f}$
   A. Test modality between modes
   B. If H_0 cannot be rejected — merge

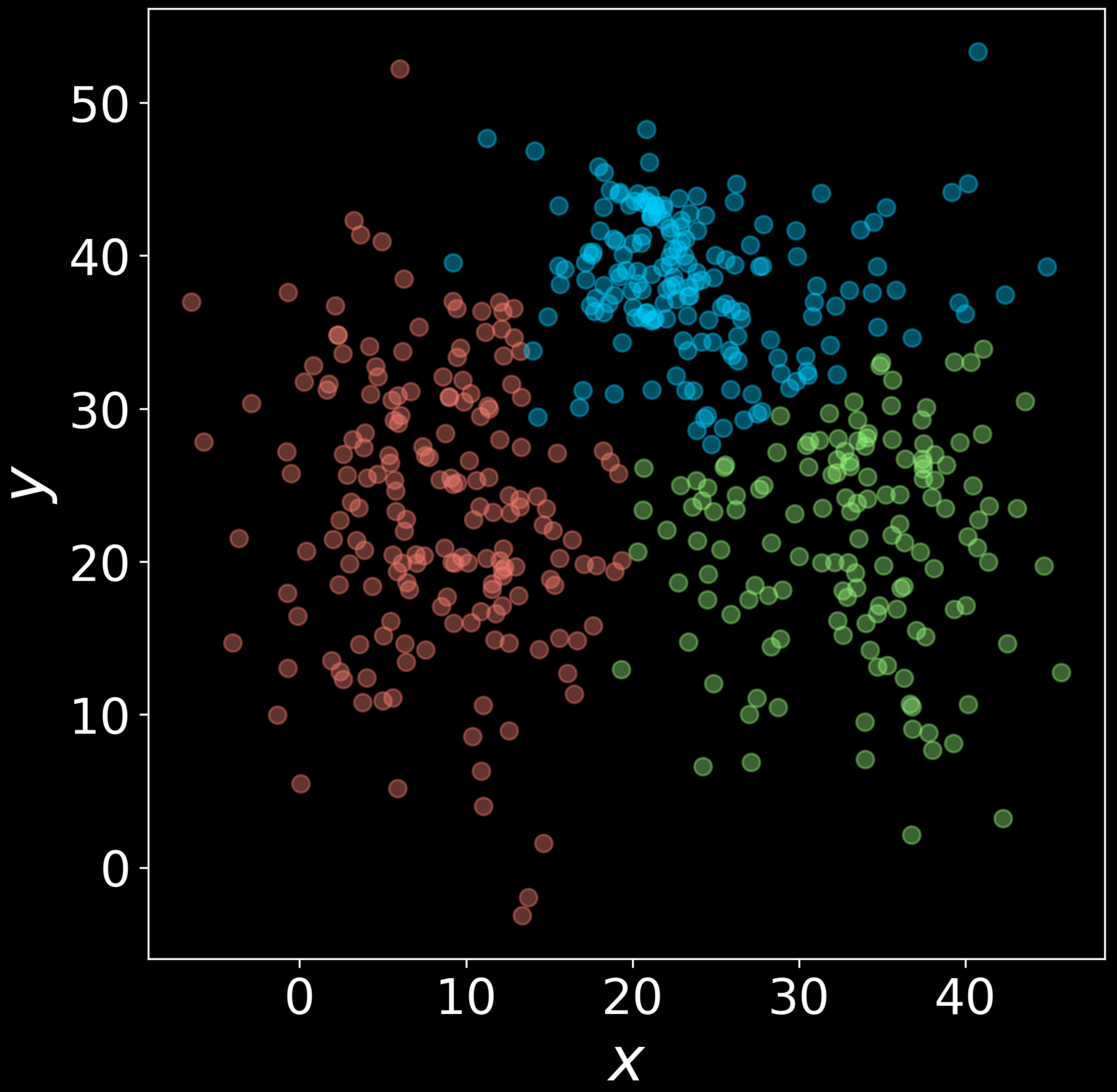
Next saddle point



Clustering pipeline: SigMA

1. Gradient ascent step
2. Scan saddle points: max $\hat{f} \rightarrow$ min $\hat{f}$
    A. Test modality between modes
    B. If H_0 cannot be rejected — merge

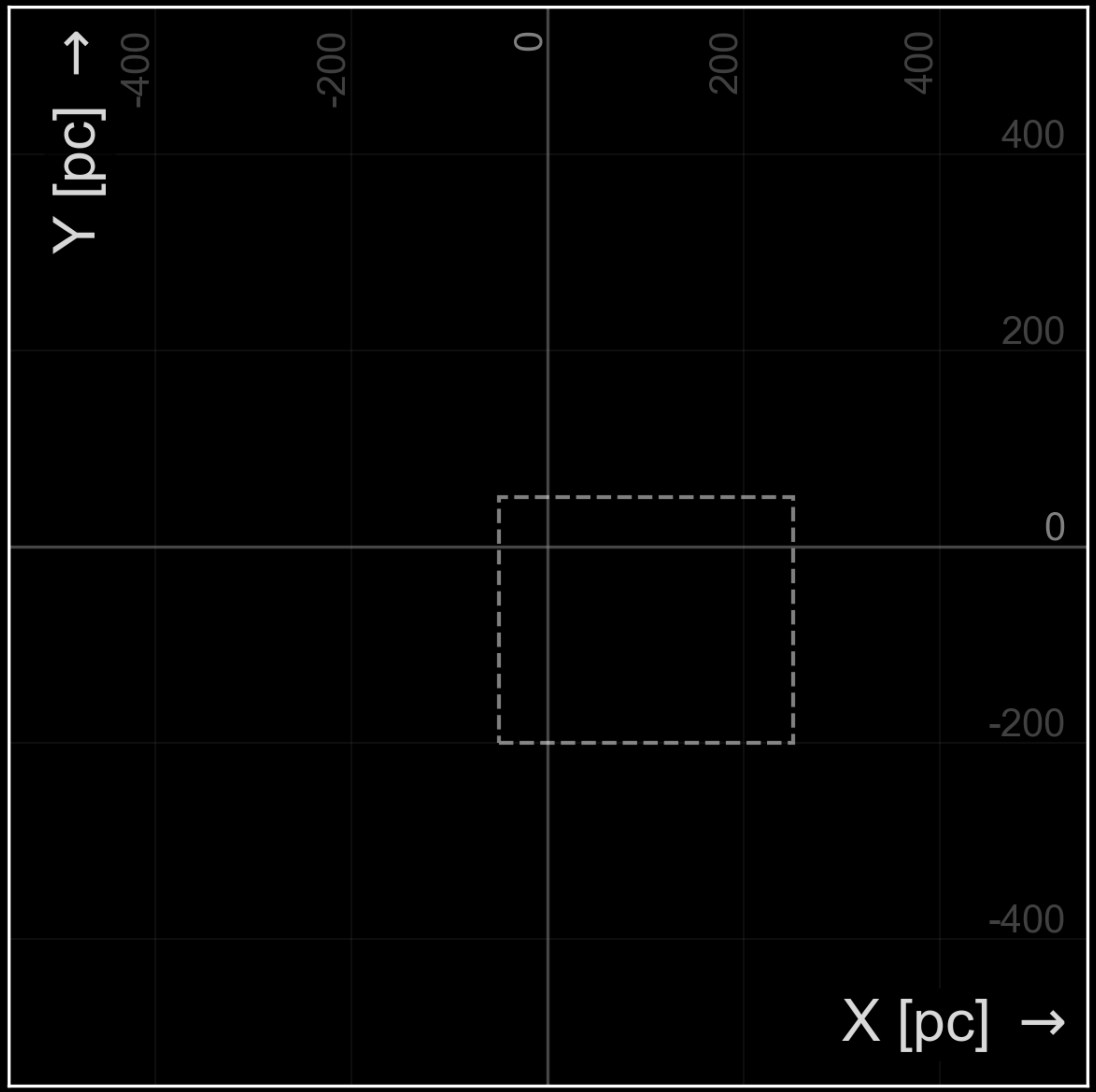




Application to search volume

Selection criteria

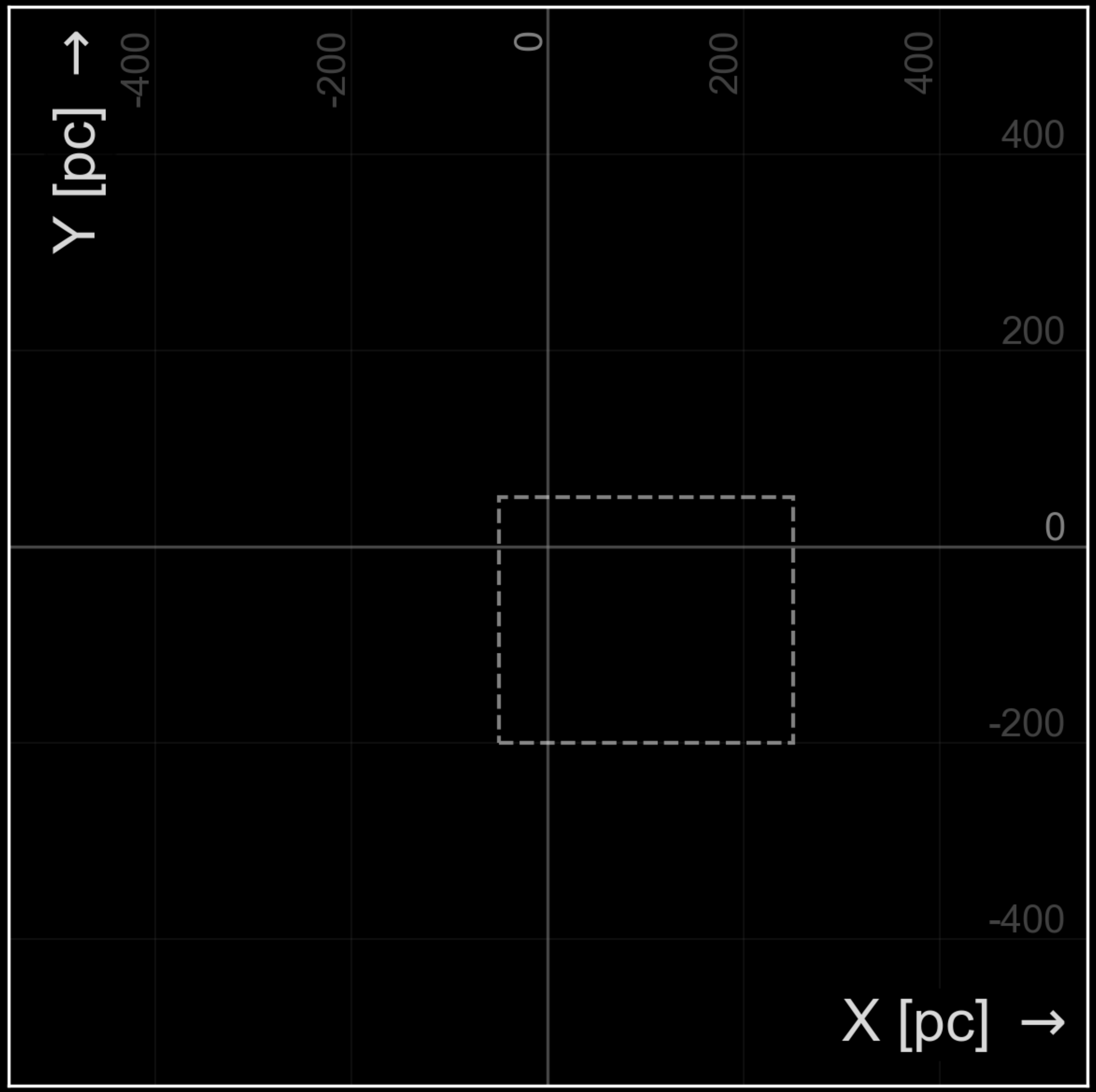
1. Aspect ratio **>3:1**



Application to search volume

Selection criteria

1. Aspect ratio **>3:1**
2. Detectable in search volume alone
   ➡ Follow-up search to recover remaining members



Results: Disk stream candidates

Disk stream candidates

Disk stream candidates (1)

Disk stream candidates (2)

Disk stream candidates (3)

Disk stream candidates (4)

Disk stream candidates (5)

Disk stream candidates (6)

Disk stream candidates (7)

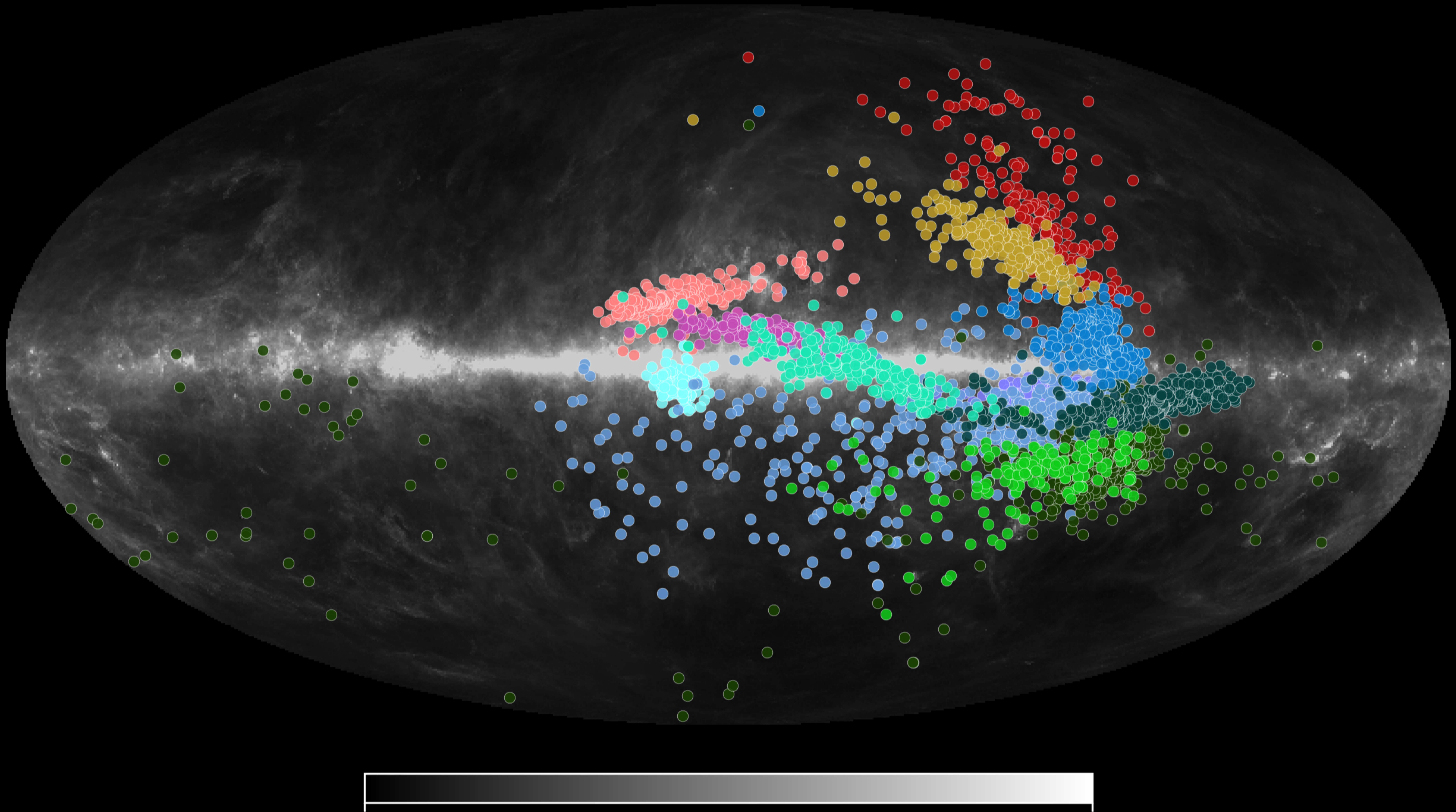
Disk stream candidates (8)

Disk stream candidates (9)

Disk stream candidates (10)

Disk stream candidates (11)

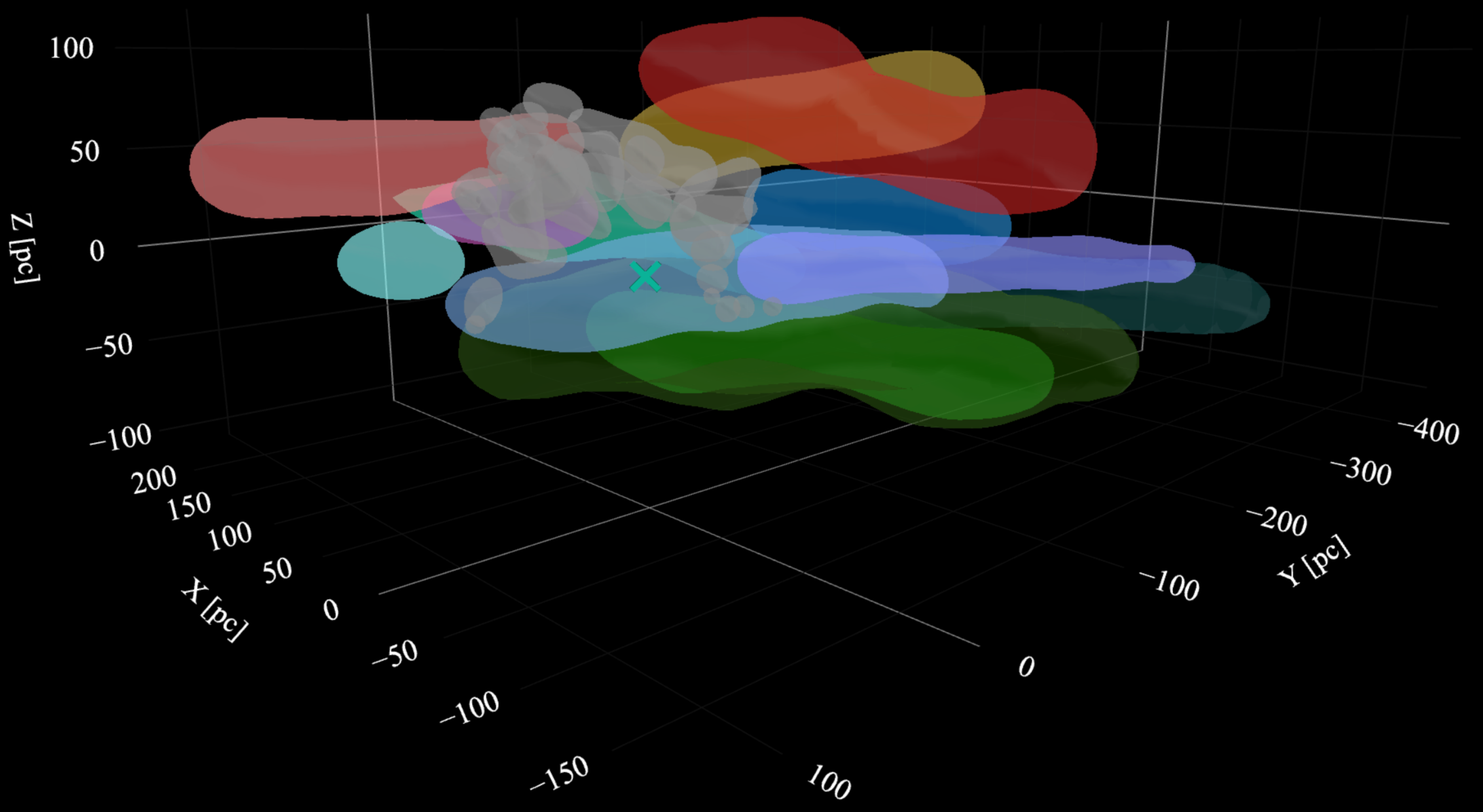
Disk stream candidates (12)



Planck Collaboration (2014)

12 disk stream candidates

- Lengths between ~200 – 400 pc
- Coeval & dynamically cold $\sigma_{3D} = 2 - 7\,\mathrm{km\,s^{-1}}$
- Densities as low as 0.2 stars / $10^3\,\mathrm{pc}^3$
- **820 streams /** kpc^3

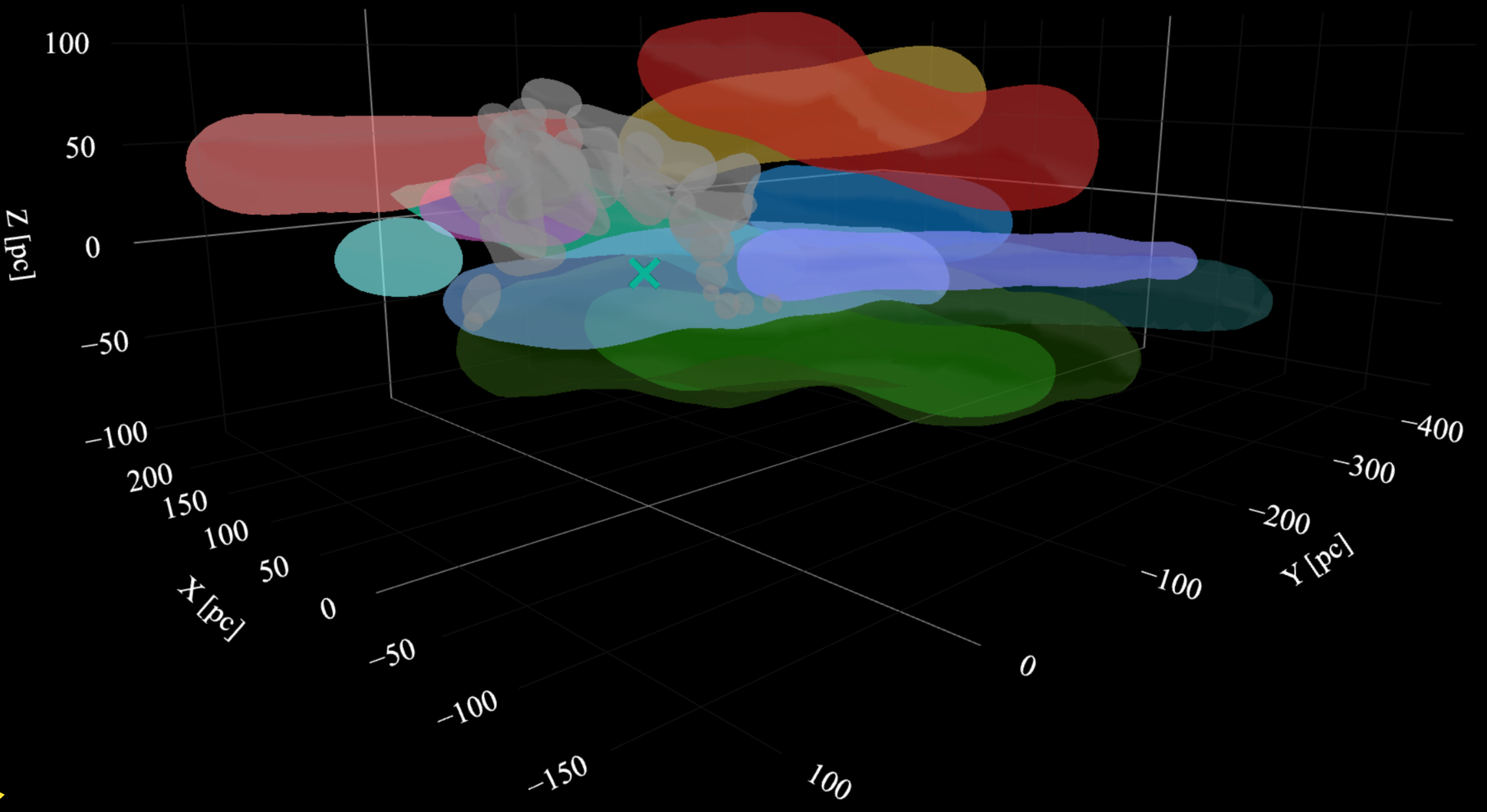


12 disk stream candidates

- Lengths between ~200 — 400 pc
- Coeval & dynamically cold $\sigma_{3D} = 2 - 7\,\mathrm{km\,s^{-1}}$
- Densities as low as 0.2 stars / 10^3 pc^3
- **820 streams /** kpc^3

Prediction (Kamdar+2021)

~5–80 streams / kpc^3



12 disk stream candidates

- Lengths between ~200 — 400 pc
- Coeval & dynamically cold $\sigma_{3D} = 2 - 7\,\mathrm{km\,s^{-1}}$
- Densities as low as 0.2 stars / 10^3 pc^3
- **820 streams /** kpc^3

Prediction (Kamdar+2021)

~**5–80** streams / kpc^3

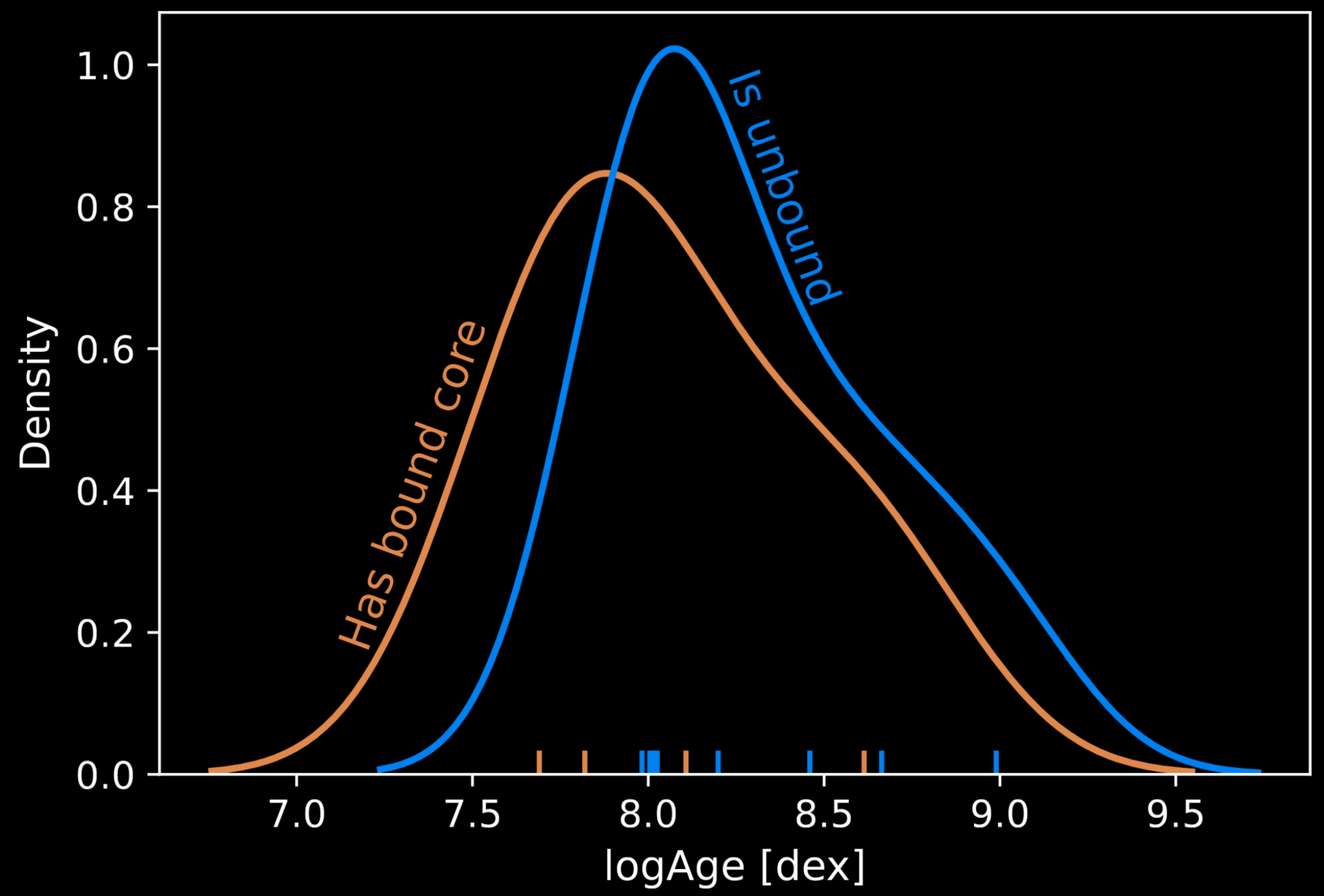
Points to either

i) **higher** cluster/stream **formation efficiency**, or

ii) **lower destruction efficiency** via e.g. GMCs

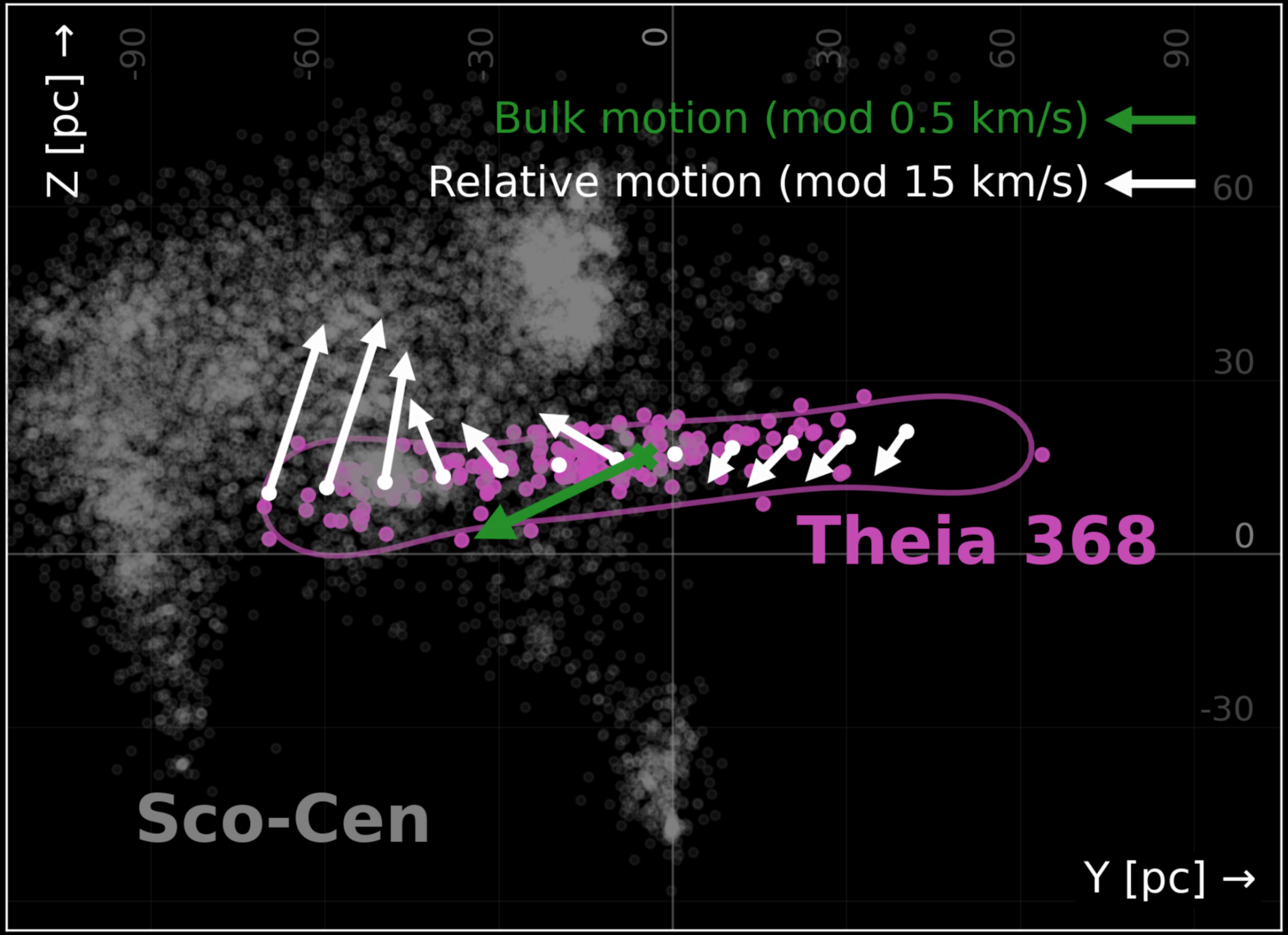
Further findings I

- Most streams are $\lesssim 200\,\mathrm{Myr}$
- Some have bound cores
  - unbound older $\sim 100\,\mathrm{Myr}$



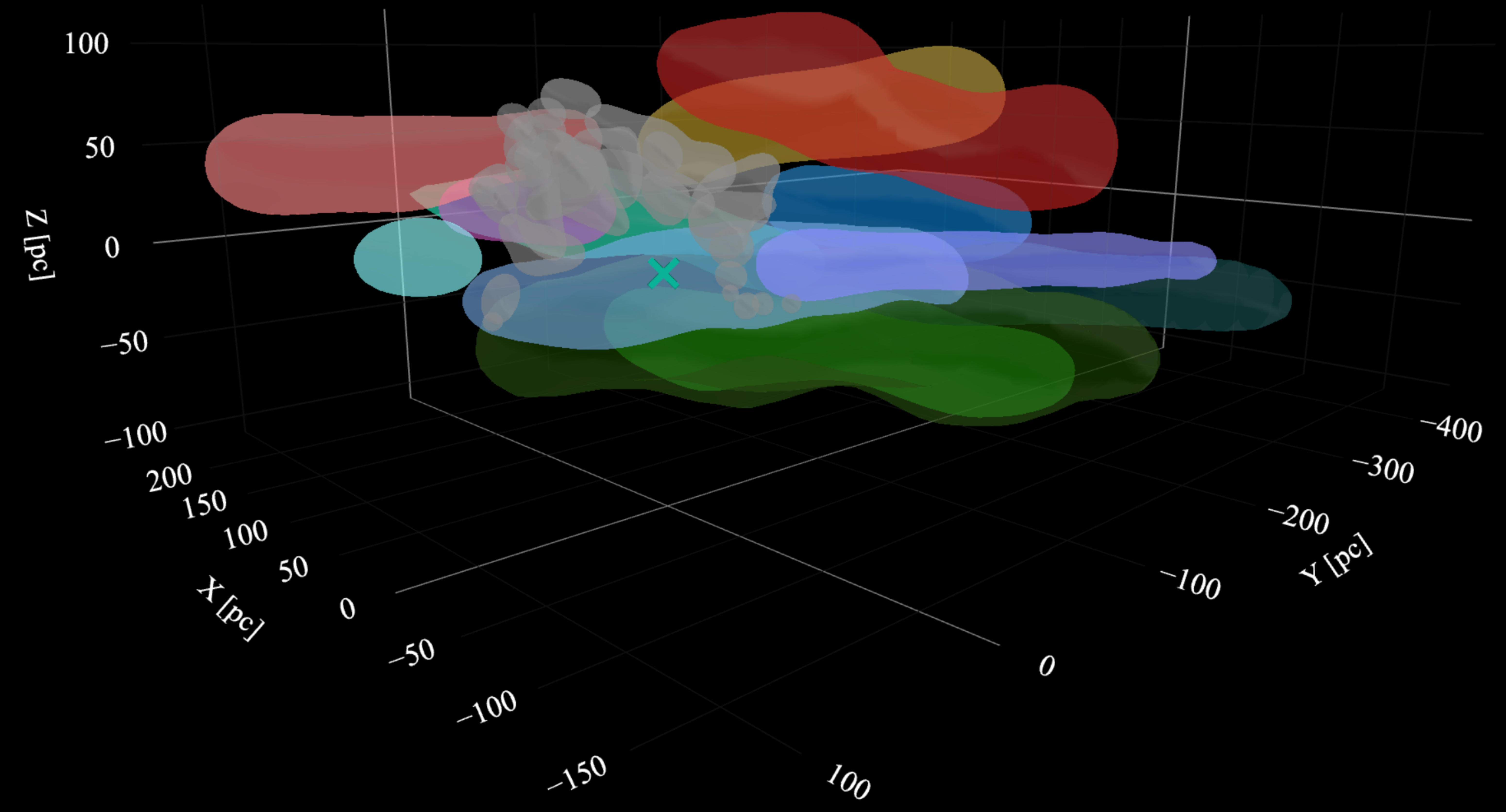
Further findings II

- Theia 368 may have undergone disruption from Sco-Cen
- Possibly interactions w/ primordial gas mass of the OB association



Results

- 12 disk stream candidates
- Lengths between ~200 – 400 pc
- Coeval & dynamically cold $\sigma_{3D} = 2 - 7\,\mathrm{km\,s^{-1}}$
- Densities as low as 0.2 stars / $10^3\,\mathrm{pc}^3$
- **820 streams /** kpc^3



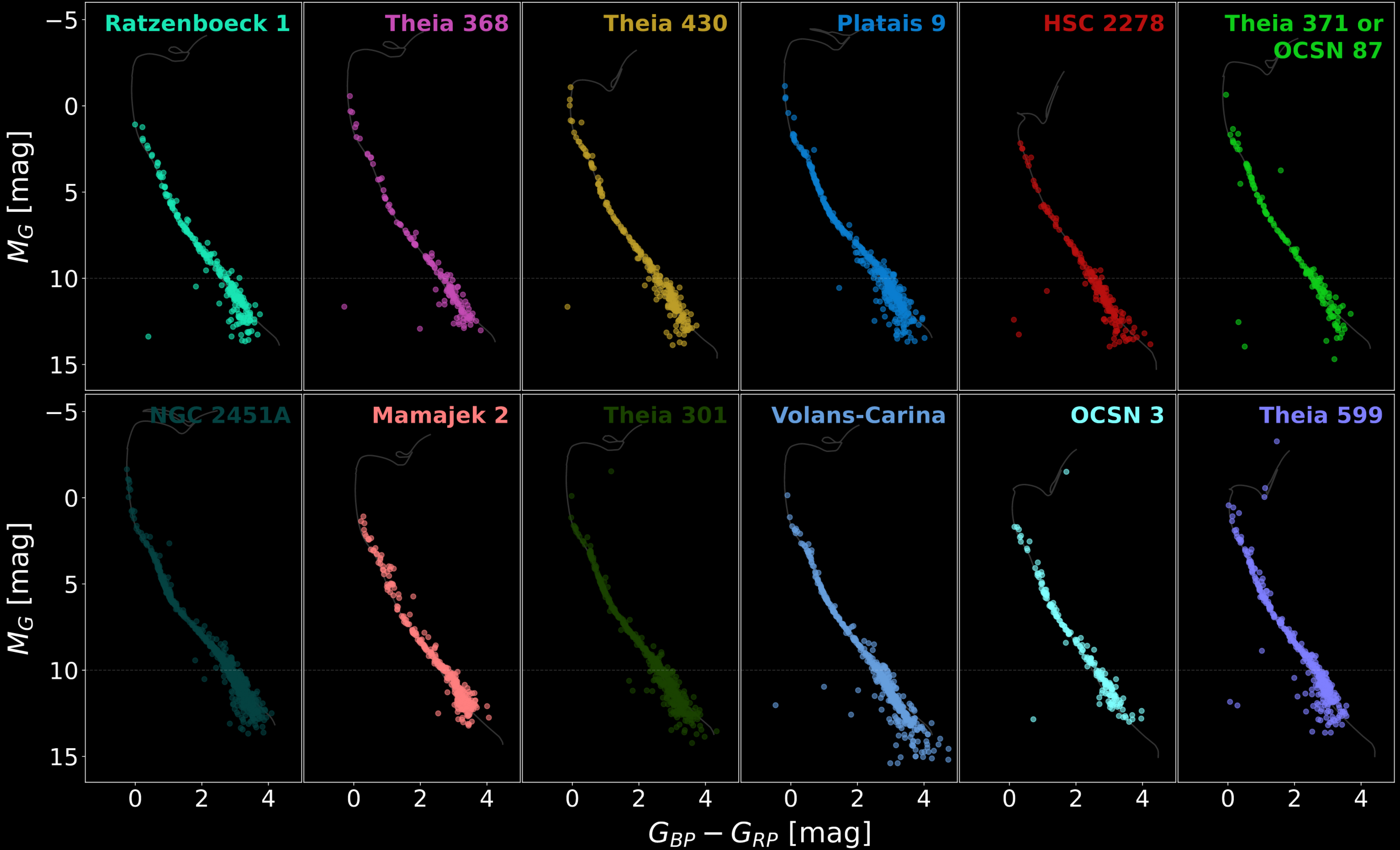
Thank you!

Backup

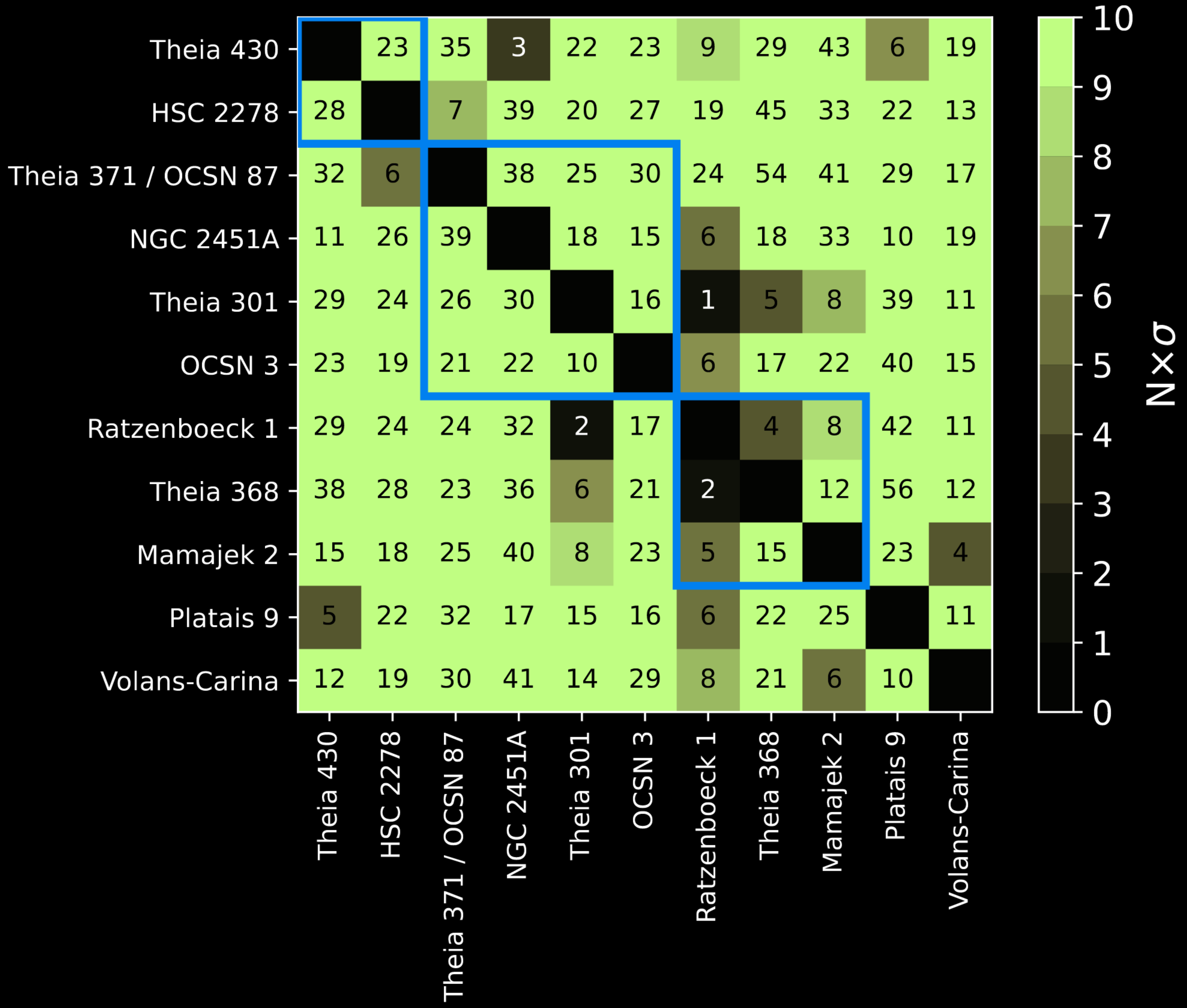
Set-up by Kamdar et al. (2019)

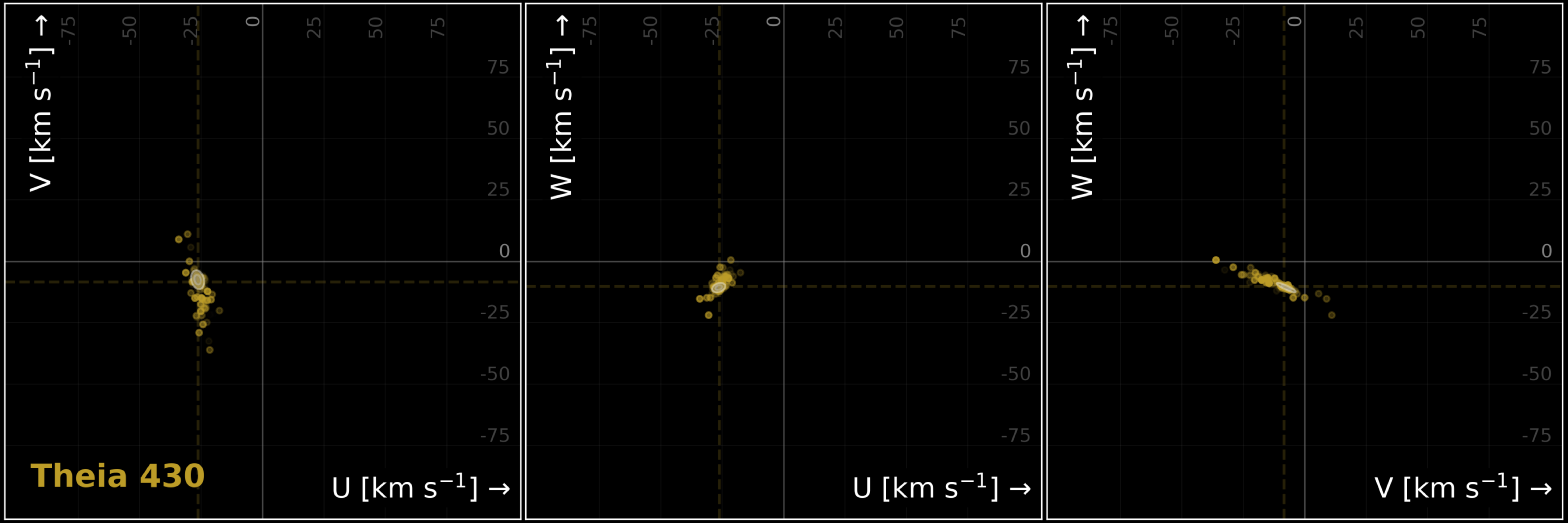
Star-by-star model evolving 4 billion stars over the last 5 Gyr.

- Time-varying Galactic potential w/ bar (two-armed) spiral structure
- Live GMCs & star formation (observationally constrained)
  - Cluster formation efficiency
  - Initial boundedness / virial state
  - Gas expulsion and disruption
  - >5 Gyr phase-mixed into smooth background
  - Gaia DR2 & End-of-Mission (EOM) uncertainties

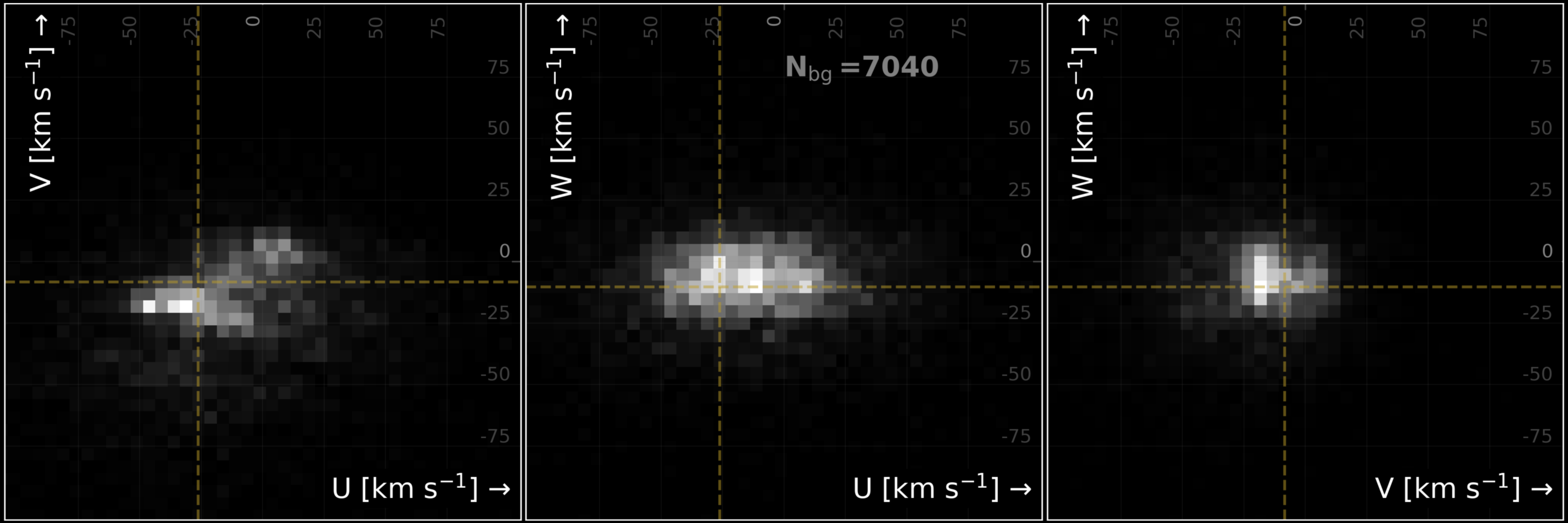


Ratzenboeck 1
Theia 368
Theia 430
Platais 9
HSC 2278
Theia 371 or OCSN 87
NGC 2451A
Mamajek 2
Theia 301
Volans-Carina
OCSN 3
Theia 599
M_G [mag]
$G_{BP} - G_{RP}$ [mag]

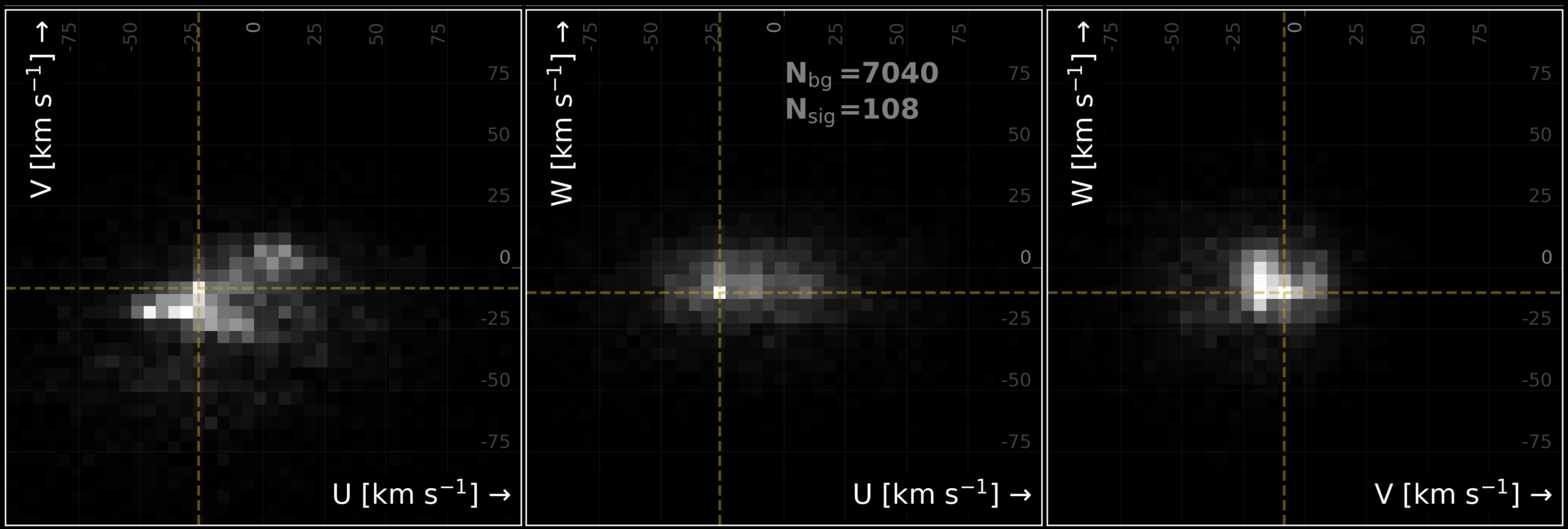

Theia 430
HSC 2278
Theia 371 / OCSN 87
NGC 2451A
Theia 301
OCSN 3
Ratzenboeck 1
Theia 368
Mamajek 2
Platais 9
Volans-Carina
23 35 3 22 23 9 29 43 6 19
28 7 39 20 27 19 45 33 22 13
32 6 38 25 30 24 54 41 29 17
11 26 39 18 15 6 18 33 10 19
29 24 26 30 16 1 5 8 39 11
23 19 21 22 10 6 17 22 40 15
29 24 24 32 2 17 4 8 42 11
38 28 23 36 6 21 2 12 56 12
15 18 25 40 8 23 5 15 23 4
5 22 32 17 15 16 6 22 25 11
12 19 30 41 14 29 8 21 6 10
10
9
8
7
6
5
4
3
2
1
0
N×σ



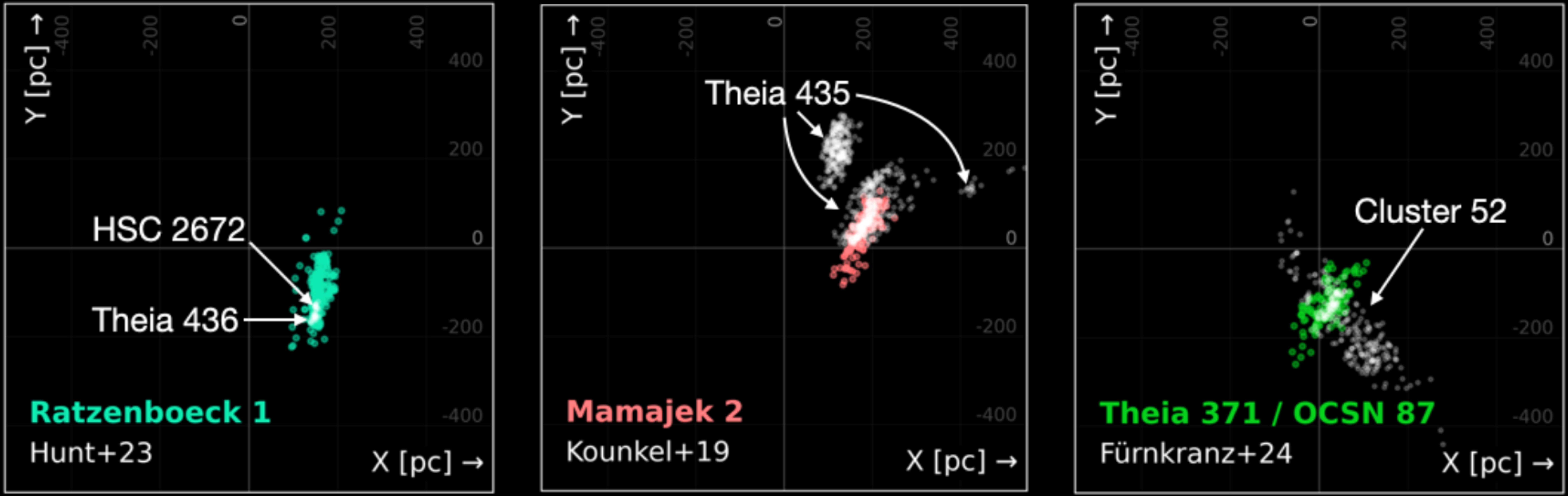
Theia 430
V [km s−1] →
W [km s−1] →
U [km s−1] →
V [km s−1] →
-75
-50
-25
0
25
50
75



V [km s^{-1}] →
U [km s^{-1}] →
W [km s^{-1}] →
U [km s^{-1}] →
N_{bg} =7040
W [km s^{-1}] →
V [km s^{-1}] →
-75
-50
-25
0
25
50
75



N_{bg} =7040
N_{sig} =108
V [km s^{-1}] →
U [km s^{-1}] →
W [km s^{-1}] →
U [km s^{-1}] →
W [km s^{-1}] →
V [km s^{-1}] →



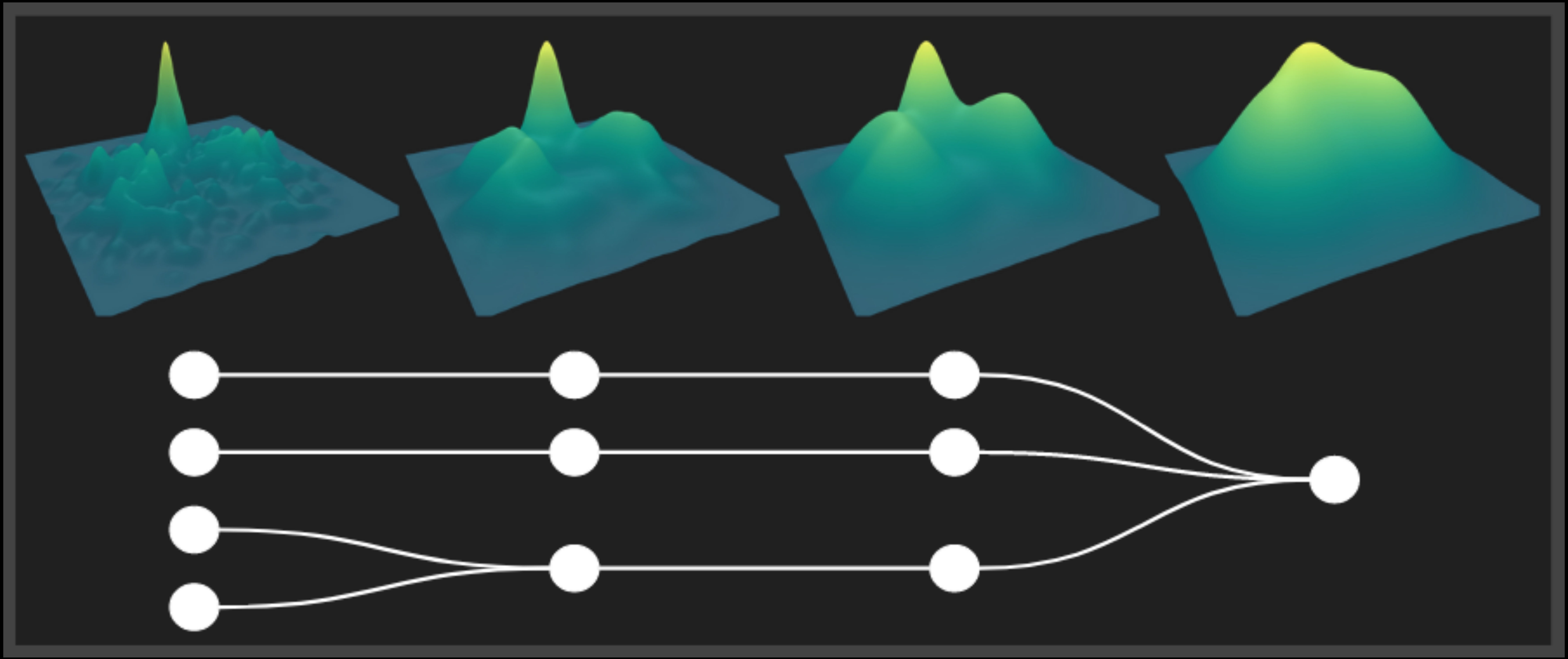
Y [pc] →
X [pc] →
HSC 2672
Theia 436
Ratzenboeck 1
Hunt+23
Theia 435
Mamajek 2
Kounkel+19
Cluster 52
Theia 371 / OCSN 87
Fürnkranz+24

How to set parameters?

`SigMA(k,`α`)`

Choosing k

$$\hat{T}_n(t) \sim \mathcal{N}(0,1) \iff \log N < k < N^{4/(4+p)}$$



Choosing α

- Many hypotheses tests increases chance of false positives
- Limit proportion of **false positives** among all positives
  - Apply **Benjamini & Hochberg procedure**

➡ Data driven way of choosing significance α

Problem #1

Distance metric

Distance metric

- Mixed meaning of dimensions
- 3 positional features
  - In Cartesian space
- 2 velocity features
  - Measurements “on sky”
  - spherical coordinates

Problem #2

Uncertainties

Time complexity

Density computation
(k-d tree)

mode & saddle
search (union find)

$$\mathcal{O}(pN\log N) + \mathcal{O}(pN\log N) + \mathcal{O}(Nk) + \mathcal{O}(|\mathcal{S}|)$$

Graph construction

Cluster tree
pruning

Robustness of $\hat{T}_n(t)$