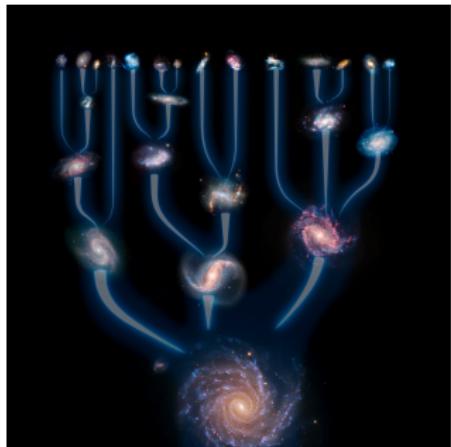


Unveiling the Milky Way's Formation History: Resolving Chemo-Dynamical Substructures in APOGEE and GALAH

Jacob Tutt¹

¹Department of Physics, University of Cambridge, UK

Galactic Archaeology's Motivation



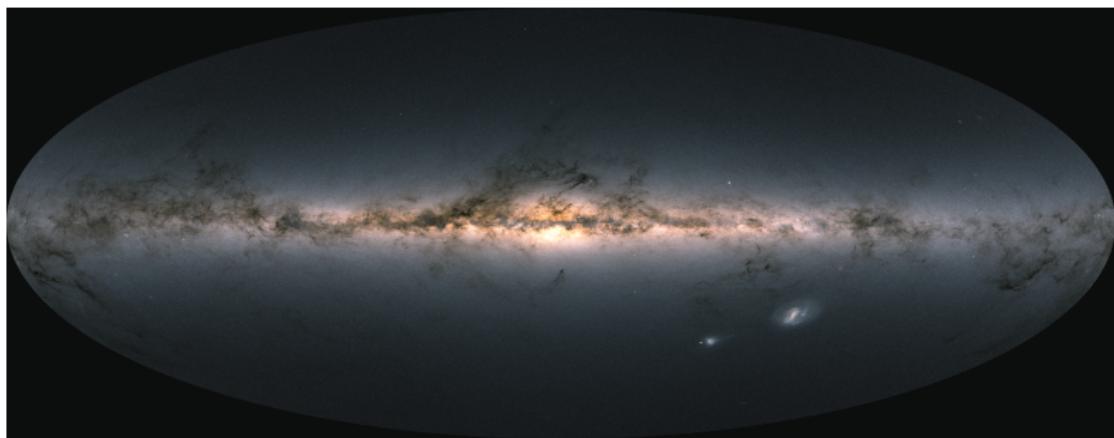
Credit: ESO/L. Calçada

- ▶ Uncovering hierarchical galaxy formation.
- ▶ Complements higher redshift galaxy formation surveys.
- ▶ Probe the Λ CDM model and dark matter distribution.

| Merger Type | Number | Mass Ratio |
|---------------|--------|-------------|
| Minor Mergers | ~30 | 1:3 – 1:100 |
| Major Mergers | ~3 | >1:3 |

N-Body Simulations from Fakhouri et al. (2010)

The Era of ‘Big Data Astronomy’



Credit: ESA/Gaia/DPAC, A. Moitinho.

| Survey | Gaia EDR3 | APOGEE | GALAH |
|----------------|------------------------|-----------------|----------------------|
| Focus | Astrometry, Photometry | IR spectroscopy | Optical spectroscopy |
| Sources | $\sim 1.4 \times 10^9$ | $\sim 734,000$ | $\sim 918,000$ |

The Evolving Picture: The GS/E Merger

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS 478, 611–619 (2018)

doi:10.1093/mnras/kyy982

Co-formation of the disc and the stellar halo*

V. Belokurov,^{1,2} D. Erkal,^{1,3} N. W. Evans,¹ S. E. Koposov,^{1,4} and A. J. Deason⁵

¹Institute of Astronomy, Madingley Road, Cambridge CB3 0HA

²Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

³Department of Physics, University of Surrey, Guilford GU2 7XH

⁴Department of Physics, McWilliams Center for Cosmology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

⁵Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

Accepted 2018 April 17. Received 2018 April 16; in original form 2018 February 9



ABSTRACT

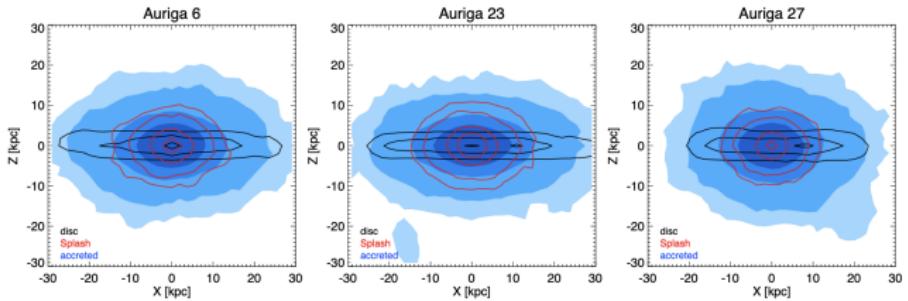
Using a large sample of main sequence stars with 7D measurements supplied by *Gaia* and SDSS, we study the kinematic properties of the local (within ~ 10 kpc from the Sun) stellar halo. We demonstrate that the halo's velocity ellipsoid evolves strongly with metallicity. At the low-[Fe/H] end, the orbital anisotropy (the amount of motion in the radial direction compared with the tangential one) is mildly radial, with $0.2 < \beta < 0.4$. For stars with $[\text{Fe}/\text{H}] > -1.7$, however, we measure extreme values of $\beta \sim 0.9$. Across the metallicity range considered, namely $-0.5 < [\text{Fe}/\text{H}] < -1$, the stellar halo's spin is minimal, at the level of $20 < \tilde{\epsilon}_3(\text{km s}^{-1}) < 30$. Using a suite of cosmological zoom-in simulations of halo formation, we deduce that the observed acute anisotropy is inconsistent with the continuous accretion of dwarf satellites. Instead, we argue, the stellar debris in the inner halo was deposited in a major accretion event by a satellite with $M_{\text{ej}} > 10^8 M_{\odot}$ around the epoch of the Galactic disc formation, between 8 and 11 Gyr ago. The radical halo anisotropy is the result of the dramatic radialization of the massive progenitor's orbit, amplified by the action of the growing disc.

Key words: galaxies: dwarf—Local Group—galaxies: structure.

Credit: V. Belokurov et al. (2018)

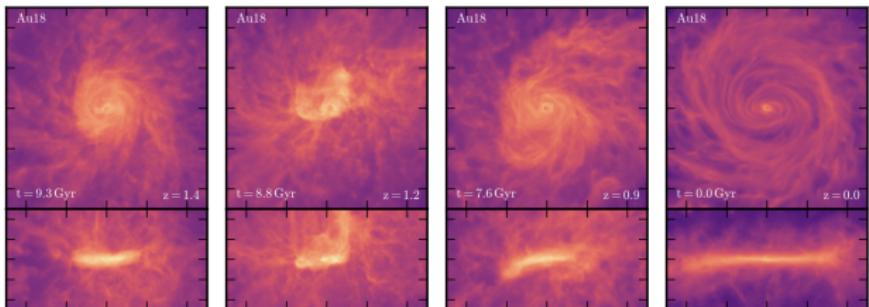
The Evolving Picture: Downstream Effects

Splash



Credit: V. Belokurov et al. (2020)

Eos



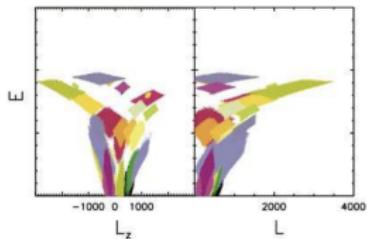
Credit: R. Grand et al. (2020)

Unbiased Detection of Halo Substructures

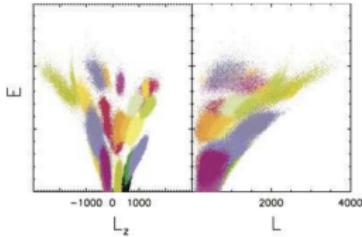
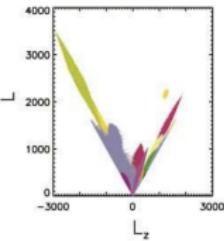
Goal:

- ▶ Unbiased Decomposition of the Milky Way's Halo's (stellar neighbourhood) Substructures.
- ▶ **Approach:**
 - ▶ Ensure the reproducibility of Myeong et al. (2022).
 - ▶ Apply dimensionality reduction to provide insights into clustering stability.
 - ▶ Alternative clustering approaches to improve the convergence and computational efficiency.

Integral of Motion Space



(a) Initial distribution of simulated merger events



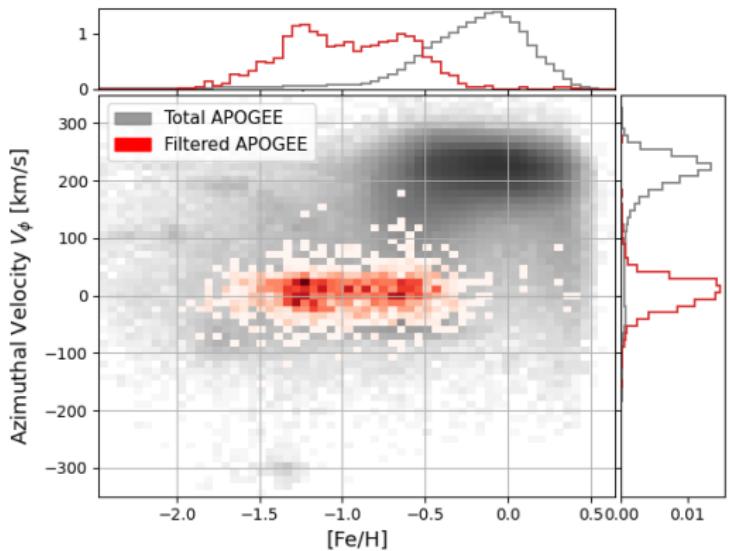
(b) Distribution after 12 Gyr (with observational errors)

Credit: Helmi et al. (2000)

For axisymmetric potentials:

| Symbol | Description |
|--------|----------------------------------------------|
| E | Orbital energy |
| L_z | Angular momentum (along z -axis) |
| L | Angular momentum (Total, quasi-conserved) |

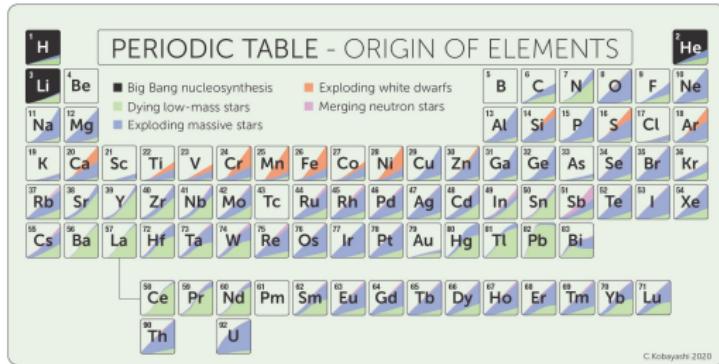
Biases of Halo Selection



Dynamical Cuts:

- ▶ Eccentricity, $e > 0.8$
- ▶ Apocenter, $> 5\text{kpc}$
- ▶ Energy, $< 0\text{km}^2\text{s}^{-2}$

Chemical Tagging



Credit: Kobayashi et al. (2020)

| Group | Elements | Traces |
|--------------------|----------------|------------------------------------------|
| Iron-peak | Fe, Mn, Ni | Overall Metallicity - Type Ia and II SNe |
| α -elements | Mg, Si, Ca, Ti | Core-collapse (Type II) SNe |
| Odd-Z elements | Na, Al | Similar Core-collapse (Type II) SNe |
| s-process | Y, Ba, Ce | Slow neutron capture - AGB stars |
| r-process | Eu | Neutron star mergers/ Rare CC-SNe |

Data Acquisition

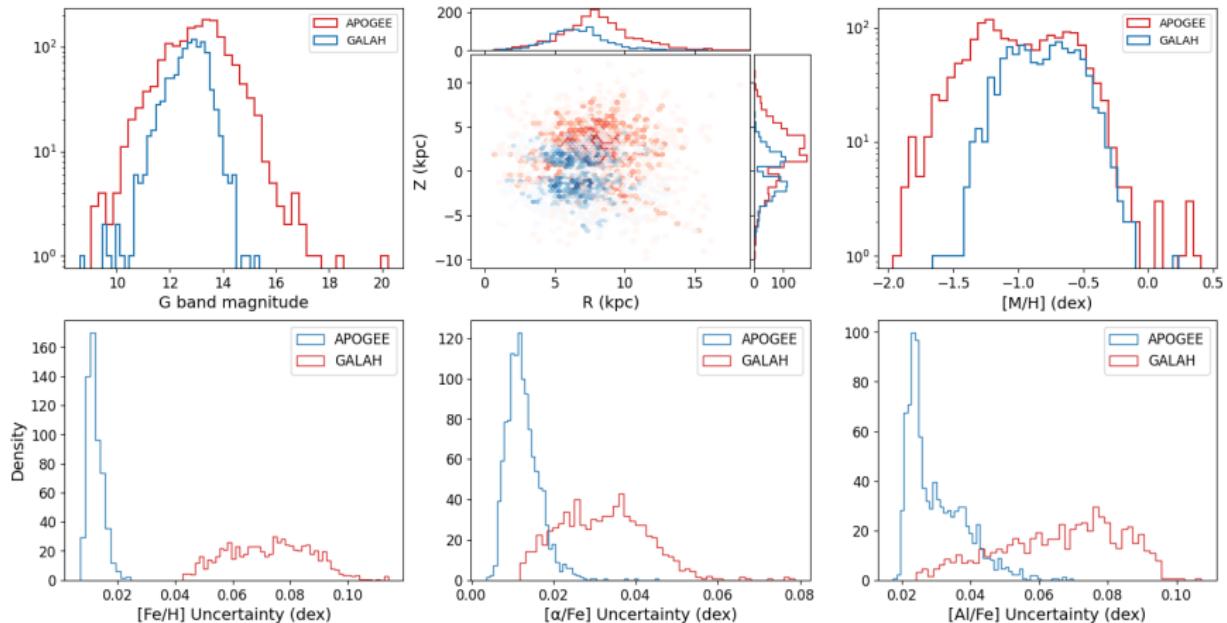
APOGEE

- ▶ **Sample Size:** 1612
- ▶ **Dimensions:** Energy, [Fe/H], [α /Fe], [Al/Fe], [Ce/Fe], [Mg/Mn]

GALAH

- ▶ **Sample Size:** 1061
- ▶ **Dimensions:** Energy, [Fe/H], [α /Fe], [Na/Fe], [Al/Fe], [Mn/Fe], [Y/Fe], [Ba/Fe], [Eu/Fe], [Mg/Cu], [Mg/Mn], [Ba/Eu]

Dataset Comparison



Extreme Deconvolution

The latent distribution of true values \mathbf{v} is modeled as a mixture of K Gaussians:

$$p(\mathbf{v}) = \sum_{j=1}^K \alpha_j \mathcal{N}(\mathbf{v} | \mathbf{m}_j, \mathbf{V}_j), \quad (1)$$

Likelihoods of noisy observations \mathbf{w}_i are computed by marginalising over \mathbf{v} :

$$p(\mathbf{w}_i | \theta) = \sum_j \int d\mathbf{v} p(\mathbf{w}_i | \mathbf{v}) p(\mathbf{v}|j, \theta) p(j|\theta). \quad (2)$$

Where:

$$p(\mathbf{w}_i | \mathbf{v}) = \mathcal{N}(\mathbf{w}_i | \mathbf{v}, \mathbf{S}_i), \quad (3)$$

$$p(\mathbf{v}|j, \theta) = \mathcal{N}(\mathbf{v} | \mathbf{m}_j, \mathbf{V}_j), \quad (4)$$

$$p(j|\theta) = \alpha_j. \quad (5)$$

Credit: Bovy et al (2011)

Extreme Deconvolution

As a result, the total likelihood of \mathbf{w}_i simplifies to another mixture of Gaussians:

$$p(\mathbf{w}_i | \theta) = \sum_j \alpha_j \mathcal{N}(\mathbf{w}_i | \mathbf{m}_j, \mathbf{T}_{ij}), \quad (6)$$

where the effective covariance \mathbf{T}_{ij} accounts for both the Gaussian component and the measurement uncertainty:

$$\mathbf{T}_{ij} = \mathbf{V}_j + \mathbf{S}_i. \quad (7)$$

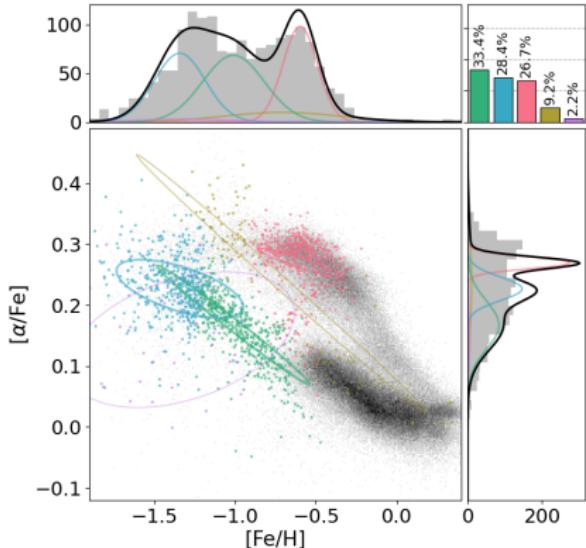
The log-likelihood across all N data points becomes:

$$\phi = \sum_i \ln p(\mathbf{w}_i | \theta) = \sum_i \ln \sum_{j=1}^K \alpha_j \mathcal{N}(\mathbf{w}_i | \mathbf{m}_j, \mathbf{T}_{ij}). \quad (8)$$

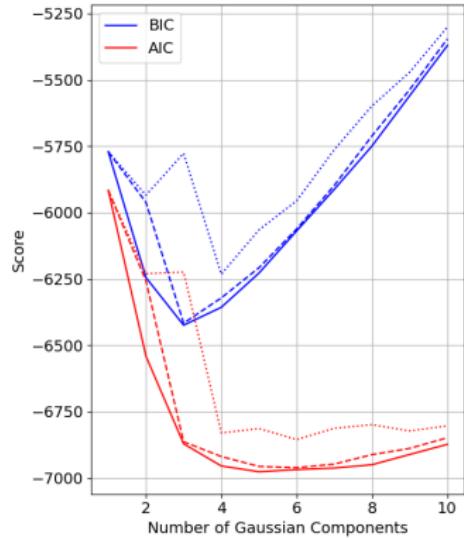
Credit: Bovy et al (2011)

Extreme Deconvolution Pipeline

Motivation of Scaling Scheme

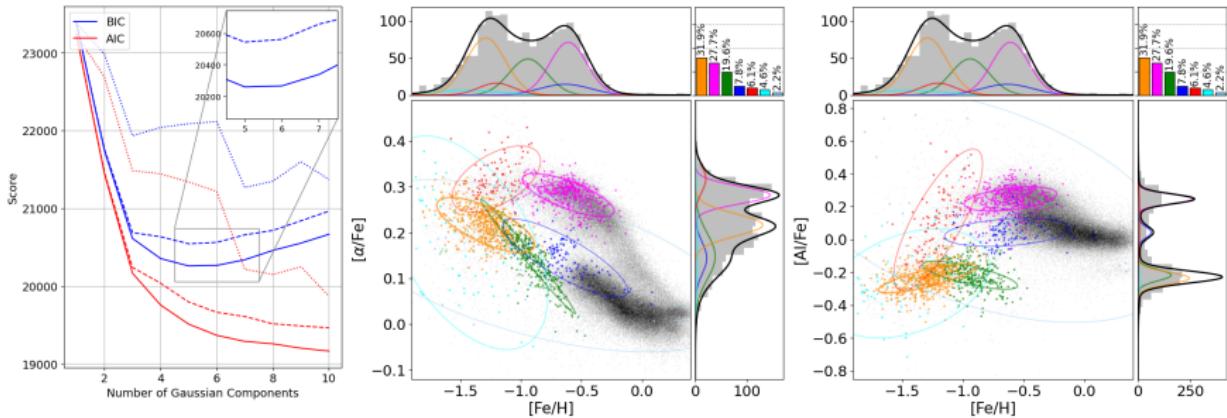


Unscaled XD Clustering Results in $[Fe/H]$ -
 $[\alpha/Fe]$ Plane



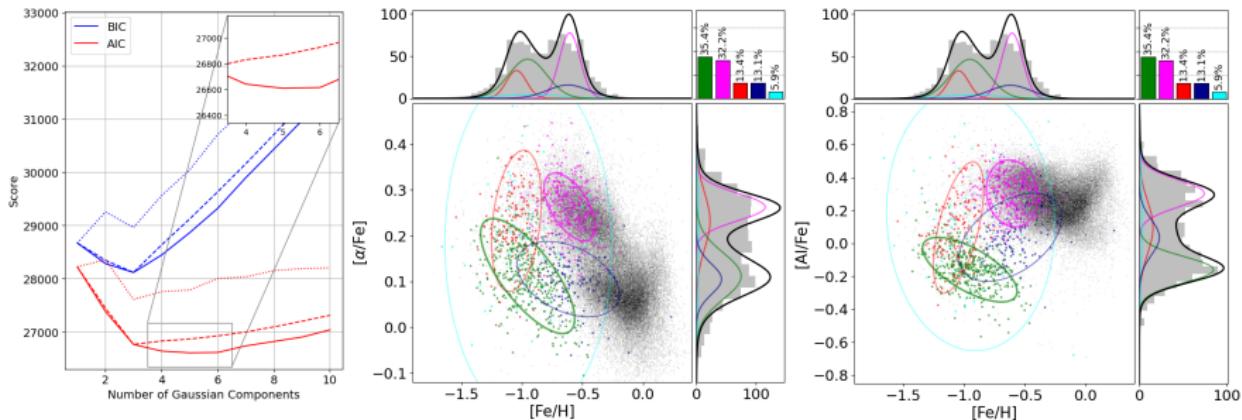
Unscaled XD Clustering
Quantitative Model Comparison

Extreme Deconvolution - APOGEE



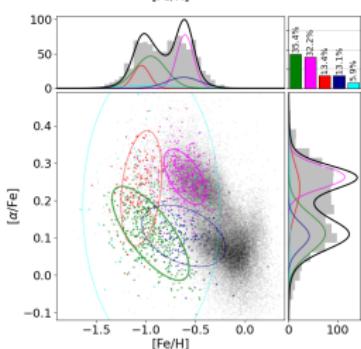
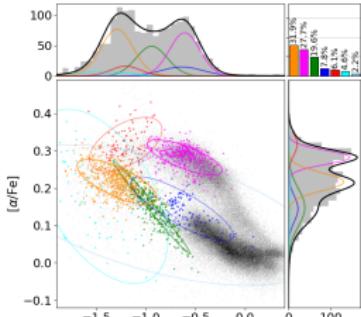
- Subtle BIC Discrepancy with original work (favouring 5 over 7)
- ‘Loss’ of Aurora Detection in 5 Component Model
- Trivial differences between GS/E split

Extreme Deconvolution - GALAH



- ▶ AIC providing correct isolation of 5 gaussian components
- ▶ ‘Exact’ agreement with original work

Key Scientific Recoveries



- ▶ Singular Accreted Component
 - ▶ GS/E - Less Efficient Progenitor
- ▶ Eos Population
 - ▶ Bridges GS/E and thin disk
- ▶ Aurora's Rapid Evolution

Dimensionality Reduction

Goals:

- ▶ Investigate the cohesion and isolation of the substructures identified during high-dimensional clustering
- ▶ Understand sensitivity of Aurora's detection.

Dimensionality Reduction

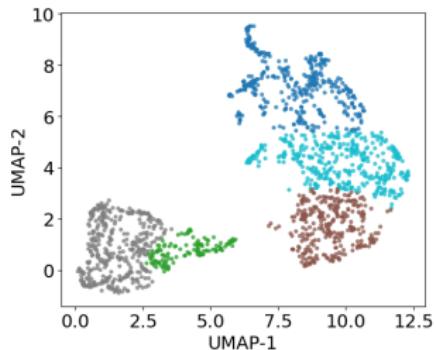
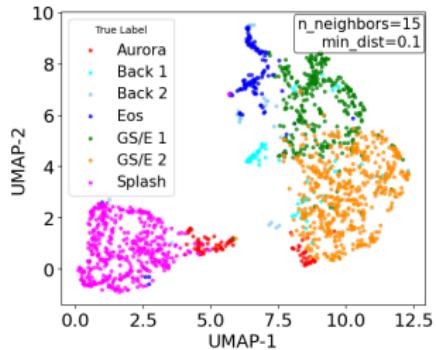
Goals:

- ▶ Investigate the cohesion and isolation of the substructures identified during high-dimensional clustering
- ▶ Understand sensitivity of Aurora's detection.

UMAP:

- ▶ Non-Linear dimensionality reduction
- ▶ Advantages over T-SNE:
 - ▶ Increased speed
 - ▶ Better preservation of global structure
- ▶ Hyperparameters:
 - ▶ `n_neighbours`: Balance of the local and global structure
 - ▶ `min_dist`: Controls lower dimensional projection

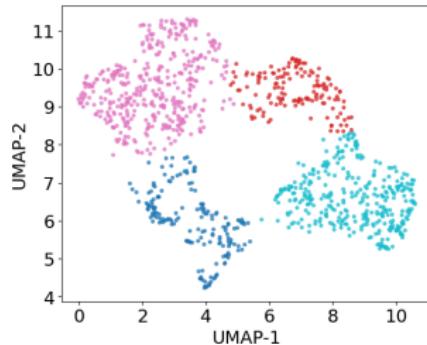
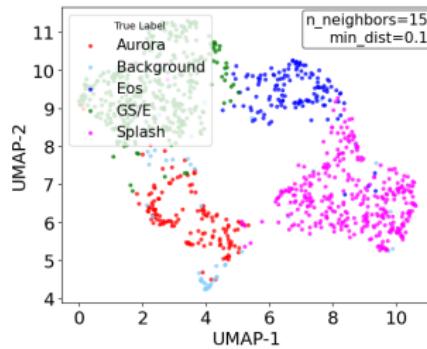
UMAP - 6D APOGEE



Key Results:

- ▶ Proof of concept
- ▶ A split in Aurora?
- ▶ GMM used to demonstrate cohesion and isolation

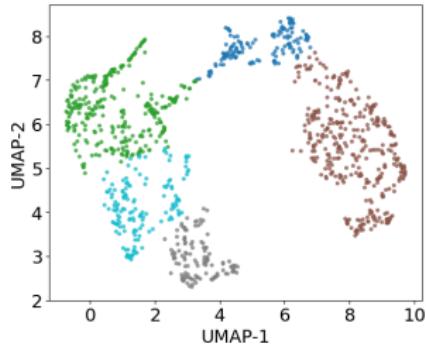
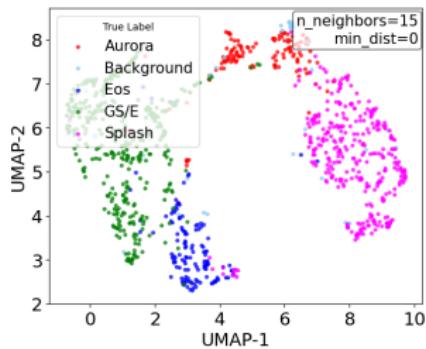
UMAP - 12D GALAH



Key Results:

- ▶ Near ‘perfect’ GMM recovery
- ▶ Greater Cohesion and Isolation:
 - ▶ 12D → Nucleosynthetic discrimination
 - ▶ Is this the only reason ?

UMAP - 6D GALAH



Key Results:

- ▶ Still Greater Cohesion and Isolation
- ▶ Attributed to:
 - ▶ Exclusion of low metalicity
 - ▶ Region of 'confusion' with GS/E

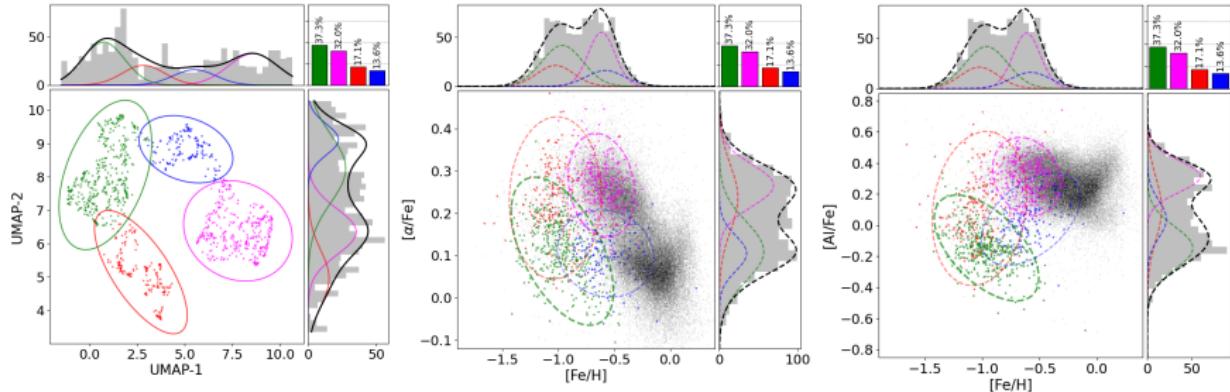
Scalable Methodologies

A Low-Dimensional Pipeline:

- ▶ Dimensionality Reduction
- ▶ Clustering in Embedding Space
- ▶ Re-projection back into Original Space
- ▶ Approximate Devolution:

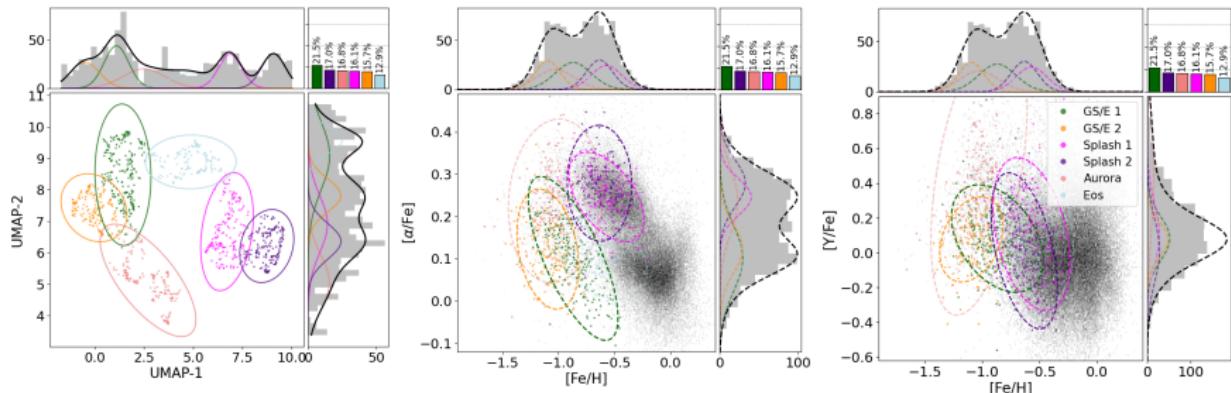
$$\boldsymbol{\Sigma}_{\text{intr}} = \boldsymbol{\Sigma}_{\text{obs}} - \langle \boldsymbol{\Delta} \rangle \quad (9)$$

4 Component Re-identification



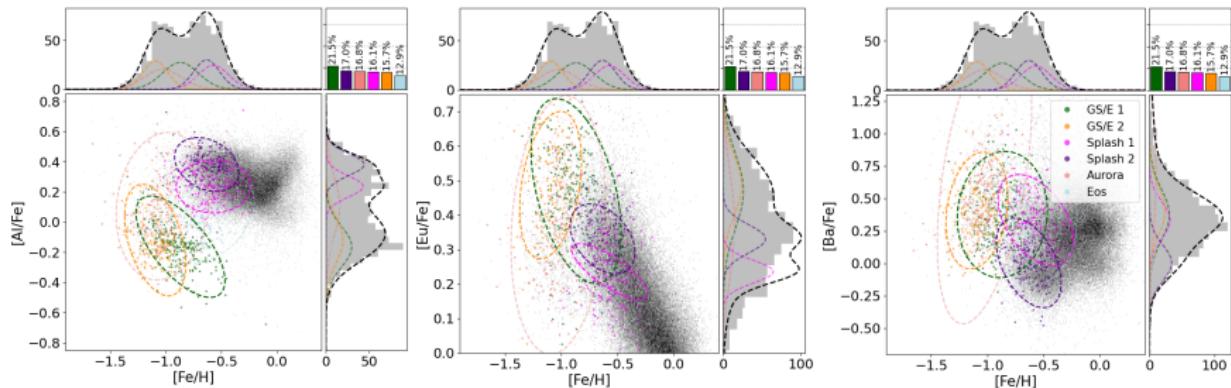
- ▶ 4.7 hours → 7 seconds ($2500 \times$ Speed-Up)
- ▶ Entirely Consistent Results
- ▶ 29.3% Increase in Uncertainties

6 Component Re-identification



- ▶ A rough recovery of the GS/E Split
- ▶ Not achievable in high dimensional clustering

6 Component Re-identification



- ▶ A potential split in splash?
- ▶ A simple division of a large components in embedding space
- ▶ Or ... an astrophysical distinction

Key Results

1. **Recovered key populations:** Confirming objectivity and reproducibility.
2. **GALAH's higher dimensionality:** Provides greater halo substructure separation despite higher uncertainties.
3. **Clustering in Embedding space:**
 - ▶ Near-identical results to high-dimensional analysis
 - ▶ Achieved in 0.04% of the time (29% higher uncertainty)
 - ▶ A future stable and scalable alternative

Future Work

1. **Aurora's Split:** Test for bimodality in Aurora structure
2. **Splash's Split:** Explore physical basis for two Splash subpopulations (simulations)
3. **Hybrid Pipeline:**
 - ▶ Fast (low-D) clustering for initial grouping
 - ▶ Uncertainty-aware (high-D) clustering for accuracy

Questions

Thank you for your attention!

Jacob Tutt

Department of Physics, University of Cambridge

jlt67@cam.ac.uk

<https://github.com/jacbtutt>

Extreme Deconvolution

Expectation-step

$$q_{ij} = \frac{\alpha_j \mathcal{N}(\mathbf{w}_i | \mathbf{m}_j, \mathbf{T}_{ij})}{\sum_k \alpha_k \mathcal{N}(\mathbf{w}_i | \mathbf{m}_k, \mathbf{T}_{ik})} \quad (10)$$

Maximisation-step

$$\alpha_j = \frac{1}{N} \sum_i q_{ij} \quad (13)$$

$$\mathbf{b}_{ij} = \mathbf{m}_j + \mathbf{V}_j \mathbf{T}_{ij}^{-1} (\mathbf{w}_i - \mathbf{m}_j) \quad (11)$$

$$\mathbf{m}_j = \frac{1}{q_j} \sum_i q_{ij} \mathbf{b}_{ij} \quad (14)$$

$$\mathbf{B}_{ij} = \mathbf{V}_j - \mathbf{V}_j \mathbf{T}_{ij}^{-1} \mathbf{V}_j \quad (12)$$

$$\mathbf{V}_j = \frac{1}{q_j} \sum_i q_{ij} \left[(\mathbf{m}_j - \mathbf{b}_{ij})(\mathbf{m}_j - \mathbf{b}_{ij})^\top + \mathbf{B}_{ij} \right] \quad (15)$$

Model Comparison

Akaike Information Criterion

- ▶ Favors models with best predictive accuracy

Bayesian Information Criterion

- ▶ Favors models with best overall fit

$$\text{AIC} = 2k - 2 \ln \mathcal{L},$$

$$\text{BIC} = k \ln n - 2 \ln \mathcal{L},$$

where:

- ▶ k : number of free parameters,
- ▶ n : number of data points,
- ▶ \mathcal{L} : maximum likelihood of the model.

UMAP Algorithm:

1. Compute local distances

- ▶ For each point, find distance to the n-th nearest neighbor
($n_{\text{neighbours}}$)

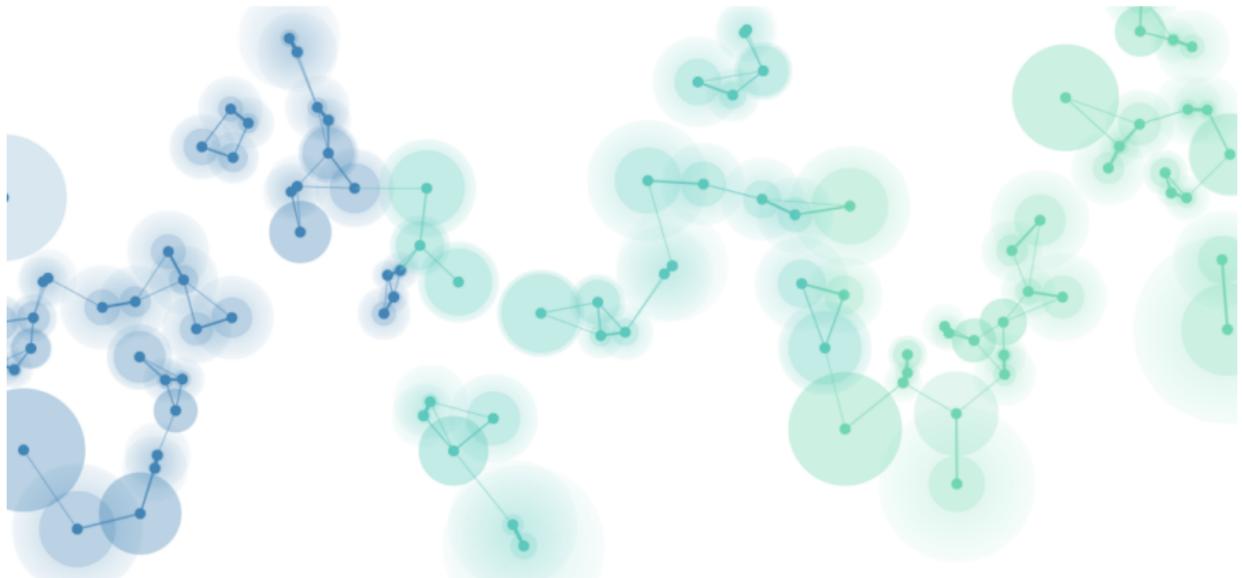
2. Construct Representation

- ▶ Build a weighted graph representing connection probabilities
- ▶ Done using local radii (scaled by nth nearest neighbor)
- ▶ Ensures mutual relationships are captured

3. Optimise low-dimensional embedding

- ▶ Initialise points in low-dimensional space (`min_dist`)

UMAP Visualisation



Splash Decomposition

| Feature | Splash 1 | Splash 2 | Tracer |
|----------|----------|----------|------------------|
| Colour | Magenta | Purple | |
| Fraction | 16.1% | 17.0% | |
| [Eu/Fe] | Lower | Higher | r-process |
| [Al/Fe] | Lower | Higher | Core-collapse SN |
| [Ba/Fe] | Higher | Lower | s-process (AGB) |

Table: Comparison of chemical properties between Splash 1 and Splash 2 populations.