

StarStream: Automatic detection algorithm for stellar streams

YINGTIAN CHEN ¹, OLEG Y. GNEDIN ¹, ADRIAN M. PRICE-WHELAN ², AND COLIN HOLM-HANSEN ¹

¹Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

²Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA

Abstract

The *Gaia* mission has led to the discovery of over 100 stellar streams in the Milky Way, most of which likely originated from globular clusters (GCs). As the upcoming wide-field surveys can potentially continue to increase the number of known streams, there is a growing need to shift focus from manual detection of individual streams to automated detection methods that prioritize both quality and quantity. Traditional techniques rely heavily on the visual expectation that GC streams are dynamically cold and thin. This assumption does not hold for all streams, whose morphologies and kinematics can vary significantly with the progenitor’s mass and orbit. As a result, these methods are biased toward a subset of the whole stream population, with often unquantified purity and completeness. In this work, we present *StarStream*, an automatic stream detection algorithm based on a physics-inspired model rather than visual expectation. Our method provides a more accurate prediction of stream stars in the multi-dimensional space of observables, while using fewer free parameters to account for the diversity of streams. Applied to a mock GC stream catalog tailored for the *Gaia* DR3 dataset, our algorithm achieves both purity and completeness of at least 65% at Galactic latitudes $|b| > 30^\circ$.

Keywords: [Stellar streams \(2166\)](#); [Globular star clusters \(656\)](#); [Stellar dynamics \(1596\)](#); [Galaxy dynamics \(591\)](#)

1. Introduction

Stellar streams are elongated tidal structures originating from either an existing or fully dissolved progenitor system, such as a globular cluster (GC) or a dwarf galaxy (D. Lynden-Bell & R. M. Lynden-Bell 1995). Compared to the high background density of Milky Way (MW) field stars, streams have an extremely low signal-to-noise ratio (S/N). As a result, only a few streams were identified prior to the past decade, including the Sagittarius stream by R. A. Ibata et al. (1994) and the Palomar 5 (Pal 5) stream by M. Odenkirchen et al. (2001). These pioneering efforts attempted to increase S/N by selecting a subset of stars that are more likely to belong to streams than to the field, and then visually searching for stream-like structures. This selection was typically performed by applying a matched filter in the color–magnitude diagram, usually a window function centered around the progenitor’s isochrone (e.g., C. M. Rockosi et al. 2002; C. J. Grillmair 2009; E. J. Bernard et al. 2014; N. Shipp et al. 2018).

The launch of the *Gaia* mission (Gaia Collaboration et al. 2016) has revolutionized the discovery of stellar streams by providing an all-sky map of stars in the six-dimensional phase space, particularly offering high-precision proper motions down to $G \approx 20$ since Data Release 2 (DR2, Gaia Collab-

oration et al. 2018). This enables additional matched filters based on astrometric measurements, significantly increasing the number of detected streams (see the review by A. Bonaca & A. M. Price-Whelan 2025). The all-sky coverage and homogeneity of the *Gaia* data also motivate the development of automatic stream detection methods to replace visual inspection (e.g., C. Mateu et al. 2011; D. Shih et al. 2021). One such method is STREAMFINDER (K. Malhan & R. A. Ibata 2018), which uses a mixture model in the multi-dimensional space of observables to automatically detect clusters of stellar orbits within a Gaussian tube. STREAMFINDER successfully identified 87 thin streams in *Gaia* Data Release 3 (DR3, Gaia Collaboration et al. 2023), including 28 new discoveries (R. Ibata et al. 2024).

As the number of stream detections grows, more evidence shows that streams have density structure such as *fans* (B. Sesar et al. 2016), *gaps* (D. Erkal et al. 2016), *spurs* (A. M. Price-Whelan & A. Bonaca 2018), and *cocoons* (K. Malhan et al. 2019; M. Valluri et al. 2024). These features are likely produced by perturbations in the host galaxy’s potential, including bar rotation (K. Hattori et al. 2016; A. M. Price-Whelan et al. 2016; S. Pearson et al. 2017), disk rotation (J. Nibauer et al. 2024), and close encounters with other objects (R. G. Carlberg et al. 2012; W. H. W. Ngan & R. G. Carlberg 2014; D. Erkal et al. 2016, 2017; N. Banik et al. 2018). On the other hand, recent works proposed that stream density can trace the mass loss history of their progenitors (M. Gieles et al. 2021; Y. Chen et al. 2025a). These breakthroughs

emphasize the need to quantify the purity and completeness of stream detection, in order to accurately characterize their density structures.

Previous studies have revealed density structures in individual streams using flexible density models (e.g., [D. Erkal et al. 2017](#); [K. Tavangar & A. M. Price-Whelan 2025](#)). However, these models involve many free parameters, making them computationally expensive when applied to all-sky data and better suited for precisely characterizing stream membership after discovery. Even the simpler model used in *STREAMFINDER* requires millions of CPU hours on *Gaia* Early Data Release 3 ([Gaia Collaboration et al. 2021](#)). This will greatly limit the usage of these methods for next-generation wide-field photometric surveys, such as those conducted by the *Vera C. Rubin Observatory* ([Rubin, LSST Science Collaboration et al. 2009](#)) and the *Nancy Grace Roman Space Telescope* ([Roman, D. Spergel et al. 2015](#)).

Furthermore, these models usually represent a stream as a tube surrounding a well-defined track, based on the visual expectation that streams are dynamically cold and thin. However, this assumption is inaccurate, as even GC streams can be dynamically hot or spatially complex depending on the progenitor’s mass and orbit ([N. C. Amorisco 2015](#)). As a result, such visually inspired models are inefficient at detecting more “irregular” streams.

Compared to the visually-inspired models mentioned above, recent theoretical advances in stream formation have enabled physics-inspired models that achieve higher accuracy with fewer free parameters. Specifically, given the host Galactic potential and the progenitor’s mass, position, and velocity, particle spray methods ([A. Varghese et al. 2011](#); [R. R. Lane et al. 2012](#); [A. H. W. Küpper et al. 2012](#); [A. Bonaca et al. 2014](#); [M. A. Fardal et al. 2015](#); [D. Roberts et al. 2025](#); [Y. Chen et al. 2025b](#)) can efficiently generate tracer particles that follow the expected distribution of stream stars in six-dimensional phase space, typically requiring only one or even zero free parameters. In particular, the method of [Y. Chen et al. \(2025b\)](#) is calibrated to match N-body simulations within 10% error for typical GC streams, without introducing any additional parameters. As a result, stream models based on this approach can potentially reduce computational cost by a significant amount, while also being able to detect hotter and wider streams.

In this work, we present *StarStream*, an automatic detection algorithm for stellar streams using a physics-inspired stream model based on [Y. Chen et al. \(2025b\)](#). We employ kernel density estimation (KDE) to construct smooth probability density functions (PDFs) for both the stream and background populations. The algorithm is applied to the mock dataset from [C. Holm-Hansen et al. \(2025, hereafter H25\)](#), tailored for *Gaia* DR3, to quantify its purity and completeness in detecting streams originating from existing GCs. The paper is

structured as follows. In §2, we describe the methodology of *StarStream* in detail. We then present validation tests on the mock dataset in §3. In §4, we discuss the motivation for applying the method to upcoming surveys (§4.1) and improvements over other algorithms (§4.2). Finally, we summarize our findings in §5.

2. Method

We distinguish stream members from the background stars using a mixture model, which is a powerful tool for identifying faint structures such as ultra-faint dwarf galaxies (e.g., [A. B. Pace & T. S. Li 2019](#)) and stellar streams (e.g., [K. Malhan & R. A. Ibata 2018](#); [K. Tavangar & A. M. Price-Whelan 2025](#)). Specifically, we construct the joint probability density function (PDF) of the stream and background populations as

$$p(\mathbf{x}) = f_s p_s(\mathbf{x}) + (1 - f_s) p_{bg}(\mathbf{x})$$

where $\mathbf{x}$ denotes a point in the multi-dimensional observable space, including positions, velocities, colors, magnitudes, and other properties. Traditional methods define PDFs of the stream $p_s(\mathbf{x})$ and the background $p_{bg}(\mathbf{x})$ as parametric functions with several fixed or adjustable parameters. The best-fit values of these parameters, together with the stream fraction f_s , are often estimated by maximizing the log-likelihood,

$$\ln \mathcal{L} \equiv \sum_{i=1}^N \ln [f_s p_{s,i} + (1 - f_s) p_{bg,i}] \quad (1)$$

where $p_{s,i} \equiv p_s(\mathbf{x}_i)$ and $p_{bg,i} \equiv p_{bg}(\mathbf{x}_i)$ are the probability densities of the i -th star being a member of the stream and background, respectively.

Since parameter estimation is exponentially more computationally expensive as the number of adjustable parameters grows, many methods tend to simplify the stream model by assuming it to be a thin tube along a predefined track. Similarly, the background model is often approximated as uniform or only slowly varying across observables. However, recent advances in the theory of GC stream formation now allow for more accurate stream modeling with even fewer parameters ([Y. Chen et al. 2025b](#)). Additionally, we can employ a nonparametric KDE to model the nonuniform background without introducing extra model parameters. In this work, we develop a new stream detection method by incorporating these improvements into the mixture model. In the following sections, we detail our approach for accurately estimating $p_s(\mathbf{x})$ and $p_{bg}(\mathbf{x})$.

2.1. Stream probability density

To approximate the PDF of streams, we first generate simulated streams around progenitor GCs using the particle spray algorithm. Specifically, we use the *agama* ([E. Vasiliev 2019](#))

implementation of the Y. Chen et al. (2025b) algorithm,³ which initializes the positions and velocities of stream tracer particles from a multivariate Gaussian distribution, calibrated using N-body simulations of disrupting GCs. This algorithm accurately reproduces the width and length of simulated streams across a wide range of cluster masses and orbital types.

To obtain the probability density in the color–magnitude space, we first assign a stellar mass to each tracer particle by drawing from the P. Kroupa (2001) initial mass function (IMF). Although the mass function (MF) may evolve due to energy equipartition that preferentially ejects low-mass stars, the high-mass end above *Gaia*’s detection limit ($\gtrsim 0.5 M_\odot$) remains largely consistent with the IMF (see H25). We then compute the colors and magnitudes using the MESA Isochrones and Stellar Tracks (MIST, A. Dotter 2016; U. Meštrić et al. 2022), taking the progenitor GC’s age and metallicity as input. For this study, we adopt *Gaia*’s *G* magnitude and BP – RP color. In §3.4, we also test an alternative isochrone model, PARSEC (A. Bressan et al. 2012), which has negligible effect on the final detection quality.

For each stream, we release tracer particles over the last 1 Gyr assuming a uniform ejection rate. The ejection rate can be set time-varying if needed. However, the uniform rate suffices to produce a realistic stream density distribution that is distinguishable from the background, and variations in the ejection rate only slightly affect the density along most streams (see §2.2 in Y. Chen et al. 2025a). We generate 4000 tracer particles per stream, which is sufficient to fully sample the multi-dimensional parameter space. The minimum stellar mass used in sampling the mass function is set to the lowest possible mass of the closest tracer particle that remains above the detection limit. Since the heliocentric distances of stream stars vary along the stream, this minimum mass is a conservative choice to ensure that all regions of the stream are well-sampled above the local detection limit.

We use Gaussian KDE to estimate the stream PDF from tracer particles in the M -dimensional parameter space, including positions, velocities, colors, magnitudes, and other properties. By denoting the M -dimensional coordinate of the j -th tracer as $\mathbf{x}_j \equiv (x_j^1, x_j^2, \dots, x_j^M)$, the probability density at any arbitrary point $\mathbf{x}$ is

$$\begin{aligned} p_s(\mathbf{x}) &\approx \frac{1}{N_{\text{tr}}} \sum_{j=1}^{N_{\text{tr}}} p_{\text{KDE}}(\mathbf{x}|\mathbf{x}_j, \boldsymbol{\sigma}) \\ &\equiv \frac{1}{N_{\text{tr}}} \sum_{j=1}^{N_{\text{tr}}} \prod_{k=1}^M \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{(x^k - x_j^k)^2}{2\sigma_k^2}\right] \end{aligned}$$

where N_{tr} is the total number of tracers. It is straightforward to verify that integrating $p_s(\mathbf{x})$ over the full M -dimensional space yields unity. We define $\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2, \dots, \sigma_M)$ as the array of KDE kernel bandwidths, with no correlation between each dimension. In practice, we set σ_k to 0.1 times the standard deviation of all tracer particles in the k -th dimension when k refers to positions or velocities. For magnitudes, we use $\sigma = 0.1$; and for colors, $\sigma = 0.02$. We have verified that varying these values by a factor of 0.5–2 has a negligible effect on our results.

Note that the KDE approach naturally captures the fact that most stream stars are faint, since we sample stream particle masses from the P. Kroupa (2001) IMF. As a result, more tracer particles are gathered toward the faint end, leading to higher probability densities in that region.

Before applying KDE to estimate the probability density, it is helpful to rotate the equatorial coordinate system so that the new latitude of the stream center is zero. This is particularly important for streams at high declination, where the metric tensor deviates significantly from identity. In this work, we always work in a rotated coordinate frame (ϕ_1, ϕ_2) , where the progenitor GC is located at $(0, 0)$ and the proper motion is in the positive ϕ_1 direction. In this case, the diagonal elements of the metric tensor $\mathbf{g} = \text{diag}(1, \cos^2 \phi_2)$ only deviate from those of the identity tensor by 3% even with $\phi_2 = 10^\circ$. This coordinate system is similar to the great circle frame commonly used to describe stellar streams, but is less ambiguous when the stream is wide or has strong curvature.

In practice, the observables of each star often have significant observational uncertainties, denoted by $\boldsymbol{\sigma}_0 \equiv (\sigma_{1,0}, \sigma_{2,0}, \dots, \sigma_{M,0})$, which may exceed the corresponding KDE kernel widths. As we show in Appendix A, the effective PDF for such a star is the convolution between the original PDF and a Gaussian kernel with standard deviations $\boldsymbol{\sigma}_0$. This convolution results in a modified KDE evaluated at the same location, with kernel width replaced by $\sigma_k'^2 \equiv \sigma_k^2 + \sigma_{k,0}^2$. Therefore, for a star with M -dimensional coordinates $\mathbf{x}$ and uncertainties $\boldsymbol{\sigma}_0$, the stream probability density becomes

$$p_s(\mathbf{x}, \boldsymbol{\sigma}_0) \approx \frac{1}{N_{\text{tr}}} \sum_{j=1}^{N_{\text{tr}}} p_{\text{KDE}}(\mathbf{x}|\mathbf{x}_j, \boldsymbol{\sigma}')$$

where $\boldsymbol{\sigma}'$ incorporates both the KDE kernel width and observational uncertainty.

For *Gaia* data specifically, the uncertainties in positions and magnitudes are almost always smaller than the corresponding KDE bandwidths. For simplicity and computational efficiency, we therefore ignore the uncertainties in these parameters in our subsequent analysis and consider only the uncertainties in proper motions and color.

It is worth noting that *Gaia* astrometric uncertainties have nonzero correlations. These correlations influence the error propagation from the original equatorial frame to the rotated

³ Tutorials for this algorithm are available at https://github.com/ybillchen/particle_spray and are preserved on Zenodo at Y. Chen (2024).

(ϕ_1, ϕ_2) frame. In Appendix B, we explicitly calculate the linear uncertainty propagation associated with this coordinate transformation. However, to simplify our calculations, we do not include these correlations when performing the convolution with the Gaussian kernel, allowing us to treat each dimension independently during KDE evaluation. This simplification has only a minor effect on the inferred probability density, as the correlations are generally weak ($r < 0.5$).

2.2. Background probability density

Similarly to the stream probability density, we also use Gaussian KDE to estimate the background PDF around a stream. We directly use the observed stars in the same spatial region as the stream to construct the KDE estimator. However, there are typically 10^7 observed stars in these regions. Constructing and evaluating a multi-dimensional Gaussian KDE with so many data points can be extremely inefficient. To address this, we perform an initial selection around the isochrone at the progenitor’s distance to exclude stars that are too red or too blue to be plausible stream members. For *Gaia* specifically, we select stars within a color offset $\Delta(\text{BP} - \text{RP}) = 0.5$ around the main sequence, red-giant branch (RGB), and horizontal branch of the isochrone. We also extend the isochrone with $\Delta G = 1.5$ above the tip of the RGB and around the horizontal branch to include stars clustered in those regions. This selection is rather conservative, given that the typical spread around the isochrone is $\sigma_{\text{BP-RP}} \lesssim 0.1$ (M. Riello et al. 2021) even considering the distance spread of stream stars. Nevertheless, it still reduces the number of background stars by a factor of 2 – 10.

To further speed up the calculation, we use a grid interpolation technique, given the fact that the background population is relatively homogeneous and uncorrelated across most observables. For *Gaia*, we account for correlations only in the two-dimensional position space $\mathbf{x}_{\text{pos}}^{2\text{D}}$ and the two-dimensional color–magnitude space $\mathbf{x}_{\text{cm}}^{2\text{D}}$. The proper motions $\mu_{\phi 1} \equiv \dot{\phi}_1 \cos \phi_2^4$ and $\mu_{\phi 2} \equiv \dot{\phi}_2$ are assumed to be uncorrelated with other observables. Under these assumptions, the background PDF becomes

$$p_{\text{bg}}(\mathbf{x}) = p_{\text{bg}}^{\text{pos}}(\mathbf{x}_{\text{pos}}^{2\text{D}}) p_{\text{bg}}^{\mu_{\phi 1}}(\mu_{\phi 1}) p_{\text{bg}}^{\mu_{\phi 2}}(\mu_{\phi 2}) p_{\text{bg}}^{\text{cm}}(\mathbf{x}_{\text{cm}}^{2\text{D}})$$

where $\mathbf{x} \equiv (\mathbf{x}_{\text{pos}}^{2\text{D}}, \mu_{\phi 1}, \mu_{\phi 2}, \mathbf{x}_{\text{cm}}^{2\text{D}})$. We evaluate the PDF in each subspace independently using Gaussian KDE on a rectangular grid large enough to cover all background stars in that subspace. Each subspace is kept at most two-dimensional, since the number of grid points grows exponentially with dimensionality. For efficiency, we randomly select 10^4 background stars near the stream to construct the KDE estimators.

We adopt bandwidths of 0.5° for position, 1 mas yr^{-1} for proper motions, and 0.1 for the color–magnitude space. The grid spacings are set equal to the corresponding bandwidths. The final results are not sensitive to the exact choice of bandwidths or grid spacing, as these values sufficiently resolve the density structure in observable space. Quantitatively, we have verified that multiplying or dividing the bandwidths and grid spacings by a factor of 2 results in only $\lesssim 10\%$ variation in detection purity and completeness for a typical stream.

Note that the PDF estimation can be biased in regions with $\phi_2 > 10^\circ$, where the metric tensor deviates from the identity tensor by more than 3% (see §2.1). Our stream generation method may also deviate from the actual stream track in regions far from the GC if the adopted Galactic potential model is inaccurate. For these reasons, the KDE approach is best suited for relatively small regions, such as a 10° cone around the GC. Nevertheless, this region is still sufficiently large to enclose the half-number radius for 2/3 of the simulated streams.

Finally, we perform linear interpolation on the grids and compute the product of the independent PDFs to estimate the background probability density p_{bg} at any point of interest. Unlike our estimation of the stream probability density, where the simulated stream has no observational uncertainty, the real *Gaia* data already include measurement uncertainties. As a result, the background PDFs obtained above are already the convolution of the original PDFs with Gaussian kernels characterizing uncertainties. Therefore, we do not need to apply the convolution again when evaluating the background PDF at a given point. This simplification is particularly helpful, as performing convolution would be highly inefficient within the grid interpolation framework.

An alternate method is first deconvolving the real *Gaia* data to obtain the actual underlying PDF of the background, and then using the same approach as our stream model to compute the PDF. The first step can be achieved using techniques such as “extreme deconvolution” (J. Bovy et al. 2011). Although this method provides a coherent approach to obtain PDFs for both the stream and the background, it is more computationally expensive by ~ 1000 times compared to grid interpolation as this alternate method requires performing the full KDE evaluation over background stars.

Although we demonstrate our method using *Gaia* data as an example, our method can be readily adapted to other datasets. For instance, metallicity and radial velocity can be included as additional observables when dealing with spectroscopic surveys. Since we perform grid fitting separately for most dimensions, adding extra observables only increases the computational cost linearly.

⁴ Some works refer to $\dot{\phi}_1 \cos \phi_2$ as $\mu_{\phi 1}^*$. We use $\mu_{\phi 1}$ for simplicity, as this does not lead to confusion.

2.3. Stream detection

We optimize our mixture model for the stream and background populations by varying the stream fraction f_s to maximize the log-likelihood function in Eq. (1). The best-fit stream and background probability densities for star i are then given by $f_s p_{s,i}$ and $(1 - f_s) p_{bg,i}$, respectively. Following the standard definition used in mixture models, the membership probability that star i belongs to the stream is given by

$$P_{s,i} \equiv \frac{f_s p_{s,i}}{f_s p_{s,i} + (1 - f_s) p_{bg,i}}. \quad (2)$$

We consider stars with membership probability greater than a chosen threshold P_{th} to be identified as stream members. In this work, we adopt the standard setup of the mixture model with $P_{th} = 0.5$. However, we emphasize that P_{th} is an adjustable parameter that can be tuned depending on whether the analysis prioritizes completeness or purity.

3. Method validation

We validate our new method by applying it to a mock catalog of stream and background stars tailored for *Gaia* DR3, based on the H25 stream catalog.

3.1. Mock observational data

Our mock catalog consists of both a stream population and a background population, each described by six observables: positions (ϕ_1, ϕ_2), proper motions ($\mu_{\phi_1}, \mu_{\phi_2}$), color (BP – RP), and magnitude (G). We do not include radial velocity or metallicity, as these quantities are available for only a small subset of stars in *Gaia* DR3. Parallax is also excluded because of its large uncertainty at the typical heliocentric distances of stellar streams.

The stream population is from the mock catalog of GC streams in a simulated MW-like galaxy (ID 523889) by H25. This catalog generates synthetic streams around a mock GC population based on the GC formation model of Y. Chen & O. Y. Gnedin (2024). It fits the host potential of simulated galaxies at each snapshot using basis function expansion (BFE), accounting for the time evolution of the potential by linearly interpolating between snapshots. The catalog then explicitly integrates the orbit of each GC in this potential over the last ~ 3.5 Gyr and computes the mass loss rate based on the GC mass and the local tidal field. H25 initializes each GC with the P. Kroupa (2001) IMF and releases stars according to the time-varying mass loss rate. Stars are released probabilistically, with the ejection probability inversely proportional to the square root of stellar mass. The released stars form stellar streams using the particle spray algorithm by Y. Chen et al. (2025b). The catalog provides the initial mass, age, and metallicity [Fe/H] of each star, allowing us to assign synthetic *Gaia* photometry directly using the MIST isochrone model.

Ideally, we should generate our simulated streams in the same potential used in H25 for full consistency. However, because it is challenging to constrain the time evolution of the Galactic potential in practice, we use only the static potential at the final snapshot to avoid over-idealization.

The stream duration in H25 is longer than that of most observed GC stream segments ($\lesssim 1$ Gyr, Y. Chen et al. 2025a). In this work, we only use the portions of streams that were released in the last 1 Gyr to mimic streams that are currently observable. Since the main goal of this method is to detect streams around GCs, it is also unnecessary to include stream stars released more than 1 Gyr ago or located outside the 10° cone centered on the GC, as these stars tend to be more sensitive to the time evolution of the Galactic potential.

To create a more realistic stream population, we also add observational errors to the mock stream stars. Similarly to §2.1, we neglect errors in positions and magnitudes. For proper motion uncertainties, we adopt the following parametric form from *Gaia*’s performance website,⁵ which describes the dependence of the uncertainty on apparent magnitude,

$$\sigma_\mu = \frac{\sqrt{40 + 800z + 30z^2}}{1000} \text{ mas yr}^{-1} \quad (3)$$

where $z = 10^{0.4[\max(G, 13) - 15]}$. This expression increases from 0.01 to 0.6 mas yr^{−1} with $G = 13 - 20$. These errors can be significant, especially given that the intrinsic proper motion spread of a Pal 5-like stream is much smaller.

Gaia’s performance website suggests multiplying Eq. (3) by a fudge factor of 1.03 and 0.89 for the proper motions of right ascension and declination, respectively. In the rotated frame (ϕ_1, ϕ_2), we have verified that simply setting this factor to unity for both coordinates also reproduce the actual proper motion uncertainties with sufficient accuracy. We then add Gaussian noise to the original proper motions using Eq. (3).

Similarly, we parameterize the dependence of color uncertainty on apparent magnitude using the following expression,

$$\sigma_{BP-RP} = 10^{[\max(G, 14) - 23]/3} \quad (4)$$

which increases from 0.001 to 0.1 for $G = 14 - 20$. This expression accurately reproduces the mean BP – RP color uncertainty in *Gaia* DR3 by M. Riello et al. (2021).

For the background population, we directly use observational data from *Gaia* DR3 within the same 10° cone centered on the progenitor GC of each stream in the H25 catalog. We only select stars with $G < 20$. Based on the *Gaia* DR3 selection function from *gaiaunlimited* (T. Cantat-Gaudin et al. 2023), we find that the mean completeness exceeds 99% in regions where streams are located. Even for the most incomplete case, the completeness remains above 90%. Thus, this

⁵ <https://www.cosmos.esa.int/web/gaia/science-performance>

magnitude cut provides near-complete coverage for our mock dataset. We also apply the same color–magnitude selection as in §2.2 to reduce the number of background stars that are extremely unlikely to be misidentified by the method.

This results in between 1 million (near the Galactic pole) and 30 million (near the Galactic center) background stars in a selected region, compared to 10 – 10,000 stream stars in the same region. This dramatic contrast between the two populations (up to ~ 4 orders of magnitude) underscores the challenge of detecting streams using traditional methods.

We restrict our analysis to streams originating from surviving GCs with $M > 10^3 M_\odot$ and containing at least 10 stars after applying the above selection criteria. This leads to 158 streams from our chosen catalog from H25. For each stream, we construct a mock dataset following the procedure described above and evaluate our detection algorithm on it, assuming no prior knowledge of how the dataset is constructed.

3.2. Method performance

First, we illustrate key concepts of our method by applying it to an example mock stream originating from a GC with mass $M = 5.6 \times 10^5 M_\odot$, located at a Galactocentric radius $r_{\text{Gal}} = 11$ kpc and a Galactic latitude $b = 21^\circ$. The tidal radius is 125 pc, corresponding to an angular size of 0.63° at its heliocentric distance $d_\odot = 8.7$ kpc. In the *top row* of Fig. 1, we show the distributions of simulated tracer particles in position space, proper motion space, and color–magnitude space. The stream PDF $p_s(\mathbf{x})$ is constructed via KDE using these simulated tracers. Although we display $p_s(\mathbf{x})$ as 2D contours in each of the three subspaces, we emphasize that the KDE is constructed in the full six-dimensional space of all observables and projected onto 2D subspaces.

We then apply our method to the mock dataset, which includes 194 mock stream stars from H25 and approximately four million background stars from *Gaia* DR3. We detect 244 stream members, of which 140 are true members (see the *bottom row* of Fig. 1). Most of the false positives and missed detections are either along the RGB, where background contamination is high, or near $G = 20$, where observational uncertainties become significant according to Eqs. (3) and (4). As we show later in §4.1 and Fig. 8, both of these regions have lower S/N than the main sequence turnoff. We also note that our simulated stream in Fig. 1 samples the RGB less densely than the main sequence due to the lower abundance of RGB stars. This, however, only weakly affects the detection quality, as RGB stars only contribute a small fraction of the total and are intrinsically hard to detect in any case due to their low S/N.

It is worth noting that for faint stars ($G \gtrsim 19$), the uncertainties in color and proper motions can exceed the intrinsic spread of the entire stream. Our method accounts for this by convolving the KDE with a Gaussian blob centered on each

star, as described in §2.1. Without this convolution, many stars would be shifted outside the effective selection window in proper motion and color space, leading to significantly fewer detections than expected. In contrast, while our detection accuracy does decrease for these faint stars, the method still tends to recover the correct total number of stream members by balancing false positives and missed detections. This is expected as our mixture model tends to reproduce the correct number of stream stars to recover the correct density ratio between the stream and the background.

Next, we examine the statistical performance of our method by applying it to all mock streams. We quantify detection quality using three metrics: detection ratio, purity, and completeness. The *detection ratio* is defined as the ratio of the total number of detected stream members N_{detect} to the total number of true members N_{true} . In the *upper panel* of Fig. 2, we show $N_{\text{detect}}/N_{\text{true}}$ as a function of the threshold probability P_{th} . Starting from a large value $\gg 1$ at $P_{\text{th}} = 0$, the median detection ratio of all streams rapidly drops to 1 at $P_{\text{th}} \approx 0.07$, with an interquartile range of approximately 0.3 dex. It then gradually declines to 0.2 at $P_{\text{th}} = 0.8$, followed by a rapid decrease to 0 as P_{th} approaches 1.

Furthermore, we define the *purity* as the ratio of correctly identified stream members N_{correct} to the total number of detected members N_{detect} . The *completeness* is defined as the ratio of N_{correct} to N_{true} . By definition, both purity and completeness range from 0 to 100%, and the ratio between completeness and purity equals the detection ratio. In the *lower panel* of Fig. 2, we show the median purity and completeness as functions of P_{th} . The median purity increases rapidly from 0 to about 70% for $P_{\text{th}} = 0 - 0.1$, and then gradually approaches 100% as $P_{\text{th}} \rightarrow 1$ (note that purity is not well defined at $P_{\text{th}} = 1$). In contrast, completeness drops from 100% to around 60% over $P_{\text{th}} = 0 - 0.1$, and then continues to decrease to 0 at $P_{\text{th}} = 1$. The two curves intersect at $P_{\text{th}} \approx 0.07$, where the detection ratio also reaches unity.

When divided by Galactic latitude b , the high-latitude streams with $|b| > 30^\circ$ reach a detection ratio of unity at $P_{\text{th}} = 0.5$. However, a broad range of $P_{\text{th}} = 0.2 - 0.6$ yields detection ratios that deviate from unity by less than 0.2 dex. For these high-latitude streams, purity and completeness also intersect at $P_{\text{th}} \approx 0.5$, reaching 80% and 72% respectively. On the other hand, the majority of streams are located at low Galactic latitudes $|b| < 30^\circ$. Therefore, their detection quality more closely follows the overall trend of the full sample, which is worse than the high-latitude streams. This difference comes from the distinct background densities between the two populations. Compared to the high- b streams near the Galactic poles, the low- b streams closer to the Galactic plane are contaminated by roughly 10 times more background stars. In these regions, the signal from actual stream stars is more

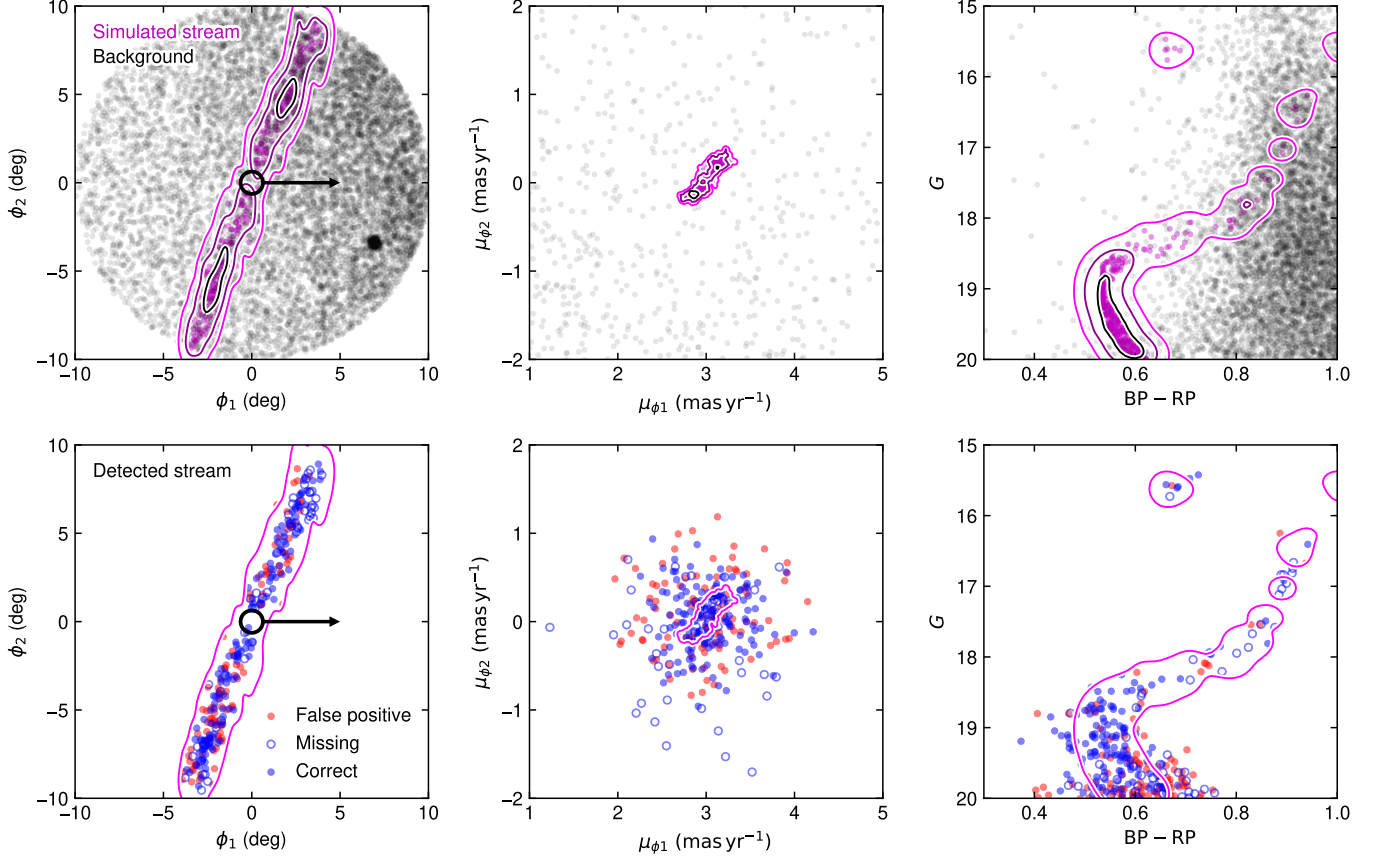


Figure 1. Demonstration of a test of the method on a mock stream. *Top row:* distributions of simulated tracer particles (magenta) and background stars (black) in position space (*left*), proper motion space (*middle*), and color-magnitude space (*right*). We plot the stream PDF $p_s(\mathbf{x})$ from Gaussian KDE as gray contours. The three contours from dark to light represent three values of $\ln p_s(\mathbf{x}) - \ln p_{s,\max}(\mathbf{x}) = -0.5, -2, -4.5$, corresponding to the $1\text{-}\sigma$, $2\text{-}\sigma$, and $3\text{-}\sigma$ ranges of the standard Gaussian distribution, respectively. We only show 10^4 background stars randomly chosen from the total of 4×10^6 for visual clarity. *Bottom row:* application of the method on this mock stream using $P_{\text{th}} = 0.4$. We show stars that are false positives (red circle), missed by the method (blue open circle), or correctly detected (blue solid circle) in the same subspaces as the *top row*. The mock stream members already have uncertainties added and are mixed with background stars from real *Gaia* DR3 data. We also plot the contours of $\ln p_s(\mathbf{x}) - \ln p_{s,\max}(\mathbf{x}) = -4.5$ for comparison. For the *left column* of both rows, we show the location of the progenitor GC as the circle of tidal radius. The velocity of the GC is represented by the arrow.

easily washed out by the strong background noise, leading to a lower probability to be identified as stream members.

Based on the above tests, our default $P_{\text{th}} = 0.5$ yields high purity and completeness $> 70\%$ for streams with relatively low contamination at $|b| > 30^\circ$. The similarity between purity and completeness also ensures that N_{detect} serves as an unbiased estimator of N_{true} . This is particularly important for inferring properties of the progenitor GC, such as the mass loss rate. Streams at $|b| < 30^\circ$ should be treated more carefully since the median completeness can drop to about 40%. However, we emphasize that P_{th} is a user-defined parameter that can be set anywhere between 0 and 1, depending on whether the application prioritizes purity or completeness.

To further study the dependence of detection quality metrics on Galactic latitude, we plot them against $|b|$ in Fig. 3. We apply Gaussian kernel smoothing (similarly to §3.2 in Y. Chen & O. Y. Gnedin 2023) to estimate the median and interquartile

ranges of the three metrics as smooth functions of $|b|$, using a kernel bandwidth that grows linearly from 5° to 15° over the range $|b| = 0 - 75^\circ$. We find that the detection ratio $N_{\text{detect}}/N_{\text{true}}$ remains consistent with unity for $|b| > 30^\circ$. However, the ratio deviates unity near the Galactic plane, where the dispersion is also larger since $\sim 20\%$ of streams with detection ratios < 0.1 are all located at $|b| < 30^\circ$.

The new method reports zero detections for only 4% of all streams. This suggests that the method almost always identifies some stream members as long as a stream is present. However, this conclusion requires further validation by addressing two key questions: 1) How often does the method report false detections when no stream is present? 2) To what extent does the idealization of our mock data benefit detection quality? We explore these questions in the following subsections.

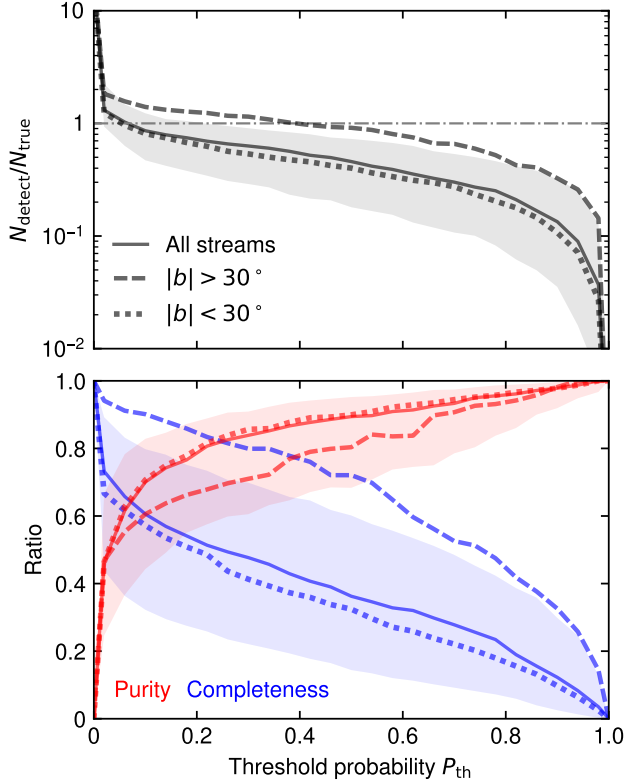


Figure 2. Detection ratio (*upper panel*) and completeness/purity (*lower panel*) of detected stream members as functions of probability threshold. The solid lines stand for the median values among all test streams, while the shaded ranges show the interquartile ranges. Detection ratio = 1 is highlighted as the dot-dashed line. We also show the median detection ratio, completeness, and purity for streams with progenitor GCs at low Galactic latitude ($|b| < 30^\circ$, dotted curves) and high Galactic latitude ($|b| > 30^\circ$, dashed curves) separately. Since the purity at $P_{\text{th}} = 1$ is not well defined, we extrapolate the values at $P_{\text{th}} = 0.99$ out to 1 for visual clarity.

3.3. Null test

In addition to the above tests that focus on how many stream member stars can be correctly detected, it is also important to perform the null test to ensure that the method does not report false detections when there is no stream. Therefore, we design the null test by removing the signal (i.e., mock stream stars) from the test dataset and applying the method only to the background stars in the same region around each GC. A perfect method should yield zero detections in the null test. However, due to random fluctuations in the background that are not captured by the KDE, it is possible that a small fraction of background stars are accidentally better described by the stream model. This can lead to the detection of “false streams”.

In Fig. 4, we show the number of detections in the null test, N_{null} , as a function of Galactic latitude $|b|$. For comparison, we also plot the true number of stream stars N_{true} and

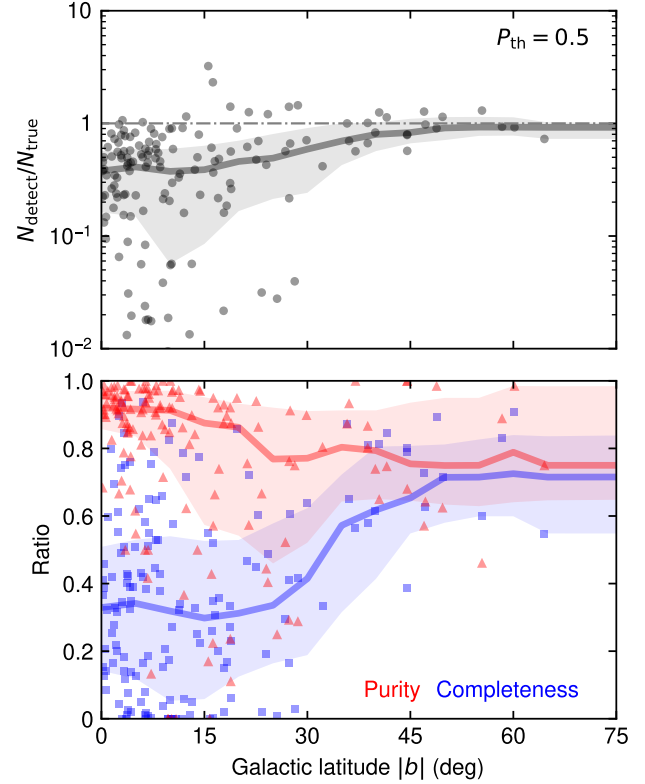


Figure 3. Detection ratio (*upper panel*) and completeness/purity (*lower panel*) of detected stream members as functions of the progenitor GC’s absolute Galactic latitude $|b|$ with $P_{\text{th}} = 0.5$. Individual streams are shown as black circles (detection ratio), blue diamonds (completeness), and red triangles (purity). The solid lines stand for the median values among all test streams, while the shaded ranges show the interquartile ranges. Detection ratio = 1 is highlighted as the dot-dashed line. We calculate the percentiles at any b using nearby streams smoothed by the Gaussian kernel, with bandwidth varying linearly from 5° to 15° from the Galactic plane to the poles.

the number of detections N_{detect} when the signal is included. Our method reports exactly zero detections for 60% of the mock GCs. For the remaining cases, N_{null} is still generally much smaller than the corresponding N_{true} and N_{detect} . Only 11% of cases yield $N_{\text{null}} \geq 10$. Conversely, 16% of cases have $N_{\text{detect}} < 10$. Based on this, we recommend excluding detections below a threshold $N_{\text{detect}} \approx 10$ to avoid most false positives while not discarding too many true stream members.

3.4. Dependence on isochrone models

It should be noted that the H25 catalog uses the same MIST isochrone model adopted in this work to generate the the mock photometry. Before applying our method to real observational data, it is important to assess how sensitive the detection quality is to the choice of isochrone model. To quantify this, we perform a test where we switch to the PARSEC isochrones (A. Bressan et al. 2012) for constructing the simulated stream PDF. Specifically, we set $[M/H]$ in PARSEC equal to the

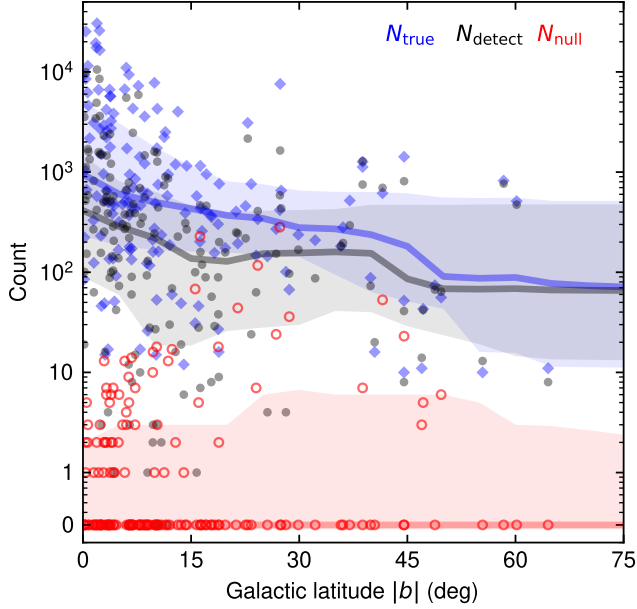


Figure 4. Number of detected stars in the null test N_{null} (red open circles) as a function of the progenitor GC’s absolute Galactic latitude $|b|$. For comparison, we also show the true number of stream stars N_{true} (blue solid diamonds) and the number of detections N_{detect} (black solid circles) without removing the signal. Similarly to Fig. 3, we show the median values as solid lines, while the shaded ranges represent the interquartile ranges.

metallicity $[\text{Fe}/\text{H}]$ of the mock GC and fix the age to 10 Gyr. These settings are intentionally approximate to mimic real observational conditions, where GC ages are often uncertain.

In the *left column* of Fig. 5, we show the detection ratio, purity, and completeness as functions of $|b|$ when switching to the PARSEC isochrones. Despite using a different isochrone model with approximate parameters, these metrics show no statistically significant differences compared to those in Fig. 3. This is because our KDE-based approach provides sufficient flexibility to accommodate variations between isochrone models. Furthermore, the variance introduced by changing isochrone models is negligible compared to the typical color dispersion from observational errors and magnitude dispersion due to the spread in heliocentric distances. Thus, our method is robust to the choice of isochrone model.

3.5. Dependence on Galactic potential models

Our simulated streams are generated in the same Galactic potential model as the final snapshot corresponding to our selected catalog from H25. When applied to real data, however, we do not know the Galactic potential exactly. Most existing Milky Way potential models predict a total mass differing by $\lesssim 10\%$ (see the review by J. A. Hunt & E. Vasiliev 2025). To quantify the impact of using an inaccurate potential, we test three alternative models. The first two are based on the same BFE framework as H25, but with all expansion coefficients

scaled by $\pm 20\%$. This results in a proportional change in the enclosed mass at all radii. We refer to these cases as $0.8M_{\text{Gal}}$ and $1.2M_{\text{Gal}}$. The third model is MilkyWayPotential2022 from gala (A. M. Price-Whelan 2017; A. Price-Whelan et al. 2024), which has been validated against MW mass measurements out to ~ 150 kpc (J. A. Hunt & E. Vasiliev 2025).

The *right column* of Fig. 5 shows the detection ratio, purity, and completeness as functions of $|b|$ for these models. In almost all cases, these detection metrics decrease only slightly. An exception is MilkyWayPotential2022, which shows a $\lesssim 10\%$ increase in purity at $|b| > 60^\circ$. This is likely a stochastic effect given the small number of streams at such high latitudes. All metrics remain within the interquartile range (which is even narrower than the 1σ range) of the original values, indicating that even a 20% variation in the potential has little effect on detection quality. Since recent measurements of the MW potential have uncertainties of only $\approx 10\%$ (e.g., R. Ibata et al. 2024), these tests are conservative.

The MilkyWayPotential2022 model performs almost identically to the original BFE model. This is encouraging as the former is designed to match the real MW instead of the simulated galaxy. While both models have similar halo structures beyond 10 kpc, their disk components differ significantly, with enclosed masses near 1 kpc differing by $\approx 30\%$. This suggests that the performance of our method is insensitive to the exact choice of Galactic potential model within 10° around the GC.

We emphasize, however, that our method has not been tested beyond 10° , where the Galactic potential becomes more important. Even within 10° , the slight decrease in purity for the alternative models is largely driven by stars located between 5° and 10° , indicating that the influence of the potential grows with distance from the progenitor. Moreover, the LMC may introduce a larger perturbation to the MW’s potential (see review by E. Vasiliev 2023) than that in the simulated galaxy. Since the H25 catalog does not include a galaxy with a realistic LMC analog, we are unable to quantify the effect of LMC in this work.

3.6. Dependence on stream generation algorithms

The mock stream catalog also uses the same Y. Chen et al. (2025b) particle spray algorithm as adopted in this work. To further validate our method, we conduct an additional test by generating a mock stream using an N-body simulation, which is generally considered more accurate than most particle spray methods. Specifically, we initialize the simulation using the same initial conditions as the example stream in §3.2 and Fig. 1. Following Y. Chen et al. (2025b), we model the progenitor GC using the I. R. King (1966) model with $W = 8$, a typical value for Galactic GCs. We set the particle mass to $10 M_\odot$, the softening length to 1 pc, and the simulation time step to 2^{-13} kpc km $^{-1}$ s ≈ 0.1 Myr. We then backtrack

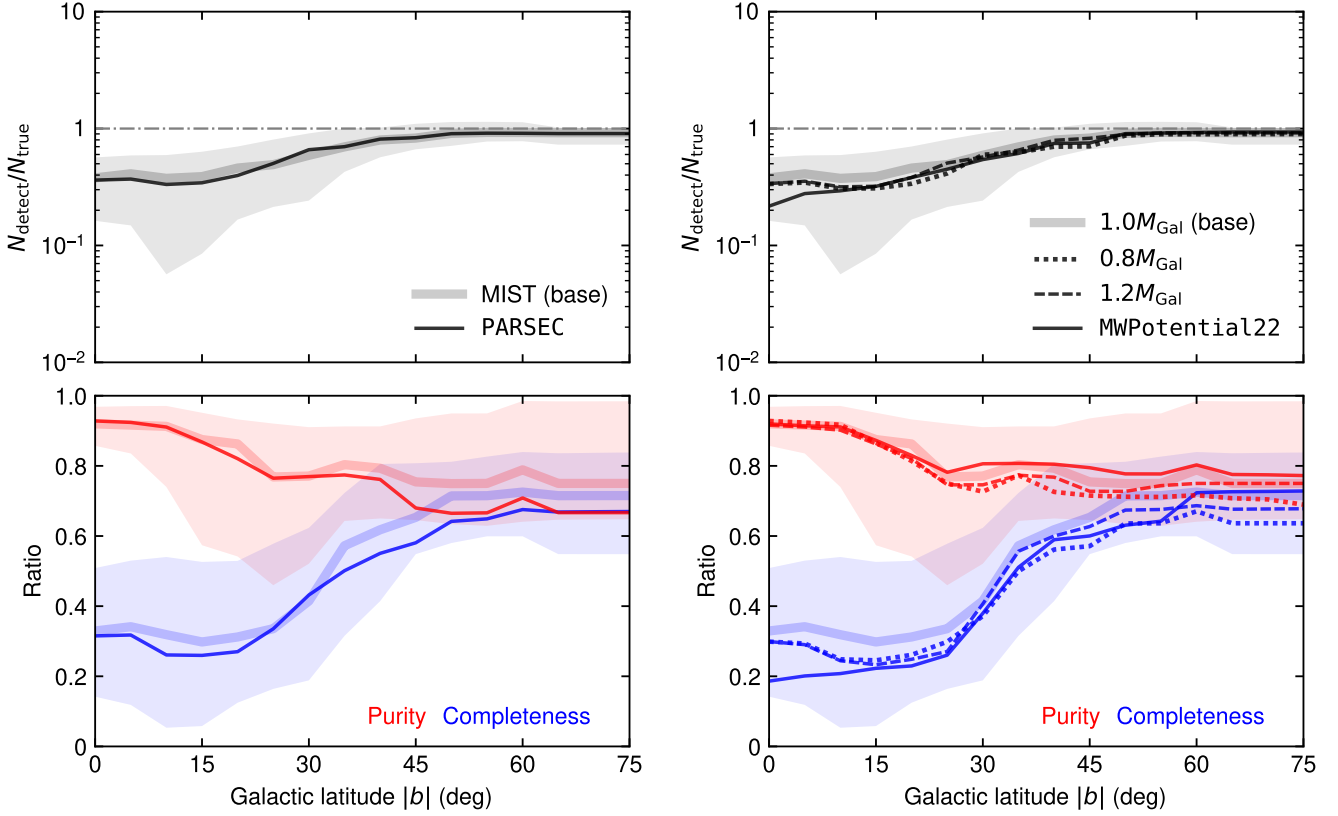


Figure 5. Same as Fig. 3, but with the alternative isochrone model PARSEC (*left column*) and three alternative Galactic potential models (*right column*): 1) the base potential model scaled down by 20% (dotted curves), 2) the base potential model scaled up by 20% (dashed curves), and 3) the MilkyWayPotential2022 model from gala. We show the interquartile ranges for the base case for comparison.

the GC orbit for 1 Gyr in the same static Galactic potential described in §3.1, and run the N-body simulation forward to the present day using the fast-multipole gravity solver *falcon* (W. Dehnen 2000, 2002). Although the N-body simulation includes only collisionless dynamics and omits close stellar encounters, it is sufficiently distinct from the particle spray model to serve our purpose of validating the detection method with an alternative algorithm.

Since the particles in the simulation are different from individual stars, we randomly select a subsample of escaped particles and re-sample their masses using the stellar mass distribution of stars with $G < 20$ from H25. We then assign G magnitudes and BP – RP colors to these particles using the MIST isochrone. This ensures that the new mock stream from the N-body simulation has the same number of stars and mass distribution as its counterpart in H25, which is important for a fair comparison of detection quality metrics.

Finally, we add background stars to the new mock stream following §3.1 and apply our stream detection method. Among the 194 mock stream stars, we detect 230 members, with 133 correct detections. These numbers are only about 5 – 10% lower than those reported in §3.2. This slight decrease is expected, as we attempt to recover stream members

generated using a different algorithm. However, since our method evaluates membership probability by comparing the stream probability density to the background, the detection performance is not strongly sensitive to the specific stream generation algorithm as long as the S/N, defined as the ratio between the densities of actual stream members and background stars in the multi-dimensional observable space, is greater than 1.

3.7. Influence of dust extinction

When applying *StarStream* to real data, it is important to account for dust extinction, particularly near the Galactic plane where the color excess reaches $E(B - V) > 1$. Extinction reduces the number of observable stream stars above the detection limit by up to a factor of 10 near the mid-plane, while also shifting the stream to redder colors, where the background density is higher (see the *upper right* panel of Fig. 1). In this section, we investigate the influence of these effects by incorporating dust extinction into our mock dataset.

We use the Python package *dustmaps* (G. M. Green 2018) to compute the extinction A_V of each stream star in H25, based on the map of D. J. Schlegel et al. (1998, hereafter SFD). This map is recalibrated by E. F. Schlafly & D. P. Finkbeiner (2011) with $R_V = 3.1$. The A_V values are then passed to

the MIST bolometric correction interpolation table⁶ to obtain the G -band extinction and $BP - RP$ color excess. Since the table only covers $A_V = 0 - 6$, we remove stars with $A_V > 6$, as they are also likely too faint to be observable. Note that the SFD map provides extinction along the full line of sight from the solar system to infinity. This overestimates the true extinction, since streams are located at finite distances. While the overestimation is probably modest, this test should be regarded as an extreme case that assumes maximum possible extinction.

The detection quality would decline significantly if we directly apply the same method to the new dataset, since the simulated isochrone no longer aligns with the reddened distribution of stream stars in the color–magnitude space. Fortunately, in practice we have access to extinction values for most MW GCs from catalogs such as W. E. Harris (1996), allowing us to account for realistic extinction when simulating the stream. This is equivalent to replacing the original isochrone with a reddened one, where the extinction and color excess are calculated from the progenitor’s A_V . Here, we take each GC’s A_V directly from the SFD map at its location. As before, we exclude streams whose progenitor GCs have $A_V > 6$. Since extinction makes streams fainter, the number of streams with at least 10 stars also decreases, leaving 123 valid streams out of the original 158.

Stars from the same stream may have different A_V values in the mock dataset, since extinction is not constant within the search radius. However, we redden the simulated isochrone using only a single A_V value at the center of the stream. This reflects the practical difficulty of obtaining precise extinction for every individual stream star. We emphasize that our tests are designed to reproduce the actual detection quality of *StarStream* when applied to real data. Thus, it is important to avoid any over-idealization.

In Fig. 6, we show the detection ratio, completeness, and purity after applying *StarStream* to the new mock dataset that includes extinction. As expected, both completeness and purity decrease substantially near the Galactic plane ($|b| < 30^\circ$). Close to the mid-plane, purity falls from $\sim 90\%$ to below 10% , and completeness drops to a similar level. This decline is primarily due to the high background density near the reddened isochrone and the lower number of stream members that remain observable. We find that completeness can be improved by lowering the probability threshold to $P_{\text{th}} < 0.1$, whereas varying P_{th} between $0.01 - 0.99$ does not significantly improve purity. In contrast, the high-latitude regions are only slightly affected. For streams at $|b| > 30^\circ$, the median completeness and purity decrease by only $\sim 10\%$ to 62% and

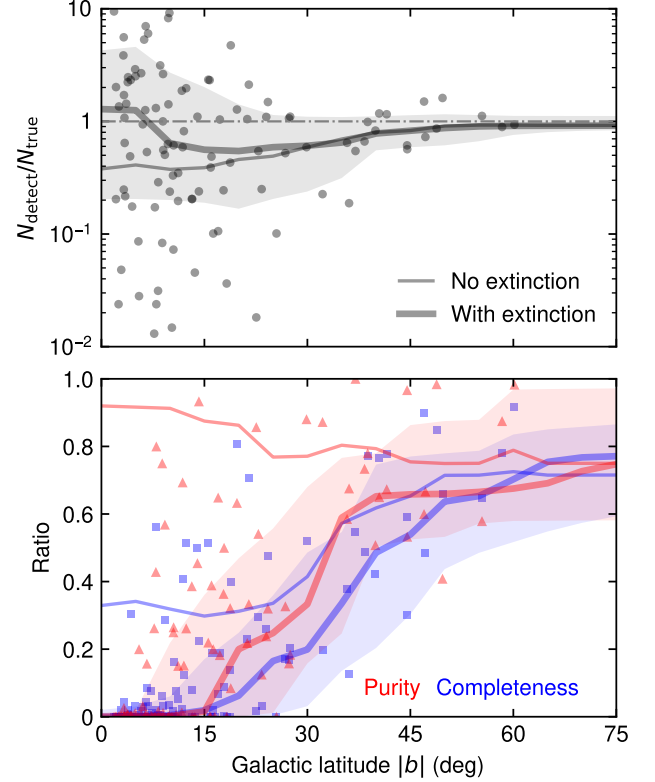


Figure 6. Same as Fig. 3, but including dust extinction from the SFD map. We also show the original cases (no extinction) as thin curves for comparison.

67% , respectively. The detection ratio in this latitude range is nearly unchanged.

We also perform the null test described in §3.3 on the new dataset, and show it in Fig. 7. In this case, only 13% of detections yield $N_{\text{null}} = 0$. This reduction is mainly due to the higher false detection rate at low latitudes, where N_{null} is nearly the same as N_{detect} for $|b| < 15^\circ$. However, the high-latitude region again remains almost unaffected. At $|b| > 30^\circ$, false and true detections can still be cleanly separated using the same threshold of $N_{\text{detect}} \approx 10$ recommended in §3.3.

Therefore, although both completeness and purity of *StarStream* decrease at low latitudes when accounting for extreme extinction, the method still achieves high values $\approx 65\%$ at $|b| > 30^\circ$, where extinction is less significant. Since the extinction adopted in this section overestimates the true values, these results represent lower limits of the actual detection quality. As discussed later in §4.1, future spectroscopic and deep photometric surveys are necessary to reveal low- b streams with high extinction.

⁶ <https://waps.cfa.harvard.edu/MIST/model-grids.html>

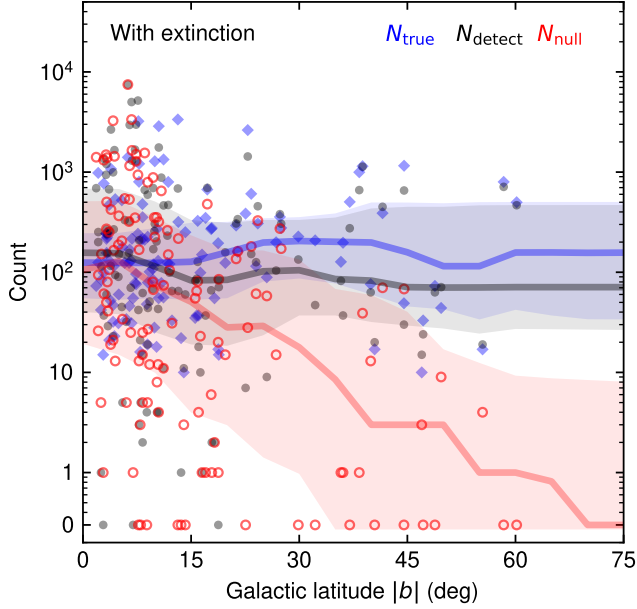


Figure 7. Same as Fig. 4, but including dust extinction from the SFD map.

4. Discussion

4.1. Application to other surveys

Although this work is framed with Gaia DR3, it can be straightforwardly extended to other surveys. In this subsection, we discuss the importance of spectroscopic surveys and deep photometric surveys in further enhancing detection quality.

Our method operates in the six-dimensional space of positions, proper motions, color, and magnitude. We find that the median detection ratio drops significantly to 18% when color and magnitude are excluded. In this case, the median purity and completeness also decrease to 50% and 6%, respectively. If proper motions are excluded, we cannot even detect any stars for more than 80% of streams. These results demonstrate that six-dimensional information of positions, proper motions, colors, and magnitudes is important for recovering most streams in *Gaia* DR3. For streams with extremely low density, strong background contamination, or high extinction we may need even more independent observables. This highlights the importance of spectroscopic surveys, such as the Apache Point Observatory Galactic Evolution Experiment (APOGEE, S. R. Majewski et al. 2017), the Southern Stellar Stream Spectroscopic Survey (S⁵, T. S. Li et al. 2019), and the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2022) Milky Way Survey (MWS, A. P. Cooper et al. 2023), which provide radial velocities and metallicities and may help reveal very faint streams.

Our example stream in Fig. 1 shows that most correctly identified members lie near the main sequence turnoff, where

the S/N is high. In Fig. 8, we present the distribution of S/N in the color-magnitude diagram, where S/N is defined as the ratio between the densities of actual stream members and background stars. We estimate both densities in the six-dimensional space using KDE approaches similar to those described in §2.1 and §2.2. However, we multiply the stream kernel widths in §2.1 by a factor of 3, since the number of member stars is smaller than the number of simulated tracer particles.

It is remarkable that a significant number of stars have S/N values larger than unity in six-dimensional space, despite the number of stream members being orders of magnitude smaller than the background contaminants. The region with S/N > 1 coincides with the area of high purity and completeness near the main sequence turnoff at $G \approx 19$. Brighter stars in the RGB suffer from higher background contamination, resulting in lower S/N. Stars in the horizontal branch show relatively high S/N because they are bluer than most background stars; however, the horizontal branch contributes only a few members to the stream. The majority of stars lie below the main sequence turnoff and have lower S/N due to large observational uncertainties, as described by Eqs. (3) and (4). This issue is not unique to the example stream. Since the typical main sequence turnoff is at absolute magnitude $M_G = 3 - 4$, corresponding to $G = 20$ at a heliocentric distance ≈ 20 kpc, *Gaia* can barely detect stars fainter than the turnoff for most MW streams. This highlights the importance of deep photometric surveys such as LSST and *Roman*, which are expected to detect significantly more stream members thanks to their much deeper detection limits (S. Pearson et al. 2024; C. Holm-Hansen et al. 2025). Moreover, *Roman* can also provide high-quality astrometry measurements for faint stars, significantly improving S/N below the main sequence turnoff.

4.2. Improvements to existing methods

The new detection method constructs stream KDE via a particle spray algorithm, which only requires the progenitor GC’s mass, position, and velocity as input. We then assign colors and magnitudes to the tracers based on an isochrone model, which depends only on the progenitor’s metallicity and age. In practice, most of these input parameters are available from existing catalogs (e.g., M. Hilker et al. 2019). The only exception is age, which, however, has a weak impact on our results (see §3.4). Therefore, our method avoids making unnecessary and unrealistic assumptions about the stream’s morphology and kinematics. Additionally, we also construct the background KDE directly from a subsample of observed stars.

As a result, both the stream and background models have no free parameters. The mixture model that combines these two components includes only one free parameter: the stream fraction f_s . This minimal parameterization offers a direct ad-

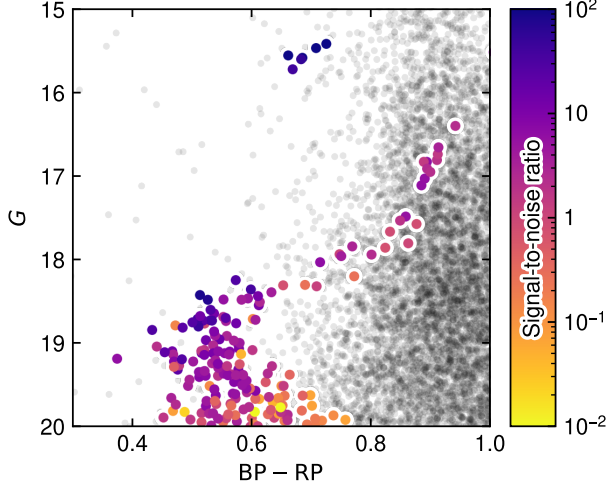


Figure 8. Distribution of S/N in the color–magnitude diagram for actual members of the mock stream in Fig. 1. The signal and noise densities are estimated using a similar KDE method in §2.1 and §2.2, respectively. We multiply the stream kernel width by a factor of 3 to account for the lower number density of stars that do not sample the parameter space as well as simulated tracers. The same 10^4 background stars as in Fig. 1 are also shown for reference.

vantage in computational efficiency. On average, our Python implementation takes ~ 10 minutes to detect a single stream when running on 32 cores of an Intel Haswell CPU. The total computation time to analyze all mock streams is approximately 1000 CPU hours of computation time, which is orders of magnitude faster than typical mixture models. For instance, STREAMFINDER requires millions of CPU hours on a similar dataset (R. Ibata et al. 2021).

Although the particle spray method used here takes no free parameters, our stream model is more accurate than simply assuming the stream is an elongated structure along its orbital track. A representative example of the latter is STREAMFINDER, which detects clustering of stellar orbits within a Gaussian tube. Using STREAMFINDER, R. Ibata et al. (2024) successfully detected 16 streams originating from known GCs in *Gaia* DR3. To compare performance, we apply our method to the same 16 GCs in *Gaia* DR3, using GC properties from the M. Hilker et al. (2019) catalog and the Galactic potential model MilkyWayPotential2022. Excluding M68, ω Cen, and M5, whose streams in R. Ibata et al. (2024) are not connected to the progenitors, all other streams extend at least 5° within our 1 Gyr integration time. In this region, our method detects on average 5 times as many stream members. Even for the Pal 5 stream, which is widely thought to be one of the most complete, we detect 131 members compared to 76 reported in R. Ibata et al. (2024), which a meaningful improvement. Notably, the actual Pal 5 stream extends beyond 5° because it formed over approximately 6 Gyr (Y. Chen et al. 2025a), much longer than our

integration time. For the remaining streams that extend farther within our 1 Gyr integration time, we detect on average 4 times as many member stars inside 10° .

StarStream is not the first physics-motivated attempt to search for GC streams. For example, C. J. Grillmair (2022); Y. Yang et al. (2023); C. J. Grillmair (2025) have successfully identified tidal features around individual GCs by comparing observations with simulated streams. Compared to these works, our approach automates this technique using KDE. In addition, we provide quantitative metrics for detection quality, demonstrating the broader potential of this method for identifying more GC streams.

5. Summary

In this work, we present *StarStream*, an automatic detection algorithm for stellar streams based on a physics-inspired stream model. We construct a mixture model in the multidimensional space of observables, including positions, velocities, colors, magnitudes, etc. The model consists of background and stream components, whose PDFs are represented using KDE. For the background, we build the KDE from a subsample of observed stars; for the stream, we construct the KDE from tracers generated using the particle spray algorithm of Y. Chen et al. (2025b). We illustrate the method using an example stream in Fig. 1.

We quantitatively assess the detection quality of our method around existing GCs using the mock stream catalog from H25, which is tailored to *Gaia* DR3 and includes six observables: sky coordinates (ϕ_1, ϕ_2), proper motions ($\mu_{\phi_1}, \mu_{\phi_2}$), color (BP – RP), and magnitude (G). Our mock dataset incorporates magnitude-dependent uncertainties for each observable, and we include all *Gaia* DR3 stars as the background population. The method achieves both purity and completeness around 65% even with extreme dust extinction ($> 70\%$ without extinction). The detection ratio is near unity for high-latitude streams (Fig. 3 and Fig. 6). For low-latitude streams, however, high background contamination and extinction can significantly reduce both purity and completeness to $< 10\%$.

Next, we perform a series of tests to examine the robustness of the method. We begin with a null test to evaluate the frequency of false positive detections. After removing the signal (i.e., mock stream stars) from the dataset, our method correctly reports $N_{\text{null}} = 0$ for 13% of all streams with extreme extinction (60% without extinction), while a threshold $N_{\text{detect}} \approx 10$ cleanly separates true and false detections for high-latitude streams (Fig. 4 and Fig. 7). We further test the method using a different isochrone model and different Galactic potential models (Fig. 5), and a different stream generation algorithm. These alternative configurations do not significantly weaken the detection quality. Being robust to alternate isochrone models is important in application to real

data as the predicted isochrone can vary significantly among different models.

We find that both purity and completeness drop significantly when proper motions or color and magnitude are excluded from the input dataset. This emphasizes the importance of incorporating multiple independent observables for stream detection. With the full six-dimensional input, however, the S/N can exceed unity even when the number of stream members is orders of magnitude smaller than the background contaminants (Fig. 8). Stars with high S/N are primarily located near the main sequence turnoff, coinciding with those that exhibit high purity and completeness. In contrast, fainter stars near $G = 20$ and brighter stars on the RGB both have lower S/N due to large observational uncertainties or strong background contamination, respectively.

Finally, we compare our new method to existing methods such as STREAMFINDER. Our method is several orders of magnitude more computationally efficient, primarily because the physics-inspired stream model requires no free parameters. This greatly accelerates the model optimization process. At the same time, for streams associated with existing GCs in R. Ibata et al. (2024), our method detects on average 5 times as many member stars within the 5° circle. It is also worth noting that the method may uncover additional streams when applied to the full set of GCs in *Gaia* DR3.

We have published the package *StarStream* on GitHub via <https://github.com/ybillchen/StarStream>, where we also provide example Python notebooks for running the code. The

code requires the mass and six-dimensional phase-space coordinates of the progenitor to generate stream tracers using the particle spray algorithm. It also requires an isochrone that is fit to the progenitor to compute mock photometry for the tracers. The input dataset should be a multi-dimensional array of observables, coupled with another array of observational uncertainties of the same shape. Users can also specify the threshold probability P_{th} , tracer particle ejection rate, KDE kernel widths, interpolation grid spacings for the background PDF, and the Galactic potential, if different values from the default of this paper are preferred.

Acknowledgments

We thank Monica Valluri, Eric Bell, Katya Gozman, and Jacob Nibauer for insightful discussions. YC, OYG, and CHH were supported in part by National Aeronautics and Space Administration through contract NAS5-26555 for Space Telescope Science Institute programs HST-AR-16614 and JWST-GO-03433. This research benefited from the Gravity in the Local Group conference hosted by the McWilliam's Center for Cosmology and Astrophysics, Carnegie Mellon University.

Software: agama (E. Vasiliev 2019), numpy (C. R. Harris et al. 2020), matplotlib (J. D. Hunter 2007), scipy (P. Virtanen et al. 2020), astropy (The Astropy Collaboration et al. 2018), gala (A. M. Price-Whelan 2017; A. Price-Whelan et al. 2024), pandas (The pandas development team 2024), dustmaps (G. M. Green 2018), falCON (W. Dehnen 2000, 2002), gaiaunlimited (T. Cantat-Gaudin et al. 2023)

Appendix

A. Kernel density estimation with uncertainties

Given a sample of N points $\{x_j\}$, we can estimate the probability density function $p(x)$ at any point of interest using Gaussian KDE,

$$\begin{aligned} p(x) &\approx \frac{1}{N} \sum_{j=1}^N p_{\text{KDE}}(x|x_j, \sigma) \\ &\equiv \frac{1}{N} \sum_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(x-x_j)^2}{2\sigma^2} \right]. \end{aligned}$$

However, if the location of the point of interest is uncertain and follows a distribution function $f(x)$, the expected probability density p is the average of $p(x)$ weighted by $f(x)$,

$$p = \int_{-\infty}^{\infty} p(x) f(x) dx \approx \sum_{j=1}^N \int_{-\infty}^{\infty} p_{\text{KDE}}(x|x_j, \sigma) f(x) dx.$$

If we assume $f(x)$ is a Gaussian function centered at x_0 with uncertainty σ_0 , the last integral is the convolution of two Gaussian distributions $\mathcal{N}(x_i, \sigma_i^2) \star \mathcal{N}(0, \sigma_0^2)$ with the independent variable x_0 . This convolution is simply another Gaussian distribution $\mathcal{N}(x_i, \sigma'^2) \equiv \mathcal{N}(x_i, \sigma^2 + \sigma_0^2)$. We can thus obtain the expected probability density as

$$p \approx \frac{1}{N} \sum_{j=1}^N p_{\text{KDE}}(x|x_j, \sigma').$$

Therefore, the expected probability density of a point with Gaussian uncertainty σ_0 equals the standard Gaussian KDE at the same location with bandwidth replaced by $\sigma'^2 \equiv \sigma^2 + \sigma_0^2$ for every sample point $\{x_i\}$.

B. Astrometry uncertainty propagation

The uncertainties of astrometry measurements in sky coordinate system $(\alpha, \delta, \mu_\alpha, \mu_\delta, \varpi, v_r)$ are commonly quantified as a 6×6 covariance matrix

$$\mathbf{C} = \begin{pmatrix} V_\alpha & C_{\alpha\delta} & \cdots & C_{\alpha v_r} \\ C_{\alpha\delta} & V_\delta & \cdots & C_{\delta v_r} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\alpha v_r} & C_{\delta v_r} & \cdots & V_{v_r} \end{pmatrix}$$

where $V_i = \sigma_i^2$ is the variance of quantity i and $C_{ij} = \sigma_i \sigma_j \rho_{ij}$ is the covariance between quantities i and j , in which $\rho_{ij} = \rho_{ji}$ is the correlation coefficient.

To obtain the covariance matrix in the rotated frame $(\phi_1, \phi_2, \mu_{\phi_1}, \mu_{\phi_2}, \varpi', v'_r)$,

$$\mathbf{C}' = \begin{pmatrix} V_{\phi_1} & C_{\phi_1\phi_2} & \cdots & C_{\phi_1 v'_r} \\ C_{\phi_1\phi_2} & V_{\phi_2} & \cdots & C_{\phi_2 v'_r} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\phi_1 v'_r} & C_{\phi_2 v'_r} & \cdots & V_{v'_r} \end{pmatrix}$$

We use linear uncertainty propagation to approximate $\mathbf{C}'$

$$\mathbf{C}' \approx \mathbf{J} \mathbf{C} \mathbf{J}^T \quad (\text{B1})$$

where $\mathbf{J}$ is the Jacobian matrix

$$\mathbf{J} \equiv \frac{\partial(\phi_1, \phi_2, \mu_{\phi_1}, \mu_{\phi_2}, \varpi', v'_r)}{\partial(\alpha, \delta, \mu_\alpha, \mu_\delta, \varpi, v_r)}.$$

Note that $(\varpi', v'_r) = (\varpi, v_r)$ in coordinate rotation. Also, other new coordinates do not explicitly depend on parallax and radial velocity. Therefore, we directly know $\partial\varpi'/\partial i = \delta_{i\varpi}$, $\partial v'_r/\partial i = \delta_{iv_r}$, $\partial i/\partial\varpi = \delta_{i\varpi'}$, and $\partial i/\partial v_r = \delta_{iv'_r}$, where the Kronecker symbol $\delta_{ij} = 1$ only if $i = j$, otherwise 0. This greatly simplifies our calculation as we only need to take into account the rotation of angular coordinates $(\alpha, \delta, \mu_\alpha, \mu_\delta)$ and $(\phi_1, \phi_2, \mu_{\phi_1}, \mu_{\phi_2})$ independent from ϖ and v_r .

Rotation in the spherical coordinate system is nonlinear, which makes it challenging to derive $\mathbf{J}$ analytically. However, by first transforming the sky coordinates to the Cartesian system (x, y, z, v_x, v_y, v_z) , we can more easily deal with rotation as it becomes a linear coordinate transformation, which can be described by the 6×6 rotational matrix,

$$\mathbf{R} \equiv \begin{pmatrix} \mathbf{R}_{3 \times 3} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{3 \times 3} \end{pmatrix}$$

Here, $\mathbf{R}_{3 \times 3} \in \text{SO}(3)$ is the standard 3D rotational matrix. After the rotation, we can transform the rotated Cartesian frame back to the great circle frame. The combined Jacobian matrix thus equals the product of the Jacobian matrices for the three transformations. Since rotation is a linear translation, the Jacobian matrix of rotation is simply $\mathbf{R}$ itself. We derive the two remaining Jacobian matrices as follows.

We consider the rotation of angular coordinates in the unit sphere. Such a simplification does not affect our calculation as the parallax and radial velocity are independent coordinates. The transformation from sky coordinates to the cartesian system is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \alpha \cos \delta \\ \sin \alpha \cos \delta \\ \sin \delta \end{pmatrix}$$

and

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} -\mu_\alpha \sin \alpha - \mu_\delta \cos \alpha \sin \delta \\ \mu_\alpha \cos \alpha - \mu_\delta \sin \alpha \sin \delta \\ \mu_\delta \cos \delta \end{pmatrix}.$$

Recall that we define $\mu_\alpha \equiv \alpha \cos \delta$. Therefore, the corresponding 6×4 Jacobian matrix is

$$\mathbf{J}_1 \equiv \frac{\partial(x, y, z, v_x, v_y, v_z)}{\partial(\alpha, \delta, \mu_\alpha, \mu_\delta)} = \begin{pmatrix} -\sin \alpha \cos \delta & -\cos \alpha \sin \delta & 0 & 0 \\ \cos \alpha \cos \delta & -\sin \alpha \sin \delta & 0 & 0 \\ 0 & \cos \delta & 0 & 0 \\ -\mu_\alpha \cos \alpha + \mu_\delta \sin \alpha \sin \delta & -\mu_\delta \cos \alpha \cos \delta & -\sin \alpha & -\cos \alpha \sin \delta \\ -\mu_\alpha \sin \alpha - \mu_\delta \cos \alpha \sin \delta & -\mu_\delta \sin \alpha \cos \delta & \cos \alpha & -\sin \alpha \sin \delta \\ 0 & -\mu_\delta \sin \delta & 0 & \cos \delta \end{pmatrix}.$$

We can also calculate the inverse transformation to the great circle frame,

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \arctan \frac{y'}{x'} \\ \arcsin z' \end{pmatrix}$$

and

$$\begin{pmatrix} \mu_{\phi 1} \\ \mu_{\phi 2} \end{pmatrix} = \frac{1}{\sqrt{x'^2 + y'^2}} \begin{pmatrix} -v'_x y' + v'_y x' \\ v'_z \end{pmatrix}.$$

The 4×6 Jacobian matrix of this transformation is

$$\mathbf{J}_2 \equiv \frac{\partial(\phi_1, \phi_2, \mu_{\phi 1}, \mu_{\phi 2})}{\partial(x', y', z', v'_x, v'_y, v'_z)} = \begin{pmatrix} -\frac{\sin \phi_1}{\cos \phi_2} & \frac{\cos \phi_1}{\cos \phi_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\cos \phi_2} & 0 & 0 & 0 \\ -\frac{\mu_{\phi 2} \sin \phi_1 \sin \phi_2}{\cos \phi_2} & \frac{\mu_{\phi 2} \cos \phi_1 \sin \phi_2}{\cos \phi_2} & 0 & -\sin \phi_1 & \cos \phi_1 & 0 \\ 0 & 0 & \frac{\mu_{\phi 2} \sin \phi_2}{\cos^2 \phi_2} & 0 & 0 & \frac{1}{\cos \phi_2} \end{pmatrix}$$

For clarity, we already write $\mathbf{J}_2$ in terms of the great circle coordinates.

The combined Jacobian matrix is given by

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_2 \mathbf{R} \mathbf{J}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{2 \times 2} \end{pmatrix}.$$

The identity matrix in the lower right accounts for the transformation of (ϖ, v_r) . We have verified that $\mathbf{J}_2 \mathbf{R} \mathbf{J}_1$ also becomes the identity matrix in the case of no rotation, $\mathbf{R} = \mathbf{I}_{6 \times 6}$. Finally, we can insert $\mathbf{J}$ to Eq. (B1) to obtain the covariance matrix in the great circle frame.

References

- Amorisco, N. C. 2015, MNRAS, 450, 575, doi: [10.1093/mnras/stv648](https://doi.org/10.1093/mnras/stv648)
- Banik, N., Bertone, G., Bovy, J., & Bozorgnia, N. 2018, JCAP, 2018, 061, doi: [10.1088/1475-7516/2018/07/061](https://doi.org/10.1088/1475-7516/2018/07/061)
- Bernard, E. J., Ferguson, A. M. N., Schlafly, E. F., et al. 2014, MNRAS, 443, L84, doi: [10.1093/mnras/ltu089](https://doi.org/10.1093/mnras/ltu089)
- Bonaca, A., Geha, M., Küpper, A. H. W., et al. 2014, ApJ, 795, 94, doi: [10.1088/0004-637X/795/1/94](https://doi.org/10.1088/0004-637X/795/1/94)
- Bonaca, A., & Price-Whelan, A. M. 2025, New Astronomy Reviews, 100, 101713, doi: [10.1016/j.newar.2024.101713](https://doi.org/10.1016/j.newar.2024.101713)
- Bovy, J., Hogg, D. W., & Roweis, S. T. 2011, Ann. Appl. Stat., 5, doi: [10.1214/10-AOAS439](https://doi.org/10.1214/10-AOAS439)
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127, doi: [10.1111/j.1365-2966.2012.21948.x](https://doi.org/10.1111/j.1365-2966.2012.21948.x)
- Cantat-Gaudin, T., Fouesneau, M., Rix, H.-W., et al. 2023, A&A, 669, A55, doi: [10.1051/0004-6361/202244784](https://doi.org/10.1051/0004-6361/202244784)
- Carlberg, R. G., Grillmair, C. J., & Hetherington, N. 2012, ApJ, 760, 75, doi: [10.1088/0004-637X/760/1/75](https://doi.org/10.1088/0004-637X/760/1/75)
- Chen, Y. 2024, ybillchen/particle_spray: v0.1.0, Zenodo, doi: [10.5281/zenodo.13923250](https://doi.org/10.5281/zenodo.13923250)
- Chen, Y., & Gnedin, O. Y. 2023, MNRAS, 522, 5638, doi: [10.1093/mnras/stad1328](https://doi.org/10.1093/mnras/stad1328)
- Chen, Y., & Gnedin, O. Y. 2024, MNRAS, 527, 3692, doi: [10.1093/mnras/stad3345](https://doi.org/10.1093/mnras/stad3345)

- Chen, Y., Li, H., & Gnedin, O. Y. 2025a, *ApJL*, 980, L18, doi: [10.3847/2041-8213/adaf93](https://doi.org/10.3847/2041-8213/adaf93)
- Chen, Y., Valluri, M., Gnedin, O. Y., & Ash, N. 2025b, *ApJS*, 276, 32, doi: [10.3847/1538-4365/ad9904](https://doi.org/10.3847/1538-4365/ad9904)
- Cooper, A. P., Koposov, S. E., Allende Prieto, C., et al. 2023, *ApJ*, 947, 37, doi: [10.3847/1538-4357/acb3c0](https://doi.org/10.3847/1538-4357/acb3c0)
- Dehnen, W. 2000, *ApJL*, 536, L39, doi: [10.1086/312724](https://doi.org/10.1086/312724)
- Dehnen, W. 2002, *JCoPh*, 179, 27, doi: [10.1006/jcph.2002.7026](https://doi.org/10.1006/jcph.2002.7026)
- DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, *AJ*, 164, 207, doi: [10.3847/1538-3881/ac882b](https://doi.org/10.3847/1538-3881/ac882b)
- Dotter, A. 2016, *ApJS*, 222, 8, doi: [10.3847/0067-0049/222/1/8](https://doi.org/10.3847/0067-0049/222/1/8)
- Erkal, D., Belokurov, V., Bovy, J., & Sanders, J. L. 2016, *MNRAS*, 463, 102, doi: [10.1093/mnras/stw1957](https://doi.org/10.1093/mnras/stw1957)
- Erkal, D., Koposov, S. E., & Belokurov, V. 2017, *MNRAS*, 470, 60, doi: [10.1093/mnras/stx1208](https://doi.org/10.1093/mnras/stx1208)
- Fardal, M. A., Huang, S., & Weinberg, M. D. 2015, *MNRAS*, 452, 301, doi: [10.1093/mnras/stv1198](https://doi.org/10.1093/mnras/stv1198)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016, *A&A*, 595, A2, doi: [10.1051/0004-6361/201629512](https://doi.org/10.1051/0004-6361/201629512)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1, doi: [10.1051/0004-6361/201833051](https://doi.org/10.1051/0004-6361/201833051)
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, *A&A*, 650, C3, doi: [10.1051/0004-6361/202039657e](https://doi.org/10.1051/0004-6361/202039657e)
- Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, *A&A*, 674, A1, doi: [10.1051/0004-6361/202243940](https://doi.org/10.1051/0004-6361/202243940)
- Gieles, M., Erkal, D., Antonini, F., Balbinot, E., & Peñarrubia, J. 2021, *Nat Astron*, 5, 957, doi: [10.1038/s41550-021-01392-2](https://doi.org/10.1038/s41550-021-01392-2)
- Green, G. M. 2018, *JOSS*, 3, 695, doi: [10.21105/joss.00695](https://doi.org/10.21105/joss.00695)
- Grillmair, C. J. 2009, *ApJ*, 693, 1118, doi: [10.1088/0004-637X/693/2/1118](https://doi.org/10.1088/0004-637X/693/2/1118)
- Grillmair, C. J. 2022, *ApJ*, 929, 89, doi: [10.3847/1538-4357/ac5bd7](https://doi.org/10.3847/1538-4357/ac5bd7)
- Grillmair, C. J. 2025, *ApJ*, 979, 75, doi: [10.3847/1538-4357/ada2ea](https://doi.org/10.3847/1538-4357/ada2ea)
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
- Harris, W. E. 1996, *AJ*, 112, 1487, doi: [10.1086/118116](https://doi.org/10.1086/118116)
- Hattori, K., Erkal, D., & Sanders, J. L. 2016, *MNRAS*, 460, 497, doi: [10.1093/mnras/stw1006](https://doi.org/10.1093/mnras/stw1006)
- Hilker, M., Baumgardt, H., Sollima, A., & Bellini, A. 2019, *Proc. IAU*, 14, 451, doi: [10.1017/S1743921319006823](https://doi.org/10.1017/S1743921319006823)
- Holm-Hansen, C., Chen, Y., & Gnedin, O. Y. 2025, *arXiv:2510.09604 [astro-ph]*, doi: [10.48550/arXiv.2510.09604](https://doi.org/10.48550/arXiv.2510.09604)
- Hunt, J. A., & Vasiliev, E. 2025, *New Astronomy Reviews*, 100, 101721, doi: [10.1016/j.newar.2024.101721](https://doi.org/10.1016/j.newar.2024.101721)
- Hunter, J. D. 2007, *CSE*, 9, 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Ibata, R., Malhan, K., Martin, N., et al. 2021, *ApJ*, 914, 123, doi: [10.3847/1538-4357/abfcc2](https://doi.org/10.3847/1538-4357/abfcc2)
- Ibata, R., Malhan, K., Tenachi, W., et al. 2024, *ApJ*, 967, 89, doi: [10.3847/1538-4357/ad382d](https://doi.org/10.3847/1538-4357/ad382d)
- Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, *Nature*, 370, 194, doi: [10.1038/370194a0](https://doi.org/10.1038/370194a0)
- King, I. R. 1966, *AJ*, 71, 64, doi: [10.1086/109857](https://doi.org/10.1086/109857)
- Kroupa, P. 2001, *MNRAS*, 322, 231, doi: [10.1046/j.1365-8711.2001.04022.x](https://doi.org/10.1046/j.1365-8711.2001.04022.x)
- Küpper, A. H. W., Lane, R. R., & Heggie, D. C. 2012, *MNRAS*, 420, 2700, doi: [10.1111/j.1365-2966.2011.20242.x](https://doi.org/10.1111/j.1365-2966.2011.20242.x)
- Lane, R. R., Küpper, A. H. W., & Heggie, D. C. 2012, *MNRAS*, 423, 2845, doi: [10.1111/j.1365-2966.2012.21093.x](https://doi.org/10.1111/j.1365-2966.2012.21093.x)
- Li, T. S., Koposov, S. E., Zucker, D. B., et al. 2019, *MNRAS*, 490, 3508, doi: [10.1093/mnras/stz2731](https://doi.org/10.1093/mnras/stz2731)
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, *arXiv:0912.0201 [astro-ph]*, doi: [10.48550/arXiv.0912.0201](https://doi.org/10.48550/arXiv.0912.0201)
- Lynden-Bell, D., & Lynden-Bell, R. M. 1995, *MNRAS*, 275, 429, doi: [10.1093/mnras/275.2.429](https://doi.org/10.1093/mnras/275.2.429)
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, 154, 94, doi: [10.3847/1538-3881/aa784d](https://doi.org/10.3847/1538-3881/aa784d)
- Malhan, K., & Ibata, R. A. 2018, *MNRAS*, 477, 4063, doi: [10.1093/mnras/sty912](https://doi.org/10.1093/mnras/sty912)
- Malhan, K., Ibata, R. A., Carlberg, R. G., Valluri, M., & Freese, K. 2019, *ApJ*, 881, 106, doi: [10.3847/1538-4357/ab2e07](https://doi.org/10.3847/1538-4357/ab2e07)
- Mateu, C., Bruzual, G., Aguilar, L., et al. 2011, *MNRAS*, 415, 214, doi: [10.1111/j.1365-2966.2011.18690.x](https://doi.org/10.1111/j.1365-2966.2011.18690.x)
- Meštrić, U., Vanzella, E., Zanella, A., et al. 2022, *MNRAS*, 516, 3532, doi: [10.1093/mnras/stac2309](https://doi.org/10.1093/mnras/stac2309)
- Ngan, W. H. W., & Carlberg, R. G. 2014, *ApJ*, 788, 181, doi: [10.1088/0004-637X/788/2/181](https://doi.org/10.1088/0004-637X/788/2/181)
- Nibauer, J., Bonaca, A., Lisanti, M., Erkal, D., & Hastings, Z. 2024, *ApJ*, 969, 55, doi: [10.3847/1538-4357/ad4299](https://doi.org/10.3847/1538-4357/ad4299)
- Odenkirchen, M., Grebel, E. K., Rockosi, C. M., et al. 2001, *ApJ*, 548, L165, doi: [10.1086/319095](https://doi.org/10.1086/319095)
- Pace, A. B., & Li, T. S. 2019, *ApJ*, 875, 77, doi: [10.3847/1538-4357/ab0ace](https://doi.org/10.3847/1538-4357/ab0ace)
- Pearson, S., Bonaca, A., Chen, Y., & Gnedin, O. Y. 2024, *ApJ*, 976, 54, doi: [10.3847/1538-4357/ad8348](https://doi.org/10.3847/1538-4357/ad8348)
- Pearson, S., Price-Whelan, A. M., & Johnston, K. V. 2017, *Nature Astronomy*, 1, 633, doi: [10.1038/s41550-017-0220-3](https://doi.org/10.1038/s41550-017-0220-3)
- Price-Whelan, A., Souchereau, H., Wagg, T., et al. 2024, *Zenodo*, doi: [10.5281/zenodo.593786](https://doi.org/10.5281/zenodo.593786)
- Price-Whelan, A. M. 2017, *JOSS*, 2, 388, doi: [10.21105/joss.00388](https://doi.org/10.21105/joss.00388)
- Price-Whelan, A. M., & Bonaca, A. 2018, *ApJL*, 863, L20, doi: [10.3847/2041-8213/aad7b5](https://doi.org/10.3847/2041-8213/aad7b5)
- Price-Whelan, A. M., Sesar, B., Johnston, K. V., & Rix, H.-W. 2016, *ApJ*, 824, 104, doi: [10.3847/0004-637X/824/2/104](https://doi.org/10.3847/0004-637X/824/2/104)
- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, *A&A*, 649, A3, doi: [10.1051/0004-6361/202039587](https://doi.org/10.1051/0004-6361/202039587)
- Roberts, D., Gieles, M., Erkal, D., & Sanders, J. L. 2025, *Monthly Notices of the Royal Astronomical Society*, 538, 454, doi: [10.1093/mnras/staf321](https://doi.org/10.1093/mnras/staf321)

- Rockosi, C. M., Odenkirchen, M., Grebel, E. K., et al. 2002, *AJ*, 124, 349, doi: [10.1086/340957](https://doi.org/10.1086/340957)
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, 737, 103, doi: [10.1088/0004-637X/737/2/103](https://doi.org/10.1088/0004-637X/737/2/103)
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525, doi: [10.1086/305772](https://doi.org/10.1086/305772)
- Sesar, B., Price-Whelan, A. M., Cohen, J. G., et al. 2016, *ApJL*, 816, L4, doi: [10.3847/2041-8205/816/1/L4](https://doi.org/10.3847/2041-8205/816/1/L4)
- Shih, D., Buckley, M. R., Necib, L., & Tamanas, J. 2021, *MNRAS*, 509, 5992, doi: [10.1093/mnras/stab3372](https://doi.org/10.1093/mnras/stab3372)
- Shipp, N., Drlica-Wagner, A., Balbinot, E., et al. 2018, *ApJ*, 862, 114, doi: [10.3847/1538-4357/aacdab](https://doi.org/10.3847/1538-4357/aacdab)
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757 [astro-ph]. <http://arxiv.org/abs/1503.03757>
- Tavangar, K., & Price-Whelan, A. M. 2025, arXiv:2502.13236 [astro-ph], doi: [10.48550/arXiv.2502.13236](https://doi.org/10.48550/arXiv.2502.13236)
- The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- The pandas development team. 2024, pandas-dev/pandas: Pandas, Zenodo, doi: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134)
- Valluri, M., Fagrelus, P., Koposov, S. E., et al. 2024, arXiv:2407.06336 [astro-ph]. <http://arxiv.org/abs/2407.06336>
- Varghese, A., Ibata, R., & Lewis, G. F. 2011, *MNRAS*, 417, 198, doi: [10.1111/j.1365-2966.2011.19097.x](https://doi.org/10.1111/j.1365-2966.2011.19097.x)
- Vasiliev, E. 2019, *MNRAS*, 482, 1525, doi: [10.1093/mnras/sty2672](https://doi.org/10.1093/mnras/sty2672)
- Vasiliev, E. 2023, *Galaxies*, 11, 59, doi: [10.3390/galaxies11020059](https://doi.org/10.3390/galaxies11020059)
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Yang, Y., Zhao, J.-K., Tang, X.-Z., Ye, X.-H., & Zhao, G. 2023, *ApJ*, 953, 130, doi: [10.3847/1538-4357/acdee2](https://doi.org/10.3847/1538-4357/acdee2)