

Maximising Survey Volume for Large Area Multi-Epoch Surveys with Voronoi Tessellation

Marco. C. Lam¹*

¹*Institute for Astronomy, University of Edinburgh, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*

original form 2017 January 16

ABSTRACT

The survey volume of a proper motion-limited sample is typically much smaller than a magnitude-limited sample because of the noisy astrometric measurements from detectors that are not dedicated for astrometric missions. The traditional analysis method limits the survey depth to the worst parts of the sky, which in turn limits the maximum potential of a survey. This work proposes a method to dissect the survey into small parts with Voronoi tessellation using the objects of interest as generating points, such that each part defines a ‘mini-survey’ that has its own properties. The new method is demonstrated through the application of a maximum volume density estimator to a mock catalogue of a white dwarf-only solar neighbourhood with Pan-STARRS1-like characteristics. The bias due to the choice of model as well as the accuracy due to resolution are discussed. Towards the end of this work, we demonstrate a method to increase the resolution for such treatment, which is useful for analysis of rare objects where artificial generating points are needed to provide sufficient resolution.

Key words: proper motions – surveys – stars: luminosity function, mass function – white dwarfs – solar neighbourhood – methods: data analysis.

1 INTRODUCTION

Transient, variable and moving objects are not rare, but their detections require multiple observations of the same area of sky at different time. This was not possible over a large sky area until the era of digital astronomy. The highly automated observing runs and efficient digital detectors allow efficient data collection, while faster processors and the automated data reduction pipelines produce high volume of output.

Early efforts producing proper motion object catalogues involved eyeball identification with two epochs of images, e.g. Luyten Half-Second catalogue (Luyten 1979), New Luyten Ten-Tenth catalogue (Luyten 1979). The comparatively large size of these catalogues enabled much more robust statistical analyses of objects in the solar neighbourhood. However, the completeness is unquantifiable on its own, especially in the low proper motion regime, which is essential for deriving, for example, the density functions and mass functions. The man-hours behind these projects were enormous and the repetition of measurements required just as much time. The situation changed significantly when the photographic plates from large sky area surveys were digitised in the PPM catalogue (Röser & Bastian 1991; Bastian & Röser 1993), Automated Plate Machine Project (Evans 1992; Evans & Irwin 1995), USNO A 1.0, A 2.0 & B 1.0 (Monet

1996, 1998; Monet et al. 2003), SuperCOSMOS (Hambly et al. 2001; Hambly, Irwin, & MacGillivray 2001; Hambly et al. 2001), UCAC 1, 2, 3 & 4 (Zacharias et al. 2000, 2004, 2010, 2013) and SUPERBLINK (Lépine, Shara, & Rich 2002, 2003; Lépine 2005, 2008), all of which had several epochs and simple tiling strategies.

In the current era of digital astronomy, some surveys continue to use simple tiling patterns where multiple pawprints are combined immediately to produce full coverage over a sky cell, for example in the UKIDSS (Lawrence et al. 2007), four pawprints can cover a cell and six for VISTA (Sutherland et al. 2015). In other cases, SDSS Stripe 82 had nine epochs on average (Bramich et al. 2008), ALLWISE has a coverage from 12 to over 200 frames (Kirkpatrick et al. 2014, 2016), the Pan-STARRS1 (PS1) 3 π Steradian Survey (3SS) typically has 60 epochs (Magnier et al. 2013), The Dark Energy Survey will scan $\sim 5,000$ sq. degrees 10 times (The Dark Energy Survey Collaboration 2005), Gaia will have on average 81 transits, with over 140 at the deepest parts at the end of the 5-year nominal mission (Gaia Collaboration et al. 2016), LSST will provide close to 1,000 epochs for half of the sky towards the end of the 10-year survey mission (LSST Science Collaboration et al. 2009)¹. The tiling strategies/overlapping patterns for these surveys are, however, extremely complicated in order to maximise coverage due to losses from chip

* E-mail: mlam@roe.ac.uk

¹ <https://github.com/LSSTScienceCollaborations/>

gaps between CCDs and unfavourable observing conditions. Even though the survey configurations can be accurately described by the observing log, with trillions of pixels to cover the sky once (~ 1.5 Gigapixels for ~ 7 sq. degrees in the case of a single pointing of PS1), it is not feasible to process all sky data down to pixel-by-pixel level for most studies. This problem becomes even more complex when different surveys are combined to expand the wavelength coverage and/or maximum epoch difference, for example SDSS is combined with USNO-B 1.0 to derive the proper motions (Munn et al. 2004; Gould & Kollmeier 2004), the survey typically has 5 epochs (4 from USNO-B, 1 from SDSS).

The traditional methods to analyse such huge data sets with large inhomogeneity are to limit the analysis to the shallower part of the survey in order not to run into unaccountable incompleteness. In view of this problem, when deriving the white dwarf luminosity function (WDLF) with a maximum volume density estimator, Rowell & Hambly (2011, hereafter RH11) measured the empirical photometric and astrometric uncertainty for different skycells as defined by the tiling of the photographic plates. Lam, Rowell, & Hambly (2015, hereafter LRH15) improved the completeness correction due to kinematic selections. The methods described in these works allow an analysis to probe deeper, but in order to maximise use of the data, we propose a new method that can further maximise the analytical survey volume that completeness and other biases can be corrected. Voronoi tessellation has been employed in defining simulation grid, clustering analysis, visualisation etc. However, its property that partitions objects into well defined grids has not been used in all domains of Astrophysics. By dividing the sky through Voronoi tessellation into a number of cells that is equal to the number of candidate objects, we can treat each cell as a mini-survey that has well-defined local properties. Analysis can also be performed at lower or higher resolution if needed.

In Section 2, we will describe the mathematical framework and the construction of the simulated solar neighbourhood in Section 3. The new method is applied to the simulated data in Section 4 under different selection criteria and we describe one procedure to through which the resolution can be increased. The bias due to the choice of model is briefly discussed. In the final section, we discuss other possible way to increase the resolution for the analysis and conclude this work.

2 MATHEMATICAL FRAMEWORK ON THE IMPLEMENTATION OF 1/VMAX METHOD

The maximum volume density estimator (Schmidt 1968) tests the observability of an object by finding the maximum volume in which it can be observed by a survey (e.g. at a different part of the sky at a different distance). It is proven to be unbiased (Felten 1976) and can easily combine multiple different surveys (Avni & Bahcall 1980). In a sample of proper motion objects, we need to consider both the photometric and astrometric properties (see LRH15 for details). The number density is found by summing the number of objects weighted by the inverse of the maximum volumes. For surveys with small variations in quality from field to field and from epoch to epoch, or with small survey footprint areas, the survey limits can be defined easily. However, in modern surveys, the variations are not small; this is especially true for ground-based observations. Therefore, properties have to be found locally to analyse the data most accurately. Through the use of Voronoi tessellation, objects can be partitioned into individual 2D cells within which we assume the

sky properties are defined by the governing object. Each of these cells has a different area depending on the projected density of the population.

2.1 Voronoi Tessellation

A Voronoi tessellation is by partitioning a plane with n points into n convex polygons such that each polygon contains one point. Any position in a given polygon is closer to its generating point than to any other for the case of Voronoi tessellation using Euclidean distance ($D_E = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$), as opposed to Manhattan distance ($D_M = |x_1 - x_2| + |y_1 - y_2|$). For the use in Astronomy, such tessellation has to be constructed on a spherical surface (2-sphere).

In the follow work, the tessellation is constructed with the SCIPY package `spatial.SphericalVoronoi` where each polygon is given a unique ID which is combined with the vertices to form a dictionary. The areas are calculated by first decomposing the polygons into spherical triangles with the generating points and their vertices², and then using L'Huilier's Theorem to find the spherical excess. For a unit-sphere, the spherical excess is equal to the solid angle of the triangle. The sum of the constituent spherical triangles provides the solid angle of each cell.

The major assumption for the using of Voronoi Tessellation is that for a statistical sample, the probability distribution of the survey depth/completeness at where objects are found is the same as that for the entire survey, so the volume is very similar to the exact solution.

2.2 Object Properties

For a given object x in Voronoi cell i , the properties of the cell area are assumed to be represented by the given object parameters. The epoch of the measurement is labelled by j , which is essential in calculating the photometric and proper motion uncertainties as functions of distance when tested for observability. The major difference in the following approach is that the proper motion uncertainty is found from the formal propagation of errors instead of measuring the empirical form as a function of magnitude, $\sigma_\mu(mag)$, which limits the survey to the worst part of a tile (RH11). This new approach does not need to take into account the scatter in $\sigma_\mu(mag)$, which is mostly caused by different local sky properties of the object and its neighbours as well as their colours. Different types of object can differ by up to a few magnitudes in the optical/infrared colours: when two objects with similar magnitudes in one filter can have very different proper motion uncertainties if one is close to the detection limit in another filter. The modelling of the photometric uncertainties from CCD detectors is simpler than that for photographic plates where the photometric response is far from linear at both the faint and bright ends. Thus, the uncertainties can be estimated with relatively simple equations.

2.2.1 Photometric Uncertainty

The noise, N , of a photometric measurement can be estimated by

$$N = \sqrt{(F + d + s) \times t + rn^2} \quad (1)$$

² https://github.com/tylerjereddy/py_sphere_Voronoi

where F is the instrumental flux, d is the dark current, s is the sky background flux, t is the exposure time and rn is the read noise. Among these quantities, d , s , t and rn are fixed quantities in a given epoch. When the observability is tested, only F varies as a function of distance, D . Therefore, in a Voronoi cell i at epoch j , the photometric noise of object x is

$$N_{x,i,j}(D) = \sqrt{(F_{x,i}(D) + d_{i,j} + s_{i,j}) \times t_{i,j} + rn^2} \quad (2)$$

where

$$F_{x,i}(D) = F_x \times \left(\frac{D_x}{D}\right)^2 \times c_i. \quad (3)$$

The index i for the $F_{x,i}$ and the correction factor, c_i , are redundant in this work, so c_i is set to 1. However, they are needed if interstellar reddening or other positional dependant corrections are applied. The random photometric uncertainty of an object at an arbitrary distance in a given epoch is therefore,

$$\delta f_{x,i,j}(D) = \frac{N_{x,i,j}(D)}{F_{x,i}(D)}. \quad (4)$$

The total uncertainty combining with the systematic uncertainty, σ_s , coming from the absolute calibration of the detector is therefore

$$\sigma_{x,i,j}(D) = \sqrt{\delta f_{x,i,j}^2(D) + \sigma_s^2}, \quad (5)$$

which represents the photometric uncertainty as a function of the distance to the object. This is to model the change in the uncertainties due to varying apparent magnitude with distance.

2.2.2 Astrometric Uncertainty

The computation of proper motion in one co-ordinate by the least-square method can be expressed in the following matrix form for object x (omitted in the matrix for clarity), and the epoch is labelled by the subscript 1 to j .

$$\underbrace{\begin{pmatrix} \frac{1}{\sigma_0} & \frac{\Delta t_0}{\sigma_0} \\ \frac{1}{\sigma_1} & \frac{\Delta t_1}{\sigma_1} \\ \vdots & \vdots \\ \frac{1}{\sigma_j} & \frac{\Delta t_j}{\sigma_j} \end{pmatrix}}_{\mathbf{A}} \times \begin{pmatrix} \bar{\alpha} \\ \mu_\alpha \end{pmatrix} = \begin{pmatrix} \frac{\Delta \alpha_0}{\sigma_0} \\ \frac{\Delta \alpha_1}{\sigma_1} \\ \vdots \\ \frac{\Delta \alpha_j}{\sigma_j} \end{pmatrix} \quad (6)$$

where Δt_j is the time difference between the mean epoch and epoch j , $\Delta \alpha_j$ is the positional offset from the mean position in the direction of the right ascension and μ_α is the proper motion in the direction of the right ascension. The associated uncertainties can be found from the diagonal terms of the normal matrix $\mathbf{A}^T \mathbf{A}$ because the position and proper motion are uncorrelated,

$$\begin{pmatrix} \frac{1}{\sigma_\alpha^2} & 0 \\ 0 & \frac{1}{\sigma_{\mu_\alpha}^2} \end{pmatrix} = \mathbf{A}^T \mathbf{A} \quad (7)$$

i.e. for each cell,

$$\frac{1}{\sigma_{\mu_\alpha}^2} = \sum_j \left(\frac{\Delta t_j}{\sigma_j}\right)^2 \quad (8)$$

and the total proper motion uncertainty is

$$\sigma_\mu = \sqrt{\sigma_{\mu_\alpha}^2 + \sigma_{\mu_\delta}^2} = \sqrt{2} \sigma_{\mu_\alpha} \quad (9)$$

because the uncertainties in the α and δ directions are symmetrical in 4-parameter astrometric solution (two positions and two proper motions). In the case of 5-parameter solution where parallax is solved and for the 7-parameter solution where, in addition, the acceleration terms in both directions are solved for, the uncertainties will not be symmetrical due to the parallactic term, so the off-diagonal terms have to be taken into account which would otherwise be negligible compared to the diagonal terms.

2.3 Consequence to Vmax Calculation

There is only one minor adjustment to the volume integral – the lower proper motion limit. Instead of finding the limit by measuring from a number of nearby objects which includes mostly objects with different colours, the limit is defined by the properties of the Voronoi cell which comes only from the generating object of the cell. For object x , the maximum volume has to be tested in each Voronoi cell i , the expression is almost identical to that in LRH15, except for the i and k indexes

$$V_{\max} = \sum_i \Omega_i \int_{D_{\min,i}}^{D_{\max,i}} \frac{\rho(D)}{\rho_\odot} \times D^2 \times \left[\int_{a(D)}^{b(D)} P_{k(i)}(v_T) dv_T \right] dD \quad (10)$$

where P is the tangential velocity distribution, $k(i)$ denotes the HEALPix ID corresponding to Voronoi cell i with area Ω_i , v_T is the tangential velocity, D_{\min} and D_{\max} are the minimum and maximum photometric distances, and $\sigma_\mu(D)$ is the proper motion uncertainty as a function of the distance to the object. This is to model the change in the proper motion uncertainties due to varying apparent magnitude with distance (i.e. at greater distance the proper motion uncertainty will be larger because the object becomes fainter which increases the single-epoch positional uncertainty).

The inner integral can vanish before reaching the distance limits so the integrator must use a small step size or the distances at which the inner integral vanish have to be calculated explicitly. The cell ID i and HEALPix ID k can be set as a one to one mapping by calculating the tangential velocity distribution for each of the Voronoi cell, but since the tangential velocity correction is fixed on the sky, while the Voronoi tessellation is dependent on the sample, the use of a tangential velocity correction look-up table reduces computation time. The lower tangential velocity limit in the inner integral, $a(D)$, is

$$a(D) = \max[v_{\min}, 4.74 \times s \times \sigma_\mu(D) \times D] \quad (11)$$

where the factor 4.74 comes from the unit conversion from arc-second per year to kilometer per second at distance D , v_{\min} is the global lower tangential velocity limit and s is the significance of the proper motions. The expression is identical to that in LRH15 but the $\sigma_\mu(D)$ is calculated from a completely different way.

3 SIMULATED DATA SET

To demonstrate the power of the Voronoi tessellation method described in Section 2, we apply it to a data set of simulated solar neighbourhood. This section details the construction of the Monte Carlo simulation.

We generated snapshots of white dwarf (WD)-only solar neighbourhoods containing six dimensional phase space information. The procedure is very similar to that described in LRH15,

Table 1. Physical properties of the Galaxy used in the Monte Carlo simulation.

Parameter	Thin disk	Thick disk	Stellar Halo
$\langle U \rangle / \text{km s}^{-1}$	-8.62 ^a	-11.0 ^d	-26.0 ^d
$\langle V \rangle / \text{km s}^{-1}$	-20.04 ^a	-42.0 ^d	-199.0 ^d
$\langle W \rangle / \text{km s}^{-1}$	-7.10 ^a	-12.0 ^d	-12.0 ^d
$\sigma_U / \text{km s}^{-1}$	32.4 ^a	50.0 ^d	141.0 ^d
$\sigma_V / \text{km s}^{-1}$	23.0 ^a	56.0 ^d	106.0 ^d
$\sigma_W / \text{km s}^{-1}$	18.1 ^a	34.0 ^d	94.0 ^d
H/pc	250 ^b	1000 ^e	∞
n/pc ⁻³	0.00310 ^c	0.00064 ^c	0.00019 ^c

^a Fuchs, Jahreiß, & Flynn 2009^b Mendez & Guzman 1998^c RH11^d Chiba & Beers 2000^e Sandage & Fouts 1987

however we introduce changes to the noise model of the system. The volume probed is assumed to be small such that the simulation is performed in a Cartesian space, instead of a plane polar system centred at the Galactic Centre (GC). The Galaxy has three distinct kinematic components: a thin disc, a thick disk and a stellar halo, all of which we model with no density variations along the co-planar direction of the Galactic plane. All vertical structures follow exponential profiles, with scale height $H_{\text{population}}$. None of the three components are tilted relative to each other. The velocity components, U , V and W , of each object are drawn from the Gaussian distributions constructed from the measured means and standard deviations of the three sets of kinematics that describes the three populations in the solar neighbourhood (Table 1). To generate WD populations, the thin and thick disk populations are assigned with constant star formation rates since look back time, $\tau = 8$ Gyr and $\tau = 10$ Gyr respectively, while the halo has a star burst of duration 1 Gyr at $\tau = 13$ Gyr. The initial mass function has an exponent of -2.3 in the mass range of interest (Kroupa 2001), and the initial-final mass relation follows the one in Kalirai et al. (2009) without including globular clusters in the analysis where the final WD mass is

$$\mathcal{M}_f = 0.101\mathcal{M}_i + 0.463 \mathcal{M}_{\odot}. \quad (12)$$

The main sequence (MS) life time has to be added in order to calculate the cooling time, and hence the magnitude of a WD. We have adopted the stellar evolution tracks from the Padova group (PARSEC; Bressan et al. 2012) with a metallicity of $Z = 0.019$ and $Y = 0.30$ (Girardi et al. 2000). To generate a theoretical luminosity function (LF), we couple this WD population with a cooling model based upon the assumption that all WDs have surface gravity $\log g = 8.0$ and pure hydrogen atmosphere (DA)³ (Holberg & Bergeron 2006; Kowalski & Saumon 2006; Tremblay, Bergeron, & Gianninas 2011; Bergeron et al. 2011). This LF is used as the probability distribution function (PDF) in the simulation, and we take the PDF normalisations from the WD densities found in RH11. The input parameters are assumed to be invariant with time and are summarised in Table 1.

From the distance and bolometric magnitude drawn from the PDF, we calculate the apparent magnitudes in the Pan-STARRS g_{P1} , r_{P1} , i_{P1} , z_{P1} and y_{P1} filters⁴ (Tonry et al. 2012; Schlafly et al.

Table 2. Parameters for the sky background count.

Filter	Mean (photon s ⁻¹)	Standard Deviation (photon s ⁻¹)
g	39.972	12.355
r	135.728	47.981
i	262.725	72.499
z	257.930	90.175
y	272.195	72.493

2012; Magnier et al. 2013; Chambers et al. 2016; Waters et al. 2016; Magnier et al. 2016; Flewelling et al. 2016; Magnier et al. 2016,?). The uncertainties in those filters, σ_{m_i} , are found by calculating from the sky background flux, exposure time, dark current and read noise that are representative of the PS1 3SS at Processing Version 2 (PV2). The sky background flux is drawn from a Gaussian distribution measured from the 3SS in each of the filters (Table 2). The means and standard deviations⁵ were measured from 100 fields drawn randomly across the survey footprint area.

To simulate the variations in the observing properties of the sky, HEALPix is used to pixelate the sky using a resolution of $N_{\text{side}} = 256$, i.e. each has a size of ~ 0.0534 sq. degrees, which is sufficiently small compared to the projected density of white dwarfs which is less than 1 per sq. degree (e.g. RH11). Each pixel is given the sky background noise at each epoch of the measurement. If the sky noise is less than 1 photon s⁻¹, we reassign it to 1 photon s⁻¹. An Analogue-to-Digital Unit (ADU) of 1 e⁻ per photon is assumed. The 3SS has 12 epochs on average in each of the filters, so the number of epochs for each object in the simulation is drawn from a distribution⁶ that follows $1 + P(11)$ where $P(11)$ is Poisson distribution with a mean of 11, the epochs are drawn from a random distribution with a period of 3 years. When objects are distributed over the sky, they will take the set of values defined by the nearest pixel. Using the treatments from Section 2.2.1, with a dark current of 0.2 s⁻¹, exposure times of 43, 40, 35, 30 and 30 seconds in the five filters, a constant read noise of 5.5 photons (Metcalf et al. 2013), each object is assigned with proper motion uncertainty using Eq. 9. These values produce an all sky survey that has 10 σ detections (and standard deviation) in g_{P1} , r_{P1} , i_{P1} , z_{P1} and y_{P1} at 22.43 ± 0.06 mag, 21.83 ± 0.09 mag, 21.52 ± 0.06 mag, 21.33 ± 0.10 mag and 21.27 ± 0.05 mag. These values are similar to PV2 values, but the distribution is much narrower because the noise model is noiseless (e.g. no diffraction spikes, optical ghosts, cosmic rays or other effects that lead to large uncertainties).

4 APPLICATION TO WDLFS

This section describes the application of our survey dissection method (presented in Section 2) to simulated PS1-like WD catalogues generated using the recipe described in Section 3. The bright limits at all filters are set at 15 mag. The faint limits are at 21.5, 21.0, 20.5, 20.0 & 19.5 mag in g_{P1} , r_{P1} , i_{P1} , z_{P1} and y_{P1} filters respectively, which are the typical magnitudes at which the 3SS is complete. The lower proper motion limit is set to five times the

⁵ 1.4826 times the median absolute deviation is used for robust estimation of the standard deviation

⁶ This study only focuses on tessellation, the effect of non-detection is another huge step in the optimisation of analysis

³ <http://www.astro.umontreal.ca/~bergeron/CoolingModels/>

⁴ <http://panstarrs.stsci.edu/>

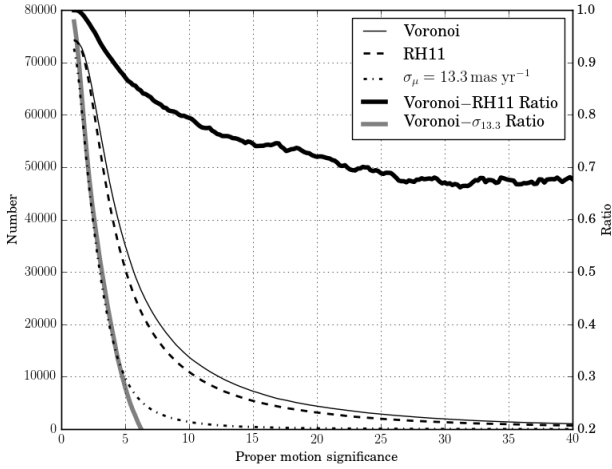


Figure 1. The number of objects recovered using the Voronoi and RH11 methods in a mixed thin disc, thick disc and halo simulation as a function of proper motion significance. The Voronoi method is plotted as a thin solid line, RH11 method is plotted as a dashed line, the using of a global proper motion uncertainty at 95th percentile (i.e. 13.3 mas yr⁻¹) instead of RH11 tiling is in dot-dashed line. The ratios between the two methods to the Voronoi method are plotted as thick solid lines and correspond to the ordinate axis on the right.

proper motion uncertainties, σ_μ , unless specified otherwise; and the upper proper motion limit is set at $0.5'' \text{ yr}^{-1}$. The tangential velocity limits are different in each analysis to explore different scenarios. The photometric parallaxes were not derived, the real distances and bolometric magnitudes are used. The volume is found by integrating Eq. 10 from D_{\min} to D , and the maximum volume from D_{\min} to D_{\max} .

4.1 Comparison with the RH11 selection

The RH11 method increases the survey volume by restricting shallow survey depths only over areas that are severely limited by a small number of poor observations. In this section we illustrate how the Voronoi method can further increase the number of objects that can be recovered while rigorous completeness correction can be performed. In Fig. 1 & 2, the Voronoi method is plotted as a solid line, RH11 method as a dashed line and the ratio between the two method in thick solid line. With the Voronoi method, more objects can be recovered (Fig. 1) and this ratio is fairly constant over the range of absolute bolometric magnitudes (Fig. 2). Under a very high proper motion significance selection, the ratio appears to converge at ~ 0.7 , this value is survey- and kinematics model-dependent, so it should not be interpreted as the fraction of discoverable high proper motion objects left out in the previous analyses.

Due to the noiseless nature of the simulation (i.e. no bad pixels, saturation, diffraction, optical ghosts and other effects that affect the photometric and astrometric precision significantly), the size of the proper motion uncertainties in the simulation is typically smaller than real measurements. Nevertheless, at 5σ level the Voronoi method can recover $\sim 15\%$ more objects than the RH11 method and much more objects than applying a global lower proper motion limit (Fig. 2). At lower proper motion significances, the RH11-to-Voronoi ratio slowly increases with magnitude. That is because faint objects can only be seen from a small distance so they tend to have large proper motions. At high proper motion significance, the same effect applies but is extended to brighter objects,

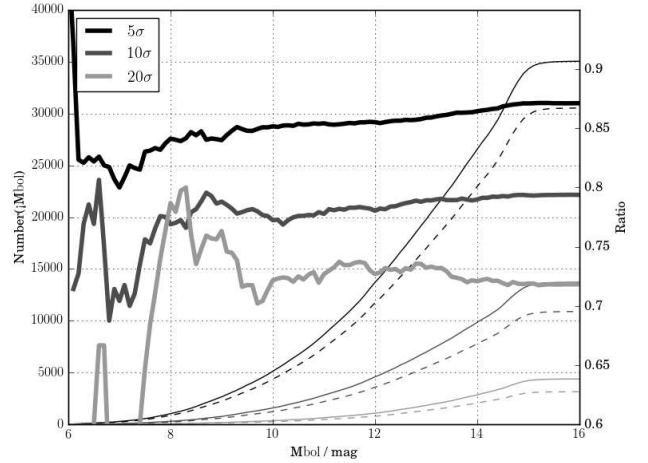


Figure 2. The number of object discovered as a function of absolute bolometric magnitude with the Voronoi method (solid) and RH11 method (dashed) at 5σ (black), 10σ (grey) and 20σ (light grey) level, values correspond to the ordinate axis on the left. The ratios between the two methods are in thick solid lines and correspond to the ordinate axis on the right.

hence moving from 5σ to 10σ and then to 20σ , the gradient flattens. At the bright end, the noisy features are due to small number of objects.

4.2 Thin Disc and Combined Discs

Study of the thin disc WDLF requires a selection of the low-velocity population in order to minimise the contaminations from older populations, which typically possess higher velocities. In this section, we show the observed WDLFs from a thin disc-only simulation and from a mixed thin disc, thick disc and halo simulation. The WDLF comparison plots are displayed with the WDLF in the top panel, differences between the input and calculated WDLFs and the $\langle V/V_{\max} \rangle$ as a function of bolometric magnitude are in the middle and at the bottom panels respectively.

4.2.1 Thin disc-only sample

In an analysis selecting only thin-disc WDs, the observed WDLF agrees very closely with the input function down to $M_{\text{bol}} \approx 15$ when the number of objects drops significantly in real term (Fig. 3). From the $\langle V/V_{\max} \rangle$ (mag) distribution, the derived solution is very stable throughout, except at the brightest and the faintest ends where the $\frac{1}{V_{\max}}$ method is known to become less reliable as the number of objects decreases. $\langle V/V_{\max} \rangle$ tends to 0.5: the expected value of a uniform distribution between 0 and 1. With a small number of objects, it is expected that the value may not necessarily lie close to 0.5 and only $\sim 60\%$ of the time it lies within the error bar (the uncertainty in $\langle V/V_{\max} \rangle$ is $\frac{1}{\sqrt{12N}}$). The small oscillation about the line at $\langle V/V_{\max} \rangle$ (mag) = 0.5 is a good indication that the sample is unbiased over a large dynamic range of magnitudes. The outliers at the extreme ends result from the application of the density estimator to a small number of objects, and so likely do not represent the true values. Taking 40 and 60 km s⁻¹ as the lower and upper tangential velocity limits of the inner integral (Eq. 10 & 11 and the equivalent set of the upper limit), the total integrated number density of the work is $3.23 \pm 0.13 \times 10^{-3}$

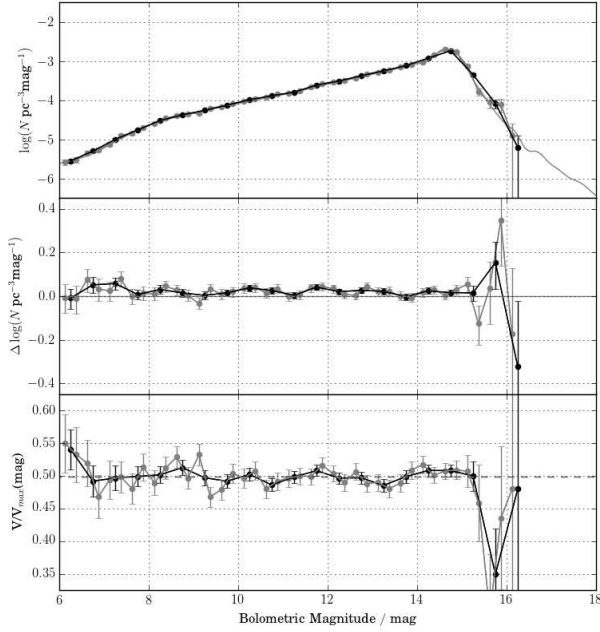


Figure 3. *Top:* The WDLFs for a thin disc-only realisation using the new method, where 0.5 mag and 0.25 mag binning are displayed in black and grey respectively. The light grey line shows the input function. *Middle:* the deviations of the WDLF as a function of absolute bolometric magnitude. *Bottom:* the $\langle V/V_{\max} \rangle$ as a function of absolute bolometric magnitude.

pc^{-3} with half-magnitude binning and $3.23 \pm 0.18 \times 10^{-3} \text{ pc}^{-3}$ with quarter-magnitude binning, compared to the input $3.10 \times 10^{-4} \text{ pc}^{-3}$, both show very good agreement. The larger uncertainty in the small binning is due to large uncertainties at the faintest magnitude bins attributed from small number of objects. For the entire sample, $\langle V/V_{\max} \rangle = 0.4991 \pm 0.0025$ which indicates an unbiased analysis.

4.2.2 Mixed Population (40 – 60 km s^{-1})

The modification to the density estimator itself is small, it is not expected that the effect due to contamination should differ from the previous analysis in LRH15. The extra depth enabled by the new method could have lead to a significant increase in the measured density due to a combination of two effects: (1) an increase in contamination fraction as the thin disc contribution drops rapidly with distance: at the Galactic poles, the thin disc and thick disc densities equate at 525 pc; and (2) the kinematic completeness correction applied on contaminants, which are more common at fainter magnitudes.

The kinematics of the two discs are well measured, however, the relative density of WDs in them are much less studied – only measured once in RH11. To understand the effects of contamination, a better understanding of the two populations is needed. Nevertheless, we can compare the WDLFs from the last section to a mixed population with the same set of upper and lower tangential velocity limits (40 and 60 km s^{-1} , Fig. 4). The total integrated number density is $3.97 \pm 0.18 \times 10^{-3} \text{ pc}^{-3}$ as compared to $3.10 \times 10^{-3} \text{ pc}^{-3}$ for the thin disc and $0.64 \times 10^{-3} \text{ pc}^{-3}$ for the thick disc, which sum to $3.74 \times 10^{-3} \text{ pc}^{-3}$. If it is treated as a pure thin disc WDLF, there is a roughly constant overestimation of 0.1 dex at all magnitude range. When both discs are considered, the small over-density is due to contamination from the thick disc where the using

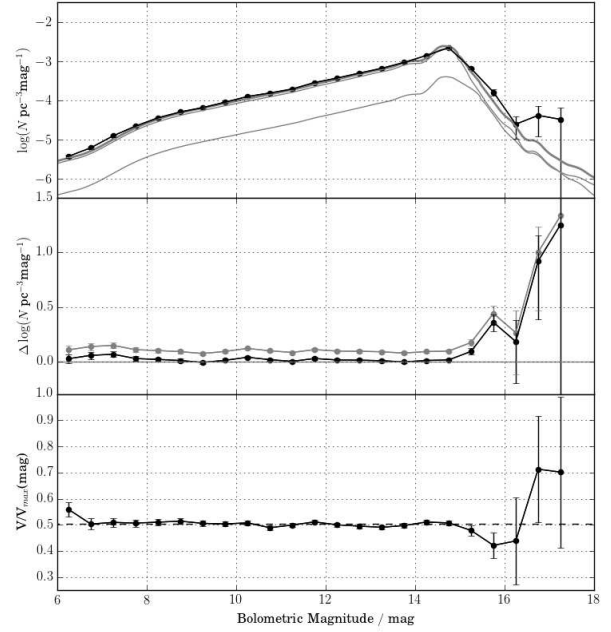


Figure 4. *Top:* WDLF for all objects with tangential velocities between 40 and 60 km s^{-1} from a mixed-population catalogue using the new method assuming thin disc properties bins (black), the input thin and thick disc luminosity functions (thin grey line) and the combined thin and thick disc luminosity function (thick grey line). *Middle:* the deviations of the WDLF as a function of absolute bolometric magnitude from the thin disc luminosity function (grey) and from the combined luminosity function (black). *Bottom:* the $\langle V/V_{\max} \rangle$ as a function of absolute bolometric magnitude.

of a thin disc scale-height on these objects will lead to an overestimation of the maximum volume. In this simulation, 16.0% of the data are from the thick disc.

The $\langle V/V_{\max} \rangle$ distribution is very similar to the clean sample, and for the entire sample $\langle V/V_{\max} \rangle = 0.5023 \pm 0.0023$ which is just within one standard deviation of statistical noise. We believe this velocity range is a good choice for driving an upper limit of the thin disc white dwarf density in the solar neighbourhood. The feature at the faintest end has to be discarded when the WDLF is used for analysis that is sensitive to all luminosities. This is also a good indication that if we fail to see such features in the observed sample, the reported thick disc density may be too high.

4.2.3 Mixed Population (40 – 120 km s^{-1})

In M16, a choice of 40 and 120 km s^{-1} are used as the limits to study both discs together, with a scale-height of 300 pc instead of 250 pc. From H06, it is known that the effect of scale-height is larger at the bright end because objects can be seen at a larger distance hence the density-correction is larger in Eq. 10. The effect on the total normalisation is small because faint objects dominate after density correction. However, studies in, for example, star formation history (Rowell 2013) or high energy exotic particles (Isern et al. 2008) are sensitive to the whole range of magnitudes. This cannot simply be assumed to be negligible. Fig. 5 investigates this effect by comparing the cases of 250 and 300 pc scale-heights, it shows that the choice of scale-height has almost no effect to the WDLFs except at the brightest magnitudes, and for the distribution of $\langle V/V_{\max} \rangle$, the differences are negligible. However, the absolute normalisation of both scale-heights are consistently overestimated by ~ 0.1 dex

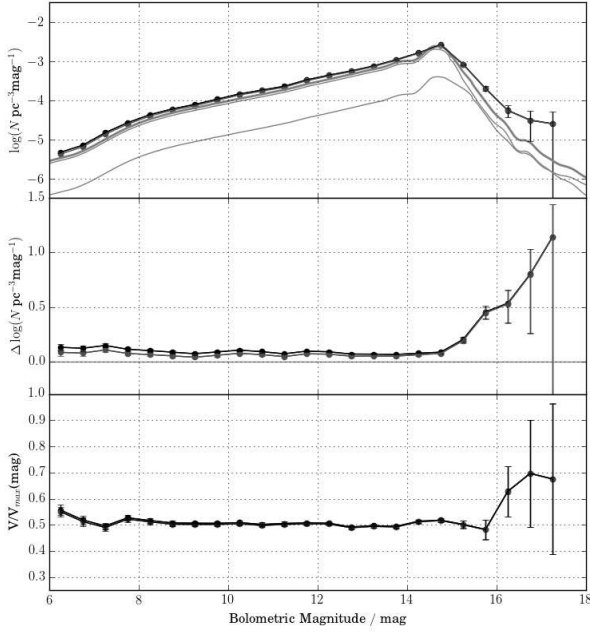


Figure 5. *Top:* WDLFs for all objects with tangential velocity between 40 and 120 km s^{-1} from a mixed-population catalogue using the new method assuming thin disc properties. Black uses a scale-height of 250 pc, grey uses 300 pc and the light grey line is the input function. *Middle:* the deviations of the WDLF as a function of magnitude. *Bottom:* the $\langle V/V_{\text{max}} \rangle$ as a function of magnitude.

and diverge further beyond 15 mag when the number density of the thick disc becomes comparable to that of the thin disc. Because of the larger scale-height correction to the maximum volume of the thin disc objects as compared to the thick disc and halo objects (see Fig. 5 of LRH15), the non-negligible amount of contaminations causes an overestimation in the number density.

In this simulation, 28.3% of the objects are from the thick disc and the halo in the range of $40 - 120 \text{ km s}^{-1}$, in comparison, only 16.0% of objects are not from the thin disc in the $40 - 60 \text{ km s}^{-1}$ selection, the derived WDLF is only discrepant fainter than $\sim 16 \text{ mag}$. From M16, we do not see any excess at the faint magnitudes and their WDLF shows a rapid decrease in the number density. The most likely explanation is that the thick disc number density reported in RH11 is too high so the contamination rate from the thick disc is much lower than in this simulation, hence the plateau at 16 – 17 mag is not observed in M16. Another possibility is that

4.3 Halo

The study of the halo WDLF requires a selection of high velocity population in order to minimise the contaminations from the thick disc which carry high velocities. In this section, we will show the observed WDLFs from a halo-only simulation and from a mixed thin disc, thick disc and halo simulation.

4.3.1 Halo-only sample

In the halo-only simulation, the observed WDLF agrees very well with the input function (Fig. 6). From the $\langle V/V_{\text{max}} \rangle(\text{mag})$ distribution, the derived solution is stable throughout, except at the faintest bin where there is only one object, as indicated by the lower

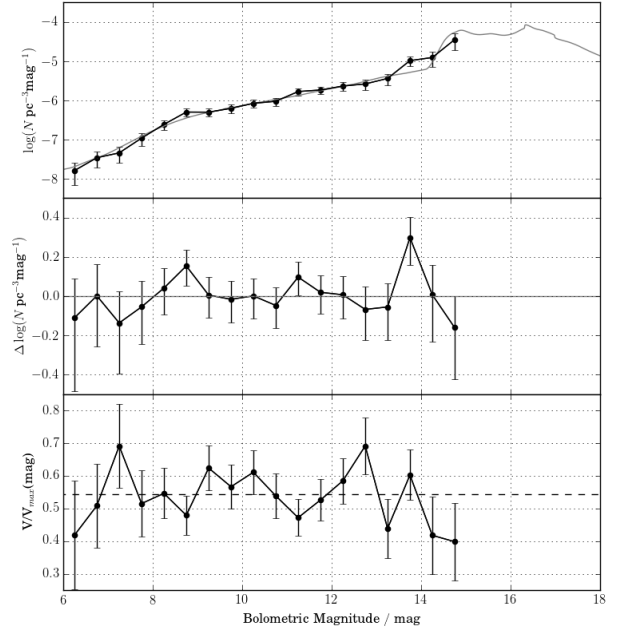


Figure 6. *Top:* observed WDLFs (black) for a halo-only realisation using the new method and the input function (grey). *Middle:* the deviations of the WDLF as a function of absolute bolometric magnitude. *Bottom:* the $\langle V/V_{\text{max}} \rangle$ as a function of absolute bolometric magnitude.

bound of the error bar that goes to negative infinity in logarithmic space. The small oscillation about the line at $\langle V/V_{\text{max}} \rangle = 0.5$ is a good indication that the sample is unbiased over the entire range of magnitudes. The upper and lower tangential velocity limits are set at 200 and 500 km s^{-1} which define the survey limits of the inner integral (Eq. 10 & 11 and the equivalent set of the upper limit), the total integrated number density of the work is $1.47 \pm 0.56 \times 10^{-4} \text{ pc}^{-3}$, compared to the input $1.90 \times 10^{-4} \text{ pc}^{-3}$ (the integrated density up to 15 mag is $1.68 \times 10^{-4} \text{ pc}^{-3}$). These are in very good agreement given that the faintest observed bin is the at where the number density is 10 times denser than the bin half a magnitude brighter.

4.3.2 Mixed Population ($200 - 500 \text{ km s}^{-1}$)

In 10 realisations of the halo-only simulation, under the selection of $200 - 500 \text{ km s}^{-1}$, there is a mean contamination rate of 6.4% (minimum at 3.1% and maximum at 9.7%). These few contaminants have little effect on the WDLFs so the samples should be representative of the halo, we therefore do not consider them further.

4.4 Sensitivity to resolution

The purpose for this new method is to tackle complex survey that has small scale variations which reduces the maximum survey volume that is available to a study. The way this method divides the sky has avoided the detailed treatment of the survey and approximates it with the object properties that goes into the analysis, which raises a concern that whether the resolution for a small sample would be sufficient, and not causing significant systematic bias. We suggest a method to increase the resolution by a factor of ~ 3 , which can be repeated to further increase the resolution if needed, as it would be useful if there are only a few objects over the sky. A higher resolution can be achieved by using the vertices as new generating

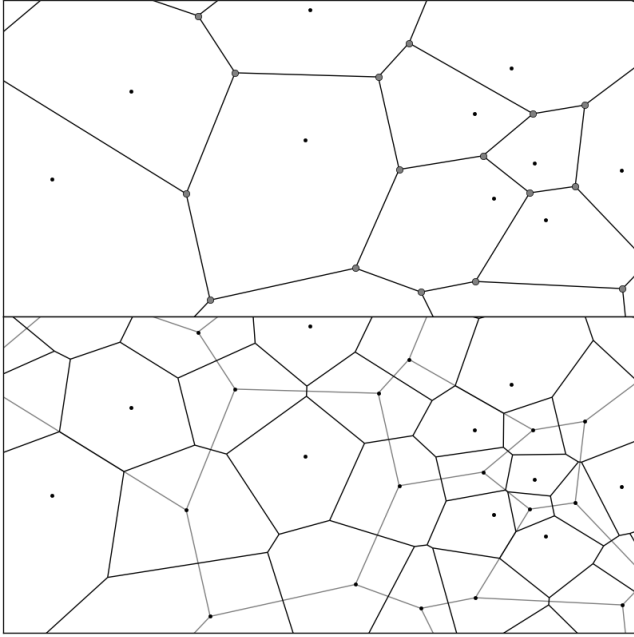


Figure 7. *Top:* a Voronoi diagram with 15 cells. The generating points (only 10 are shown) and the tessellation vertices, are displayed in black and grey respectively. *Bottom:* a Voronoi diagram with increased resolution generated with the points and vertices from the top panel, grey lines show the original cells.

points, which the properties at the new cell can be approximated by those carried by the nearest object (which does not need to be the same kind of object because only the survey properties are required). Alternatively, the local properties can be derived from the raw data. Fig. 8 demonstrate with 10 halo analyses that for a very well-behaved survey (small differences in survey depths), a resolution in the order of ~ 30 sq. degrees per cell compared to ~ 100 sq. degrees per cell will only lead to an increase of $\sim 1\%$ in number density (top panel). The increase arises from the deeper part of the survey that the standard resolution always underestimates (as it is statistically less likely to land on the deeper parts). To understand the effect at lower resolutions, we simulate this by using 1 in 3 objects to generate the Voronoi tessellation. The change in number density is also in the range of $\sim 1\%$, but it can go either way (bottom panel). Over a large number of simulations, the ratios should average to one.

5 CONCLUSION

In this work, we have demonstrated the use of Voronoi tessellation can increase the survey volume to more optimally retrieve objects from a large sky area multi-epoch survey. The assumption behind this is not ideal, but it is not possible to take an average value of the number of epochs and their properties for each cell. Further subdivisions will take much more time to compute the volumes as this algorithm scales as $x \times N^2$, where x is the number of subdivision of each cell and N is the number of objects. Nevertheless, this method is one big step in approaching the optimal sampling of the survey footprint with limited computing power.

From a mixed population simulation, we find that under the framework of our galactic models, the new method recovers $\sim 10 - 15\%$ more objects than the RH11 method under a typical lower

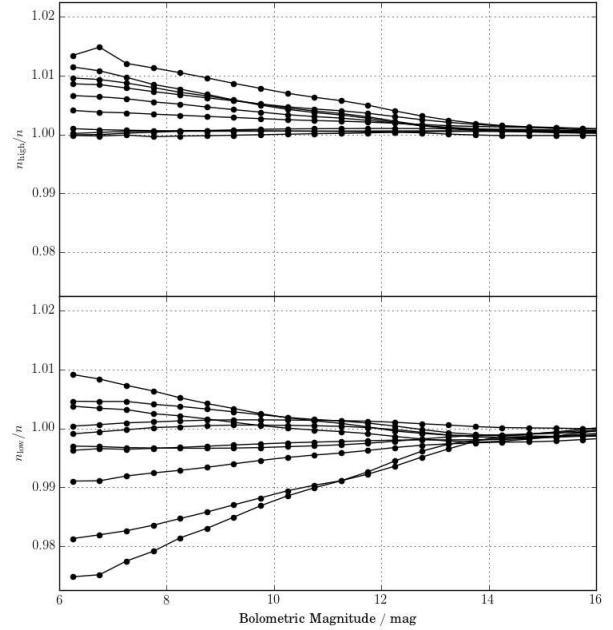


Figure 8. *Top:* the ratio of number density between the high resolution ($\sim 3\times$) analysis to the standard resolution as a function of absolute bolometric magnitude. *Bottom:* the ratio of number density between the low resolution ($\sim \frac{1}{3}\times$) analysis to the standard resolution as a function of absolute bolometric magnitude.

proper motion selection. When considering a restricted tangential velocity selection ($40 - 60 \text{ km s}^{-1}$), we do not observe any bias in the WDLFs brighter than 16 mag, and the results are still mostly consistent with the error bars due to the large lower bounds of the uncertainties. In the case of $40 - 120 \text{ km s}^{-1}$, the WDLFs will become unreliable at 15 mag. However, if the thick disc density is lower than the value adopted in the simulation, the contamination rate will be much lower, hence the bias at faint magnitudes will be much smaller. In the high velocity regime, at a given thick disc to halo density ratio, a $200 - 500 \text{ km s}^{-1}$ selection will only contain a small fraction of contaminations so it is a good sample for studying the halo.

We have demonstrated one way to increase the resolution of the tessellation and it shows that for a well behaved survey, a low resolution only limits the volume by $\sim 1\%$. An adaptive way that only subdivides cells larger than a certain solid angle can provide a grid of cells that have similar areas should it be more useful in some certain scenarios. When applying this method to real surveys, careful treatment at the boundaries is needed because the area of the survey is important. Leaving the boundaries untreated one will always end up with 4π steradian of sky area. In order to have a correct boundary that defines the survey, artificial points have to be added to create a layer of bounding cells surrounding the survey area such that the boundaries of the second last layer of cells overlap the survey footprint. One can identify the artificial points by using the survey boundary as a cell boundary, and then locate a generating point that can produce such condition.

ACKNOWLEDGMENTS

The Pan-STARRS 1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the Univer-

sity of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under Grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

We thank the PS1 Builders and PS1 operations staff for construction and operation of the PS1 system and access to the data products provided. ML acknowledges financial support from the consolidation grant of the Institute for Astronomy, University of Edinburgh.

REFERENCES

- Avni Y., Bahcall J. N., 1980, *ApJ*, 235, 694
 Bastian U., Röser S., 1993, *psc.book*, III,
 Bergeron P., et al., 2011, *ApJ*, 737, 28
 Bramich D. M., et al., 2008, *MNRAS*, 386, 887
 Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127
 Chambers K. C., et al., 2016, *arXiv*, arXiv:1612.05560
 Chiba M., Beers T. C., 2000, *AJ*, 119, 2843
 Evans D. W., 1992, *MNRAS*, 255, 521
 Evans D. W., Irwin M., 1995, *MNRAS*, 277, 820
 Felten J. E., 1976, *ApJ*, 207, 700
 Flewelling H. A., et al., 2016, *arXiv*, arXiv:1612.05243
 Fuchs B., Jahreiß H., Flynn C., 2009, *AJ*, 137, 266
 Gaia Collaboration, et al., 2016, *A&A*, 595, A1
 Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, *A&AS*, 141, 371
 Gould A., Kollmeier J. A., 2004, *ApJS*, 152, 103
 Hambly N. C., Davenhall A. C., Irwin M. J., MacGillivray H. T., 2001, *MNRAS*, 326, 1315
 Hambly N. C., Irwin M. J., MacGillivray H. T., 2001, *MNRAS*, 326, 1295
 Hambly N. C., et al., 2001, *MNRAS*, 326, 1279
 Holberg J. B., Bergeron P., 2006, *AJ*, 132, 1221
 Isern J., García-Berro E., Torres S., Catalán S., 2008, *ApJ*, 682, L109
 Kalirai J. S., Saul Davis D., Richer H. B., Bergeron P., Catelan M., Hansen B. M. S., Rich R. M., 2009, *ApJ*, 705, 408
 Kirkpatrick J. D., et al., 2016, *ApJS*, 224, 36
 Kirkpatrick J. D., et al., 2014, *ApJ*, 783, 122
 Kowalski P. M., Saumon D., 2006, *ApJ*, 651, L137
 Kroupa P., 2001, *MNRAS*, 322, 231
 Lépine S., 2008, *AJ*, 135, 2177
 Lépine S., 2005, *AJ*, 130, 1247
 Lépine S., Shara M. M., Rich R. M., 2003, *AJ*, 126, 921
 Lépine S., Shara M. M., Rich R. M., 2002, *AJ*, 124, 1190
 Lam M. C., Rowell N., Hambly N. C., 2015, *MNRAS*, 450, 4098
 Lawrence A., et al., 2007, *MNRAS*, 379, 1599
 LSST Science Collaboration, et al., 2009, *arXiv*, arXiv:0912.0201
 Luyten W. J., 1979, *nlcs.book*,
 Luyten W. J., 1979, *lccs.book*,
 Magnier E. A., et al., 2013, *ApJS*, 205, 20
 Magnier E. A., et al., 2016, *arXiv*, arXiv:1612.05240
 Magnier E. A., et al., 2016, *arXiv*, arXiv:1612.05244
 Magnier E. A., et al., 2016, *arXiv*, arXiv:1612.05242
 Mendez R. A., Guzman R., 1998, *A&A*, 333, 106
 Metcalfe N., et al., 2013, *MNRAS*, 435, 1825
 Monet D., 1996, *AAS*, 28, 54.04
 Monet D. G., 1998, *AAS*, 30, 120.03
 Monet D. G., et al., 2003, *AJ*, 125, 984
 Munn J. A., et al., 2004, *AJ*, 127, 3034
 Röser S., Bastian U., 1991, *psc..book*, I,
 Roeser S., Demleitner M., Schilbach E., 2010, *AJ*, 139, 2440
 Rowell N., 2013, *MNRAS*, 434, 1549
 Rowell N., Hambly N. C., 2011, *MNRAS*, 417, 93
 Sandage A., Fouts G., 1987, *AJ*, 93, 74
 Schlafly E. F., et al., 2012, *ApJ*, 756, 158
 Schmidt M., 1968, *ApJ*, 151, 393
 Sutherland W., et al., 2015, *A&A*, 575, A25
 The Dark Energy Survey Collaboration, 2005, *astro*, arXiv:astro-ph/0510346
 Tonry J. L., et al., 2012, *ApJ*, 750, 99
 Tremblay P.-E., Bergeron P., Gianninas A., 2011, *ApJ*, 730, 128
 Waters C. Z., et al., 2016, *arXiv*, arXiv:1612.05245
 Zacharias N., et al., 2010, *AJ*, 139, 2184
 Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2013, *AJ*, 145, 44
 Zacharias N., et al., 2000, *AJ*, 120, 2131
 Zacharias N., Urban S. E., Zacharias M. I., Wycoff G. L., Hall D. M., Monet D. G., Rafferty T. J., 2004, *AJ*, 127, 3043