Mon. Not. R. Astron. Soc. **000**, 1–11 (2015) Printed 10 February 2015 (MN LATEX style file v2.2)

A maximum volume density estimator generalised over a proper motion limited sample

Marco C. Lam⋆, Nicholas Rowell, Nigel C. Hambly
Institute for Astronomy, University of Edinburgh, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

Accepted . Received ; in original form

ABSTRACT
The traditional Schmidt density estimator has been proven to be unbiased and effective in a magnitude limited sample. Previously, efforts have been made to generalise it for populations with non-uniform density and proper motion limited cases. This work shows that the generalisation method breaks down for a proper motion limited sample due to over–simplistic assumptions. Populations with larger differences in the kinematics as compared to the Local Standard of Rest are most severely affected. We show that this systematic bias can be removed by treating the discovery fraction inseparable from the generalised maximum volume integrand. The treatment can be applied to any proper motion limited sample with good knowledge of the kinematics. This work demonstrates the method through application to a mock catalogue of a white dwarf-only solar neighbourhood for various scenarios and compared against the traditional treatment using a survey with Pan-STARRS–like characteristics.

Key words: proper motions – methods: statistical – stars: luminosity function, mass function – white dwarfs – stars: kinematics and dynamics – solar neighbourhood.

1 INTRODUCTION

The use of a maximum volume as a density estimator began when Schmidt (1968) introduced the V/V_{max} technique for analysing the luminosity function of quasars, ϕ_s, where V and V_{max} are the volume enclosed by the object and the volume in which the object can be found at the given survey limits respectively and ϕ_s is the number density. The estimator can be used as a means of testing the completeness of a sample simultaneously. Felten (1976) showed that ϕ_s is unbiased and superior to the older "classical" estimator (N/V) when the magnitude or luminosity bins are not small. He further described the procedure of combining the density estimator from two non-overlapping areas.

Information from a single survey is limited due to the relatively small number of objects. When several complete samples are combined, more fundamental parameters of the population of objects can be determined, and with smaller associated uncertainties. In the light of this problem, Avni & Bahcall (1980) investigated different ways of combining catalogues, namely the incoherent region-independent method, the incoherent domain-independent method and the coherent method. All of these are superior over the original Schmidt method, with the coherent method being most accurate. When the effects of a space-density gradient were corrected for (Stobie, Ishida, & Peacock 1989; Tinney, Reid, & Mould 1993; Lam & Hambly 2015) this method was extended to estimate stellar density where the density profile of the Galaxy varies significantly along different lines of sight. Because of the small distances probed, only the scale height effects are considered while the scale length is assumed to be constant.

In order to consider a sample of proper motion objects, Schmidt (1975) extended his estimator to cope with both photometric and proper motion detection limits. The new estimator considers the tangential velocity as an intrinsic property of an object such that it can be kept as a constant. Then, the distance limits can be found easily by applying the upper and lower proper motion limits of the survey to a simple relation between tangential velocity, proper motion and distance: $v_{tan} = 4.74\mu D$ km s^{-1}, where μ is the proper motion in arcseconds per year and D is the distance to the object in parsecs.

Cool white dwarfs (WDs) and subdwarfs (sds) have similar optical colours to the main sequence stars (MSS) while those of brown dwarfs (BDs) are similar to the giants. Therefore, it is difficult to distinguish them in colour-colour space. Due to their small radii, WDs, sds and BDs are located far from MSSs and giants in the HR diagram. However, when objects are only detected in a few broadband filters, it is impossible to classify them reliably which would lead to poor object selections and distance estimates. To overcome this problem, it is common to use reduced proper motion (RPM) as a crude estimate of absolute magnitude to separate sam-

⋆ E-mail: mlam@roe.ac.uk

ples of subluminous objects from higher luminosity contaminants. In order to obtain a clean sample of WDs (for example an extreme subdwarf would easily be confused with a WDs with low tangential velocity) a lower tangential velocity has to be applied to remove the ambiguous objects. This procedure introduces an incompleteness which has to be corrected for. This problem was identified by Bahcall & Casertano (1986) and Evans (1992) separately. The former adopted a Monte Carlo (MC) simulation approach to correct for the incompleteness, while the latter was done analytically. However, the two methods evolved separately. In the simulation front, Liebert et al. (1999) constructed some simulations based on different Galactic models to study the incompleteness due to proper motion selection after some strong arguments (Oswalt & Smith 1995, Oswalt et al. 1996) pointing towards an incomplete LHS catalogue used in earlier studies (eg. Winget et al. 1987, Liebert, Dahn, & Monet 1988). This correction, known as the discovery fraction, χ, was then applied by Harris et al. (2006, hereafter H06). On the other hand, Knox, Hawkins, & Hambly (1999) used the analytical approach to correct for the incompleteness. Instead of calculating the discovery fractions from integrating over the density profile, Digby et al. (2003) arrived at the discovery fractions by integrating over the Schwarzschild distribution functions. Rowell & Hambly (2011, hereafter RH11) further generalised the technique to cope with an all sky survey as opposed to the individual fields of view employed in earlier works.

This work studies the discovery fraction in detail and shows that the discovery fraction has to be incorporated into the volume integral in order to arrive at a correct density estimation. In the next Section, we discuss the method of the simulation of the solar neighbourhood. In Section 3, different maximum volume estimators and discovery fractions and how their shortcomings can be removed by a new approach are discussed. This new method is then applied to a simulated sample of white dwarfs in Section 4, where we choose survey parameters typical of the state–of–the–art Pan–STARRS optical sky survey (Hambly et al. 2013 and references therein). In the last Section, we discuss the possible extension of the method and conclude this work.

2 POPULATION SYNTHESIS

Monte Carlo simulations are used to produce snapshots of WD-only solar neighbourhoods which carry six dimensional phase space information. The volume probed in this work is assumed to be small such that the simulation is done in a Cartesian space, instead of a plane polar system centred at the Galactic Centre (GC). The Galaxy is further assumed to have three distinct kinematic components: a thin disc, a thick disc and a stellar halo, all of which have no density variations along the co-planar direction of the Galactic plane. All vertical structures follow exponential profiles, with scale height H. 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. The thin and thick disc populations are assigned with constant star formation rates since look back time, $\tau = 8\,\mathrm{Gyr}$ and $\tau = 10\,\mathrm{Gyr}$ respectively while the halo has a star burst of duration 1 Gyr at $\tau = 12.5\,\mathrm{Gyr}$. The initial mass function has an exponent of -2.3 (Kroupa 2001), and the initial-final mass function(IFMF) follows the ones in Kalirai et al. (2009)

$$m_f = \begin{cases} 0.101 m_i + 0.463, & 0.5 M_\odot < m_i \leqslant 4.0 M_\odot \\ 0.047 m_i + 0.679, & 4.0 M_\odot < m_i \leqslant 7.0 M_\odot. \end{cases} \quad (1)$$

The 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). Together with the pure hydrogen WD (DA) atmosphere cooling models with constant surface gravity $\log g = 8.0$ and synthetic colours[1] (Holberg & Bergeron 2006, Kowalski & Saumon 2006, Tremblay, Bergeron, & Gianninas 2011 and Bergeron et al. 2011), a theoretical luminosity function (LF) is produced which can then be used as the PDF in the Monte Carlo simulation. The normalisations of the PDFs are adopted from the WD densities found in RH11. The volume in which objects are distributed is limited to half a magnitude deeper than the maximum distance at which the survey can probe given its brightness. The half magnitude is to allow for random fluctuations near the detection limit after noise is added. The input parameters are assumed to be invariant with time and are summarised in Table 1.

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

Parameter	Thin Disc	Thick Disc	Stellar Halo
$\langle U \rangle$/km s^{-1}	-8.62[a]	-11.0[d]	-26.0[d]
$\langle V \rangle$/km s^{-1}	-20.04[a]	-42.0[d]	-199.0[d]
$\langle W \rangle$/km s^{-1}	-7.10[a]	-12.0[d]	-12.0[d]
σ_U/km s^{-1}	32.4[a]	50.0[d]	141.0[d]
σ_V/km s^{-1}	23.0[a]	56.0[d]	106.0[d]
σ_W/km s^{-1}	18.1[a]	34.0[d]	94.0[d]
H/pc	250[b]	780[e]	∞
n/pc^{-3}	0.00310[c]	0.00064[c]	0.00019[c]

a Fuchs, Jahreiß, & Flynn 2009
b Mendez & Guzman 1998
c Rowell & Hambly 2011
d Chiba & Beers 2000
e Girardi et al. 2000

From the true distance and true bolometric magnitude drawn from the PDF, the true apparent magnitudes in the Pan-STARRS g_{p1}, r_{p1}, i_{p1}, z_{p1} and y_{p1} filters are assigned (Tonry et al. 2012; Schlafly et al. 2012; Magnier et al. 2013). The uncertainties in those filters, σ_{m_i}, are assumed to scale exponentially with magnitude and are described by

$$\sigma_{m_i} = a_i \times \exp^{(m_i - 15.0)} + b_i \quad (2)$$

where a_i and b_i are constants measured from the first Pan-STARRS survey (PS1; magnitude subscript 'p1') at Processing Version (PV) 1.1 and m_i is the magnitude in filter i (Fig. 1). Realistic dispersion is added to the uncertainties by resampling σ_{m_i} with a Gaussian distribution with standard deviations of $0.1 \times \sigma_{m_i}$ centred at the noiseless σ_{m_i}.

[1] http://www.astro.umontreal.ca/∼bergeron/CoolingModels/

Table 2. Parameters for the noise model.

Filter or Proper Motion	a_i/mag	b_i/mag
g	0.000125	0.00065
r	0.000120	0.00075
i	0.000175	0.00065
z	0.000325	0.00089
y	0.000750	0.00120
μ	0.000300	0.00050

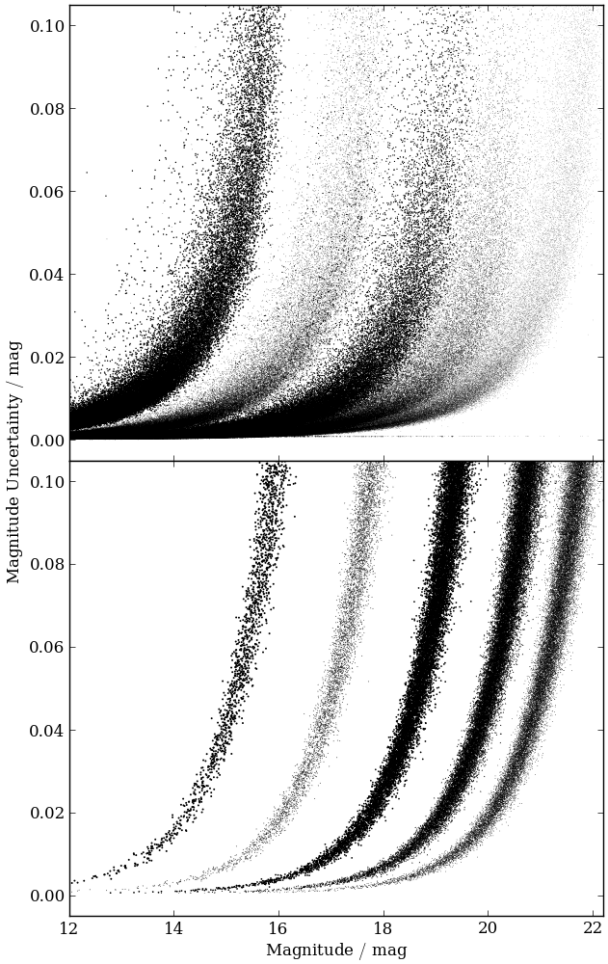

Figure 1. Top: the magnitudes and the associated uncertainties of point sources with 5σ proper motions in the PV1.1 of Pan-STARRS, from right to left g_{p1}, r_{p1}, i_{p1}, z_{p1} and y_{p1} filters. Each successive filter is offset by one magnitude. Bottom: the magnitudes and uncertainties distribution of WDs reproduced in the simulation.

The magnitudes in each filter are then drawn from a Gaussian distribution with a standard deviation σ_{m_i}. The proper motion uncertainty is based on the r_{p1} magnitude.

3 GENERALISED MAXIMUM VOLUME DENSITY ESTIMATOR

The classical estimator $\Phi = N/\mathrm{V}$ for a volume-limited sample is of little practical use for analysing small numbers of objects or strongly localised and kinematically biased groups of stars that are selected by proper motion and apparent magnitude. For these samples, the $1/\mathrm{V}_{max}$ method is generally regarded as a superior estimator of the LF (Felten 1976). The contribution of each object to the LF is weighted by the inverse of the maximum volume in which an object could be observed by the survey. For example, for a given bin of objects with index k, the space density is the sum of all $1/\mathrm{V}_{max}$

$$\Phi_k = \sum_{i=1}^{N_k} \frac{1}{\mathrm{V}_{max,i}} \tag{3}$$

for N_k objects in the k^{th} bin. The uncertainty of each star's contribution is assumed to follow Poisson statistics. The sum of all errors in quadrature within a luminosity bin is therefore,

$$\sigma_k = \left[\sum_{i=1}^{N_k} \left(\frac{1}{\mathrm{V}_{max,i}} \right)^2 \right]^{\frac{1}{2}} . \tag{4}$$

The traditional $1/\mathrm{V}_{max}$ technique assumes that objects are uniformly distributed in space. However, in reality, stars in the solar neighbourhood are concentrated in the plane of the disc. The effects of space-density gradient can be corrected by assuming a density law and defining a maximum generalised volume V_{gen} (Stobie, Ishida, & Peacock 1989; Tinney, Reid, & Mould 1993) which is calculated by integrating the appropriate stellar density profile $\rho/\rho_\odot$ along the line of sight between the minimum distance, d_{min}, and maximum distance, d_{max}. This leads to the integral

$$\mathrm{V}_{gen,S89} = \Omega \int_{d_{min}}^{d_{max}} \frac{\rho(r)}{\rho_\odot} \, r^2 \, dr \tag{5}$$

where Ω is the size of the solid angle of the survey. To minimise the contamination from extreme subdwarfs scattered into the WD regime in RPM-colour space, a lower tangential velocity limit, $v_{tan,lower}$, is applied to remove most of the contaminants. Traditionally, the discovery fraction, χ_v which is the fraction of objects with tangential velocities larger than the lower tangential velocity limit, is only Galactic model and survey-footprint dependent (Bahcall & Casertano 1986, Liebert et al. 1999, Harris et al. 2006) such that the maximum volume density estimator can be written in the form

$$\mathrm{V}_{gen,H06} = \chi_v(v_{tan,lower}) \, \Omega \int_{d_{min}}^{d_{max}} \frac{\rho(r)}{\rho_\odot} \, r^2 \, dr \tag{6}$$

where the distance limits are derived from both photometric and proper motion limits of the survey by calculating

$$d_{max} = d \times min \left[10^{\frac{(m_{max,i} - M_i)}{5}}, \frac{\mu}{\mu_{min}} \right] \tag{7}$$

$$d_{min} = d \times max \left[10^{\frac{(m_{min,i} - M_i)}{5}}, \frac{\mu}{\mu_{max}} \right] \tag{8}$$

where $m_{min,i}$, $m_{max,i}$ and M_i are the photometric limits and the absolute magnitudes of the object in filter i respectively. The proper motion terms, μ/μ_{max} and μ/μ_{min}, are rationalised by assuming an object would carry the same tangential velocity if it were placed closer to or farther from the observer (analogous to the absolute magnitude) and/or in an arbitrary line of sight.

3.1 Attempt to Modify the Discovery Fraction

RH11 extended the χ_v to include a directional dependence in order to account for the varying survey properties and stellar tangential velocity distribution across the sky. In RH11, the Schwarzschild distribution function is used instead to calculate the tangential velocity distribution, $P(v_{tan})$, analytically. The discovery fraction can be found by projecting the velocity ellipsoid onto the tangent plane of observation (Murray 1983). Thus it allows one to arrive at a precise $\chi(\alpha, \delta)$ without taking the average properties over a large area. Therefore, the volume integral can be modified to

$$\mathrm{V}_{gen,RH11} = \sum_i \Omega_i \, \chi_v(i, v_{tan,lower}) \int_{d_{min}}^{d_{max}} \frac{\rho_i(r)}{\rho_\odot} \, r^2 \, dr \quad (9)$$

where i denotes each sky cell covered by a Schmidt survey field employed in the production of the catalogue used by RH11. This should, in theory, have taken into account all the small scale variations which a positional-independent χ_v would not be able to deal with.

Problems

We have identified two biases present in this formulation of the generalised volume. These can be described as follows:

(i) Consider a spatially uniform population like the stellar halo, and an all-sky survey where the proper motion limits are the same over the whole sky. The tangential velocity distribution varies along different lines of sight due to the solar motion. For stars at a given magnitude, a different fraction of the population will pass the proper motion limits along different lines of sight due to the differences in the tangential velocities. The consequence is that different numbers of stars would be detected in different regions of sky even for identical survey limits and population spatial density. The original V_{max} method only considers the population spatial density and survey limits, so stars in different parts of the sky will have the same maximum volume within the same size of solid angle. The result is that a different density will be measured along different lines of sight, with lines of sight that observe a greater fraction of the tangential velocity distribution measuring a higher density for the stellar population.

(ii) The kinematics of an object is a property of the Galaxy. An object at a given magnitude at any given position should not carry the same velocity at some other position when tested for observability. In the most extreme cases, on average, a halo WD observed in the direction of the Anti-Galactic Center (AGC) would appear to have a large velocity due to reflex motion imparted by the Sun in its orbit within the Galaxy. However, if one is observed in the direction to the solar apex instead, the motion would be much smaller on average. The assumption of holding a constant tangential velocity is thus invalid. This assumption is only good over small angles and small range of line of sight distances where an acceptable size of the smallness is very difficult to access if not unquantifiable.

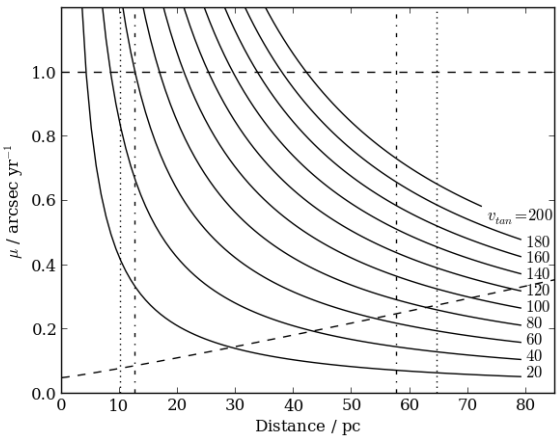

Figure 2. This plot illustrates how tangential velocity and proper motion limits behave in the $\mu - D$ space. The solid lines are the contours of the tangential velocities from 20 to 200 km s^{-1} in steps of 20 km s^{-1}. The photometric distance limits mark the range of distances in which an object can be placed and stay within the detection limits. The dot-dashed lines are the distance limits calculated from Equations 8 and 9, which is by fixing the tangential velocity of an object such that a proper motion limit correspond to a fixed distance limit. The dotted lines are the distance limits calculated from Equations 14 and 15, where the distance limits are not functions of the kinematics. The dashed lines are the upper and lower proper motion limits; in the latter case, this limit rises as astrometric errors rise for increasingly faint flux levels as distance increases.

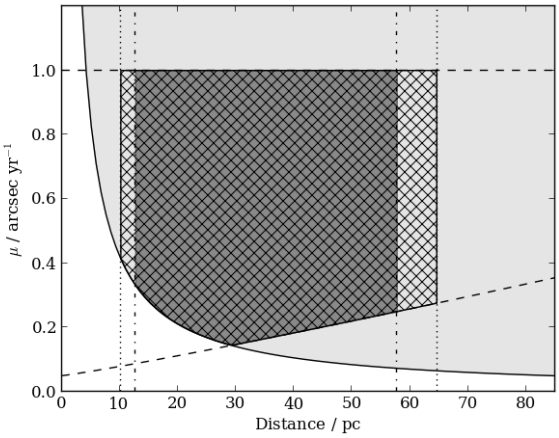

Figure 3. The light grey area is where the discovery fraction is calculated according to traditional methods, while in dark grey is the area over which the volume is integrated. The hatched area is where both the discovery fraction and the volume are integrated in the new method. The weighted area gives the discovery fraction in the new method (see Figure 4 for the weight maps). It is obvious that the discovery fraction in the traditional method is self-inconsistent and thus always over-estimates the completeness which leads to an under-estimation in the number density.

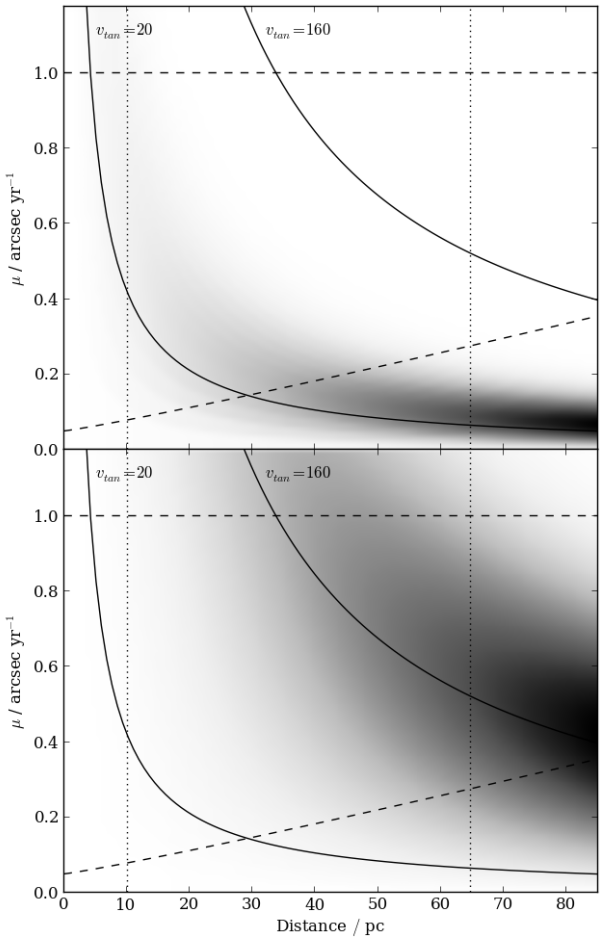

Figure 4. The top panel shows the probability distribution as a function of proper motion and distance of a volume limited thin disc population over-plotted with the proper motions, tangential velocities and distance limits. The discovery fraction is the weighted area within the specified limits where the probability distribution function is used as the weights. The lower panel shows those of a halo population. The heaviest part of the thin disc weight function is excluded by the proper motion limit, but much more of the heavily weighted region of the stellar halo is within the allowed parameter space, so the tangential velocity limit at $\sim$ 20kms^{-1} has a much smaller effect in the thin disc than in the halo.

3.2 A Closer Look at the Discovery Fraction

Proper motions, tangential velocities and distances are related by a simple equation

$$v_{tan} \approx 4.74\,\mu\,D. \tag{10}$$

All the survey limits can be shown in the proper motion-distance space (Fig. 2). A valid approach should not impose any assumptions that restrict an object in this two dimensional space. The discovery fraction is equal to the weighted area restricted by the survey limits, where the weight map (Fig. 4) is generated from dividing the tangential velocity distribution by the distances,

$$W(\mu, D) = \frac{P(v_{tan})}{4.74 \times D}. \tag{11}$$

In the traditional approach, v_{tan} is fixed and is thus separable from μ, which follow the lines of constant tangential velocity at v_{tan} = 20 - 200 km s^{-1} in 20 km s^{-1} interval in Fig. 2. However, the two variables are related through distance as shown in Equation 10. This implies a selection criterion in one parameter would lead to a selection effect in the other. The traditional way only deals with the tangential velocity limits but ignores the consequential effect of the change in the proper motion limits bounding the discovery fractions. A similar effect also appears in the treatment of the distance limits. The volume integral is bounded by the maximum and minimum distances in which an object can be found but the discovery fraction in RH11 includes everything above the line $v_{tan} = 20\,\mathrm{km\,s^{-1}}$ as represented by the light grey area in Fig. 3. The true discovery fraction should be bounded by the same limits as applied to the volume integral (hatched area). Thus, χ_v is always overestimated which translates to an underestimation in the luminosity function. This effect is stronger for:

i) a survey with small upper proper motion limit because the lower this limit, the larger the overestimation of the discovery fraction. This can be seen from Fig. 3 where a smaller upper proper motion limit would lead to a smaller cross-hatched area, which means the discovery fraction would be overestimated.

ii) a population with large differences in the kinematics compared to the observer because objects tend to have large proper motions. This shifts the region with the highest probability in $W(\mu, D)$ (Fig. 4) to larger proper motions. In the case of the Galaxy, the stellar halo is the most susceptible to this effect.

iii) intrinsically faint objects that carry small maximum observable distance. It is clear from Fig. 4 that the relative $W(\mu, D)$ at low proper motion increases with distance when the $P(v_{tan})/D$ peaks at smaller proper motion as distance increases. This affects the discovery fraction more severely when the lower proper motion limit is large since more heavy-weighted area would be included in, for example, both the H06 and RH11 methods.

As an example of a current survey which can be employed in WDLF studies (Hambly et al. 2013), the Pan-STARRS upper proper motion limit is O(1)" yr^{-1} which means the first problem may affect the analysis of WDLF without proper treatment. Halo WDs have higher velocities and are older (i.e. fainter) than those of the discs and hence the effect on the halo population would be much larger than that in the discs. As it is of great interest to probe the halo WDLF at such low luminosity to explore the possible scenarios of the star formation history of the Galaxy, it is necessary to correct this bias.

3.3 A New Approach

In order to compute the discovery fraction properly, the parameter space in which an object could be observed by the survey has to be identical to that used in the discovery fraction integral. For each step of the numerical integration, it is necessary to calculate the instantaneous discovery fraction that which is limited by the upper and lower proper motion limits, as well as the tangential velocity limits. It is worth mentioning that in the effective volume method in RH11 a

similar approach was adopted that would have corrected for the bias noted in the last Section although the bias was not explicitly identified and discussed in that work. Instead of dealing in an object by object basis, their correction was applied statistically. The strength of that method is that the WDLFs of the three components could be untangled. However, binning objects by their bolometric magnitudes before membership association would lead to a loss of information. Furthermore, it loses the generality so that it cannot be applied to other luminosity estimators. In order to keep the analysis in an object by object basis, one should consider this modified volume integral,

$$\mathrm{V}_{new} = \sum_i \Omega_i \int_{d_{min}}^{d_{max}} \chi_{\mu,v}(i,r)\, \frac{\rho_i(r)}{\rho_\odot}\, r^2\, dr \qquad (12)$$

where the discovery fraction can be found by

$$\chi_{\mu,v}(i,d) = \chi_v(i,d,\mu_{lower}(d)) - \chi_v(i,d,\mu_{upper}(d)) \qquad (13)$$

$$= \int_{4.74\mu_{min}r}^{4.74\mu_{max}r} P(v_{tan}) dv_{tan} \qquad (14)$$

with distance limits calculated from the photometric limits only,

$$d_{max} = d \times min\left[10^{\frac{(m_{max,i}-M_i)}{5}}\right] \qquad (15)$$

$$d_{min} = d \times max\left[10^{\frac{(m_{min,i}-M_i)}{5}}\right] \qquad (16)$$

with the same notation as in Equations 7 and 8. The new volume integral has the distance limits decoupled from the kinematics, which are completely absorbed into the discovery fraction. The decoupling simultaneously means that regardless of how the kinematic behaviour changes with respect to the direction of observation, the entire μ-D-v_{tan} parameter space in which an object can be found is explored. When the discovery fraction has a dependence on distance, it cannot be separated from the integrand. This means that the discovery fraction varies from object to object in any given direction, so it is sufficiently general to take a more realistic form of velocity distribution. This in turn allows for a distance dependent tangential velocity distribution that describes the Galaxy more realistically.

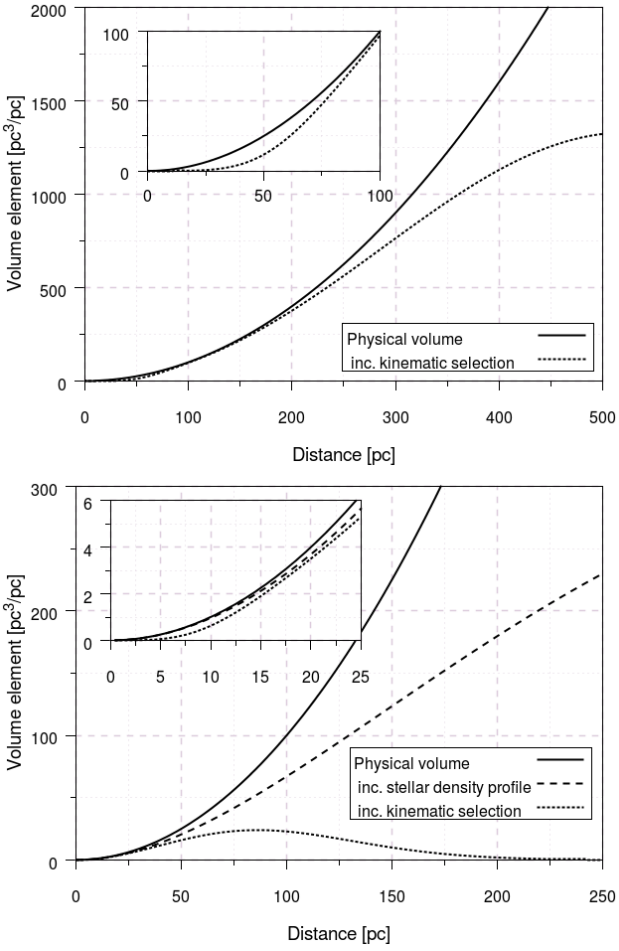

Figure 5. Differential survey volume as a function of line of sight distance, for the stellar halo (top) and the thin disc (bottom). The lines represent the true physical volume (solid), the volume generalised for the stellar density profile (dashed) and the inclusion of the effects of kinematic selection (dotted) as explained in the text. The inset boxes reveal the behaviour at small distances, where the range of tangential velocities that pass the proper motion limits is reduced.

Consequence on the survey volume

It is instructive to examine how the survey volume as a function of distance is modified when the effects of kinematic selection are taken into account. This amounts to computing the fraction of objects that pass the proper motion limits at each distance step, by integrating the tangential velocity distribution between velocities corresponding to the lower and upper proper motion limits. The total generalised survey volume out to a distance r is therefore calculated by Equation 12. Note that $P(v_{tan})$ and $\rho(r)$ depend on the line of sight direction, so this model only holds for small fields. Now that all selection effects are taken into account, this function provides a model of the distance distribution of objects in the survey. The differential survey volume as a function of distance, for the stellar halo and the thin disk, is presented in Fig. 5. These plots are computed for a line of sight towards the NGP, in order to maximise the effect of the non-uniform population density profile in the case of the thin disc where the density profile falls off most rapidly in the direction perpendicular to the plane. Note that in the case of the halo, the density profile is uniform so the volume generalised for density profile matches the true volume. The lower and upper proper motion limits are set at 0.1 and 1.0" yr^{-1} and the size of the solid angle is 0.01 sr. These figures reveal that the effects of kinematic selection always act to reduce the generalised survey volume at any given distance. The strength of the effect varies as a function of distance according to the interplay of the proper motion limits and tangential velocity distribution; for example, at small distances the halo generalised volume (inset, top) is significantly reduced due to the combined effects of the large average motion placing many objects above the upper proper motion limit, and the narrow range of tangential velocities that pass the proper motion limits at small distances.

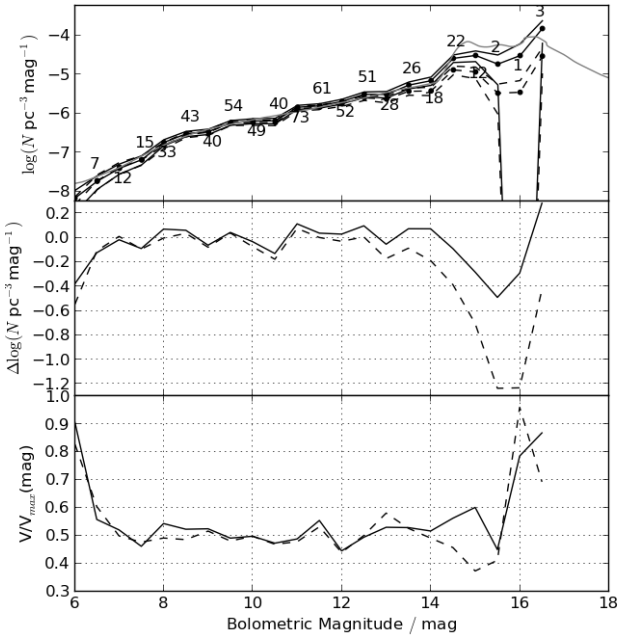

Figure 6. Top: WDLFs for a population with halo kinematics with the measured halo WD density. The two WDLFs are derived from the new approach (solid) and the RH11 method (dashed) overplotted with the input true WDLF (grey). The lines above and below the WDLFs are the 1-sigma upper and lower limits. Middle: the deviations of the WDLF as a function of magnitude. The RH11 solution departs from the true LF at ∼13.0 mag while the new approach at ∼14.5 mag. Bottom: the $\langle V/V_{max}\rangle$ as a function of magnitude from the two methods. The uncertainties in $\langle V/V_{max}\rangle$ is $\frac{1}{\sqrt{12N}}$, so in the $\langle V/V_{max}\rangle$s are within the statistical fluctuations up to 16 mag.

4 APPLICATION TO WHITE DWARF LUMINOSITY FUNCTIONS

We have used the simulation outlined in Section 2 to generate mock Pan-STARRS catalogues to compare the differences in applying the RH11 generalised volume and the new generalised volume described in this work. The bright limits in all filters are set at 15.00 mag, while the faint limits in g_{p1}, r_{p1}, i_{p1}, z_{p1} and y_{p1} filters are at 21.50 mag, 21.00 mag, 20.50 mag, 20.00 mag and 20.00 mag respectively. The lower proper motion limit is defined as five times the proper motion uncertainty at the given magnitude as defined in Equation 2. The upper proper motion limit is set at 0.3" yr^{-1} unless specified otherwise. Photometric parallaxes are found by fitting the magnitudes to the WD synthetic atmosphere models at fixed surface gravity $\log g = 8.0$. Tests have shown that there exists local minima for effective temperature below 6000K so a Markov-Chain Monte Carlo method is used for minimisation. The implementation adopted the Python programme emcee[2] (Foreman-Mackey et al. 2013). A lower tangential velocity limit is set at 20 km s^{-1}.

[2] http://dan.iel.fm/emcee/

4.1 Stellar halo

As discussed in Section 3.2, faint objects and objects from a population with large difference in the kinematics from the observers are most susceptible to underestimation of the number density when using the traditional technique. Thus, the largest systematic errors are expected to be found in the faint end of the stellar halo WDLF. In the top panel of Fig.6, the differences in the WDLF constructed by the two methods are shown. The two LFs agree with each other up to M_{bol} ∼12.5 mag. Beyond that, the traditional method consistently underestimates the number density and the deviation increases as the objects get fainter. The maximum difference is more than 1.0 dex at around 15.5 mag. The PanSTARRS filter g_{p1} reaches ∼400 nm at the blue edge, so it is expected that the photometric parallax solutions become unreliable at the bright end, while with small numbers of objects in the faint end the statistical noise is significant. The $\langle V/V_{max}\rangle$s in both cases are within 1σ from 0.5, which is a necessary condition for an unbiased sample.

4.2 Stellar Halo with 10 times the Observed Density

The stellar halo has a very low number density and it is well known that maximum volume estimators are prone to systematic bias with small number statistics. Thus, we have generated a stellar halo with 10 times the observed density to reduce such uncertainties. The WDLFs produced are more easily compared when the systematic uncertainties are much smaller and the corresponding plots are shown in Fig. 7. The two LFs agree with each other up to M_{bol} ∼ 13.0 mag again. With 10 times the density, faint objects down to 16.0 mag ($T_{\rm eff}$ ∼ 3000 K) can be generated easily. The LF produced with the new method agrees with the input LF to luminosities down to ∼14.5 while the RH11 method consistently underestimates the density, as predicted in Section 3.2. The $\langle V/V_{max}\rangle$ shows an interesting behaviour - the distribution is fluctuating about 0.5 regardless of the bias in the WDLF. We suggest that the V/V_{max} distribution is not a good indicator of bias: $\langle V/V_{max}\rangle \approx 0.5$ is a necessary but not sufficient condition for both bias and completeness.

4.3 Different upper proper motion limits

As mentioned in Section 3.2, a small upper proper motion limit would lead to a strong bias. This is illustrated in Fig 8: as the upper proper motion limit increases, the old and new approaches converge. However, when the limit decreases, the underestimation in the number density increases when adopting the old method. From the bottom panel, even at an upper limit of 0.5" yr^{-1}, the traditional method fails to recover the WDLF beyond ∼15 mag. With the new method, the number density estimation is recovered under any restricted proper motion selection within the statistical noise. In the modern and future surveys with the rapid photometric systems, the upper proper motion limit would be in the order of arcseconds so this effect would be small. However, when the pairing of the bright objects includes older photographic plate data, the upper proper motion limit would be severely hampered due to a large maximum epoch difference (e.g. Digby et al. 2003 has an upper limit of 0.18" yr^{-1}

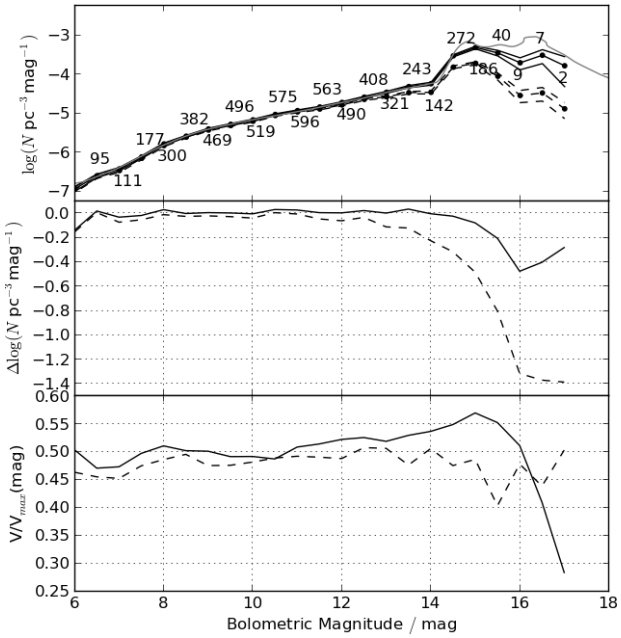

Figure 7. Same as Fig.6 except the WD density is artificially inflated by 10 times to reduce random noise. With the increase number density, the WDLF departs a magnitude fainter in both cases. The $\langle V/V_{max}\rangle$ of the old method is still fluctuating about 0.5 even when the WDLF departs further and further away from the true LF. In the new method, the $\langle V/V_{max}\rangle$ is consistent within statistical uncertainties.

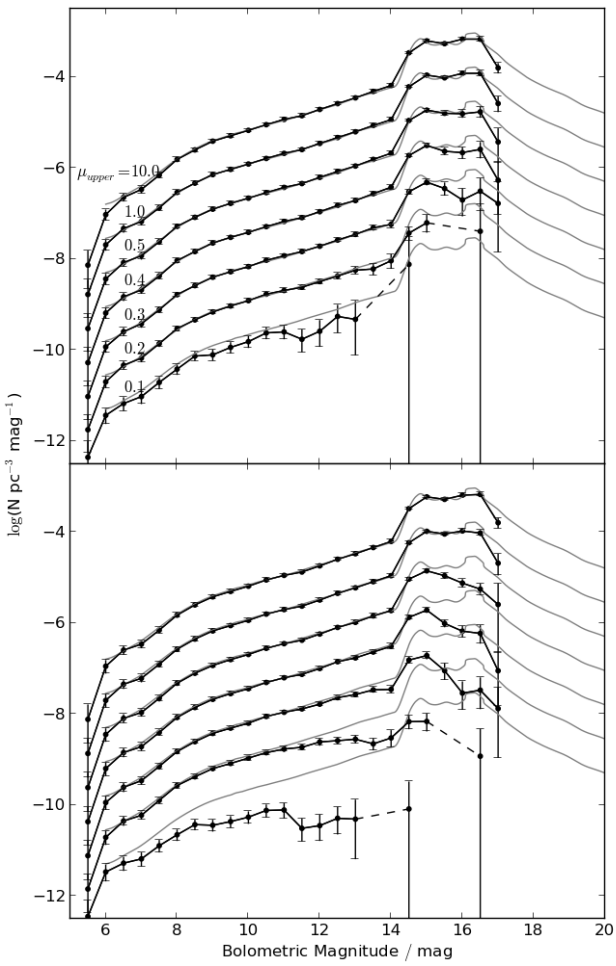

Figure 8. Top: WFLD constructed with different upper proper motion limits at 10 times the measured density with the new generalised volume method. From top to bottom, the upper proper motion limit is set at 10.0, 1.0, 0.5, 0.4, 0.3, 0.2 and 0.1" yr^{-1}. Bottom: same procedure repeated with the traditional generalised volume. In the two sets with the smallest limits, dashed lines are used to join up the LFs at the faintest bin where the area under the dashed line should not be included in determining the total number density.

from pairing between POSS-I, POSS-II and SDSS). Another case is when samples at different velocity ranges are analysed separately (e.g. the effective volume method in RH11 where the discovery fraction was treated correctly) the restricted velocity range would amplify the shortcoming of the old method.

4.4 Thin Disc, Thick Disc and Stellar Halo

In the study of the Galactic WDLFs it is very difficult to assign population membership to individual objects. The consensus is that thin disc objects dominate the solar neighbourhood so that when studying the LF for the thin disc it is possible to assume a thin disc characteristic for all objects. This does not happen to be a good assumption as shown in the top panel of Fig. 9 where for the purposes of this exercise we apply only the new method. The solid line is the WDLF constructed from all stars regardless of the population. It is overestimated by $\sim$0.2 dex at all magnitudes. An overestimation is expected but a consistent overdensity of 0.2 dex is not negligible. Due to the spatial density correction, the maximum volume of an object integrated over a disc profile has smaller volume than a halo profile. While objects are weighted by the inverse maximum volume, the apparently negligible contribution from the older populations would by amplified by density correction such that each contaminating object would have a volume underestimated by tens of percentage points. The discovery fraction for a disc population is always smaller than that of the halo, so this amplifies the $1/V_{max}$ by another few tens of percentage points. The V/V_{max} distributions are consistently at around 0.55 as a consequence of the halo objects contribution. The ratio of the volume element between an exponential disc density profile and a uniform halo increases with distance. Therefore, for a group of uniformly distributed objects being assigned to follow an exponential profile in the integrals, the resulting $\langle V/V_{max}\rangle$ is expected to be larger than 0.5. The dashed line is constructed with an addition of a fixed upper tangential velocity limit at 80 km s^{-1} which would eliminate most of the halo objects, but not the ones from the thick disc. This reduces the contamination down to roughly 0.1 dex at all magnitudes, with a smaller effect on the $\langle V/V_{max}\rangle$ distribution. The dotted line is computed using only thin disc objects from that simulation for comparison. Everything agrees to within statistical uncertainties at all magnitudes in the thin disc-only analysis. Since the discovery fraction requires the kinematics of the populations, a good membership association is very important.

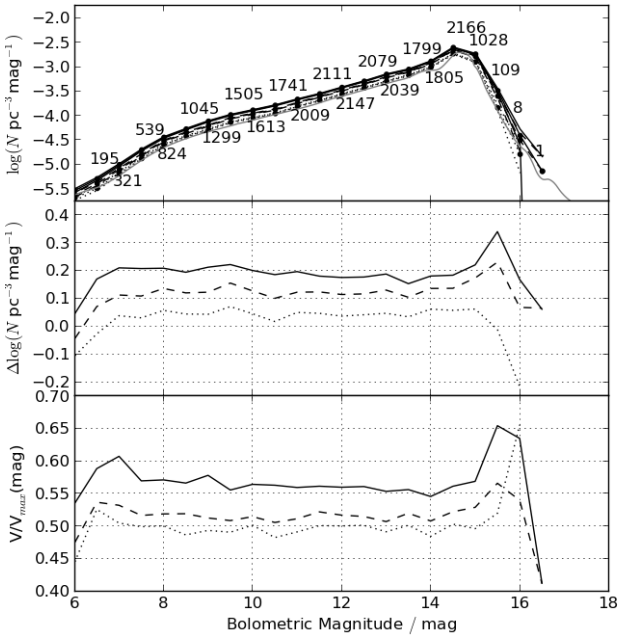

Figure 9. Top: WDLFs of the Galaxy constructed by assuming negligible halo contamination (solid), with an additional fixed upper tangential velocity limit at 80 km s^{-1} (dashed) and by picking up only objects tagged as "thin disc object" from the simulation (dotted). Middle: the differences in the WDLF from the true LF as a function of magnitude. The overdensities are expected from the wrong association of kinematics model with the objects. Bottom: the $\langle V/V_{max}\rangle$ distributions as a function of magnitude. $\langle V/V_{max}\rangle$s are consistently larger than 0.5 implying objects tend to be found at larger distances than on average. This is expected from the objects with a uniform spatial density being associated with ones that have an exponential profile. The larger discovery fraction of the thin disc with respect to the halo amplifies the the effect further (see text for detailed explanation).

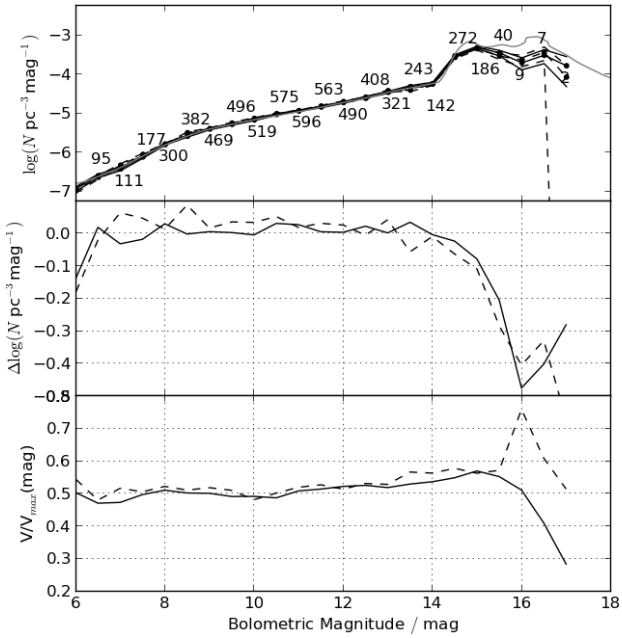

Figure 10. Top: WDLFs for the stellar halo with different input Galactic models. The solid line is the original halo WDLF identical to the one in Fig 7, the dashed line is the one that follows a simplified Besancon model but then analysed with the original set of kinematics. Middle and bottom panels are the same as before.

4.5 Different Galactic Models

The accuracy of all LF estimators is sensitive to the assumed population kinematics and density profiles. Therefore it is important to investigate the effects of applying a different kinematic model in the analysis of a population simulated with another model. To demonstrate the effect, we have adopted a simplified version of the Besancon galaxy model (Robin et al. 2003, 2004). A halo population at 10 times the measured density with $(U, V, W) = (0.0, -220.0, 0.0)$ km s^{-1} and velocity dispersion $(\sigma_U, \sigma_V, \sigma_W) = (131.0, 106.0, 85.0)$ km s^{-1} was generated using the same recipe described in Section 2. The halo population is then analysed using the original set of kinematic properties. The same WDLF from Section 4.2 is overplotted for comparison in Fig. 10. The two WDLFs agree at all magnitudes and even the trends in underdensities at faint magnitudes are very similar. The differences in the $\langle V/V_{max}\rangle$ distribution at faint magnitudes can be explained by the same argument as in Section 4.4. Due to the similar kinematic properties between the halo kinematic models applied in this work and the Besancon model, the size of this effect is much smaller. From this simple experiment we conclude that small differences between the assumed and underlying kinematics results in little differences in both the derived WDLF and the $\langle V/V_{max}\rangle$ distribution.

5 CONCLUSION

We have presented an improved treatment of the proper motion and tangential velocity limits to arrive at an unbiased luminosity function estimator for stellar populations selected on the basis of both magnitude and proper motion. Our simulations have shown that the assumption of setting the tangential velocity of an object as its intrinsic property would be invalid when a sample is drawn from a survey that has restricted proper motion limits and for populations where the tangential velocity distribution function varies over the sky. In order to select high quality proper motion objects when analysing real survey data the lower proper motion limit and tangential velocity limit are usually set at such a level that the velocity space is very restricted. Therefore, it is necessary to correlate the tangential velocity with the proper motion depending on the properties of each individual object. Two biases were identified from dissecting the traditional V_{max} integral. The first one is an inconsistent density measurement of any non-uniform sample in which the density depends on the line of sight direction. The traditional method only considers the population spatial density and survey limits, so different densities would be measured at different line of sight directions as the distribution of proper motion and tangential velocity vary with line of sight. The other arises from the assumption that any given object would carry a constant tangential velocity. The choice of the upper proper motion limit has to be carefully determined from the properties of the sur-

vey. An arbitrarily large upper limit has negligible effects in the analysis as the survey volume is sensitive only to the lower proper motion limit. The discovery fraction is, however, sensitive to both upper and lower proper motion limits. From our simulations, it is found that surveys with upper proper motion limits smaller than ∼0.5" yr^{-1} require more attention. This value seems small to the modern and future surveys (e.g. PanSTARRS, Gaia, LSST etc.) but when a pairing is done with the early epoch Schmidt survey plate data for the bright objects for example the upper proper motion limits would shrink rapidly when the maximum epoch differences can be up to more than half a century (the pairing radius cannot be increased indefinitely when associating catalogues with very different depth and resolution without introducing many spurious source matches).

Population membership association has to be done carefully because with a wrong set of kinematics the derived set of discovery fractions would become completely meaningless. However, in the case that a population completely dominates the density budget, for example the thin disc objects in the solar neighbourhood, the wrong membership associations have negligible effects to the luminosity functions. Finally, we find that the usefulness of the V/V_{max} distribution and the $\langle V/V_{max}\rangle$ are overvalued. A flat V/V_{max} distribution with $\langle V/V_{max}\rangle \sim 0.5$ only indicates if the analysis is biased towards the further or the closer objects, they have no indication of the completeness of the sample. For a given flat distribution, one can conclude, at best, that the survey is equally sensitive to objects at any distance. However, given a complete sample a flat distribution and $\langle V/V_{max}\rangle \sim 0.5$ are expected.

Application to other density estimators

The treatment of detection limits due to proper motion selection have never been included in other density estimators. The generalised Schmidt estimator is the only one that has been extended to incorporate such constraints. However, as detailed earlier, the treatment breaks down if we have restricted proper motion selection criteria. Geijo et al. (2006) compared the properties of three different estimators: Generalised Schmidt, Stepwise Maximum Likelihood and Choloniewski method. It is clear that the treatment of the proper motion limits were only applied to the Schmidt estimator by considering the tangential velocity as an intrinsic property of an object, while the other two considered the case of a magnitude limited sample. However, their analysis is not flawed because the sample adopted in their simulation has no upper and lower proper motion limits (ie. $\mu \in [\,0, \infty\,]$). It should be possible to generalise the discovery fraction (as a completeness correction arising from tangential velocity and/or proper motion limit) to other density estimators as the proper motion is now decoupled from the detection limit.

Small Scale Variations

Modern digital surveys are probing larger volumes with multi-epoch measurements over large areas of sky (e.g. SDSS, VST, Pan-STARRS, LSST and Gaia). Astrometric precision is, however, a strong function of the direction of the lines of sight due to varying observing qualities and imperfect camera fill factor arising from CCD chip gaps. In order to fully utilise the sample volume probed in any single or combined surveys, it is necessary to generalise the density estimators to cope with small scale fluctuations of the survey properties. Otherwise, the volume of the survey must be limited to the shallowest part of the sky. One possible solution is to pixelise the sky into equal area pixels using HEALPix[3]. The photometric and astrometric precision of each pixel are then limited to the worst observing condition within the pixel.

Galactic model

The current Schwarzschild distribution assumes a Cartesian system centred at the Sun. This approximation is good for surveys that probe only small distances (a few hundred parsecs). However, in the future surveys, most notably the Gaia and LSST, where the lower proper motion limits at the detection limits would be as low as $0.2\,\mathrm{mas\,yr^{-1}}$ ($G_{Gaia} \sim 20\,\mathrm{mag}$), and $1\,\mathrm{mas\,yr^{-1}}$ ($r_{LSST} \sim 24\,\mathrm{mag}$) respectively, an object with a tangential velocity of $20\,\mathrm{km\,s^{-1}}$ could be detected with 5σ confidence at ∼17 kpc and ∼3.5 kpc respectively. A WD with $T_{\rm eff} \sim 3000\,\mathrm{K}$ and $M_{bol} \approx 16\,\mathrm{mag}$, which translate to $G_{\rm Gaia} \sim 15.8\,\mathrm{mag}$ and $g_{\rm LSST} \sim 17.3\,\mathrm{mag}$ at 10 pc, would be detectable at ∼1-2 kpc in both cases. The overdensity at the spiral arms is not accounted for in any work to date because the current surveys have not reached such distances where the nearest spiral arm lies. However, when the faintest objects can be detected at a distance of over a thousand parsecs, the brighter objects would lie easily inside the spiral arm region which is ∼1 kpc from the Sun. In such a situation it is necessary to employ a more detailed density profile to account for the varying density in not just the vertical direction but also in the planar directions. With the decoupling of the proper motion limit from the photometric limits, the incorporation of a more sophisticated model would simply be a translation of the Galactic density and velocity profiles from the galactocentric frame to the geocentric frame. The key difference is that a Schwarzschild distribution function gives a constant tangential velocity distribution function as a function of the line of sight (ie. $\chi = \chi(\alpha, \delta)$), but in the adoption of a complex model, the density and velocity distribution functions have to be found as functions of both distance and the direction of line of sight such that $\chi = \chi(\alpha, \delta, D)$.

ACKNOWLEDGEMENTS

The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the University 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 Na-

[3] http://healpix.jpl.nasa.gov/

tional 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). ML acknowledges financial support from a UK STFC PhD studentship.

We thank the PS1 Builders and PS1 operations staff for construction and operation of the PS1 system and access to the data products provided.

REFERENCES

Avni Y., Bahcall J. N., 1980, ApJ, 235, 694
Bahcall J. N., Casertano S., 1986, ApJ, 308, 347
Bergeron P., et al., 2011, ApJ, 737, 28
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Chiba M., Beers T. C., 2000, AJ, 119, 2843
Digby A. P., Hambly N. C., Cooke J. A., Reid I. N., Cannon R. D., 2003, MNRAS, 344, 583
Evans D. W., 1992, MNRAS, 255, 521
Felten J. E., 1976, ApJ, 207, 700
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fuchs B., Jahreiß H., Flynn C., 2009, AJ, 137, 266
Geijo E. M., Torres S., Isern J., García-Berro E., 2006, MNRAS, 369, 1654
Girard T. M., Korchagin V. I., Casetti-Dinescu D. I., van Altena W. F., López C. E., Monet D. G., 2006, AJ, 132, 1768
Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, A&AS, 141, 371
Hambly N., Rowell N., Tonry J., Magnier E., Stubbs C., 2013, ASPC, 469, 253
Harris H. C., et al., 2006, AJ, 131, 571
Holberg J. B., Bergeron P., 2006, AJ, 132, 1221
Kalirai J. S., Saul Davis D., Richer H. B., Bergeron P., Catelan M., Hansen B. M. S., Rich R. M., 2009, ApJ, 705, 408
Knox R. A., Hawkins M. R. S., Hambly N. C., 1999, MNRAS, 306, 736
Kowalski P. M., Saumon D., 2006, ApJ, 651, L137
Kroupa P., 2001, MNRAS, 322, 231
Lam M. C., Hambly N. C., 2015, PASP, in press
Liebert J., Dahn C. C., Harris H. C., Legget S. K., 1999, ASPC, 169, 51
Liebert J., Dahn C. C., Monet D. G., 1988, ApJ, 332, 891
Magnier E. A., et al., 2013, ApJS, 205, 20
Mendez R. A., Guzman R., 1998, A&A, 333, 106
Murray C. A., 1983, veas.book,
Oswalt T. D., Smith J. A., 1995, LNP, 443, 24
Oswalt T. D., Smith J. A., Wood M. A., Hintzen P., 1996, Natur, 382, 692
Robin A. C., Reylé C., Derrière S., Picaud S., 2004, A&A, 416, 157
Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523
Rowell N., Hambly N. C., 2011, MNRAS, 417, 93
Schlafly E. F., et al., 2012, ApJ, 756, 158
Schmidt M., 1975, ApJ, 202, 22
Schmidt M., 1968, ApJ, 151, 393
Stobie R. S., Ishida K., Peacock J. A., 1989, MNRAS, 238, 709
Tinney C. G., Reid I. N., Mould J. R., 1993, ApJ, 414, 254
Tonry J. L., et al., 2012, ApJ, 750, 99
Tremblay P.-E., Bergeron P., Gianninas A., 2011, ApJ, 730, 128
Winget D. E., Hansen C. J., Liebert J., van Horn H. M., Fontaine G., Nather R. E., Kepler S. O., Lamb D. Q., 1987, ApJ, 315, L77