Probabilistic selection of high-redshift quasars

Daniel J. Mortlock,1* Mitesh Patel,1 Stephen J. Warren,1 Paul C. Hewett,2 Bram P. Venemans,3 Richard G. McMahon2 and Chris Simpson4

1Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ
2Institute of Astronomy, Madingley Road, Cambridge CB3 0HA
3European Southern Observatory, Karl-Schwarzschild Strasse 2, 85748 Garching bei München, Germany
4Astrophysics Research Institute, Liverpool John Moores University, Twelve Quays House, Egerton Wharf, Birkenhead CH41 1LD

Accepted 2011 August 25. Received 2011 August 11; in original form 2010 December 29

ABSTRACT
High-redshift quasars (HZQs) with redshifts of $z \gtrsim 6$ are so rare that any photometrically selected sample of sources with HZQ-like colours is likely to be dominated by Galactic stars and brown dwarfs scattered from the stellar locus. It is impractical to re-observe all such candidates, so an alternative approach was developed in which Bayesian model comparison techniques are used to calculate the probability that a candidate is a HZQ, $P_q$, by combining models of the quasar and star populations with the photometric measurements of the object. This method was motivated specifically by the large number of HZQ candidates identified by cross-matching the UKIRT (United Kingdom Infrared Telescope) Infrared Deep Sky Survey (UKIDSS) Large Area Survey (LAS) to the Sloan Digital Sky Survey (SDSS): in the $\sim 2270 \text{ deg}^2$ covered by the LAS in the UKIDSS Eighth Data Release (DR8) there are $\sim 9 \times 10^3$ real astronomical point sources with the measured colours of the target quasars, of which only $\sim 10$ are expected to be HZQs. Applying Bayesian model comparison to the sample reveals that most sources with HZQ-like colours have $P_q \lesssim 0.1$ and can be confidently rejected without the need for any further observations. In the case of the UKIDSS DR8 LAS, there were just 107 candidates with $P_q \gtrsim 0.1$; these objects were prioritized for re-observation by ranking according to $P_q$ (and their likely redshift, which was also inferred from the photometric data). Most candidates were rejected after one or two (moderate-depth) photometric measurements by recalculating $P_q$ using the new data. That left 12 confirmed HZQs, six of which were previously identified in the SDSS and six of which were new UKIDSS discoveries. The high efficiency of this Bayesian selection method suggests that it could usefully be extended to other HZQ surveys (e.g. searches by the Panoramic Survey Telescope And Rapid Response System, Pan-STARRS, or the Visible and Infrared Survey Telescope for Astronomy, VISTA) as well as to other searches for rare objects.

Key words: methods: statistical – surveys – quasars: general.

1 INTRODUCTION
Quasars are the most luminous non-transient astronomical sources to redshifts of at least $z \simeq 6.5$ and have, ever since their discovery (Hazard, Mackey & Shimmins 1963; Schmidt 1963), been key cosmological probes (e.g. Schneider 1999). Observations of high-redshift quasars (HZQs) with $z \simeq 6$ have revealed a marked increase in the optical depth to neutral hydrogen (H I) at redshifts of $z \gtrsim 5.7$ (Becker et al. 2001; Fan et al. 2002; Fan et al. 2006), which appears to mark the end of cosmological reionization (see e.g. Barkana & Loeb 2001). Measurements of the quasar luminosity function (QLF) at $z \simeq 6$ also constrain the growth of structure and the early formation of super-massive black holes in the first billion years of the Universe (e.g. Jiang et al. 2008; Willott et al. 2010a). There is thus a strong motivation to discover any new HZQs, and there is a particular premium on finding the most luminous quasars because most HZQ science requires high signal-to-noise ratio (S/N) spectroscopic data. The possibility of making extensive spectroscopic observations of the brightest HZQs also differentiates them from other high-redshift sources: the $z \simeq 7$ field galaxies found in deep surveys are too faint to obtain high S/N spectra (e.g. Stark et al. 2010); and gamma-ray bursts remain sufficiently bright for spectroscopy only for a few days (e.g. Gehrels, Ramirez-Ruiz & Fox 2009).

*E-mail: mortlock@ic.ac.uk
Despite the strong motivations for identifying bright HZQs, only \( \sim 50 \) are known at present (e.g. Fan et al. 2006; Jiang et al. 2008; Willott et al. 2010b; Mortlock et al., in preparation). This is primarily because HZQs are so rare: the results of Jiang et al. (2008) imply there are only \( \sim 450 \) redshift \( z \geq 6.0 \) quasars brighter than \( z = 21.0 \) over the whole sky. A direct implication is that the bright quasars at \( z \gtrsim 6 \) are only likely to be found in fairly shallow wide-area surveys. The first such HZQ search was based on the Sloan Digital Sky Survey (SDSS; York et al. 2000), which has a typical single-scan magnitude limit of \( z_{\text{lim}} \simeq 20.8 \), and has yielded 19 redshift \( z \gtrsim 5.8 \) quasars in \( 6600 \text{deg}^2 \) (Fan et al. 2006). Similar numbers of lower luminosity HZQs have been found in the deeper SDSS Stripe 82 region (Jiang et al. 2009) and the Canada–France High-z Quasar Survey (CFHQs; Willott et al. 2007). Continuing and future optical surveys will be able to increase the number of known \( z \geq 6 \) quasars but will not be able to probe past a redshift limit of \( z \gtrsim 6.5 \). Sources beyond this redshift are undetectable in optical surveys as their Ly\( \alpha \) emission is redshifted out of the optical bands and all shorter wavelength photons are absorbed by intervening H\( \text{I} \) (Gunn & Peterson 1965).

Quasars with \( z \gtrsim 6.5 \) will eventually be identified by future radio surveys (e.g. Wylee 2008), but the most immediate progress will be made by observing in the near-infrared (NIR). The largest completed NIR survey, the Two Micron All-Sky Survey (2MASS; Skrutskie et al. 2006), has a detection limit of \( z_{\text{lim}} \simeq 16.6 \) and so is too shallow to detect any HZQs. The partially complete UKIRT (United Kingdom Infrared Telescope) Infrared Deep Sky Survey (UKIDSS; Lawrence et al. 2007) includes a Large Area Survey (LAS) which reaches typical depths of \( Y = 20.2 \) and \( J = 19.6 \) (Warren et al. 2007) and has already yielded several new \( z \gtrsim 6 \) quasars (Venemans et al. 2007; Mortlock et al. 2009b, 2011; Venemans et al., in preparation; Patel et al., in preparation). In the future, the Panoramic Survey Telescope And Rapid Response System (Pan-STARRS; Kaiser et al. 2002) and the Visible and Infrared Survey Telescope for Astronomy (VISTA; Emerson et al. 2004) should extend the UKIDSS results in both redshift and numbers, and various planned satellite missions could increase the size of the HZQ samples by an order of magnitude (e.g. Willott et al. 2010b).

The existence of surveys with the appropriate combinations of area, depth and wavelength coverage is not, however, a sufficient condition for discovering HZQs. It is also necessary to be able to separate these rare objects of interest from the far more numerous main-sequence stars and brown dwarfs. Aside from being a good pragmatic selection option, the simplicity of hard colour cuts makes it very easy to quantify the completeness of the resultant quasar samples (e.g. Fan et al. 2003; Willott et al. 2010b). Cut-based approaches do, however, have several shortcomings: the conversion from photometric measurements to a binary selection represents a significant loss of information; and the most promising high S/N candidates are inevitably grouped with more numerous marginal candidates near the edges of the selection region. Even if there are sufficient observational resources to follow-up all the objects identified in this way, it is inevitable that some worthy candidates will be rejected.\(^1\)

An alternative to applying hard data cuts is to adopt a probabilistic approach to quasar selection, replacing the construction of a definite candidate list with a calculation of the probability, \( P_q \), that each source is a quasar. While this idea has not been applied to \( z \gtrsim 6 \) quasar searches, it has been used to generate large samples of lower-redshift quasars (Richards et al. 2004, 2009; Bailer-Jones et al. 2008; Bovy et al. 2011). Most relevantly, Richards et al. (2004) applied kernel density estimation (KDE) to training sets of spectroscopically confirmed stars and quasars, giving estimates of the observed distribution of both populations in SDSS colour space. Bovy et al. (2011) used extreme deconvolution in place of KDE to estimate the intrinsic distribution. In both cases, the second step was to apply Bayes’s theorem to calculate \( P_q \) for each source in turn. The power of these methods is adequately illustrated by the first result obtained by Richards et al. (2004): a photometric sample of \( \sim 10^3 \) quasars that is 95 per cent complete to \( g \leq 21.0 \) with only 5 per cent contamination.

The use of a prior to account for the fact that quasars are outnumbered by Galactic stars is important to all the above probabilistic selection methods, although most sources would have been classified decisively (and correctly) simply by comparing the normalized KDE quasar and star density estimates at the sources’ locations.

\(^1\) The limitations of cut-based candidate selection were illustrated in the case of the \( z = 6.13 \) quasar ULAS J1319+0950 (Mortlock et al. 2009b). This source was detected with \( i = 22.83 \pm 0.32 \) and \( z = 20.13 \pm 0.12 \) in SDSS, and so satisfied the Fan et al. (2001) requirements that \( i - z > 2.2 \) and \( z < 20.2 \); however it was observed in slightly worse than average conditions and hence did not meet the additional requirement that the \( z \)-band noise be \( \sigma_z < 0.1 \).
in colour space. More critical was the availability of significant numbers of confirmed stars and quasars from which their distributions in the four-dimensional SDSS colour space could be inferred. Unfortunately, the need for large training samples makes it problematic to use this selection method to search for rare objects as, by definition, very few are known. In the case of HZQs, which have reasonably distinctive and predictable colours, one option would be to generate a synthetic training set of simulated quasars. However that would still not overcome the more fundamental problem that most sources with the observed colours of HZQs are extreme outliers from the stellar locus rather than distant quasars. The fact that most such objects will be close to the survey’s detection limit could be used to down-weight faint sources with large photometric errors, although the algorithm described by Richards et al. (2004) would require non-trivial modifications to account for this. From an inferential point of view the problem is that, by ignoring the photometric errors, KDE of the observed colour distribution does not utilize all the information contained in the data. For brighter sources this should not be too important, as there is still sufficient information to make a confident classification in most cases; but the inclusion of the photometric errors in the analysis of fainter sources would prevent the overly optimistic identification of stellar outliers as strong HZQ candidates.

In principle, the ideal method of HZQ candidate selection is to adopt a fully self-consistent Bayesian method which combines all the information available for each source in an optimal way. This idea is explored in this paper, starting from first principles by adapting standard Bayesian model selection techniques to astronomical classification (Section 2). The resultant formalism is then applied to the UKIDSS–SDSS HZQ search in Section 3. These results and some future extensions to this technique are summarized in Section 4. Some technical issues relating to the evaluation of the likelihood for photometric data are explored in Appendix A, and the method used to model the stellar population is detailed in Appendix B.

All photometry is given in the native system of the relevant survey. Thus SDSS i and z photometry is on the AB system, whereas UKIDSS Y and J photometry is Vega-based. Under the assumption that Vega has zero magnitude in all passbands, the Vega to AB conversions for these two UKIDSS filters are $Y_{\text{AB}} = Y_{\text{Vega}} + 0.634$ and $J_{\text{AB}} = J_{\text{Vega}} + 0.938$ (Hewett et al. 2006). All SDSS and follow-up photometry in the i and z bands is reported using airmass magnitudes (Lupton, Gunn & Szalay 1999); as a result model colours in these bands depend on the overall flux level assumed. Photometric measurements are denoted with the symbol hat (i.e. $\hat{\ }$) to emphasize that these are purely data-derived statistics. All detection limits are given as the magnitude of a point source which would, on average, be measured with an S/N of 5 in such observations. The rest-frame absolute magnitudes of quasars are given as $M = M_{\text{AB,1450}}$ (i.e. on the AB system at a restframe wavelength of $\lambda = 1450 \mu$m). Conversions between absolute and apparent magnitudes are performed assuming a fiducial flat cosmological model with normalized matter density $\Omega_m = 0.27$, normalized vacuum density $\Omega_{\Lambda} = 0.73$ and Hubble constant $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (cf. Dunkley et al. 2009).

2 PROBABILISTIC CLASSIFICATION OF ASTRONOMICAL SOURCES

Having made some measurements of an astronomical source, what can be inferred about the type of object it is? Assuming there are $N_t$ distinct populations of astronomical objects, $t = \{t_1, t_2, \ldots, t_{N_t}\}$, under consideration, the fullest answer to this question is to use the $N_t$ available measurements, $d = \{d_1, d_2, \ldots, d_{N_t}\}$, along with the fact that the object was detected in the first place, to calculate the posterior probability,\(^1\) $Pr(t|d, \det, t)$, of each hypothesis $t$. Applying Bayes’s theorem yields the standard model comparison result (e.g. Jaynes 2003; Sivia & Skilling 2006) that

$$Pr(t|d, \det, t) = \frac{Pr(d|\det, t) Pr(d, \det|t)}{\sum_{t'=1}^{N_t} Pr(t'|\det, t) Pr(d, \det|t')}$$

where $Pr(t|\det, t)$ is the prior probability that a detected source is of type $t$ and $Pr(d, \det|t)$ is the probability of detecting the source and obtaining the observed data under the $t$th hypothesis. This is known as the evidence or the model likelihood, and is given by

$$Pr(d, \det|t) = \int Pr(\theta|t) Pr(d, \det|\theta, t) \, d\theta_1 \, d\theta_2 \cdots \, d\theta_{N_t}$$

where $Pr(\theta|t)$ is the unit-normalized prior distribution of the $N_t$ model parameters, $\theta$, that describe objects of type $t$, and the likelihood, $Pr(d, \det|\theta, t)$, is the probability of detecting the source and obtaining the observed data given a particular value of those parameters.

Equation (1) is a standard application of Bayes’s theorem but for the explicit statement that the source under consideration has been detected. The reason for its inclusion here is to ensure that the prior distribution of each population’s parameters can be normalized unambiguously, as well as to avoid the meaningless notion of an unconditional prior probability of the nature of a source. Asked out of context, the question “what is the probability that a source is a quasar?” is ill-posed and has no sensible answer. This immediately implies that it is impossible to determine the prior probability of a source being of a certain type without at least some constraining information, such as a range of fluxes or colours. Thus the similar question “what is the probability that a source with $z \leq 21.0$ is a quasar?” does have a well-defined answer, the numerical value of which is given approximately by the observed numbers of quasars and stars down to the specified limit. This would then be a reasonable empirical value for the quasar prior, although even here the answer depends on various other factors, such as Galactic latitude. The implication of the above arguments is that the prior would have to be calculated independently for surveys with, e.g., different footprints on the sky or different depths, a far from satisfactory situation.

These potentially troublesome ambiguities can be resolved by combining the model and parameter priors with the likelihood into

\(^2\) It would also be possible to include various non-astronomical noise processes (e.g. bad pixels, cross-talk, noise peaks, etc.) amongst the models that might explain the data, a possibility which is especially relevant when searching for rare objects. The difficulty in implementing this idea is that, whereas most astrophysical populations are at least reasonably well constrained, the huge variety of poorly understood noise processes make it far more difficult to quantify these processes. None the less, it is a useful reminder that all probabilities are conditional, and the model selection approach followed here is always predicated on the source being drawn from one of the astronomical populations that have explicitly been included in the calculation.

\(^3\) The notation $Pr(A|B)$ is used to indicate the degree to which (the truth of) proposition $B$ implies (the truth of) proposition $A$. As such, the probability $Pr(A|B)$ is not a mathematical function in the usual sense, although if $A$ and $B$ are mathematical in nature then formal expressions such as $Pr(x = x_0|y = y_0)$ are replaced by the less cumbersome, shorthand $Pr(x|y)$. 

\(\ast\) The notation $Pr(A|B)$ is used to indicate the degree to which (the truth of) proposition $B$ implies (the truth of) proposition $A$. As such, the probability $Pr(A|B)$ is not a mathematical function in the usual sense, although if $A$ and $B$ are mathematical in nature then formal expressions such as $Pr(x = x_0|y = y_0)$ are replaced by the less cumbersome, shorthand $Pr(x|y)$.
a weighted evidence term, defined as
\[
W_t(d, \det) = \frac{\rho_t(\theta_t) \Pr(\det(\theta_t, t) \Pr(d(\theta_t, t) d\theta_{t,1} d\theta_{t,2} \ldots d\theta_{t,N_t})}{\sum_{t=1}^{N_t} W_t(d, \det)},
\]
where \(\rho_t(\theta_t)\) is the surface density of type \(t\) sources on the sky and \(\Pr(\det(\theta_t, t)\) is the probability that a source of type \(t\) with parameters \(\theta_t\) would have been detected in the survey. The main benefit of using \(\rho_t(\theta_t)\) instead of the necessarily normalized \(\Pr(\theta_t|d, \det)\) is that the source density has an empirical normalization given by the observed number of sources per unit solid angle. Not being dependent on generally arbitrary parameter space boundaries it is independent of the details of the current experiment, and need only be calculated once. The trade-off is the need to introduce the detection probability, \(\Pr(\det(\theta_t, t)\), although most sources are sufficiently brighter than the survey’s detection limit that \(\Pr(\det(\theta_t, t)\) is close to unity and can be ignored. Using the weighted evidence, equation (1) simplifies to
\[
\Pr(t|d, \det) = \frac{W_t(d, \det)}{\sum_{t=1}^{N_t} W_t(d, \det)}.
\]

Equations (3) and (4) describe a general method for probabilistic classification of an astronomical object, by explicitly combining existing knowledge of the populations from which it might have been drawn with the information contained in whatever measurements have been made of the source in question. The next steps towards adapting this general formalism to HZQ candidate selection are to examine the likelihood for photometric data (Section 2.1) and to specialize to the specific case in which the source is assumed to be either a quasar or a star (Section 2.2).

### 2.1 Photometric data

In optical and NIR astronomy, even a low-quality spectrum is usually sufficient to establish the basic nature of the source with near certainty, obviating the need for the formal statistical approach described above. Photometric measurements, however, are generally more ambiguous, and some sort of Bayesian approach is required to avoid making overly certain classifications (e.g. Mortlock, Peiris & Ivezić 2009a). For this reason only photometric measurements are considered henceforth.

Given a model described by parameters \(\theta_t\), the likelihood \(\Pr(d(\theta_t, t)\) of measuring photometric data \(d\) must include information on how the parameters relate to observables, as well as describing the stochastic aspects of the measurement process. In the case of optical or NIR survey photometry, the likelihood should account for a number of distinct effects: the background uncertainty in the images; the Poisson noise in the number of photons received from the source; possible inter-band noise correlations (e.g. Scranton et al. 2005); and non-detections, including cases in which the background-subtracted counts are negative. For the problem of assessing HZQ candidates it is reasonable to ignore some of these effects: very few ambiguous candidates are more than a magnitude or two brighter than the survey limits, so the source Poisson photon noise can be neglected; and the gain from including the inter-band noise correlations is negligible, especially given the fact that the correlations are often poorly known. It is, however, vital to allow for non-detections, particular in the case of HZQs which have negligible flux bluewards of their redshifted Lyα emission lines.

Traditional logarithmic magnitudes (Pogson 1856) cannot represent negative measured fluxes; and, while the likelihood for non-detections can be expressed in terms of asinh magnitudes (Lupton et al. 1999), the resulting expressions are cumbersome (Appendix A). The most straightforward approach is to work in flux units (i.e. calibrated and background-subtracted counts), in which case the data vector is \(d = \mathbf{F} = \{F_{\lambda_1}, F_{\lambda_2}, \ldots, F_{\lambda_N}\}\), where \(F_{\lambda_b}\) is the reported flux in the \(b\)th of the \(N_t\) bands. Ignoring inter-band correlations allows the likelihood to be separated into the form
\[
\Pr(\mathbf{F}|\theta_t) = \prod_{b=1}^{N_t} \Pr(F_{\lambda_b}|F(\theta_t)) ,
\]
where \(F_{\lambda_b}(\theta_t)\) is the true flux in band \(b\) of an object of type \(t\) described by parameters \(\theta_t\). (The explicit dependence of the true flux on \(\theta_t\) could be omitted, but it is retained here to emphasize the fact that \(F_{\lambda_b}\) is only ever an intermediate quantity.)

For sources within a few magnitudes of the survey limit (which includes almost all HZQ candidates) the photometric errors are dominated by the uncertainties in the background subtraction, which is typically very well approximated as being additive and Gaussian in flux units. However, many of the HZQ candidates under consideration will be extreme outliers from the stellar locus, and the frequency of such events in real data is almost always higher than would be predicted by Gaussian statistics. As all the probability calculations here are performed numerically there would be no significant penalty for adopting a more complicated noise distribution with stronger tails; but the small number of outliers makes it difficult to assess what distribution should be adopted. Regardless of the specific form of the non-Gaussian tails of the photometric noise distribution, the net effect would be to decrease \(P_{\theta_h}\) due to the increased likelihood of stars being scattered to have quasar-like colours. Perhaps more importantly, the probabilities of most candidates would not be changed significantly: given that only two classes are under consideration, \(P_{\theta_h}\) is determined primarily by the relative distance of a source’s measured colours from the quasar and star loci. As such, the impact of inaccurate modelling of the photometric errors on the resultant candidate samples – and on the relative ranking of the candidates – should be minimal. It is only in a few unusual cases that the relative likelihood of the measured photometry under the two different hypotheses is changed by increasing the tails of photometric noise distribution, and these tend to correspond to non-astronomical contaminants for which neither model is a good fit.

On balance it seems clearest to assume Gaussianity, for which the single-band likelihood is
\[
\Pr(F_{\lambda_b}|F(\theta_t)) = \frac{1}{(2\pi)^{1/2} \sigma_{\lambda_b}} \exp \left\{ -\frac{1}{2} \left( \frac{F_{\lambda_b} - F_{\lambda_b}(\theta_t)}{\sigma_{\lambda_b}} \right)^2 \right\} ,
\]
where \(\sigma_{\lambda_b}\) is the background uncertainty (in flux units).

If the photometric data are only given in terms of magnitudes then they can be converted into flux units, although some care is required to ensure that the correct noise level is recovered. The conversions for both logarithmic and asinh magnitudes are given in Appendix A.

It is also possible that only upper limits are reported for sources which are undetected – or, more accurately, were measured with a low S/N – in one or more bands. With access to the raw data aperture fluxes can be measured for all undetected sources, but in some cases this is not possible. With only an upper limit it is impossible to reconstruct the likelihood as given in equation (6), as no information is retained about how far below the detection limit the source’s measured flux was. In any band for which only an upper limit, of \(F_{\lambda_{b,\lim}}\) (\(\geq 5\sigma_{\lambda_b}\) in most cases), is given, the likelihood is simply the probability that a source of true flux \(F_{\lambda_b}(\theta_t)\) would be observed with a measured flux below the stated detection limit.
Integrating over the unknown measured flux gives the likelihood of a non-detection as

\[
Pr \left[ F_\theta < F_{\text{lim},b} \mid P(\theta) \right] = \frac{1}{2} \left\{ 1 + \text{erf} \left[ \frac{F_{\text{lim},b} - F_{\theta,b}(\theta)}{2\sqrt{\sigma_b^2}} \right] \right\}, \tag{7}
\]

where \(\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt\) is the error function.

The likelihood for a source with detections in some bands and only upper limits in others is an awkward combination of probability densities (for the measurements) and cumulative probabilities (for the upper limits). However, this is not problematic in the context of model comparison as the same combinations of differential and cumulative probabilities appear in both the numerator and denominator of equation (4), and so cancel out appropriately when evaluating \(Pr(t,d, \text{det})\). The seamless handling of upper limits is one of the many benefits of the Bayesian approach to such problems.

The fact that upper limits can be self-consistently included in the classification formalism does not, however, change the fact that they represent a loss of information. There are several ways in which this information loss could be characterized, but the most relevant here is how \(Pr(t,d, \text{det})\) is changed. Given that the S/N of a non-detection is inevitably low, it might seem that the effect of replacing a noisy flux estimate with an upper limit would be minimal. In the case of the HZQ candidates considered in Section 3, however, a measurement of, e.g., \(F_\theta \simeq 3\sigma_\theta\) is often sufficient to decisively reject a candidate. A simple example demonstrating why this is the case is presented in Appendix A.

While it is useful to be able to include upper limits, flux estimates should be supplied and used if possible. In particular, flux estimates are obtained for all non-detections in the UKIDSS–SDSS quasar search described in Section 3 and hence the calculations of \(P_q\) presented in Section 3.6 are based exclusively on equation (6), not equation (7).

2.2 The probability that a source is a HZQ

The probabilistic classification formalism described above was developed specifically to assess the quality of the numerous superficially promising quasar candidates that are inevitably generated by a HZQ survey. For a real point source in such a sample only two possibilities are explicitly considered here: it is a Galactic star (i.e. \(t = s\)); or it is a target HZQ (i.e. \(t = q\)). Thus the model space is reduced to \(t = [q, s]\) and equation (4) can be simplified to give the quasar probability as

\[
P_q = Pr(q|d, \text{det}) = \frac{W_q(d, \text{det})}{W_q(d, \text{det}) + W_s(d, \text{det})}. \tag{8}
\]

The notation used in equation (8) emphasizes the role of the data in calculating \(P_q\), in the case of HZQs, but the most important single factor is the degree to which quasars are out-numbered by Galactic stars (at the depths probed by current surveys, at least). For this reason, \(P_q \ll 1\) unless a source not only has the measured colours of a HZQ, but also has sufficiently precise photometry that the data represent a statistically significant deviation from the stellar locus. An implication of these criteria is that it is almost impossible for faint sources close to a survey’s detection limit to have high \(P_q\) as their measured colours are inevitably imprecise — almost all such sources with HZQ-like colours are better explained as being scattered stars. In the space of possible multi-band measurements and errors there is at best only a small region for which the quasar probability is significant — if the survey is too shallow or does not cover the appropriate wavelengths then \(P_q\) will be low for all possible photometric measurements. The Bayesian approach to candidate selection is, in principle, as exciting as possible; the next step is to see how it works in practice by applying it to a real data set.

3 SEARCHING FOR HZQS USING UKIDSS AND SDSS DATA

The probabilistic approach to HZQ selection described in Section 2 was developed to prioritize the large number of candidates generated by matching NIR UKIDSS sources to optical SDSS catalogues. The generation of the cross-matched UKIDSS–SDSS sample and the initial candidate selection is described in Section 3.1. Realistic models for the star and quasar populations are developed in Section 3.2 and Section 3.3, respectively. The dependence of the quasar probability, \(P_q\), on the measured photometry of an idealised source is explored in Section 3.4, and the possibilities for photometric redshift estimation are demonstrated in Section 3.5. Finally, the probabilistic selection method is applied to the full UKIDSS–SDSS sample in Section 3.6.

3.1 Initial candidate selection

The starting point for this search for \(z \gtrsim 6\) quasars is UKIDSS (Lawrence et al. 2007), a suite of five separate NIR surveys using the Wide Field Camera (WFCAM; Casali et al. 2007) on the UKIRT. A detailed technical description of the survey is given by Dye et al. (2006), although there have been a number of improvements in the time since (Warren et al. 2009). The most relevant of the projects is the LAS, which should eventually cover \(\sim 3800\) deg\(^2\) in the UKIDSS Y, J, H and K bands (defined in Hewett et al. 2006). As of UKIDSS’s Eighth Data Release (DR8), on 2010 September 3, the LAS had covered \(\sim 2600\) deg\(^2\) in Y and J to average depths of \(Y_{\text{lim}} = 20.2 \pm 0.1\) and \(J_{\text{lim}} = 19.5 \pm 0.2\) (Dye et al. 2006; Warren et al. 2007). Querying the WFCAM Science Archive4 (WSA; Hambly et al. 2008) reveals that the DR8 LAS sample contains \(\sim 3.2 \times 10^7\) catalogued sources that were detected in both Y and J. According to the QLF of Jiang et al. (2008), only \(\sim 15\) HZQs are expected with \(Y \lesssim 19.8\) in the DR8 LAS area; the problem, then, is to identify these few sources efficiently and reliably.

The first step in the filtering process is to match the NIR UKIDSS sample to the optical catalogues from the SDSS (York et al. 2000). As of Data Release 7 (DR7; Abazajian et al. 2009), the SDSS covers \(\sim 12000\) deg\(^2\), including \(\sim 2270\) deg\(^2\) of the UKIDSS LAS area. Observations were made in the custom \(u, g, r, i\) and \(z\) filters (Fukugita et al. 1996), and photometry was obtained in all five bands for every detected source. For point sources the photometry was based on a model of the point spread function (PSF). The SDSS main survey reaches single-scan depths of \(i_{\text{lim}} \simeq 22.5 \pm 0.2\) and \(z_{\text{lim}} \simeq 20.8 \pm 0.2\), and so HZQs with \(i - z \gtrsim 2.5\) and \(z - Y \gtrsim 0.8\) close to the UKIDSS Y-band limit are likely to be undetected in single SDSS scans. In the case of non-detections, aperture photometry in the \(i\) and \(z\) bands was obtained from the SDSS images. Aperture photometry was not obtained in the three bluest SDSS bands, however, although potential quasar candidates were rejected if they were detected in \(u, g\) or \(r\). More importantly,

4 The WSA is located at http://surveys.roe.ac.uk/wsa/.

5 Conversely, most HZQs bright enough to be detectable in the UKIDSS Y-band observations would be detected in SDSS Stripe 82, making it possible to perform a fainter search (cf. Venemans et al. 2007; Jiang et al. 2008). The importance of the \(i\) depth is explored further in Section 3.4.
approximately 30 per cent of UKIDSS LAS sources were observed more than once by SDSS, and in such cases the best flux estimates from the different scans (i.e. PSF-based if available or aperture otherwise) were combined using inverse-variance weighting. The final result is a combined UKIDSS–SDSS catalogue of sources with the best available survey photometry in the $i, z, Y$ and $J$ bands (as well as $H$ and $K$, if available).

In the absence of photometric noise, the target HZQs are expected to occupy a region of the $i, z, Y$ and $J$ UKIDSS–SDSS colour space that is well separated from other astronomical sources (cf. Hewett et al. 2006). The theoretical separation between HZQs and cool stars in colour space is illustrated in Fig. 1, which shows both the stellar locus (described in Section 3.2) and the model quasar tracks (described in Section 3.3). The single dominant factor that ensures HZQs have such distinct colours is the near-complete absorption bluewards of the Ly $\alpha$ emission line due to intervening H$_i$. In the redshift range $5.8 \lesssim z \lesssim 6.5$ Ly $\alpha$ is in the $z$ band, and so such quasars should be extremely red in $i - z$ (and $i - Y$); at higher redshifts ($6.6 \lesssim z \lesssim 7.2$), Ly $\alpha$ is in the $Y$ band, leading to extremely red $z - Y$ or $i - Y$ colours. At redshifts of $z \gtrsim 7.2$ the Ly $\alpha$ break is redwards of the $Y$ filter and so such HZQs cannot be found in $Y$-band selected surveys. By contrast, most main-sequence stars are expected to have considerably bluer colours than HZQs, although the coolest M dwarfs have $i - Y \gtrsim 2.0$. While L and T dwarfs have similar $i - Y$ colours to HZQs, they are expected to be significantly redder in $Y - J$ (Hewett et al. 2006), which is the main reason that this HZQ search only includes fields with observations in both $Y$ and $J$.

The vast majority of UKIDSS–SDSS sources can be rejected as HZQ candidates using a variety of automated cuts (which are described more fully in Mortlock et al., in preparation):

(i) Any source with an unambiguous, bright match in SDSS that gives $i - Y \lesssim 3$ is far bluer than the target HZQs and is discarded, as is any source with red optical–NIR colours but with $Y - J \lesssim 0.9$. The specific selection criteria applied were: $i - Y \gtrsim 2.80$; $Y - J \lesssim 0.88$ and $i - Y \gtrsim 1.20 - 2.40(Y - J)$. These cuts are illustrated in Fig. 1.

(ii) Most galaxies appear identifiably extended in UKIDSS, and so sources with values of the UKIDSS morphology statistic $\mid$MergedClassStat$| > 4.0$ were discarded.

(iii) Sources close to a bright star, or that have required deblending in the SDSS processing, or with significant UKIDSS error flags, have potentially unreliable measured photometry and so were discarded.

(iv) Any sources with a significantly non-zero measured proper motion (as measured either between UKIDSS and SDSS or from non-simultaneous UKIDSS observations in different bands) are assumed to be nearby Galactic objects, and hence rejected as well.

Applying these cuts to the SDSS-matched UKIDSS DR8 LAS catalogue leaves a sample of $\sim 9 \times 10^3$ apparently stationary, isolated point sources that have the measured colours of HZQs. These sources are shown in Fig. 1.

How should the HZQ search proceed from here? The ideal observational approach would be to take spectra of all of the candidates, but doing so would require prohibitive observational resources – Glikman et al. (2008) needed 25-h-long Keck observations to rule out the most promising candidates from just the 27.3 deg$^2$ covered by the UKIDSS Early Data Release (EDR; Dye et al. 2006). Obtaining independent photometry of all the candidates is more feasible, but still difficult due to limitations of telescope scheduling and the range of target right ascensions. Even if the intention was to re-observe all candidates, some means of prioritizing the most promising is needed – it is clear from Fig. 1 that a randomly selected candidate from this sample will almost certainly be a scattered Galactic star. It was this dilemma that led to the development of the Bayesian selection method described in Section 2. Before it can be implemented, however, models are needed for both the star and quasar populations.

### 3.2 The stellar population

The UKIDSS–SDSS sources described in Section 3.1 have the measured colours of HZQs, but most are clearly Galactic stars, as can

---

**Figure 1.** Two-colour diagram showing the $\sim 9 \times 10^3$ UKIDSS DR8 LAS point sources which satisfy the various quality control criteria described in Section 3.1 and, after being cross-matched to SDSS, have measured colours similar to those expected of HZQs. Also shown is the stellar locus described in Section 3.2 and the 12 quasar models described in Section 3.3, with crosses indicating redshifts of $z = 5.5$, $z = 6.0$, $z = 6.5$ and $z = 7.0$. All the model colours are calculated for a source that has $Y = 19.5$. The dashed lines denote the initial pre-selection cuts, defined in Section 3.1, that are applied before subsequent processing. The horizontal dotted line shows the maximum $i - Y$ value that a $Y = 19.5$ source could have in the absence of noise, due to the use of asinh magnitudes to represent the $i$-band photometry.

---

$^6$In the early stages of the HZQ search the basic colour selection was $Y - J \leq 0.8$, but the desire to retain consistency between candidate lists after recalibration of the UKIDSS $Y$-band observations by 0.08 led to the anomalously precise final value of the $Y - J$ cut.
be seen from Fig. 1. The optical and NIR properties of the various likely contaminants are described in detail by Hewett et al. (2006), but it is most important here to establish the dominant source of contamination as it is these source(s) which must be modelled well to get useful quasar probabilities.

Most main-sequence stars are sufficiently hot that optical and NIR filters (and hence any well-designed IR filters) only probe their Rayleigh–Jeans tails. With $i - Y \approx 1$, such stars are so much bluer than the target HZQs that they are not a significant contaminant. Cooler M stars also have bluer optical–NIR colours than the target HZQs (e.g. $i - Y \approx 2$ as compared to $i - Y \gtrsim 3$ for the quasars), but can be sufficiently faint in the $i$ band that a small fraction scatter to have $i - \bar{Y} \gtrsim 3$. Moreover, M stars have $Y - J \approx 0.5$, much like the HZQs, and so the fact that their UKIDSS $Y$- and $J$-band photometry is typically fairly accurate actually increases the chance that some have the same observed colours as HZQs in the UKIDSS and SDSS bands.

Conversely, the accurate $Y - J$ values measured for L and T dwarfs (which can be just as red in optical–NIR colours as HZQs) combine with their lower numbers to ensure that they are not the major source of contamination. As even the faintest sources considered as possible HZQ candidates have S/N of $\gtrsim 10$ in the $Y$ and $J$ bands, the number of rare L or T dwarfs with $Y - J \gtrsim 1.0$ being measured to have $\bar{Y} - J \gtrsim 0.4$ is considerably less than the number of more common M stars measured to have $i - \bar{Y} \gtrsim 3$. Candidate lists based on simple colour cuts are hence dominated by scattered cool M dwarfs (with some L and T dwarfs), as shown empirically by the results of the follow-up photometric measurements discussed in Section 3.6 and Mortlock et al. (in preparation), as well as the spectroscopic observations made by Glikman et al. (2008). Hence an accurate model of the intrinsic distribution of M dwarfs (and, to a lesser extent, L and T dwarfs) in the UKIDSS–SDSS bands is necessary to calculate useful values of $P_q$.

There are several possible approaches for modelling the stellar population. One option would be to fit the distribution of observed colours or magnitudes (cf. Richards et al. 2004; Bovy et al. 2011), although this would not correctly recover the tails of the (magnitude-dependent) noise distributions which are critical to this problem. Under the approximation that most of the relevant sources are sufficiently faint to be dominated by background noise (cf. Section 2.1), it might be possible to combine deconvolution techniques to estimate the intrinsic distribution (cf. Bovy et al. 2011), but this approach was not explored here. The other extreme would be to develop a physical model for the local population of M, L and T dwarfs (cf. Deacon & Hambly 2006; Findlay et al. 2011), although it is not clear that the gain from this extra complication in the context of the problem of quasar-selection is significant: it is only important to describe the intrinsic distribution of measurable properties of these sources.

An intermediate approach to modelling the source population was adopted here. The population of nearby M, L and T dwarfs was described by developing a parametrized function to describe their intrinsic (i.e. not noise-convolved) distribution of magnitudes and colours. The reason for modelling the intrinsic distribution is to be able to estimate the probability of stars scattering into sparsely populated regions of colour space. While the core of the observed stellar distribution could be modelled empirically using the distribution of observed colours, such an approach would not provide reliable estimates of the probability of a source being scattered into the tails of the distribution (where the plausibly HZQ candidates lie). It is better to model separately the intrinsic population and the noise distribution (Section 2.1) and to then convolve the two to provide the desired extreme scattering probability.

The adopted stellar population model has two independent parameters, the intrinsic $i$ and $Y$ magnitudes of the stars, both of which are observables. In terms of the notation of Section 2, the parameter space is defined by $\theta_i = [i, Y]$ and the number density of stars per unit solid angle is thus written as $\rho_i(\theta_i) = \rho_i(i, Y)$. With $i - Y$ serving as a proxy for stellar temperature, the other colours (i.e. $z - Y$ and $Y - J$) are, to a sufficiently good approximation, dependent only on $i - Y$. The specific form of $\rho_i(i, Y)$ was obtained by comparing the predicted distribution of observed photometry to a sample of well-measured UKIDSS–SDSS point sources (with $15.0 \leq \bar{Y} \leq 19.5$ and $i - \bar{Y} \geq 2.0$) extracted from the WSA. It was critical to ensure that the distribution describing the intrinsic population was convolved with the correct photometric noise distribution when comparing with the observed counts; a full description of the fitting procedure is given in Appendix B.

Several different functional forms for $\rho_i(i, Y)$ were investigated before adopting power-law number counts in combination with an exponential power-law colour distribution for the reddest stars. Taken together,

$$\rho_i(i, Y) = \rho_i 10^{i(i-18.0)} \times \Theta(i - Y - 2) \frac{[\ln(\beta + \gamma Y)]^{1/\beta} (\beta + \gamma Y)^{-\gamma(i-Y)^\delta}}{P} \frac{1}{\delta},$$

where $\Theta(x)$ is the (Heaviside) step function, $P(t, x) = \int_x^\infty x^{-1}e^{-t}dx$ is the complementary incomplete gamma function and the best-fitting values of the free parameters are $\rho_i = 20.9$ mag$^{-2}$ deg$^{-2}$, $\alpha = 0.45939287$, $\beta = 551.74630495$, $\gamma = -16.49969157$ and $\delta = 0.04050890$. The best-fitting distribution $\rho_i(i, Y)$ is shown as a function of $i - Y$ for several values of $Y$ in Fig. 2.

The model represents an average of the stellar population over the range of Galactic latitudes, $b$, covered by the UKIDSS LAS. The LAS was deliberately designed to avoid low-$b$ fields, so the surface density of the reddest stars (which are seen only to moderate distances) varies by less than a factor of $\sim 2$ over the whole survey area. Taking $W_i(d, \text{det}) = 2W_i(d, \text{det})$ in equation (8) would decrease $P_q$ by a factor of $\sim 5$ at most; and, for the vast majority of

![Figure 2](https://academic.oup.com/mnras/article-abstract/419/1/390/1002314/444) The best-fitting intrinsic distribution of $i - Y$ colours, $\rho_i(i + i - Y, Y)$ for different values of $Y$ as labelled.
sources which are decisively classified, $P_q$ remains essentially unchanged. In only a small fraction of cases would a poor candidate be erroneously included for follow-up due to this effect. For a survey covering a greater range of Galactic latitudes it would be important to include the $b$-dependence of $p_i(i, Y)$, most simply by multiplying $W_i(d, \det)$ by a $b$-dependent scaling, the nature of which could be inferred from the survey data.

The predicted photometry in the other relevant bands (specifically $z$ and $J$) is then given by the empirical colour relations

$$z - Y = 0.362 + 0.314 (i - Y)$$

and

$$Y - J = 0.328 + 0.088 (i - Y) + 0.0295 (i - Y)^2.$$  

(11)

The latter is shown directly in Fig. 1; taken in combination these two relations define the stellar locus as shown in other two-colour diagrams. These models accurately encode the tight colours of M dwarfs, which is the main requirement here, but the L and T dwarfs have a significant intrinsic spread in $Y - J$ for a given $i - Y$ and in some cases are significantly bluer in $Y - J$ (Hewett et al. 2006; Chiu et al. 2008; Burningham et al. 2010). These features of the dwarf population ought to be taken into account to obtain the most accurate values of $P_q$, but it is the model above that was used to generate the sample of HZQ candidates described here and analysed in Mortlock et al. (in preparation). In particular, the completeness calculation is tied to the above model (i.e. not just the functional forms adopted, but the specific parameter values as well). A more accurate model of the L and T dwarf population might result in a more efficient selection procedure (as some sources with $\hat{Y} - \hat{J}$ intermediate between the quasar and star values would presumably have lower $P_q$ values), but given that it has been possible to follow-up all the candidates generated using the above stellar model, the resultant depth and completeness calculation (Mortlock et al., in preparation) are correct.

The two colour relationships in equations (10) and (11), together with the data, combine to give the likelihood (equation 5) as a function of $i$ and $Y$. Integrating the product of the likelihood and the stellar density (given in equation 9) over these two parameters then gives the weighted evidence that a source is a star, $W_i(d, \det)$, as defined in equation (3). To be more explicit, specializing to the stellar case and the UKIDSS–SDSS filters allows equation (3) to be written as

$$W_i(\hat{i}, \hat{z}, \hat{Y}, \hat{J}, \det) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_i(i, Y) \Pr(\det|i, Y, s) \Pr(\hat{t}, \hat{z}, \hat{Y}, \hat{J}|i, Y, s) \, di \, dY.$$  

(12)

The colour relationships in equations (10) and (11) combine with the $i$ and $Y$ to give the predicted photometry in the four bands of interest, from which the conversion to flux units allows the likelihood to be written as a product of four Gaussians (Section 2.1). The detection probability is close to unity for the range of $i$ and $Y$ being considered, although in practice adopting detection cut-offs is the simplest way to ensure the integrals do not extend to such faint fluxes that confusion issues would become relevant.

### 3.3 The quasar population at high redshift

UKIDSS is the first survey to have had a significant chance of detecting $z \geq 7$ quasars, so there are no empirical constraints on much of the target HZQ population. This does not, however, mean that there is no information about the HZQ population beyond the current redshift limit – it is perfectly reasonable to extrapolate from the results of $z \sim 6$ quasar surveys. Indeed, one of the main principles of Bayesian reasoning is that any available information should be applied if possible, even if it is incomplete or imprecise. A reasonable approximation to the correct prior (given in this case by the actual, but unknown, QLF) will result in final inferences that are superior to those derived from any method which does not include any information about the likely numbers of $z \geq 7$ quasars. The alternative approach is, in the case of rare object searches, inevitably overly optimistic (i.e. $P_q$ would be unreasonably high for large numbers of sources).

The measured numbers of bright HZQs (e.g. Fan et al. 2003; Jiang et al. 2008; Willott et al. 2010b) are consistent with a comoving QLF given by a power law of the form

$$\Phi_q(M, z) = 5.2 \times 10^{-9} \times 10^{0.84M + 26.0 - 0.47(z - 6.0)} \, \text{mag}^{-1} \, \text{Mpc}^{-3}. $$

(13)

This parametrization combines the magnitude dependence measured by Fan et al. (2003) with the evolution model used by Jiang et al. (2008). Willott et al. (2010b) found a significantly lower normalization than Jiang et al. (2008), and in principle the implied uncertainty should be included in the calculation of $P_q$. However, the resultant probabilities are unchanged in almost all cases, and the higher normalization was adopted to ensure a (marginally) more inclusive candidate list. (Moreover, the above values were already being used to select UKIDSS HZQ candidates before the results of Willott et al. 2010b were available.)

It would be possible to use $M$ and $z$ to parametrize the quasar population, but it is more intuitive to convert $M$ to the observable $Y$-band magnitude, so that the quasar population is, like the stellar population defined in Section 3.2, characterized by an observable surface density. Thus the quasar parameters used are $\theta_i = (Y, z)$, which leads to the population model

$$\Phi_q(Y, z) = \frac{1}{4\pi} \frac{dV_{\text{co}}}{dz} \Phi_q[Y - D_L(z) - K_Y(z), z].$$

(14)

where $dV_{\text{co}}$ is the comoving volume of a spherical shell of thickness $dz$ at redshift $z$, $D_L(z)$ is the luminosity distance and $K_Y(z)$ is the $Y$-band quasar $K$-correction that converts the rest-frame AB magnitude at 0.1450 μm to the $Y$-band Vega magnitude.

The quasars’ $K$-corrections are, in turn, evaluated using updated and expanded versions of the model quasar spectra developed by Hewett et al. (2006) and Maddox et al. (2008). In particular, the models include the absorption bluewards of the Ly α emission line caused by the increased density of H i above $z \approx 5.8$, as measured by Fan et al. (2006), although the ionized proximity zones around the quasars are not accounted for, so the absorption extends all the way to the redshifted Ly α wavelength. The variety of the quasars’ intrinsic properties is accounted for by using colours from 12 different templates that span four line-strengths and three continuum slopes. As the main motivation of using multiple models is to ensure

\footnote{The QLF is defined such that the average number of quasars with absolute magnitudes between $M$ and $M + dM$ in a comoving volume $dV_{\text{co}}$ at redshift $z$ is $\Phi_q(M, z) \, dM \, dV_{\text{co}}$.}

\footnote{The QLF parameters are reasonably well constrained by the data with the exception of the evolution term. The value found by Fan et al. (2001) over the redshift interval $3.6 \leq z \leq 5.0$ was used, although there is some evidence that the evolution at redshifts of $z \geq 7$ might be stronger (Mortlock et al., in preparation).}
appropriately high values of \( P_q \) for the HZQs with redder continuum slopes (that are closer to the stellar locus in \( Y - J \) than the fiducial quasars), the 12 models are weighted equally, even though it would be more accurate to down-weight the more extreme templates. Using a range of models also accounts for the colour variations that result from the combination of intrinsic quasar variability and non-simultaneous measurements: most UKIDSS observations took place several years after the SDSS observations of the same fields, a time-scale on which Ivezić et al. (2004) found a typical variation of \( \sim 0.15 \) mag. When multiplied by the SDSS and UKIDSS filter profiles and integrated over wavelength, the model spectra not only give \( K_\text{F}(z) \), but also the required optical–NIR colours. All 12 colour tracks are shown in Fig. 1.

The HZQ colour relationships described above, together with the data, combine to give the likelihood (equation 5) as a function of \( z \) and \( Y \). Integrating the product of the likelihood and the quasar density (given in equation 14) over these two parameters then gives the weighted evidence defined in equation (3) as

\[
W_q(i, \hat{z}, \hat{Y}, \hat{J}, \text{det}) = \int_0^\infty \int_{-\infty}^\infty \rho_q(Y, z) \text{Pr(det|Y, z, q)} \text{Pr}(i, \hat{z}, \hat{Y}, \hat{J}|Y, z, q) dY dz.
\]

(15)

The quasar model loci give the predicted \( i, \hat{z}, \hat{Y}, \hat{J} \) photometry, from which the conversion to flux units allows the likelihood to be written as a product of four Gaussians (Section 2.1). As with the integral over the stellar population (equation 12), a simple detection cut-off is needed to ensure that the integral is not dominated by the numerous undetectable ultra-faint sources that are beyond the confusion limit in the relevant bands.

### 3.4 The probability that a source is a HZQ

Having developed quantitative models for the stellar population (Section 3.2) and HZQ population (Section 3.3), the measured \( i, \hat{z}, \hat{Y}, \hat{J} \)-band photometry of a source can then be used to calculate \( P_q \) according to equation (4). The two-dimensional (weighted) evidence integrals over the quasar and star parameters are evaluated using simple numerical quadrature as this is faster than more general Monte Carlo techniques for a problem of such low dimensionality. On a standard desktop computer one evaluation of \( P_q \) takes between a tenth and a hundredth of a second, which is sufficiently fast that even the most inclusive of candidate lists can be analysed.

The speed with which \( P_q \) can be calculated also means that it is possible to explore how \( P_q \) depends on the measured photometry and the associated errors. This is potentially important, as the reasons that some candidates have low or high probabilities are not always obvious. In this context it is useful to think of \( P_q \) not as a quantity associated with a candidate HZQ, but as a function of the information that is particular to each source (i.e. the measured photometry and the associated uncertainties), leading to \( P_q = P_q(i, \Delta i, \hat{z}, \Delta\hat{z}, \hat{Y}, \Delta\hat{Y}, \hat{J}, \Delta\hat{J}) \) for the UKIDSS–SDSS sample. This function has too many parameters to explore comprehensively, but many of its important features can be seen in two-dimensional projections. Plotting \( P_q \) in the space of measured colours also facilitates direct comparison with selection methods based on colour cuts, which can be cast into a Bayesian form by treating them as if \( P_q = 1 \) for objects satisfying the cuts and \( P_q = 0 \) otherwise. In the regions of parameter space for which \( P_q \) varies rapidly with the measured colours, the Bayesian selection reduces to a cut-based method, but with the important difference that the selection boundaries are defined optimally and objectively.

The simplest non-trivial case is that of a two-band survey for which each detected source can be treated as having one measured magnitude and a single measured colour. If the two bands are taken to be \( i \) and \( \hat{z} \), this is a reasonable approximation to the SDSS HZQ survey, for which the initial selection criteria were \( \hat{z} < 20.2 \) and \( i - \hat{z} > 2.2 \) (Fan et al. 2001). Ignoring the \( Y \) and \( J \) bands and assuming the fiducial SDSS depths in \( i \) and \( \hat{z} \) given in Section 3.1, the variation of \( P_q \) with \( \hat{z} \) and \( i - \hat{z} \) is shown in Fig. 3. The only sources which have \( P_q \geq 0.1 \) are those which have \( \hat{z} \lesssim 20 \) and \( i - \hat{z} \gtrsim 2.5 \), roughly corresponding to the region of parameter space selected by Fan et al. (2001). The seemingly counter-intuitive result that \( P_q \) does not increase monotonically with \( i - \hat{z} \) is an artefact of the asinh magnitude system.

It is also noticeable from Fig. 3 that no pair of (measured) \( \hat{i} \) and \( \hat{z} \) values would result in \( P_q \approx 1 \), a result which is independent of the depth of the observations. This is because the sources which appear red in \( i - z \) are not just scattered stars, but also L and T dwarfs which actually have these red colours (and outnumber the target HZQs). The only way to generate a sample of candidates with higher \( P_q \) is to obtain data in another band, chosen such that quasars and the potential contaminants have distinct colours. This can be done by follow-up (e.g. in the \( J \) band, as done by Fan et al. 2001) or by extending the wavelength coverage of the initial survey (e.g. UKIDSS \( Y \)-band imaging). The choice between these two strategies is sometimes difficult, as adding an extra band to a survey costs area or depth, whereas the number of follow-up observations required to complete a two-band search is potentially prohibitive. However in terms of this exploration of how \( P_q \) depends on the measured colours there is no distinction, as only the observational depths and the choice of bands is important, not the sequence of observations.

The above results imply that measurements in at least three bands are required to generate a sample of strong HZQ candidates, and in particular that two appropriately chosen colours are needed.
Probabilistic quasar selection

Figure 4. The dependence of $P_q$ on a source’s observed $i - z$ and $z - J$ colours, ranging from $P_q = 0$ (white) to $P_q = 1$ (grey), for $z = 18.5$, $z = 19.0$, $z = 19.5$ and $z = 20.0$, as labelled. In all cases the noise is as appropriate for fiducial SDSS and UKIDSS observations with $\lim_{\text{SDSS}} = 22.5$, $\lim_{\text{UKIDSS}} = 20.8$ and $J_{\text{lim}} = 22.5$, but the $Y$ band has been ignored. The quasar and star loci are shown as solid curves and the dashed lines show the SDSS HZQ selection region defined in Fan et al. (2003).

Figure 5. The dependence of $P_q$ on a source’s observed $i - Y$ and $Y - J$ colours, ranging from $P_q = 0$ (white) to $P_q = 1$ (grey), for $Y = 18.0$, $Y = 18.5$, $Y = 19.0$ and $J = 19.5$, as labelled. In all cases the noise is as appropriate for fiducial SDSS and UKIDSS observations with $\lim_{\text{SDSS}} = 22.5$, $\lim_{\text{UKIDSS}} = 20.2$ and $J_{\text{lim}} = 19.5$, but the $z$ band has been ignored. The horizontal dotted line shows the maximum theoretical value of $i - Y$ for a source with $Y = Y$ and zero flux in the $i$ band in the absence of noise; this changes with $Y$ due to the use of asinh magnitudes to represent the $i$-band photometry. The fiducial quasar and star loci are shown as solid curves and the dashed lines show the basic pre-selection made to generate the UKIDSS–SDSS candidate sample.

many HZQ searches have been based on pairs of colour cuts (e.g. Warren et al. 1994; Fan et al. 2001), and this approach is compared directly with the Bayesian results (Figs 4 and 5). In both cases the depths in the included bands are chosen to match the fiducial UKIDSS and SDSS values given in Section 3.1, but data in the missing band ($Y$ and $z$, respectively) are ignored. Fig. 4 approximates the optical SDSS search of Fan et al. (2001) (but cannot mimic the CHFQS described by Willott et al. 2007 because the star and quasar models do not go sufficiently deep), whereas Fig. 5 represents one of the obvious selection options from the matched UKIDSS–SDSS data. As expected, $P_q$ has a similar colour-dependence, being low near the stellar locus and higher where the quasars are expected to be found. There is also a strong correspondence between the region of high $P_q$ and the specific selection region defined by Fan et al. (2003), particularly close to the $z$-band selection cut at $z = 20.1$. There are, however, significant systematic differences between the Fan et al. (2003) selection region and the high-$P_q$ region. The most obvious difference is that $P_q$ also varies with magnitude for a given set of measured colours. The most important aspect of the magnitude-dependence is the decrease in the size of the high-$P_q$ region close to the detection limit in the reference band (i.e. as $z \rightarrow 20.8$ or $Y \rightarrow 20.2$). For sources well above the detection limit the photometric errors are sufficiently small that there is only a minimal chance of such bright stars being measured with HZQ-like colours. But for fainter sources close to the detection limit the effective width of the observed stellar locus is greatly increased and, in the example shown in Fig. 4, encompasses the HZQ locus.

The somewhat counter-intuitive consequence is that a sample of sources with essentially identical measured colours (but a range of ‘reference’ magnitudes) can include both near-certain HZQs and obviously uninteresting scattered stars. The dependence of $P_q$ on reference magnitude is shown for SDSS-like observations in Fig. 6 and for UKIDSS-like observations in Fig. 7, with results shown for a representative range of redshifts.

With access only to SDSS $i$- and $z$-band data (i.e. no follow-up observations in $J$) the only measured colour is $i - z$, and so $L$ and $T$ dwarfs, along with any cool dwarf which is scattered faint in $i$, would be photometrically indistinguishable from a HZQ. Hence, as shown by the lower set of curves in Fig. 6, $P_q \lesssim 0.3$ even for a reasonably bright source with the measured $i - z$ colour of a target quasar. Fainter than $i \sim 20$ the quasar probability decreases to $P_q \sim 0$ as the larger photometric errors mean that any sample of sources with red $i - z$ is increasingly dominated by scattered stars. If a selection cut-off of $P_q = 0.1$ was adopted then the approximate depth of the survey would be $z \sim 20.4$, comparable to the selection cut of $z = 20.2$ adopted by Fan et al. (2003). The apparent implication is the
0.1 used in selecting UKIDSS–SDSS candidates

The dashed horizontal line shows the cut of $P_q = \approx 0.1$ that might reasonably be used in selecting candidates for follow-up and hence would give the approximate completeness limit of the HZQ survey at different redshifts.

**Figure 6.** The dependence of $P_q$ on a source’s observed $z$-band magnitude given it has the observed colours of a HZQ with a redshift of $z = 6.0$, $z = 6.2$ and $z = 6.4$, as labelled. The lower set of curves correspond to SDSS $i$- and $z$-band data only (i.e. only the $i-z$ colour is measured); the upper set of curves correspond to UKIDSS-like $J$-band measurements as well (i.e. the $i-z$ and $z-J$ colours are measured). The dashed horizontal line shows the cut of $P_q = 0.1$ that might reasonably be used in selecting candidates for follow-up and hence gives the approximate completeness limit of the HZQ survey at different redshifts.

SDSS HZQ search could have gone deeper, but the fact that most candidates in that search were single-band detections in $i$-band means that the additional problem of spurious detections (which is not included in the statistical model used here) also had to be considered.

With access to even moderately deep $J$-band data, however, the HZQs are identified much more reliably, as shown by the upper set of curves in Fig. 6. Sources with the measured $i-z$ and $z-J$ colours of HZQs over the redshift range $6.0 \lesssim z \lesssim 6.5$ have $P_q \simeq 1$ down to $z \simeq 20.3$ and the nominal threshold of $P_q = 0.1$ is only reached close to the fiducial SDSS $z$-band detection limit of $z \simeq 20.8$.

The results for UKIDSS-like observations shown in Fig. 7 reveal that HZQs over the redshift range $6.0 \lesssim z \lesssim 7.0$ have $P_q \simeq 1$ for $\hat{Y} \lesssim 19$, after which $P_q$ falls fairly sharply due to the greater numbers of stars scattered to have quasar-like colours. For each of the three redshift values shown the source’s measured colours are constant across the plot, and so it remains perfectly consistent with being a HZQ even below the UKIDSS detection limit; the change comes about as the observed stellar locus is broader for fainter $\hat{Y}$, and for $\hat{Y} \gtrsim 19.5$ it effectively covers the quasar loci. The redshift-dependence of the effective depth comes about due to the small variations of the HZQs’ $Y-J$ colour as various emission lines appear in different filters. As can be seen in Fig. 1, the expected $Y-J$ colours of HZQs increases from $Y-J \simeq 0.4$ at a redshift of $z \simeq 6.0$ to $0.6$ at a redshift of $z \simeq 6.8$, bringing them closer to the stellar locus. Hence the maximum depth at which quasars remain well-separated from the observationally broadened stellar locus is decreased and the effective depth (defined as the $Y$ magnitude at which $P_q = 0.5$) decreases from $\hat{Y} \simeq 19.7$ at a redshift of $z \simeq 6.0$ to $\hat{Y} \simeq 19.3$ at a redshift of $z \simeq 6.5$. At higher redshifts, however, the small increase in $Y-J$ is much less important than the large increase in the HZQs’ expected $i-Y$ values. As a result the effective depth at a redshift of $z \simeq 7.0$ has increased to $\hat{Y} \simeq 19.5$.

One implication of these various subtle effects is that the selection function of the UKIDSS–SDSS HZQ search given in Mortlock et al. (in preparation) has a more complicated redshift dependence than the SDSS (Fan et al. 2003) or CFHQS (Willott et al. 2010b) selection functions.

Thus far the emphasis has been on the variation of $P_q$ with the properties of a source, but it is also revealing to investigate how $P_q$ depends on survey depth. Fig. 8 shows how $P_q$ varies with $z$-band depth (assuming a fiducial magnitude of $\hat{Y} = 19.5$). As expected, extra depth in the $i$ band increases the confidence with which $z \simeq 6$ quasars can be identified. It is possible to push closer to the stellar locus (i.e. redder in $Y-J$) with confidence, and also possible to find fainter HZQs (i.e. going deeper in $Y$). Increasing the depth in the $Y$ or $J$ bands is not nearly as useful, because the $Y-J$ colour is already sufficiently well measured for a $\hat{Y} \simeq 19$ source. However extra depth in all three bands would allow a deeper survey, and hence greater numbers of HZQs, albeit of lower intrinsic luminosity. This variation also shows the importance of calculating $P_q$ using the measured noise levels in each field of a survey, rather than just using generic survey-wide depths (cf. Mortlock et al., in preparation).

In almost all the above examples the transitions between regions of high and low $P_q$ are quite sharp, with no large areas of uncertainty. The transition scale is set by the photometric errors, although for a given observation and reference magnitude (as is the case here) the errors vary with colour. This also explains why the transition is more gradual in regions of colour space which are expanded by the decreased variation of colour with flux, which is the case for red $i-z$ or $i-\hat{Y}$ here. The one case of an obviously gradual transition is shown the right-hand panel of Fig. 5, in which the source is sufficiently faint that the measurement uncertainties in all the relevant bands are $\gtrsim 0.2$ mag.

### 3.5 Photometric redshift estimation

The most important information that can be extracted from the measurements of a candidate is the probability that it is a HZQ; but if the source is assumed to be a quasar then the photometry can also be used to estimate its likely redshift. Given that a spectrum is necessary to confirm any candidate as a quasar, there is little long-term utility of such photometric redshift estimates, but they can be useful in prioritizing follow-up observations of HZQ candidates. The essential logic that it is most efficient to pursue high-$P_q$ objects first is somewhat modified by the fact that there is a particular premium on finding the most distant sources (i.e. HZQs with $z \gtrsim 6.5$ in the case of UKIDSS). It would probably make more sense to follow-up a $z$-band drop-out $z \gtrsim 6.5$ quasar candidate with $P_q \simeq 0.1$.
than a more secure candidate with \( P_q \simeq 0.9 \) that was well-detected in the \( z \) band (a fact which might have contributed to the high \( P_q \) in the first place).

The calculation is analogous to that used in Bayesian photometric redshift estimation of galaxies (e.g. Benítez 2000) but with the important difference that most quasar spectra exhibit far less variety than those of galaxies, so that it is far easier to obtain reliable parameter estimates. The posterior distribution of a (putative) quasar’s redshift, given photometric measurements \( \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu} \) and \( \hat{\nu} \), is calculated by modifying equation (15) to marginalize only over the unknown true \( Y \)-band magnitude. That gives

\[
Pr(z|\hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}) = \int_{\nu_{\text{lim}}}^{\infty} \rho_q(\nu, z) Pr(\nu|\hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}) \frac{Pr(\hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}|\nu, z, q) d\nu}{W_\nu(\hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu})} ,
\]

(16)

where \( \rho_q(\nu, z) \) is given in equation (14), the appropriate form of the likelihood, \( Pr(\nu|\hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}) \), is discussed in Section 2.1 and the denominator \( W_\nu(\hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}, \hat{\nu}) \) ensures the probability is correctly normalized. (The posterior distribution in any of the model parameters – for either quasars or stars – could similarly be evaluated by modifying equation (3) to marginalize over nuisance variables.)

The redshift posteriors for several simulated HZQs are shown in Fig. 9. In each case the photometry is exact (i.e. no noise has been added) but the calculation is performed using uncertainties appropriate to UKIDSS–SDSS measurements of a source of \( Y = \hat{Y} = 19.5 \). The highly non-linear nature of the likelihood means that the peak of the posterior does not necessarily match the input values, even with a uniform prior in the redshift and \( Y \)-band magnitude; with a realistic prior the most probable model given the (simulated) data is inevitably a more common fainter source at a lower redshift.

Shown in Fig. 10 are the redshift posteriors for the 12 HZQs in the main UKIDSS sample (see Section 3.6). In each case the posterior distributions inferred from the UKIDSS–SDSS survey photometry (under both realistic and fiducial uniform priors in \( z \) and \( Y \)) are compared to the spectroscopically measured redshift. The similarity of the distributions for the two different priors indicates that the detailed knowledge of the HZQ population is not important: the UKIDSS–SDSS photometric data (in combination with the model of the quasars’ colours) are sufficient to give largely prior-independent redshift estimates that have uncertainties of \( \Delta z \simeq 0.1 \). These inferences are broadly verified by the subsequent spectroscopic redshift measurements. The basic success of the method is unsurprising given that Fig. 1 shows that a good measurement even
Figure 10. Posterior distributions of the redshifts of the 12 HZQs in UKIDSS DR8 inferred from the UKIDSS–SDSS survey photometry. Results for both a realistic (solid curve) and uniform (dashed curve) prior in the quasar’s redshift and $Y$-band magnitude are shown. The vertical lines show the spectroscopically measured redshift of each quasar.

of just $i - Y$ would be sufficient to obtain a reasonable redshift estimate, although the uncertainties would be considerably increased if only this one colour was used.

There is, however, some suggestion that the photometric redshift estimates of the real HZQs are systematically biased low. The spectroscopic redshift is higher than the peak of the photometric posterior (for either prior) for eight of the 12 sources, although only in the case of SDSS J1030+0524 (Fan et al. 2001) is the discrepancy significant. If the bias is real then the implication is that the quasar models described in Section 3.3 do not cover the full range of HZQ spectra. For example, the photometric posterior would be shifted to higher redshift if the absorption models accounted for the ionized proximity zones around an HZQ as the spectral break would be shifted bluewards. Including this effect would, however, only result in shifts of $\Delta z \lesssim 0.05$ (e.g. Fan et al. 2006), so cannot fully explain the apparent discrepancy seen in Fig. 10. Another possible reason for low photometric redshift estimates is that the models do not include spectra with sufficiently strong emission lines to mimic the colours of HZQs with, in particular, strong Ly$\alpha$ emission. Certainly, SDSS J1030+0524 has an unusually strong Ly$\alpha$ line (with a rest-frame equivalent width of $\sim 70$ Å; Fan et al. 2001), although its photometric redshift is determined primarily by its measured UKIDSS–SDSS $i - Y$ colour of $\sim 3.8$, which is essentially independent of its Ly$\alpha$ line (as it falls in the $z$ band). Broad absorption line (BAL) quasars can also result in mis-estimated redshifts, although in the case of ULAS J0203+0012 (Venemans et al. 2007) the initial redshift was higher than that obtained once it was confirmed as a BAL (Mortlock et al. 2009b).

Whatever the limitations of the photometric redshift estimation might be, it is clear that results are sufficiently accurate to identify the most promising $z \gtrsim 6.5$ candidates for follow-up observations. Given only the redshift posteriors shown in Fig. 10, it is clear that ULAS J112001.48+064124.3 (Mortlock et al. 2011), the one $z \gtrsim 6.5$ quasar in the sample, would have been prioritized above the other 11 candidates for follow-up photometry and, eventually, spectroscopy. Moreover, there are many other considerations that must be taken into account during the follow-up process, as described in Section 3.6.

3.6 Results

Having understood how the photometric measurements of a source combine with models of the star and quasar populations to give $P_q$ (Section 3.4), the Bayesian selection method could be applied with confidence to the UKIDSS–SDSS sample of $\sim 9 \times 10^3$ candidate HZQs described in Section 3.1. The value of $P_q$ was calculated for each source, and the high-probability tail of the resultant distribution is shown as the dotted histogram in Fig. 11. The main result was
that the vast majority of sources with HZQ-like colours had \( P_q \ll 1 \) and were not considered further. Given the low number of HZQs expected in the sample, the rejection of most candidates was both desirable and predictable, and immediately rendered the follow-up task far more manageable.

Applying a cut of \( P_q \geq 0.1 \) left 1138 promising candidates in UKIDSS DR8, a completely automated reduction by a factor of \( \sim 3 \times 10^4 \) from the initial UKIDSS–SDSS sample. Having extracted essentially all the useful information from the photometry, the next step was to inspect the SDSS and UKIDSS images of the remaining candidates. Most of the candidates were, unsurprisingly, revealed to be spurious, with a large variety of explanations (see Mortlock et al., in preparation for more details): supernovae in faint galaxies that were bright in the UKIDSS observations but absent when the SDSS observations were made; Solar system asteroids that were observed by UKIDSS close to turn-around; UKIDSS cross-talk or persistence (Dye et al. 2006); and various data artefacts that resulted in obviously incorrect photometry. In theory, all such contaminants could be included as additional models (along with stars and quasars) in the Bayesian classification scheme; but, aside from the difficulties in specifying the likely numbers and properties of these various contaminants, they are sufficiently rare that their identification does not consume significant resources. In almost all the above cases visual inspection of the SDSS and UKIDSS images resulted in the confident identification of the above contaminants, although in some cases a confirmation image was obtained to check that probable asteroids and supernovae were indeed absent.

Of the 1138 UKIDSS DR8 objects with \( P_q \geq 0.1 \), only 107 were confirmed as real, stationary, astronomical sources with no obvious data problems that might result in erroneous photometry. Fig. 12 shows the colours of these candidates both from the initial UKIDSS–SDSS survey photometry (left-hand panel) and then updated after follow-up observations (right-hand panel). The distribution of these (real) candidates’ quasar probabilities is shown as the dashed histogram in Fig. 11. The distribution of \( P_q \) values can be used as a guide to calibrate the probability calculation: if the models of the measurement process and the two populations were completely accurate then the sum of \( P_q \) over all the candidates would be approximately equal to the expected number of HZQs in the sample. That is not the case here: whereas only \( \sim 15 \) HZQs are expected, the initial probability sum is \( \sim 60 \) and the sum after follow-up observations is \( \sim 40 \). The most plausible reason for these discrepancies is the use of a Gaussian likelihood (see Section 2.1) rather than a distribution with heavier tails to account better for outliers. The fact that the sum even after follow-up is still so high indicates that the probabilities even for poor candidates with \( P_q \lesssim 0.01 \) are systematically high. It is, however, the relative probabilities (although not just their rankings) that is of primary importance – the real reason for exploring a Bayesian selection algorithm in the first place was not so much to determine the probability that the candidates are HZQs, but to answer the distinct question of which of the identified candidates were the most likely to be quasars.

Having ranked the candidates, the first stage of the follow-up campaign was to see whether any of these red UKIDSS–SDSS sources had been previously confirmed as HZQs. Six of the highest-ranked objects were known HZQs that had been identified using SDSS photometry alone: SDSS J0836+0054 at redshift \( z = 5.82 \) (Fan et al. 2001); SDSS J0841+2905 at redshift \( z = 5.96 \) (Goto 2006); SDSS J1030+0524 at redshift \( z = 6.31 \) (Fan et al. 2001); SDSS J1044−0125 at redshift \( z = 5.74 \) (Fan et al. 2000); SDSS J1411+1217 at redshift \( z = 5.93 \) (Fan et al. 2004) and SDSS J1623+3112 at redshift \( z = 6.22 \) (Fan et al. 2004). All six were very strong candidates with \( P_q \geq 0.99 \), and so clearly would have been confirmed by follow-up observations had they been required. The remaining high-\( P_q \) candidates were queued for follow-up photometric observations at the Liverpool Telescope (i filter), the Isaac Newton Telescope (i and z filters) or UKIRT (\( Y \) and J filters). The follow-up images were generally deeper than the SDSS and UKIDSS survey observations, but at least as important as the increase in photometric precision was the fact that these new measurements were independent of the candidate selection process. Whereas the initial selection could be thought of as a method of identifying stars for which the SDSS measurements are faint in \( i \) or in which the UKIDSS data are bright in \( Y \), any follow-up photometric data are unbiased, if still noisy, measurements of the sources’ true properties.

Every time a follow-up measurement was made the quasar probability was recalculated with the new photometry; a candidate was discarded if \( P_q \lesssim 0.1 \) at any stage of this process. Many candidates were hence rejected after just one photometric observation, in most cases because they were revealed to be significantly brighter in the \( i \) band than indicated by the initial survey photometry. Such candidates can be seen with \( i - \bar{Y} \approx 2.5 \) in the right-hand panel of Fig. 12; as expected, their observed colours are much more like those of the reddest M dwarfs. (Further follow-up observations in the \( Y \) band of such objects would probably reveal that the true \( i - Y \) values are bluer still – as the initial sample was selected to be red in \( i - Y \), the initial \( Y \)-band photometry of the candidates tends...
Figure 12. Two-colour diagrams showing the UKIDSS DR8 LAS point sources which, after being cross-matched to SDSS, have $P_q \geq 0.1$ from the UKIDSS–SDSS survey photometry. The colours from the survey photometry are shown in the left panel; the results of follow-up photometry are shown in the right-hand panel. The 12 confirmed HZQs in the main UKIDSS DR8 sample are shown in solid symbols; a small number of candidates rejected for reasons other than their photometry are shown as circles with crosses; the other candidates, which were rejected only because their follow-up photometry gave $P_q < 0.1$, are shown as open circles. A single follow-up measurement (generally in the $i$ band) was sufficient to reject most candidates, so one colour (most often $Y-J$) of many objects is the same in the two panels; also, there is one candidate which is yet to be re-observed and so appear at the same position in both panels (as do the previously known HZQs and the candidates rejected for reasons other than their photometry). Also shown is the stellar locus described in Section 3.2 and the 12 quasar models described in Section 3.3, with crosses indicating redshifts of $z = 5.5$, $z = 6.0$, $z = 6.5$ and $z = 7.0$. All the model colours are calculated for a source that has $Y = 19.5$. The dashed lines denote the initial pre-selection colour cuts, defined in Section 3.1, that are applied before subsequent processing. The horizontal dotted line shows the maximum $i-Y$ value that a $Y = 19.5$ source could have in the absence of noise.

It is also striking that a number of the rejected candidates still have the observed $i-Y$ and $Y-J$ colours of HZQs, again illustrating the strong role that the priors (particularly the relative numbers of stars and quasars) play in this selection process.

If, after re-observation in at least the $i$, $Y$ and $J$ bands, a candidate still had $P_q \simeq 1$ (and assuming it was not already known to be a HZQ) then a spectroscopic observation was made, in most cases using the Gemini North telescope. As detailed in Mortlock et al. (in preparation), spectra were taken of six UKIDSS DR8 candidates, all of which were confirmed as new HZQs: ULAS J0148+0600 at redshift $z = 5.98$ (Venemans et al., in preparation); ULAS J0828+2633 at redshift $z = 6.1$ (Patel et al., in preparation); ULAS J1120+0641 at redshift $z = 7.09$ (Mortlock et al. 2011); ULAS J1148+0702 at redshift $z = 6.2$ (Patel et al., in preparation); ULAS J1207+0630 at redshift $z = 6.03$ (Patel et al., in preparation) and ULAS J1319+0950 at redshift $z = 6.13$ (Mortlock et al. 2009b). At the end of this follow-up process every source with $P_q \geq 0.1$ from the UKIDSS–SDSS survey photometry was either convincingly rejected or spectroscopically confirmed as a HZQ. This separation into sources with $P_q \simeq 0$ and confirmed HZQs with $P_q = 1$ is a natural result of obtaining more information about those candidates that were initially promising but fundamentally ambiguous. The separation is not readily apparent in Fig. 12 – neither the $z$-band photometry nor the photometric errors are shown – but it is well illustrated by the solid histogram in Fig. 11.

It is also important to assess whether the minimum $P_q$ cut-off is sufficiently inclusive that HZQs are not being missed. Simulations could go some way to answering this question, but it is more robust to examine the distribution of $P_q$ values of the confirmed quasars. This is illustrated in Fig. 13, which shows probabilities (calculated to be biased bright, just as their initial $i$-band photometry is biased faint.) It is also striking that a number of the rejected candidates still have the observed $i-Y$ and $Y-J$ colours of HZQs, again illustrating the strong role that the priors (particularly the relative numbers of stars and quasars) play in this selection process.  

If, after re-observation in at least the $i$, $Y$ and $J$ bands, a candidate still had $P_q \simeq 1$ (and assuming it was not already known to be a HZQ) then a spectroscopic observation was made, in most cases using the Gemini North telescope. As detailed in Mortlock et al. (in preparation), spectra were taken of six UKIDSS DR8 candidates, all of which were confirmed as new HZQs: ULAS J0148+0600 at redshift $z = 5.98$ (Venemans et al., in preparation); ULAS J0828+2633 at redshift $z = 6.1$ (Patel et al., in preparation); ULAS J1120+0641 at redshift $z = 7.09$ (Mortlock et al. 2011); ULAS J1148+0702 at redshift $z = 6.2$ (Patel et al., in preparation); ULAS J1207+0630 at redshift $z = 6.03$ (Patel et al., in preparation) and ULAS J1319+0950 at redshift $z = 6.13$ (Mortlock et al. 2009b). At the end of this follow-up process every source with $P_q \geq 0.1$ from the UKIDSS–SDSS survey photometry was either convincingly rejected or spectroscopically confirmed as a HZQ. This separation into sources with $P_q \simeq 0$ and confirmed HZQs with $P_q = 1$ is a natural result of obtaining more information about those candidates that were initially promising but fundamentally ambiguous. The separation is not readily apparent in Fig. 12 – neither the $z$-band photometry nor the photometric errors are shown – but it is well illustrated by the solid histogram in Fig. 11.

It is also important to assess whether the minimum $P_q$ cut-off is sufficiently inclusive that HZQs are not being missed. Simulations could go some way to answering this question, but it is more robust to examine the distribution of $P_q$ values of the confirmed quasars. This is illustrated in Fig. 13, which shows probabilities (calculated...
4 CONCLUSIONS

A probabilistic approach to quasar selection can, at least in principle, make use of all the relevant information about a candidate HZQ: the knowledge of the quasar and star populations; the stochastic nature of the measurement process; and, of course, the data obtained for the source in question. Having derived a general Bayesian formalism for HZQ selection, this approach was applied to real quasar candidates taken from the cross-matched UKIDSS and SDSS data sets. The ~2270 deg^2 surveyed as of UKIDSS DR8 contained only 107 real astronomical sources with quasar probabilities of $P_q \geq 0.1$, and most of these candidates were quickly rejected with follow-up photometry (a single i-band image sufficing in many cases). Aside from six known HZQs, this follow-up process left just 12 strong photometric candidates, all of which were confirmed as redshift $z \simeq 6$ quasars. If a cut-based approach had been adopted then follow-up observations would have been required for $\sim 10^3$ candidates. Not only was the Bayesian selection method very efficient, but it was also entirely quantitative and objective, as needed to estimate the high-redshift QLF from the sample (see Mortlock et al., in preparation).

It is also possible to combine this model-selection approach with parameter estimation. In the context of HZQs it is clear that the highest redshift objects are the most important, in which case $P_q$ could be combined with an estimate of the putative quasar’s redshift to rank potentially record-breaking HZQs above those at redshifts which have already been explored. Comparing the resultant photometric HZQ redshift estimates of the 12 confirmed quasars with their spectroscopic redshifts confirms that the photometry is sufficiently informative about a quasar’s redshift that it can be used to prioritize candidates in the follow-up process.

The Bayesian HZQ selection method described here will continue to be used in the analysis of subsequent UKIDSS data releases, and may also be applied to data from future NIR surveys such as Pan-STARRS (Kaiser et al. 2002) and VISTA (Emerson et al. 2004). The utility of – and need for – such a complicated approach to HZQ selection depends on the survey details: the bands used and the depths reached determine the degree to which the target HZQs are separated in colour space from the contaminating stellar population(s). In almost all cases, however, the applications of these Bayesian selection methods would represent a step closer to extracting as much science as possible from the available data.

A corollary of the survey-dependent nature of Bayesian quasar selection is that the expected distribution of candidate probabilities could be a useful tool in survey design. This was particularly apparent by the degree to which the high-$P_q$ region of the SDSS data space matched the HZQ colour cuts adopted by Fan et al. (2001). Given that the trade-off between broader filters (giving a higher S/N) and narrower filters (giving improved colour-based diagnostics) can only be assessed properly in the context of the expected source populations, the separation of the probability distributions for simulated sources of different types would be a powerful diagnostic.

None the less, the principles behind the HZQ selection method presented here are generic, and could be usefully adapted to any astronomical classification problem in which the available data on the sources of interest do not permit decisive classifications to be made. The price is the need to model the relevant source populations, but the pay-off in the case of a search for rare objects is a massive reduction in the amount of follow-up observations required to extract the few unusual sources of interest.

ACKNOWLEDGMENTS

Xiahoui Fan, Sebastian Jester and Gordon Richards all provided invaluable insights into the subtleties of quasar selection. The careful reading by the anonymous referee led to significant improvement to this paper.

MP acknowledges support from the University of London’s Perren Fund. PCH, RGM and BPV acknowledge support from the STFC-funded Galaxy Formation And Evolution programme at the Institute of Astronomy.
REFERENCES
Abazajian K. N. et al., 2009, ApJS, 182, 543
Bailer-Jones C. A. L., Smith K. W., Sordo R., Vallenari A., 2008, MNRAS, 391, 1838
Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
Becker R. H. et al., 2001, AJ, 122, 2850
Benítez N., 2000, ApJ, 536, 571
Bovy J. et al., 2011, ApJ, 729, 141
Bingham B. et al., 2010, MNRAS, 406, 1885
Casali M. et al., 2007, A&A, 467, 777
Chiu K. et al., 2008, MNRAS, 385, L53
Deacon N. R., Hambly N. C., 2006, MNRAS, 371, 1722
Dunkley J. et al., 2009, ApJS, 180, 306
Dye S. et al., 2006, MNRAS, 372, 1227
Emerson J. P., Sutherland W. J., McPherson A. M., Craig S. C., Dalton G. B., Ward A. K., 2004, The Messenger, 117, 27
Fan X. et al., 2000, AJ, 120, 1167
Fan X. et al., 2001, AJ, 122, 2833
Fan X., Narayan V. K., Strauss M. A., Becker R. H., Pent Jerri, 2000, Monthly Notices of the Royal Astronomical Society
Goto T., 2006, MNRAS, 371, 769
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Hambly N. C. et al., 2008, MNRAS, 384, 637
Hazard C., Mackey M. B., Sheminus A. J., 1963, Nat, 197, 1037
Hewett P. C., Foltz C. B., Chaffee F. H., 1995, AJ, 109, 1498
Hewett P. C., Warren S. J., Leggett S. K., Hodgkin S. T., 2006, MNRAS, 367, 454
Ivezić Z. et al., 2004, in Storchi-Bergmann T., Ho L. C., Schmitt H. R., eds, Proc. IAU Symp. 222, The Interplay Among Black Holes, Stars and ISM. Cambridge Univ. Press, Cambridge, p. 525
Jaynes E. T., 2003, Probability Theory. Cambridge Univ. Press, Cambridge, UK
Jiang L. et al., 2008, AJ, 135, 1057
Jiang L. et al., 2009, AJ, 138, 305
Kaiser N. et al., 2002, in Tyson J. A., Wolff S., eds, Proc. SPIE Vol. 4836, Survey and Other Telescope Technologies and Discoveries. SPIE, Bellingham, p. 154
Lawrence A. et al., 2007, MNRAS, 379, 1599
Lupton R. H., Gunn J. E., Szalay A. S., 1999, AJ, 118, 1406
Maddox N., Hewett P. C., Warren S. J., Croom S. M., 2008, MNRAS, 386, 1605
Mortlock D. J., 2009, in Hobson M. P., Jaffe A. H., Liddle A., Mukherjee P., Parkinson D., eds, Bayesian Methods in Cosmology. Cambridge University Press, Cambridge, p. 193
Mortlock D. J., Peiris H. V., Ivezić Z., 2009a, MNRAS, 399, 699
Mortlock D. J. et al., 2009b, A&A, 505, 97
Mortlock D. J. et al., 2011, Nat, 474, 616
Pogson N., 1856, MNRAS, 17, 12
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2007, Numerical Recipes: The Art of Scientific Computing, 3rd edn. Cambridge Univ. Press, Cambridge
Richards G. T. et al., 2004, ApJS, 155, 257
Richards G. T. et al., 2009, AJ, 137, 3884
Schmidt-M., 1963, Nat, 197, 1040
Schmidt M., Schneider D. P., Gunn J. E., 1995, AJ, 110, 68
Schneider D. P., 1999, in Holt S., Smith E., eds, AIP Conf. Ser. Vol. 470, After the Dark Ages: When Galaxies were Young (the Universe at 2 < z < 5). Am. Inst. Phys., New York, p. 233
Scranton R., Connolly A. J., Szalay A. S., Lupton R. H., Johnston D., Budavari T., Brinkman J., Fukugita M., 2005, AJ, submitted
Sivia D. S., Skilling J., 2006, Data Analysis: A Bayesian Tutorial. Oxford University Press, Oxford, UK
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Stark D. P., Ellis R. S., Chiu K., Ouchi M., Bunker A., 2010, MNRAS, 408, 1628
Venemans B. P., McMahon R. G., Warren S. J., Gonzalez-Solares E. A., Hewett P. C., Mortlock D. J., Dye S., Sharp R. G., 2007, MNRAS, 376, 76
Warren S. J., Hewett P. C., Osmer P. S., 1994, ApJ, 421, 412
Warren S. J. et al., 2007, MNRAS, 375, 213
Willott C. J. et al., 2007, MNRAS, 375, 213
Willott C. J. et al., 2010a, AJ, 134, 2435
Willott C. J. et al., 2010a, AJ, 134, 2435
Willott C. J. et al., 2010b, AJ, 139, 906
Wyithe J. S. B., 2008, MNRAS, 387, 469
York D. G. et al., 2000, AJ, 120, 1579

APPENDIX A: THE LIKELIHOOD FOR A SINGLE PHOTOMETRIC MEASUREMENT

Some care must be taken in formulating the likelihood for photometric data. As argued in Section 2.1, there are many potential complications (e.g. the Poisson fluctuations in the number of source photons received; the influence of nearby sources; cosmic rays, bad pixels, etc. that cause the noise distribution to have non-Gaussian tails; inter-band noise correlations), but in most cases it is unnecessary to model all these effects. The focus here is on a standard optical or NIR observation of a faint source for which there is a
non-negligible chance that the background-subtracted counts are negative. The likelihood for such faint sources is very well approximated as a Gaussian in flux units (cf. equation 6).

In practice, however, optical and NIR photometric data are seldom reported in the form of the estimated flux and the associated error, and some of the limitations of the common alternative representations are described here. The most fundamental problem is the use of upper limits in the case of non-detections, which discards potentially vital information (Section A1). The use of magnitudes to report the measurements of faint sources can be anywhere from awkward to incorrect (as shown in Section A2), and so formulae for transforming magnitude data into flux units are given in Section A2.4.

A1 Non-detections and upper limits

A source is commonly considered to have been detected if, in a particular observation, it has a measured flux of $\tilde{F} \gtrsim S/N_{\text{lim}} \sigma$, where $\sigma$ is the effective background uncertainty in the image (expressed in flux units) and $S/N_{\text{lim}}$ is the threshold $S/N$ required for a source to be considered as ‘detected’. In ‘blind’ surveys (e.g. SDSS and UKIDSS), which are dominated by sources which have not previously been catalogued, it is common to choose $S/N_{\text{lim}} \approx 5$ to ensure that the vast majority of detections correspond to real astronomical objects.

When making follow-up observations of known objects or cross-matching between observations at different wavelengths, however, the aim is not the detection but characterization of a source which is (almost certainly) present with a known position. In such situations the ideal is to report both $\tilde{F}$ and $\sigma$ (albeit possibly transformed into magnitudes; see Section A2), as these two numbers are sufficient to reconstruct the likelihood of the data, assuming that the noise is sky-dominated. This is the approach used by the SDSS, where flux estimates are provided in all five bands for every source that is detected in (at least) one of the SDSS observations. However this is not done in UKIDSS, resulting in both a loss of information and requiring more complicated book-keeping to distinguish between non-detections and non-observations (as the measurements in the different UKIDSS bands are not simultaneous). Large surveys like SDSS and UKIDSS make the original images available, so aperture photometry of undetected sources can be undertaken (as described in Section 3.1), but if the follow-up measurements were made in the context of particular science projects then it is still common for upper limits to be reported (e.g. a detection limit of $S/N_{\text{lim}} = 2$ was used in the CFHQS; Willott et al. 2010b). If the reported measurements were insufficient to reconstruct the likelihood then information has been discarded; the specific issue addressed here is the significance of the information lost that results from reporting just $F_{\text{lim}}$ instead of reporting both $\tilde{F}$ and $\sigma$.

In statistical terms, the difference manifests itself in the likelihood. Given a flux measurement, the likelihood is, as in equation (6),

$$\Pr(\tilde{F} | F) = \frac{1}{(2\pi)^{1/2} \sigma} \exp \left[ -\frac{1}{2} \left( \frac{\tilde{F} - F}{\sigma} \right)^2 \right],$$

(A1)

where $\tilde{F}$ is the true flux of the source. With access to only an upper limit, the likelihood is the probability that the background-subtracted counts are below the detection threshold, which is given by

$$\Pr(\tilde{F} < F_{\text{lim}} | F) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{F_{\text{lim}} - F}{2^{1/2} \sigma} \right) \right],$$

(A2)

where $\text{erf}(x) = 2 \pi^{-1/2} \int_0^x e^{-t^2} dt$ is the error function.

What would be the result of adopting an upper limit in place of a low $S/N$ flux measurement? As the first is a probability density and the second a (cumulative) probability, it is not particularly meaningful to compare the likelihoods directly. Instead, it is more appropriate to explore whether the resultant inferences differ significantly, which can be done by performing a simplified version of the model selection problem discussed in Section 2. Assuming a surface density $\Sigma$, of stars, all of which have a flux of $F_s$ in the band in question, and a surface density $\Sigma_q$ of quasars, all of which have a flux of $F_q$ in this band, the probability that a source is a quasar, $P_q$, is given by inserting either equation (A1) or equation (A2) into equation (8). If flux measurements are used then,

$$P_q = \frac{\Sigma_q \exp \left[ -\frac{1}{2} \left( \frac{\tilde{F} - F_q}{\sigma} \right)^2 \right]}{\Sigma_q \exp \left[ -\frac{1}{2} \left( \frac{\tilde{F} - F_q}{\sigma} \right)^2 \right] + \Sigma_s \exp \left[ -\frac{1}{2} \left( \frac{\tilde{F} - F_s}{\sigma} \right)^2 \right]},$$

(A3)

if only upper limits are available then

$$P_q = \frac{\Sigma_q \left[ 1 + \text{erf} \left( \frac{F_{\text{lim}} - F_q}{2^{1/2} \sigma} \right) \right]}{\Sigma_q \left[ 1 + \text{erf} \left( \frac{F_{\text{lim}} - F_q}{2^{1/2} \sigma} \right) \right] + \Sigma_s \left[ 1 + \text{erf} \left( \frac{F_{\text{lim}} - F_s}{2^{1/2} \sigma} \right) \right]},$$

(A4)

These two expressions for $P_q$ are compared in Fig. A1 for two combinations $\Sigma_q/\Sigma_s$ and detection limit.

The top panel of Fig. A1 shows the results for $\Sigma_s = 10 \Sigma_q$ and $S/N_{\text{lim}} = 5$, corresponding to an extreme case in which the standard ‘blind’ detection limits was applied to these follow-up measurements. The quasar probability inferred from the upper limit (which is independent of $\tilde{F}$, provided it is lower than the detection limit) is similar to the prior probability given by the relative numbers of stars and quasars. Using the formal flux estimate, however, the quasar hypothesis is decisively rejected if $\tilde{F} \gtrsim 3 \sigma$. Moreover, for the priors and model fluxes chosen here, the majority of measurements would be in this regime, because most of the undetected sources would be stars for which the measured flux was only just below the detection threshold. This predicted distribution of measured fluxes is also shown in Fig. A1.

The bottom panel of Fig. A1 shows the results for $\Sigma_s = 100 \Sigma_q$ and $S/N_{\text{lim}} = 2$, approximating the situation for the CFHQSs (Willott et al. 2010b). Using only the $2\sigma$ upper limit gives $P_q \approx 0.3$, much higher than the prior value of $P_q \approx 0.009$. However, there are so many more stars than quasars in this case that almost all undetected sources are stars with $\sigma \lesssim \tilde{F} \leq 2\sigma$. Despite the low $S/N$, in most cases $P_q \lesssim 0.1$ is obtained if the flux measurement is used, and such sources could reasonably be discarded as quasar candidates.

In these simplified examples the use of an upper limit in place of a formal flux estimate results in unnecessary extra uncertainty in the classification of most ‘undetected’ sources. The equivalent calculation for the UKIDSS–SDSS sample is, of course, more complicated, but the basic result still holds: the majority of candidates which are not detected in the $i$ and $z$ bands have $\tilde{F} \gtrsim 3 \sigma$, sufficient to reject them as possible HZQs with great confidence. A broader implication of the above arguments is that measurements and uncertainties should always be reported in preference to upper limits.

A2 The use of magnitudes to represent photometric data

Photometric data are usually reported in terms of either logarithmic or asinh magnitudes. Given an estimated magnitude and its associated error it is then natural to adopt a Gaussian likelihood based on these values (or equivalently, to construct a least-squares estimate in terms of the appropriately weighted magnitude differences).

© 2011 The Authors, MNRAS 419, 390–410
Monthly Notices of the Royal Astronomical Society © 2011 RAS
where $F_0$ is the zero-point flux for which $m = 0$ by construction. This formula can be inverted to give

$$F = F_0 \exp \left( -\frac{2 \ln(10)}{5} m \right).$$

(A6)

Differentiating equation (A5) gives the Jacobian used to convert probability densities as

$$\frac{dm}{dF} = \frac{5}{2 \ln(10) F}.$$  

(A7)

A2.2 Asinh magnitudes

The asinh magnitude scheme was introduced by Lupton et al. (1999) to overcome the inability of the logarithmic magnitudes to represent negative flux estimates, while retaining the familiar behaviour for high fluxes. Asinh magnitudes are defined by

$$m = m_0 - \frac{5}{2 \ln(10)} \text{asinh} \left( \frac{F}{2 \times 10^{-2m_0/5} F_0} \right),$$

(A8)

where $m_0$ is, in the limit of high $m_0$, the zero-point asinh magnitude corresponding to zero flux. As $\lim_{x \to -\infty} \text{asinh}(x/2) = \ln(x)$, equation (A8) approaches equation (A5) for $F \gg 10^{-2m_0/5} F_0$; but for $|F| \lesssim 10^{-2m_0/5} F_0$ the asinh magnitude is proportional to flux. Inverting equation (A8) gives the flux as

$$F = 2 \times 10^{-2m_0/5} F_0 \sinh \left( \frac{2 \ln(10)}{5} (m_0 - m) \right).$$

(A9)

Differentiating equation (A8) then gives the Jacobian needed to transform variables as

$$\frac{dm}{dF} = \frac{5}{2 \ln(10) F} \frac{1}{F^2 + 4(10^{-2m_0/5} F_0)^2}^{1/2}.$$  

(A10)

A2.3 Gaussian likelihoods in magnitude units

Given an estimated magnitude, $\hat{m}$, and an error $\Delta \hat{m}$, it is common to assume that the likelihood is Gaussian in magnitude units, and hence given by

$$\Pr(\hat{m} | m) = \frac{1}{(2\pi)^{1/2} \Delta \hat{m}} \exp \left[ -\frac{1}{2} \left( \frac{\hat{m} - m}{\Delta \hat{m}} \right)^2 \right],$$

(A11)

where $m$ is the true magnitude of the source in this band. Changing variables to flux units, the implied likelihood would then be

$$\Pr(\hat{F} | F) = \frac{1}{(2\pi)^{1/2} \Delta \hat{m}} \exp \left[ -\frac{1}{2} \left( \frac{m(\hat{F}) - m(F)}{\Delta \hat{m}} \right)^2 \right] \left| \frac{dm}{dF} \right|,$$

(A12)

where $m(F)$ and $|dm/dF|$ are given in equations (A5) and (A7) for logarithmic magnitudes and equations (A8) and (A10) for asinh magnitudes, respectively.

The resultant expressions for the likelihood are straightforward, but cumbersome; they are more easily explored visually, as is done in Fig. A2. For the brighter source with a true flux of $F = 10 \sigma$ (where $\sigma$ is the background noise in flux units) the Gaussian likelihoods in terms of both magnitudes are consistent with the true Gaussian in flux units, and of course almost any subsequent inferences would be similarly accurate. For the fainter sources with true fluxes of $F = 0.1 \sigma$ and $F = 2 \sigma$, however, the differences between

Gaussians in magnitude and flux units are obviously not equivalent, but are the differences sufficient to result in significantly changed inferences?

The starting point to answering this question is the basic formulæ that relate magnitude to fluxes (Section A2.1 and A2.2). Using these definitions it is then possible to transform a Gaussian distribution from flux units to magnitudes units to see how the resultant inferences differ (Section A2.3).

A2.1 Logarithmic magnitudes

The traditional logarithmic magnitude corresponding to (positive) flux $F$ is given by Pogson (1856) as

$$m = -\frac{5}{2 \ln(10)} \ln \left( \frac{F}{F_0} \right),$$

(A5)
Given the importance of performing statistical calculations in flux units, formulae are needed to convert an estimated magnitude, \( \hat{m} \), and its associated error, \( \Delta \hat{m} \), into an estimated flux, \( \hat{F} \), and, in the case of faint sources, the background noise, \( \sigma \). The correct conversions can be derived from the fact that \( \hat{m} \) and \( \Delta \hat{m} \) are inevitably calculated from \( \hat{F} \) and \( \sigma \) in the first place. The problem would be more difficult if reported magnitudes and uncertainties were obtained using more complicated statistical arguments, but for the simple conversions commonly adopted all that is required is to invert the relationships in Sections A2.1 and A2.2.

Given \( \hat{m} \) and \( \Delta \hat{m} \) in conventional logarithmic magnitudes, a straightforward substitution into equation (A6) yields

\[
\hat{F} = F_0 \exp \left[ -\frac{2 \ln(10)}{5} \hat{m} \right].
\]

Then using equation (A7) gives

\[
\sigma = F_0 \frac{2 \ln(10)}{5} \exp \left[ -\frac{2 \ln(10)}{5} \hat{m} \right] \Delta \hat{m}
\]

\[
= \hat{F} \frac{2 \ln(10)}{5} \Delta \hat{m} \simeq \Delta \hat{m} \hat{F}.
\]

Given \( \hat{m} \) and \( \Delta \hat{m} \) in terms of asinh magnitudes, the corresponding flux estimate is given from equation (A9) as

\[
\hat{F} = 2 \times 10^{-2m_0/5} F_0 \sinh \left[ \frac{2 \ln(10)}{5} (m_0 - \hat{m}) \right].
\]

Then using equation (A10) to change variables gives the background error in flux units as

\[
\sigma = \frac{4 \ln(10)}{5} 10^{-2m_0/5} F_0 \cosh \left[ \frac{2 \ln(10)}{5} (m_0 - \hat{m}) \right].
\]

For the calculation of \( P_q \) in Section 3, equations (A13) and (A14) were used to convert the reported UKIDSS Y- and J-band photometry to flux units and equations (A15) and (A16) were used to convert the reported SDSS i- and z-band SDSS photometry to flux units.

**APPENDIX B: PARAMETER FITTING FOR THE STELLAR POPULATION MODEL**

In Section 3.2, it was necessary to fit the parameters of an empirical model of the intrinsic stellar colour and magnitude distribution to a likelihood-based calculations, the formulae for which are given below.

**A2.4 Magnitude to flux conversions**

Figure A3. The probability that a source is a quasar, \( P_q \), shown as a function of the observed \( i - Y \) colour of the source. The thin lines show the (arbitrarily normalized) likelihoods of the star and the quasar, which are assumed to be Gaussian in magnitude units; the thick vertical lines give the relative normalization of the two populations. The nonsensical rise of \( P_q \) to unity for \( i - Y \leq 0 \) is purely an artefact of the seemingly innocent assumption that the likelihood is Gaussian in magnitude units.

The probability that a source is a quasar, \( P_q \), is in Section 3, equations (A13) and (A14) were used to convert the reported UKIDSS Y- and J-band photometry to flux units and equations (A15) and (A16) were used to convert the reported SDSS i- and z-band SDSS photometry to flux units.
sample of UKIDSS–SDSS point sources extracted from the WSA. Many methods exist for tackling this problem, although the fact that the observed distribution is the result of convolving the underlying target distribution with magnitude-dependent noise makes this task non-trivial in this case. The overall theme of this paper would suggest taking a Bayesian approach, but the primary aim here is to find any function to describe the intrinsic stellar distribution that is consistent with the observed data; the actual parameter values (and their uncertainties) are not of interest. Hence a faster, if less principled, method was used.

After selecting $\sim 10^5$ bright, red UKIDSS–SDSS stars (with $I - Y \geq 2.0$ and $15.0 \leq Y \leq 19.5$), the sample was binned in $Y$ and $I - Y$, with a bin size of 0.1. The data were hence reduced to the number of objects, $n_i$, in each of $N_{\text{bin}}$ cells (where the index $i$ covers the two-dimensional parameter space). For a given choice of the free parameters, $\phi$, describing the model under consideration (as distinct from the parameters $\theta$, used to characterize a single star), the expected number of stars in each cell, $\bar{n}_i(\phi)$, was calculated by convolving the intrinsic population with the appropriate photometric error distribution. It is also important to ensure that the value of $\bar{n}_i(\phi)$ is not spuriously high due to the large number of faint undetected sources that could, at least in theory, be scattered into the bin. The potential problem with a simple treatment is that such numerous faint sources are typically beyond the confusion limit of the survey and so to treat a single noise spike as being associated with each in turn results in an unrealistically high probability of a spurious source entering the survey. The key is to adopt a self-consistent treatment in which sources at or below the confusion limit are ignored (cf. Mortlock 2009).

Irrespective of the method by which $\bar{n}_i(\phi)$ is calculated, $n_i$ is Poisson-distributed and uncorrelated between bins. The log-likelihood of the full data set is hence

$$
\ln[\Pr(n|\phi)] = \ln \left\{ \prod_{i=1}^{N_{\text{bin}}} \frac{[\bar{n}_i(\phi)]^{n_i} \exp[-\bar{n}_i(\phi)]}{n_i!} \right\} = \sum_{i=1}^{N_{\text{bin}}} n_i \ln[\bar{n}_i(\phi)] - \bar{n}_i(\phi) + \text{constant.} \quad (B1)
$$

The best-fitting parameters given in equation (9) were found by minimizing equation (B1) using the downhill simplex method (Press et al. 2007).

This paper has been typeset from a \TeX/\LaTeX\ file prepared by the author.