Testing the completeness of the SDSS colour selection for ultramassive, slowly spinning black holes

Caroline Bertemes,[1]★ Benny Trakhtenbrot,[1]★† Kevin Schawinski,[1] Chris Done[2] and Martin Elvis[3]

[1]*Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*
[2]*Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[3]*Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA*

ABSTRACT

We investigate the sensitivity of the colour-based quasar selection algorithm of the Sloan Digital Sky Survey (SDSS) to several key physical parameters of supermassive black holes (SMBHs), focusing on BH spin (a_*) at the high BH-mass regime ($M_{BH} \geq 10^9 \, M_{\odot}$). We use a large grid of model spectral energy distribution (SED), assuming geometrically thin, optically thick accretion discs, and spanning a wide range of five physical parameters: BH mass M_{BH}, BH spin a_*, Eddington ratio L/L_{Edd}, redshift z, and inclination angle inc. Based on the expected fluxes in the SDSS imaging *ugriz* bands, we find that ∼99.8 per cent of our models with $M_{BH} \leq 10^{9.5} \, M_{\odot}$ are selected as quasar candidates and thus would have been targeted for spectroscopic follow-up. However, in the extremely high-mass regime, $\geq 10^{10} \, M_{\odot}$, we identify a bias against slowly/retrograde spinning SMBHs. The fraction of SEDs that would have been selected as quasar candidates drops below ∼50 per cent for $a_* < 0$ across $0.5 < z < 2$. For particularly massive BHs, with $M_{BH} \simeq 3 \times 10^{10} \, M_{\odot}$, this rate drops below ∼20 per cent, and can be yet lower for specific redshifts. We further find that the chances of identifying any hypothetical sources with $M_{BH} = 10^{11} \, M_{\odot}$ by colour selection would be extremely low at the level of ∼3 per cent. Our findings, along with several recent theoretical arguments and empirical findings, demonstrate that the current understanding of the SMBH population at the high M_{BH}, and particularly the low- or retrograde-spinning regime, is highly incomplete.

Key words: black hole physics – quasars: general.

1 INTRODUCTION

The spins of astrophysical, supermassive black holes (SMBHs), along with their masses (M_{BH}) and the growth rates ($\dot{M}_{BH}$), are key parameters in understanding the physics of their close environments and their growth history. The black hole (BH) spin[1] affects the space–time metric in the close vicinity, by setting the distance to the innermost stable (circular) orbit (ISCO), which can vary in the range ∼6, 1, and 9 gravitational radii, for a non-rotating, maximally prograde rotating, and maximally retrograde rotating BH, respectively ($a_* = 0$, 0.998, and −1, respectively). This, in turn, may affect the total output and spectral dependence of the radiation emitted from

any gas accretion flow, since higher energy photons may be emitted from smaller radii. Indeed, within the standard geometrically thin, optically thick accretion disc model, the radiative efficiency η covers almost an order of magnitude, ranging $\eta \simeq 0.038$–0.32 (see e.g. Netzer 2013). This efficiency then determines the efficiency by which the BH accumulates mass, for any given gas accretion rate through the disc. Moreover, the spin of the BH may evolve due to accretion and coalescence events. A large number of isotropically oriented accretion events would lead to a 'spin-down' of the BH, reaching $a_* \lesssim 0.2$, while prolonged accretion and/or drastically anisotropic accretion events would 'spin-up' the BH, reaching $a_* \simeq 1$ (e.g. King & Pringle 2006; King, Pringle & Hofmann 2008; Dotti et al. 2013). The predictions of the two scenarios differ mainly in the extremely high-mass regime, $M_{BH} > 10^9 \, M_{\odot}$, as such massive SMBHs have experienced more accretion episodes than their lower mass counterparts. Understanding the demographics of BH spin among both relic and accreting SMBHs [i.e. active galactic nuclei (AGN)], and particularly the most massive ones could therefore provide key insights to SMBH growth history, and potentially the galaxy-scale processes that drive it (e.g. Volonteri et al. 2013).

★ E-mail: cberteme@phys.ethz.ch (CB); benny.trakhtenbrot@phys.ethz.ch (BT)
† Zwicky fellow.

[1]Throughout this work, we describe the spin via the dimensionless Kerr parameter $a_* = J/M_{BH}^2$ (setting $G = c = 1$), where J denotes the angular momentum of the BH and M_{BH} its mass. a_* can then vary between −1 and 0.998 (see Thorne 1974).
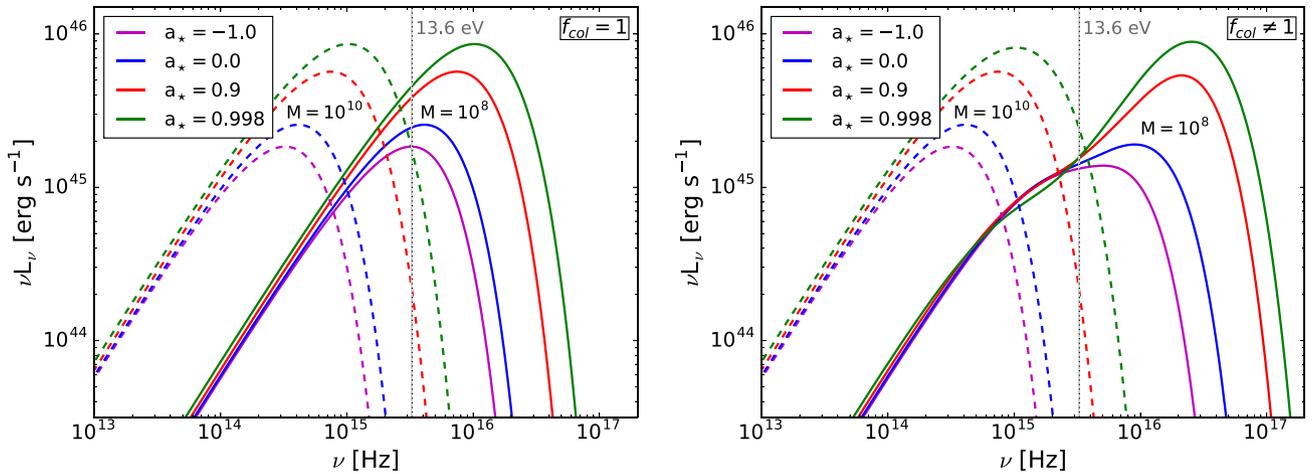
Figure 1. Example model rest-frame SEDs for different BH mass and spin values. The solid and dotted lines show models with $M_{BH} = 10^8$ and $10^{10}\,\mathrm{M_\odot}$, respectively. All models assume $\dot{M} = 1\,\mathrm{M_\odot\,yr^{-1}}$ and $\cos\,(\mathrm{inc}) = 0.8$. Left-hand panel: simplified model SEDs that ignore temperature colour corrections (i.e. setting $f_{col} = 1$). Right-hand panel: model SEDs including the self-consistent colour–temperature corrections provided by the Done et al. (2013) model.

Measuring, or even constraining spins in SMBHs is, however, challenging. One common approach is based on the reflection-dominated, gravitationally broadened component of the Kα iron emission line observed near 6.7 keV, which probes the disc region close to the ISCO. So far, this approach was applied to about 20 low-redshift AGN, providing predominantly high a_* measurements, with several objects having $a_* \geq 0.99$ (Reynolds 2014). However, these measurements may be affected by systematic uncertainties in modelling the underlying continuum and the physical interpretation of the Kα line profile (e.g. Miller & Turner 2013, and references therein). Moreover, the limited sensitivity of present-day facilities limits the usage of this method to low redshifts, with only a few exceptions of very bright and/or lensed AGN at $z \sim 1$ (Reis et al. 2014; Reynolds et al. 2014). An alternative approach, which essentially relies on fitting the (rest-frame) UV–optical spectral energy distributions (SEDs) of AGN for which M_{BH} is known, has been recently employed for several larger samples of higher redshift systems, reaching $z \sim 3.5$ (e.g. Davis & Laor 2011; Wu, Lu & Zhang 2013; Netzer & Trakhtenbrot 2014; Trakhtenbrot 2014; Capellupo et al. 2015, 2016). These studies have generally found extremely high BH spins for the most massive BHs, but typically lower spins for the lower mass objects. One notable exception, of an extremely massive BH with $a_* = 0.3$, was reported by Czerny et al. (2011). Indirect arguments involving the ensemble properties of AGN across all redshifts (e.g. Elvis, Risaliti & Zamorani 2002; Yu & Tremaine 2002, following the so-called 'Soltan argument', Soltan 1982) support an efficiency of $\eta \sim 10$–15 per cent, corresponding to $a_* \sim 0.8$–0.9. However, such calculations are very sensitive to a series of simplifying assumptions. For example, the recent revision of the relation between BH mass and host galaxy mass in the local Universe (Kormendy & Ho 2013) leads to lower efficiency and hence spin, giving either constant efficiency of $\eta \simeq 0.055$ (consistent with non-spinning BHs) or with η which increases with M_{BH}, so that the most massive BHs have extremely high spins (e.g. Ueda et al. 2014). Thus, most of the available data support high BH spins amongst the most massive BHs, favouring the 'spin-up' scenario. The iron Kα studies suggest extreme spins for the lower mass BHs as well.

With both of the aforementioned methods favouring high spin values, it is crucial to investigate the role of any selection effects and/or biases that could be affecting these analyses, particularly in

the high-mass regime ($M_{BH} \gtrsim 10^9\,\mathrm{M_\odot}$). It has already been argued that AGN selection which is based on high-ionization emission lines would be biased against BH with high M_{BH} and low (or retrograde) spins, since the lack of UV photons in their SEDs would lead to a low fraction of ionizing photons (see 13.6 eV line in Fig. 1), and thus to the observation of weak-lined quasars (WLQs; see e.g. Laor & Davis 2011). Another concern is that the prescriptions used in the continuum-fitting approach would lead to bias towards high η, and therefore high BH spins, when applied to large (optically) flux-limited surveys (such as the Sloan Digital Sky Survey, SDSS; Raimundo et al. 2012). In the present work, we will show that on top of this, the colour selection in such surveys may also be favouring high spins among high-M_{BH} systems.

The outline of the paper is as follows. In Section 2, we describe our models and the related assumptions and briefly outline the SDSS colour-based target selection algorithm. In Section 3, we present and discuss our main results regarding the ability of the SDSS target selection algorithm to identify the calculated thin accretion disc SEDs. We summarize our main findings in Section 4. Throughout this work we convert luminosities to fluxes and to magnitudes assuming a cosmological model with $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and the AB magnitude system.

2 CALCULATIONS

Our analysis focuses on testing the completeness of the colour–colour quasar selection algorithm of the SDSS (York et al. 2000; Richards et al. 2002) for SMBHs properties, and particularly BH spin (a_*). This is done by using a large grid of thin accretion disc (AD) models, calculated for different physical properties of the accreting SMBH.

2.1 Model spectra

Our calculated model SEDs are based on the assumption of a geometrically thin, optically thick AD, emitting as a sum of blackbodies, originally developed by Shakura & Sunyaev (1973), and have since been elaborated to include colour–temperature corrections and relativistic effects, among other improvements. Several recent studies demonstrate that thin AD models provide a good fit to the

Table 1. Parameters used in the grid of model SEDs.

Parameter	Min. value	Max. value	Step size
BH mass, $\log{(M_{BH}/M_\odot)}$	6	11	0.5
BH spin, a_*	-1	0.998	0.1
Accretion rate, L/L_{Edd}	0.05	1	0.05
Redshift, z	0.5	2	0.1
Inclination, inc	$10°$	$50°$	$10°$

Figure 2. Difference in the SED induced by changing a_* from -1 to 0.998 and z from 1 to 2 for a model with $M_{BH} = 10^{10}$ M$_\odot$, inc = 10, and L/L_{Edd} = 0.1 (so $\dot{M} = f(L/L_{Edd}, a_*, M_{BH})$ is not constant and the SEDs do not overlap in their power law, unlike in Fig. 1). The SEDs are plotted in our observed frame. At this mass, *ugriz* probes the turnover between power law and exponential cut-off more efficiently at higher redshift, and with that the effect of the spin.

UV–optical continuum SEDs of unobscured AGN, particularly in cases where the SMBH is accreting at a fairly high fraction of the Eddington luminosity, and reliable estimates of M_{BH} are available (e.g. Jin et al. 2012; Capellupo et al. 2015; Collinson et al. 2015, and references therein).

We note that several quasar microlensing studies (e.g. Pooley et al. 2007; Dai et al. 2010; Blackburne et al. 2014), as well as some evidence from reverberation mapping (e.g. Fausnaugh et al. 2016), have challenged the expectations of the thin AD model, with the observations suggesting significantly larger disc sizes. The present work, however, focuses on the UV–optical part of the continuum, here assumed to originate solely from the innermost disc regions, which most probably do extend towards the ISCO.

The model SEDs we use here were produced using the OPTXCONV model presented in Done et al. (2013), employed within the XSPEC package (version 12.9.01; Arnaud 1996). Here we only mention the specific parameters we used for the purposes of the present work. For a detailed discussion about the model, we refer the reader to Done et al. (2013). Our main grid of model SEDs is calculated using colour–temperature corrections, noted as f_{col}, and affecting the intrinsic (blackbody) flux originating from each annulus in the disc following $f_\nu(T) = f_{col}^{-4} B_\nu(f_{col} T)$. To verify that our results do not strongly depend on these corrections, we also consider an alternative grid of purely thin, Shakura–Sunyaev-like disc SEDs. In terms of the OPTXCONV code parameters, the latter is achieved by setting $f_{col} = 1$. Additionally, we do not consider coronal emission (setting $r_{cor} = -1$), and we assume an arbitrarily large outer AD radius, $\log{(r_{out}/r_g)} = 7$. Both should not affect the UV–optical part of the calculated SEDs.

We note that, unlike the public versions of OPTXAGN and OPTX-AGNF, the models we use here incorporate relativistic effects and an inclination dependence (see Done et al. 2013).

We have created a grid of calculated accretion disc SEDs, spanning a wide range in several input physical parameters: BH mass M_{BH}, the BH spin a_*, the Eddington ratio L/L_{Edd}, the redshift z, and the inclination angle between the polar axis of the AD and the observer's line-of-sight inc. The ranges and step sizes for all these parameters are given in Table 1. In total, the grid includes 369 600 unique model SEDs. Our choices for the range in M_{BH} and L/L_{Edd} are motivated by the results of several studies of large sample of quasars at $z \lesssim 2$ (e.g. McLure & Dunlop 2004; Fine et al. 2008; Gavignaud et al. 2008; Trump et al. 2009; Schulze & Wisotzki 2010; Shen et al. 2011; Trakhtenbrot & Netzer 2012; Schulze et al. 2015). We have verified that our results do not depend strongly on our choice to use models with $L/L_{Edd} > 0.3$, for which the original Shakura & Sunyaev (1973) model is not expected to hold. Appendix A presents some of our main findings for a limited subset of model SEDs with $L/L_{Edd} < 0.3$. The range in inclination angles is motivated by the requirement that the calculated SEDs would eventually be observed as type-I AGN. We did not consider redshifts exceeding $z = 2$, since the intergalactic medium (IGM)-related absorption is expected to dominate the rest-frame UV part
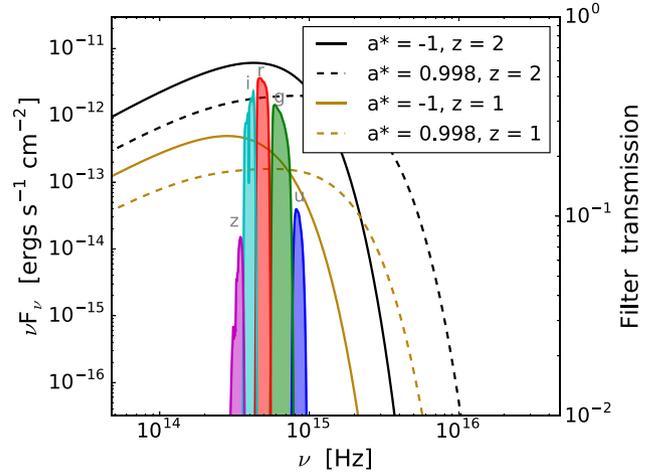
of the SEDs, observed within the SDSS imaging. Fig. 1 presents the UV–optical part of several example model SEDs, with the left-hand panel showing simplified thin disc SEDs ignoring colour–temperature corrections, and the right-hand panel showing SEDs that include self-consistent corrections. As can be seen, the corrections mainly affect SEDs of BHs with low masses and high spins. The simplified models are in excellent agreement with those presented in many other studies (see e.g. fig. 1 in Davis & Laor 2011). As can be seen, the difference in UV emission between a maximally prograde spinning BH and a stationary BH is much larger than the difference between the maximally retrograde spinning case and the stationary case, since the change in ISCO is larger (see Section 1).

Finally, we calculated synthetic magnitudes for each of the SEDs in the grid, in the five SDSS photometric bands (*ugriz*), using standard procedures. Fig. 2 displays the *ugriz* filter curves alongside some of our SEDs. We stress that, since each SED is calculated at a given redshift, the calculated synthetic magnitudes realistically reflect the *continuum* flux level of the respective model quasar (i.e. ignoring contribution from emission lines), with no need for any arbitrary scaling.

Our fiducial model SEDs do not take into account several additional effects that may alter the emergent SEDs and are expected in real AGN SEDs. These include the hard X-ray emission and soft X-ray excess; outflows from the inner parts of the AD; dust extinction in the host galaxies; and additional emission from AGN-related lines and/or from the stars of the host galaxies. These effects should be considered in light of our focus on the high (AGN) luminosity, high-M_{BH} regime, and particularly the sensitivity of the SDSS selection algorithm to BHs with low or retrograde spins. We discuss each of these below.

The fraction of power dissipated in the hard X-ray corona and soft X-ray excess components should affect the inner disc spectrum if they are ultimately powered by the mass accretion flow. The energy required to power these X-ray emission components is then not available to power the standard disc emission, therefore making the disc continuum cooler and less luminous than predicted by a pure disc model (e.g. Done et al. 2012). Our assumption of a pure disc

SED is therefore a conservative choice, as it reflects no additional losses of UV radiation, and provides the best possible case for SDSS selection. Moreover, these additional X-ray components are expected to be weakest in the high-luminosity, high-M_{BH} regime which is our main focus in this work.

Dust extinction in luminous quasars is generally found to be relatively small (see e.g. Baron et al. 2016; Capellupo et al. 2016, and references therein, but also the exceptional populations studied in e.g. Glikman et al. 2007; Banerji et al. 2012). More importantly, dust extinction will reduce the UV emission from any calculated AD SED, in a way which broadly mimics the effects of reducing the BH spin parameter (see e.g. the discussion in Capellupo et al. 2015; Collinson et al. 2015). Outflows of gas from the inner parts of the AD would have a similar effect (e.g. Slone & Netzer 2012). Thus, the inclusion of dust extinction and/or AD outflows may further limit the prospects of the SED being selected by the SDSS colour-based algorithm. Our choice to ignore these effects is, therefore, conservative in the context of the present work.

Both these may significantly reduce the radiation in the UV part of the SED, in a way which may broadly mimic the effects of reducing the BH spin parameter (see e.g. the discussion in Slone & Netzer 2012; Capellupo et al. 2015; Collinson et al. 2015). These may further complicate the emergent UV SED, but require a set of additional assumptions and parameter choices (e.g. outflow profile, extinction curve), which are beyond the scope of the present work.

Emission lines and features would affect the colour selection only if they are strong (i.e. in terms of equivalent width), and contribute differentially to two SDSS bands (or more). We have verified that our analysis is not significantly affected by the inclusion of emission lines, following the typical emission-line spectrum seen in SDSS quasars. We describe this test and the associated results in Appendix B. We also note that in the high-M_{BH}, low- (or retrograde-) spin regime, the UV emission lines are actually expected to be *weaker* than usual, due to the low number of ionizing photons, thus further limiting their possible effect on colour-based selection.

Finally, for high-luminosity SDSS quasars at $z \gtrsim 1$ the host galaxy contribution is expected to be low, of order $\lesssim 25$ per cent at rest frame 3000 Å (Schulze et al. 2015), and significantly lower at yet shorter wavelengths (e.g. Merloni et al. 2010; Collinson et al. 2015) – the spectral regime where the effects of a_* are expected to be most dramatic.

2.2 Overview of the SDSS quasar selection criteria

The original SDSS project covered about $10\,000$ deg^2 in its main imaging survey, as well as a smaller area of ~750 deg^2, in five different bands (Doi et al. 2010), *ugriz*. Quasars are selected for (multiplexed) follow-up spectroscopy through an elaborate algorithm, discussed in detail in Richards et al. (2002). The algorithm is designed to provide both high spectroscopic completeness *and* purity for quasars, building on the rich experience gained from previous optical spectroscopic quasar surveys, and incorporating the commissioning data of the SDSS itself. However, to the best of our knowledge, the algorithm was never tested against physically motivated models of quasar continuum emission.

The quasar selection algorithm is primarily based on selection in a four-dimensional colour–colour space, defined by $u - g$, $g - r$, $r - i$, and $i - z$ colours. The four-dimensional colour space is further divided into two different three-dimensional colour–colour selection subspaces: the *ugri* bands are used for identifying low-redshift quasar candidates, the *griz* bands are probing at longer wavelengths and thus more suitable for higher redshifts. A sepa-

rate selection procedure focuses on radio-emitting sources, based on a cross-match to the large-area Faint Images of the Radio Sky at Twenty-cm (FIRST) survey (Becker, White & Helfand 1995), but provides only a small minority of all quasar candidates which are not also otherwise colour selected (336 out of 8630; e.g. Schneider et al. 2010). Some (four-dimensional) regions in the colour–colour space are especially prone to contamination by white dwarfs, A stars, and M star – white dwarf binaries. The latter are defined as exclusion regions and any objects within them are immediately rejected (see red regions in Fig. 3). On the other hand, any objects lying within the defined inclusion regions (green regions in Fig. 3) are directly selected, except if they were already rejected by the exclusion boxes, which have a higher priority. We emphasize that there exists one unique inclusion box: the UV-excess ('UVX') selection region is indeed a two-dimensional region and not the projection of a four-dimensional region. The region of the colour space that is generally occupied by stars (and normal galaxies) – the so-called called 'stellar locus' (excluded from Fig. 3 for clarity) – defines another significant (multidimensional) exclusion region. The stellar locus was modelled explicitly by Richards et al. (2002) and kept constant during the survey.

Any objects that lie sufficiently far outside the stellar locus, in either the *ugri* or the *griz* projections, and do *not* fall within any of the aforementioned exclusion regions, are also selected as targets, even if are not located in any specific inclusion region.

The selected objects will undergo several quality tests and, finally, if they are within the pre-defined optical flux limits, they will be marked as possible quasar candidates. In order to be included within the SDSS flux limits for follow-up spectroscopy, the object must have an *i*-magnitude between 15 and 19.1, for the low-redshift (*ugri*) selection criteria, or between 15 and 20.2 for the high-redshift (*griz*) ones. We note that the latter criteria may, occasionally, select low-redshift sources as well.

3 RESULTS AND DISCUSSION

We now turn to analyse the location of the model SEDs in the SDSS colour–colour space, with the goal of quantifying their prospects for selection for spectroscopy by the colour-based target selection algorithm. We particularly focus on the role of a_* in the high-M_{BH} regime.

We note that testing the relevance of thin AD SEDs for *real* AGN (observed in SDSS and other surveys), for which some physical parameters are determined, is beyond the scope of the present work. This subject has been addressed in several recent works, with somewhat ambiguous results. For example, the recent studies by Capellupo et al. (2015, 2016) and Collinson et al. (2015) showed good agreement between the UV–optical spectra of luminous AGN, which have reliable estimates of M_{BH}, and the corresponding thin AD SEDs. The study of Jin et al. (2012) presented an extension of this approach to account for the X-ray emission. On the other hand, Davis, Woo & Blaes (2007) found some discrepancies between the observed UV SED shapes and what is expected from thin AD models.

3.1 Model SEDs in the colour–colour space

In this section, we discuss the position of our model SEDs in colour–colour space and with respect to the relevant inclusion and exclusion regions. We focus on specific examples where changing the BH spin parameter has a significant effect on the outcome of the selection
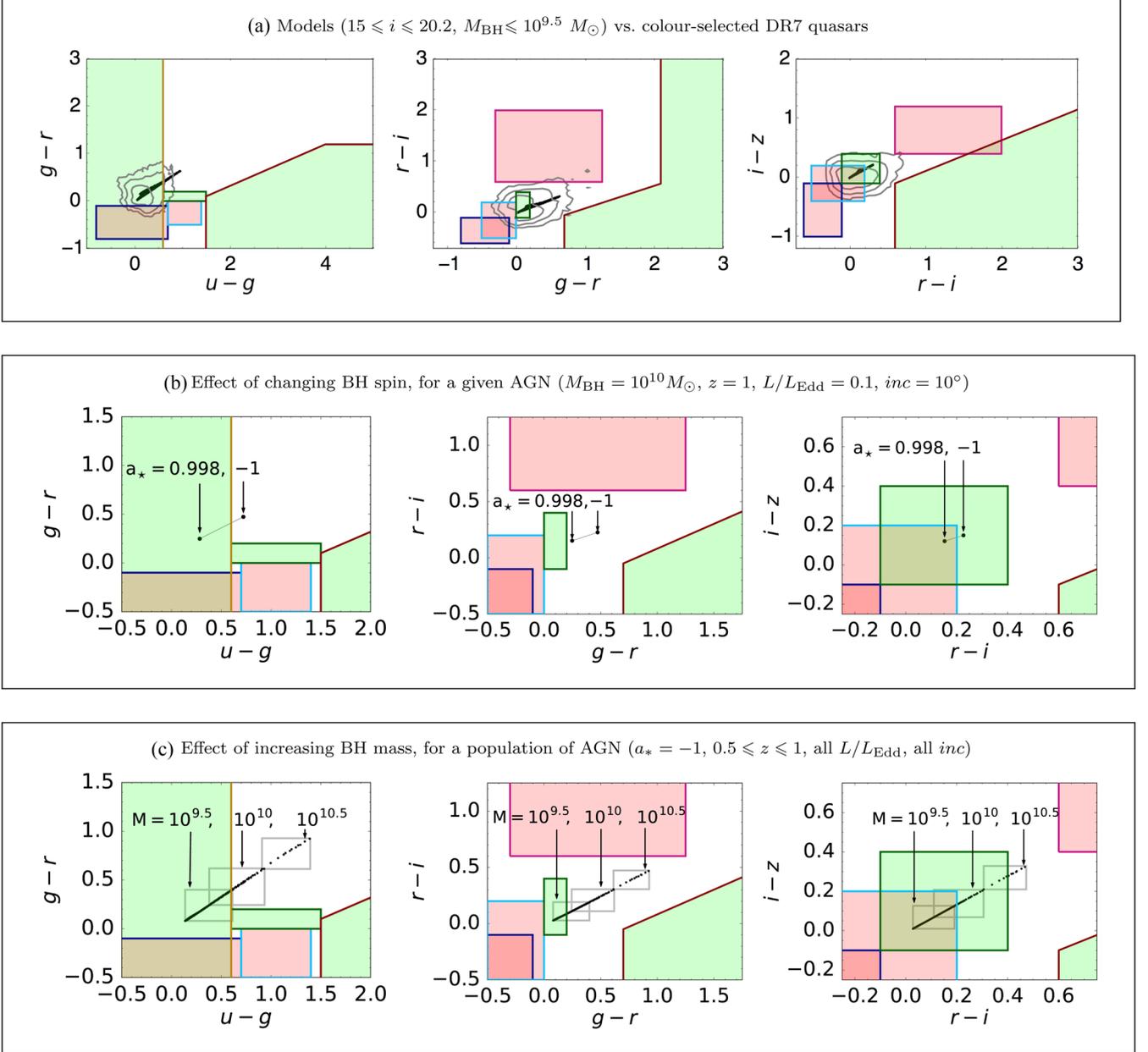
Figure 3. Position of our model SEDs in the SDSS colour–colour space. In all panels, green and red shaded regions mark the inclusion and exclusion regions, respectively, of the SDSS colour-based quasar selection algorithm. All of the model SEDs shown here include colour–temperature corrections. Note that the panels in the middle and bottom rows show an insert of the full colour–colour space (i.e. of the top row panels). Top three panels: blue points trace all the models that are within the SDSS flux limits ($15 \leq i \leq 20.2$), and with $M_{\mathrm{BH}} \leq 10^{9.5}\,\mathrm{M}_\odot$. The contours represent the colour-selected, $0.5 \leq z \leq 2$, quasars actually observed and classified in the SDSS/DR7 quasar catalogue, with subsequent contours tracing order-of-magnitude steps in the number of encircled objects. Middle three panels: the change in the position in colour–colour space caused by varying a_* from -1 to 0.998, for a model with $M_{\mathrm{BH}} = 10^{10}\,\mathrm{M}_\odot$, $L/L_{\mathrm{Edd}} = 0.1$, and inc $= 10°$, at redshift $z = 1$. In the left-hand panel, the model with $a_* = 0.998$ lies within the UV-excess inclusion region while the model with $a_* = -1$ drops out of this region, and is therefore not selected for spectroscopy. Bottom three panels: the change in the position in colour–colour space caused by varying the BH mass, while keeping BH spin fixed at $a_* = -1$. The grey rectangles bracket models with $M_{\mathrm{BH}} = 10^{9.5}$, 10^{10}, and $10^{10.5}\,\mathrm{M}_\odot$ (as illustrated), at redshifts $0.5 < z < 1$.

algorithm. In what follows, we mainly focus on the set of model SEDs that include colour temperature corrections.

The top three panels in Fig. 3(a) present the location of a large subset of our model SEDs in the SDSS colour–colour space. In these panels we only show models fulfilling $15 \leq i \leq 20.2$ and $M_{\mathrm{BH}} \leq 10^{9.5}\,\mathrm{M}_\odot$ (as blue points). We have not included the models with larger masses in order to allow for a better comparison to actual observations. We also show the observed population of

colour-selected quasars with $0.5 \leq z \leq 2$ from the SDSS/Data Release 7 (DR7) quasar catalogue (coloured contours; Abazajian et al. 2009; Schneider et al. 2010). For our models, we adopt the deeper flux limit of $15 \leq i \leq 20.2$ defined for the SDSS *griz*-selection, while most of SDSS sources at $0.5 \leq z \leq 2$ (about 90 per cent) are actually selected via the *ugri*-based selection criteria, and therefore subjected to the more limiting condition $15 \leq i \leq 19.1$.
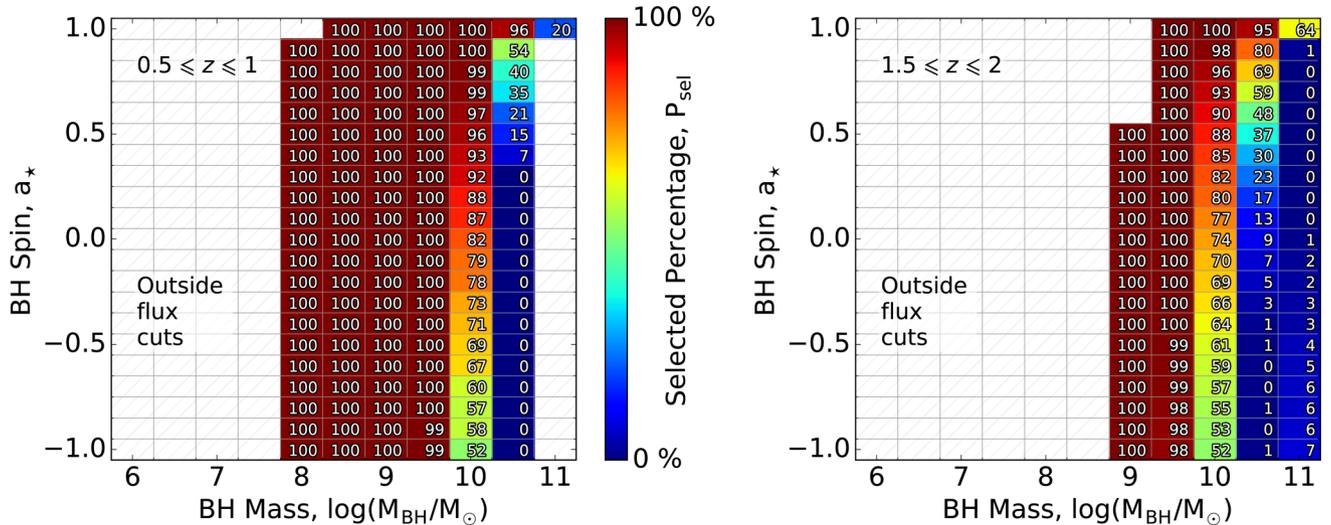
Figure 4. Percentage of *observable* sources (i.e. within the flux limits) that are selected as possible quasar candidates in each bin of (M_{BH}, a_*). Hatched bins indicate that all of the objects lie outside SDSS's flux limits. We stress that the number of observable objects varies between adjacent bins, even if they result in identical percentages of colour-selected objects (see Section 3.2).

As can be seen in the top panels of Fig. 3, our models generally lie within the space occupied by the observed SDSS quasars. Moreover, ~99.6 per cent of them are located within the UV-excess ('UVX') inclusion region ($u - g < 0.6$), and none of them is found within any (multidimensional) rejection region.

However, the colour-selected SDSS/DR7 quasars occupy a somewhat more extended region in colour–colour space, particularly in the direction perpendicular to the sequence formed by our models. The fact that our models do not account the whole span of real quasars is probably due to the varying degree of contamination from (mainly broad) emission lines, and/or the quasars' host galaxies (see e.g. Hao et al. 2013).

The centre row of panels in Fig. 3(b) illustrate the impact of the BH spin on the location of SEDs in SDSS colour space. This is illustrated by changing only the spin parameter, in the range $a_* = -1-(+)0.998$, for a subset of models where the other parameters are fixed at $M_{BH} = 10^{10}$ M$_\odot$, $L/L_{Edd} = 0.1$, and inc = $10°$. As can be seen, the change in a_* shifts the position of the model SEDs in the colour–colour plots, and in certain cases may make the difference between the model SED being selected or not, as we further discuss in Section 3.2. In this example, the model with a_*=0.998 would be selected through the UV excess inclusion region (green region with golden border in the left-hand panel), while the model with $a_* = -1$ is not selected since it is not located within *any* of the inclusion regions, and is not positioned sufficiently far outside the stellar locus.[2] The bottom three panels of Fig. 3(c) focus on the effect of increasing mass in the high-mass regime, which will be discussed further in the next section.

3.2 Completeness

We next examine what fraction of our model SEDs are selected as possible quasar candidates by the SDSS colour selection algorithm,

across the range in SMBH properties. We focus only on 'observable' model SEDs, defined through the relevant SDSS spectroscopic flux limits. We define a model as being *unobservable* either if it fails to satisfy the flux limits of the *griz*-channel [15, 20.2] or if it is not chosen by any of the *griz*-channel criteria (so can only be selected by *ugri*-space tests), but is also not included in the shallower *ugri* flux limits [15, 19.1] (see Section 2.2).

We first find that for model SEDs with $M_{BH} \leq 10^{9.5}$ M$_\odot$, the selection algorithm is extremely efficient, with a completeness rate of 99.8 per cent (when including all values of a_*, L/L_{Edd}, inc, and z, as listed in Table 1) All of the latter are selected by their UV excess through SDSS's UVX test, and about 0.014 per cent are additionally selected by SDSS's *ugri* test.

Next, we consider each bin of a given BH mass and spin, individually, to find out what percentage of the observable models within it are selected. Each such bin contains the observable model SEDs from a three-dimensional grid ranging over different values of L/L_{Edd}, inc and one of two different redshift ranges –0.5 $\leq z \leq 1$ and 1.5 $\leq z \leq 2$. Fig. 4 shows the resulting completeness rates for all the fiducial model SEDs, while Fig. 5 focuses on models ignoring colour–temperature corrections. We stress that the percentages shown in Figs 4 and Fig. 5 (and also Figs A1 and A2 in Appendix A; see below) trace the fraction of model SEDs that would have been selected by the SDSS colour-selection algorithm among all *observable* objects (i.e. those that satisfy the general *i*-band flux limits). This means that (M_{BH},a_*) bins with identical percentages generally vary in the number of observable, and thus selected model SEDs. This is of particular importance for the lowest mass, observable BHs ($M_{BH} = 10^8$ and 10^9 M$_\odot$, for the lower- and higher redshift ranges, respectively), where the number of observable model SEDs drops monotonically with increasing a_*, but the percentage of those SEDs that would be colour selected remains 100 per cent.[3]

---

[2] We note that none of the models in this example is rejected due to their position in the light blue rejection box in the right-hand panel, since the latter constitutes the projection of a four-dimensional region. In order to be rejected, an object would have to lie within the regions with a light blue frame in the other two panels as well.

[3] The decrease in the number of observable objects with an increasing a_*, for a fixed M_{BH} and (range of) L/L_{Edd}, is due to SED becoming harder, while maintaining a fixed (range of) L_{bol}. This corresponds to lower optical luminosities and therefore fainter *i*-band fluxes. See Fig. 2 for an example (cf. the solid and dashed lines).
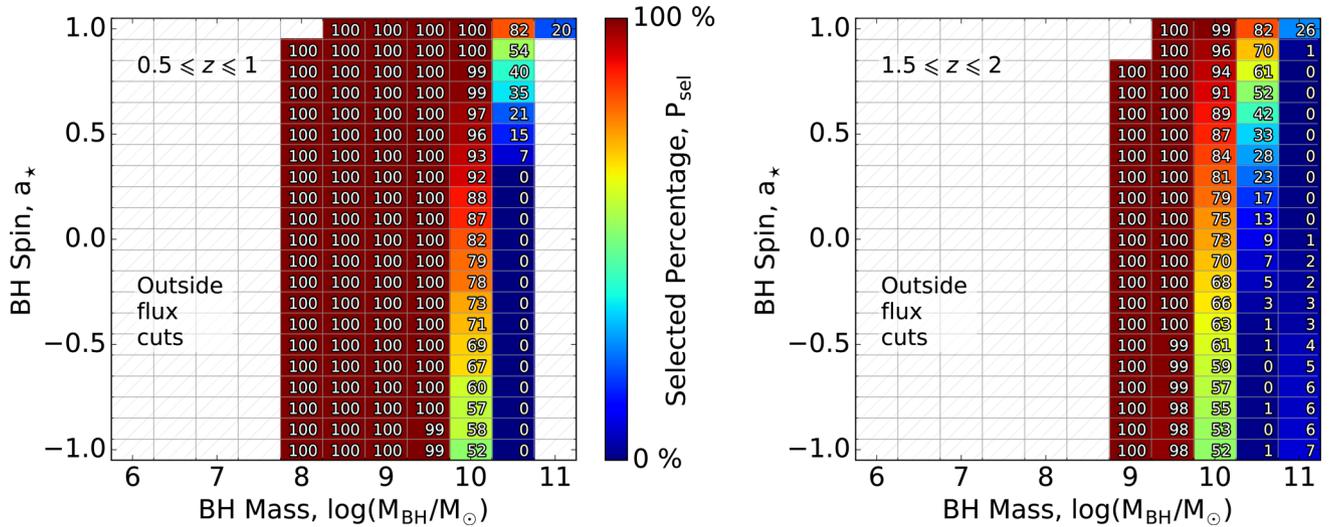
Figure 5. Same as Fig. 4, but for simplified model SEDs that ignore temperature–colour corrections (i.e. setting $f_{col} = 1$ in the Done et al. 2012 model).

As can be clearly seen in Figs 4 and 5, at the high-mass regime, where $M_{BH} \geq 10^{10} \, M_{\odot}$, the percentage of selected SEDs (i.e. the completeness) decreases significantly with mass and, more importantly, towards low or retrograde (negative) BH spin parameters, at fixed BH mass. This general trend persists both at our lower and higher redshift ranges, as it does in models with or without colour–temperature corrections, and in the subset of models with $L/L_{Edd} \leq 0.3$ (as shown in Figs A1 and A2 in Appendix A). For example, for $M_{BH} = 10^{10} \, M_{\odot}$, the fraction of selected observable models at $z \simeq 0.8$ (i.e. left-hand panels of Figs 4 and A1) drops from $\sim$100 per cent for $a_* = 0.998$ to $\sim$80 per cent and eventually $\sim$50 per cent, for models with $a_* \simeq 0$ and -1, respectively. For SEDs with lower L/L_{Edd} at higher redshifts (i.e. right-hand panel in Fig A1), the fraction drops more dramatically, reaching $\sim$20 per cent for $a_* \simeq 0$ and single-digit percentages at $a_* \leq -0.5$. At this high mass, a maximally retrograde spinning active BH would thus be only about half as likely to be selected as a candidate than a maximally prograde spinning one (at most). Such a trend in the extremely high-mass regime is relevant because this is where probing the spin could allow us to learn something about the BH's accretion history (see Section 1).

We will now get back to the position in colour–colour space for some of the most relevant models to see what causes them to fail to be selected. The bottom panels in Fig. 3(c) focus on lower redshift model SEDs which have a fixed, maximum-retrograde spin parameter $a_* = -1$ and masses of $M_{BH} = 10^{9.5}$, 10^{10}, and $10^{10.5} \, M_{\odot}$, thus sampling three adjacent bins in the left-hand panel of Fig. 4. For these horizontally adjacent bins, the percentage of selected SEDs sharply drops from 100 to 52 per cent and finally to 0 per cent with increasing mass. The bottom left-hand panel of Fig. 3 shows that this trend is due to the models dropping out of the UVX inclusion region (green region with golden border), which selects all sources with $u - g < 0.6$. However, our models with $M_{BH} = 10^{10} \, M_{\odot}$ have $u - g$ values reaching up to 0.93, and the $10^{10.5} \, M_{\odot}$ models up to 1.39. As mentioned in the previous section, the trend with decreasing spin for a given high mass $\geq 10^{10} \, M_{\odot}$ (i.e. *vertically* adjacent bins) is similarly driven by the models entering or leaving the UVX inclusion region.

In Fig. 6 we focus on the trends in selection rate with BH spin for high-mass models identified in Fig. 4, particularly for models with $M_{BH} = 10^{10}$ and $10^{10.5} \, M_{\odot}$ in both redshift ranges. Fig. 6 also
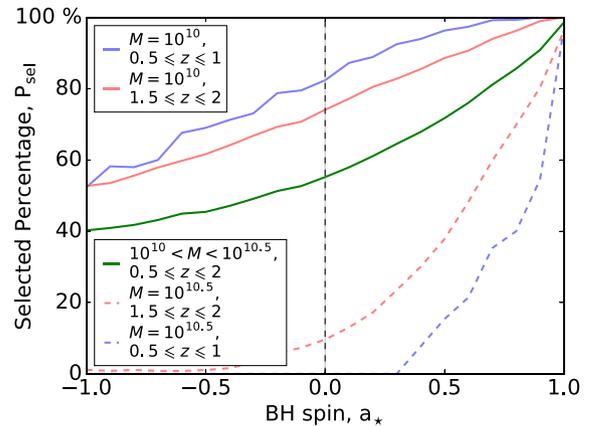

Figure 6. Fraction of selected observable models at a given mass as a function of spin. The blue curves refer to the models at the lower redshift range $0.5 < z < 1$, the red curves refer to those at $1.5 < z < 2$. The solid blue and red curves represent the models with BH mass $M_{BH} = 10^{10} \, M_{\odot}$ and the dotted ones the models with $M_{BH} = 10^{10.5} \, M_{\odot}$. The green curve includes all models with $10^{10} < M_{BH} < 10^{10.5} \, M_{\odot}$ for the whole redshift range we simulated, i.e. for $0.5 < z < 2$.

displays the fraction of selected models if we include both masses and our *entire* redshift range $0.5 < z < 2$. As can be seen, for $M_{BH} = 10^{10} \, M_{\odot}$, the spin is affecting the higher redshift models more than the lower redshift ones, but the opposite is true for $M_{BH} = 10^{10.5} \, M_{\odot}$. In fact, the effect of varying BH spin is most pronounced if the peak of the model SEDs shifts between the spectral regions covered by the u and g bands (which measure the UV excess). For the lower mass bin among the two shown in Fig. 6, $M_{BH} = 10^{10} \, M_{\odot}$, this occurs at higher redshifts, as shown in Fig. 2. However, for $M_{BH} = 10^{10.5} \, M_{\odot}$, this occurs at the lower redshift range since the higher M_{BH} shifts the SED towards longer wavelengths (see Fig. 1).

We note that the trend of decreasing selection rate with decreasing BH spin (and/or increasingly retrograde configurations) does not trivially extend to yet higher BH masses. Returning to Figs 4 and A1, we find that for models with $M_{BH} = 10^{11} \, M_{\odot}$ at higher redshifts ($1.5 \leq z \leq 2$), retrograde configurations ($a_* < 0$) have slightly higher selection rates than most of the prograde ones (see the extreme right-hand column of bins in the right-hand panels of Figs 4 and A1).
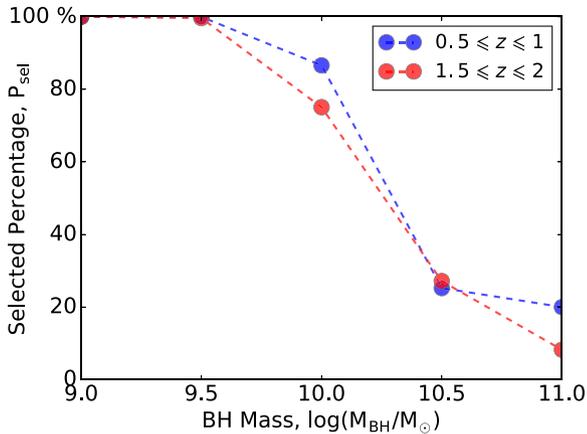
Figure 7. Fraction of selected observable models as a function of mass. The blue curves refer to the models at the lower redshift range $0.5 < z < 1$, the red curves refer to those at $1.5 < z < 2$.

This is due to the fact that, unlike the vast majority of the models in our grid, these SEDs are not selected by the UV excess, but instead by SDSS's *griz* criteria, which was designed to select objects at higher redshifts $z \geq 2.5$ (see Section 2.2). Thus, increasing the mass and/or decreasing the spin shifts the AD SED to lower frequencies in a way which is comparable to having a higher redshift. However, regardless of spin, the probability of being selected in this high-mass regime is still very low, at <10 per cent (except for the maximum spin value 0.998, where the rate is ~ 20 per cent). In Fig. 7 we plot the combined probability of selecting a model – that is, summed over all a_* values and all redshifts $0.5 < z < 2$, as a function of M_{BH}. For $M_{BH} = 10^{11}\,M_\odot$, the selection rate is ~ 3 per cent. Most of the contribution to this rate comes from models with maximum spin a_*=0.998, which have the highest selection rate. If we exclude the latter, the fraction of selected observable models in this mass regime drop to about 2 per cent.

Our analysis thus suggests that $z \sim 1.5$–2 accreting SMBHs with $M_{BH} \geq 3 \times 10^{10}\,M_\odot$, should they exist, would be virtually impossible to be selected for follow-up SDSS spectroscopy, even if they are accreting at significant fractions of the Eddington limit (i.e. having luminosities that put them above the SDSS flux limit). Subsequently, such SMBHs would be missing from any demographic analysis that is derived from the SDSS quasar content, which is the largest available (at these redshifts). As noted above, the possible effects of dust in the nuclei of the host galaxies, and/or of AD outflows, which we did not consider here, would further challenge the SDSS colour selection, as these would act to decrease the relative contribution of the (rest-frame) UV part of the emergent SEDs (e.g. Slone & Netzer 2012; Capellupo et al. 2015; Collinson et al. 2015).

Our results add to a number of other biases that work against extremely massive BHs in UV–optical surveys. It has been recently argued that above a certain threshold BH mass, of roughly $5 \times 10^{10}\,M_\odot$ (the exact limit depending on several parameters), it would be impossible to sustain a stable, rotating Keplerian AD outside the ISCO, due to the combined effects of gas and/or radiation pressure (Inayoshi & Haiman 2016; King 2016). This would imply that accenting SMBHs with such high masses cannot possibly be observed as radiatively efficient AGN. Furthermore, it has been suggested that high-mass, low- (or retrograde-) spin BHs would not be identified by high-ionization emission lines, since the fraction of ionizing photons of their continuum emission would be too low (see e.g. Laor & Davis 2011; Netzer &

Trakhtenbrot 2014). We also note that *low*-mass nuclear BHs with $M_{BH} \leq 10^6\,M_\odot$ may also escape detection, as they may not be able to produce strong broad emission lines (regardless of their intrinsically low luminosities; see Chakravorty, Elvis & Ferland 2014). Our analysis demonstrates that the difficulty in identifying such objects does not necessarily lie in the existence of some spectral features, but in fact in the practical aspects of simply obtaining a spectrum of an optical source identified in a large, multiband optical survey. From the empirical point of view, recent discoveries of ultraluminous $z > 5$ quasars within the SDSS footprint, powered by SMBHs with $M_{BH} \simeq 10^{10}\,M_\odot$ (Wang et al. 2015; Wu et al. 2015), have demonstrated the need for non-standard spectroscopic follow-up target selection criteria, which go beyond the standard colour-based procedures. All this suggests that our current understanding of the high-M_{BH}, and particularly the low- or retrograde-spin, regime of the SMBH population may be highly incomplete.

4 SUMMARY AND CONCLUSIONS

We have tested the completeness of the SDSS quasar colour-based selection algorithm by forward-modelling the SEDs of accreting, unobscured SMBHs. This was done by generating a large grid of physically motivated model SEDs of geometrically thin, optically thick accretion discs, varying several key physical parameters. In this work we focused on the dominant, accretion disc emission, thus ignoring the more complex emission features seen in real AGN spectra, which are however expected to be of minor importance. Our models include relativistic corrections and an inclination dependence. We have obtained the magnitudes of each model SED in the five photometric bands of SDSS, and fed them into the SDSS colour-based quasar selection algorithm. Our models overlap with the actual population of colour-selected quasars observed within the SDSS/DR7. However, the observed population spans a broader range in colour–colour space than our models. It is possible that a better match, and therefore a more predictive grid of SEDs, can be obtained by extending the models to include emission lines, observationally motivated noise, and/or using slim disc models. This is however beyond the scope of the present study, which focuses on identifying possible biases in the SDSS selection algorithm, particularly with regard to low or retrograde spins among the most massive BHs.

We have found a very high level of completeness across most of the BH mass range, up to $M_{BH} \simeq 10^{9.5}\,M_\odot$: the SDSS colour selection algorithm selects around above 99 per cent of the model SEDs which satisfy the survey flux limits. For higher masses, however, we have shown that the completeness drops with increasing mass, and most importantly with decreasing BH spins and for (increasingly) retrograde spin configurations. Our results imply that if any sources with $M_{BH} = 3 \times 10^{10.5}$ or $10^{11}\,M_\odot$ exist, they would be practically impossible to detect, with an average target identification rate of about 3 per cent. Therefore, large spectroscopic surveys of optically selected quasars, such as the SDSS, may be missing a significant population of such objects, even if they do exist as such (see Inayoshi & Haiman 2016; King 2016). At $M_{BH} = 10^{10}\,M_\odot$, we have shown that a maximally retrograde spinning active SMBH would be at most about half as likely to be selected for a spectroscopic follow-up observation as a maximally prograde spinning one. This could undermine any attempts to probe SMBH accretion history via spin estimates, which is in principle possible in the extremely high-mass regime. Most importantly, our analysis suggests that SMBHs which have grown to extremely high masses through a series of isotropically oriented accretion episodes, and are thus

expected to be 'spun down', might be missing from large samples of optically selected quasars. Thus, some of the recent studies that report high BH spins among the most massive BHs at $z \geq 1$ (e.g. Wu et al. 2013; Trakhtenbrot 2014; Capellupo et al. 2015, 2016) might be affected by selection biases.

We finally note that in this study we only tested the completeness of one key step in the process of quasar detection and identification. In order to quantify the overall (in-)completeness, one would need to explore possible selection effects in broad and narrow emission line diagnostics. Another natural direction for follow-up investigation would be to test to what extent do alternative AGN selection criteria, such as those based on X-ray or radio emission, depend on BH spin.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Arnaud K., 1996, in Jacoby G., Barnes J., eds, ASP Conf. Ser. Vol. 101, Astronomical Data Analysis Software and Systems V. Astron. Soc. Pac., San Francisco, p. 17
Banerji M., McMahon R. G., Hewett P. C., Alaghband-Zadeh S., Gonzalez-Solares E., Venemans B. P., Hawthorn M. J., 2012, MNRAS, 427, 2275
Baron D., Stern J., Poznanski D., Netzer H., 2016, ApJ, preprint (arXiv:1603.06948)
Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559
Blackburne J. A., Kochanek C. S., Chen B., Dai X., Chartas G., 2014, ApJ, 789, 125
Capellupo D. M., Netzer H., Lira P., Trakhtenbrot B., 2015, MNRAS, 446, 3427
Capellupo D. M., Netzer H., Lira P., Trakhtenbrot B., Mejía-Restrepo J., 2016, MNRAS, 460, 212
Chakravorty S., Elvis M., Ferland G., 2014, MNRAS, 437, 740
Collinson J. S., Ward M. J., Done C., Landt H., Elvis M., McDowell J. C., 2015, MNRAS, 449, 2174
Czerny B., Hryniewicz K., Nikołajuk M., Sądowski A., 2011, MNRAS, 415, 2942
Dai X., Kochanek C. S., Chartas G., Kozłowski S., Morgan C. W., Garmire G., Agol E., 2010, ApJ, 709, 278
Davis S. W., Laor A., 2011, ApJ, 728, 98
Davis S. W., Woo J., Blaes O. M., 2007, ApJ, 668, 682
Doi M. et al., 2010, AJ, 139, 1628
Done C., Davis S. W., Jin C., Blaes O., Ward M., 2012, MNRAS, 420, 1848
Done C., Jin C., Middleton M., Ward M., 2013, MNRAS, 434, 1955
Dotti M., Colpi M., Pallini S., Perego A., Volonteri M., 2013, ApJ, 762, 68
Elvis M., Risaliti G., Zamorani G., 2002, ApJ, 565, L75
Fausnaugh M. M. et al., 2016, ApJ, 821, 56
Fine S. et al., 2008, MNRAS, 390, 1413
Gavignaud I. et al., 2008, A&A, 492, 637
Glikman E., Helfand D. J., White R. L., Becker R. H., Gregg M. D., Lacy M., 2007, ApJ, 667, 673
Hao H. et al., 2013, MNRAS, 434, 3104
Inayoshi K., Haiman Z., 2016, ApJ, 828, 110
Jin C., Ward M. J., Done C., Gelbord J., 2012, MNRAS, 420, 1825
King A., 2016, MNRAS, 456, L109
King A., Pringle J., 2006, MNRAS, 373, L90
King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Laor A., Davis S. W., 2011, MNRAS, 417, 681
McLure R. J., Dunlop J. S., 2004, MNRAS, 352, 1390
Merloni A. et al., 2010, ApJ, 708, 137
Miller L., Turner T. J., 2013, ApJ, 773, L5
Netzer H., 2013, The Physics and Evolution of Active Galactic Nuclei. Cambridge Univ. Press, Cambridge (http://adsabs.harvard.edu/abs/2013peag.book.....N)
Netzer H., Trakhtenbrot B., 2014, MNRAS, 438, 672
Pooley D., Blackburne J., Rappaport S., Schechter P., 2007, ApJ, 661, 19
Raimundo S. I., Fabian A. C., Vasudevan R. V., Gandhi P., Wu J., 2012, MNRAS, 419, 2529
Reis R. C., Reynolds M. T., Miller J. M., Walton D. J., 2014, Nature, 507, 207
Reynolds C. S., 2014, Space Sci. Rev., 183, 277
Reynolds M. T., Walton D. J., Miller J. M., Reis R. C., 2014, ApJ, 792, L19
Richards G. T. et al., 2002, AJ, 123, 2945
Schneider D. P. et al., 2010, AJ, 139, 2360
Schulze A., Wisotzki L., 2010, A&A, 516, A87
Schulze A. et al., 2015, MNRAS, 447, 2085
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shen Y. et al., 2011, ApJS, 194, 45
Slone O., Netzer H., 2012, MNRAS, 426, 656
Soltan A., 1982, MNRAS, 200, 115
Thorne K. S., 1974, ApJ, 191, 507
Trakhtenbrot B., 2014, ApJ, 789, L9
Trakhtenbrot B., Netzer H., 2012, MNRAS, 427, 3081
Trump J. R. et al., 2009, ApJ, 700, 49
Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, ApJ, 786, 33
Vanden Berk D. E. et al., 2001, AJ, 122, 549
Volonteri M., Sikora M., Lasota J.-P., Merloni A., 2013, ApJ, 775, 94
Wang F. et al., 2015, ApJ, 807, L9
Wu S., Lu Y., Zhang F., 2013, MNRAS, 436, 3271
Wu X.-B. et al., 2015, Nature, 518, 512
York D. G. et al., 2000, AJ, 120, 1579
Yu Q., Tremaine S., 2002, MNRAS, 335, 965

APPENDIX A: LOWER EDDINGTON RATIOS

Here we present alternative versions of Figs 4 and 5, of the percentages of selected model SEDs in the M_{BH}–a_* plane, but accounting only for objects with $L/L_{Edd} \leq 0.3$ – the regime where the thin AD model is expected to be valid. As can be clearly seen in these alternative Figs A1 and A2, the main trends we identified for high-M_{BH} BHs are unchanged. Our main findings regarding a bias of the SDSS colour selection algorithm are therefore not due to our choice of range in L/L_{Edd}.

APPENDIX B: TESTING THE EFFECTS OF EMISSION LINES

Here we describe the test we performed to verify that our decision to focus on continuum emission, ignoring the emission lines and features, does not significantly affect our findings.

We added a 'typical' emission line spectrum of unobscured SDSS quasars to all the fiducial calculated AD SEDs (i.e. those *with* temperature colour corrections). The emission line spectrum was extracted from the composite SDSS quasar spectrum of Vanden Berk et al. (2001). The composite spectrum was normalized to the best-fitting power-law continuum emission ($f_\nu \propto \nu^{-0.44}$), to produce a *relative* emission line spectrum in the spectral range of $\lambda_{rest} = 1190$–5100 Å (the spectral range relevant for our purpose and where the power-law approximation is valid). Next, each of the fiducial calculated SEDs was multiplied by this relative, composite emission line spectrum. Similarly to the procedure used for the fiducial
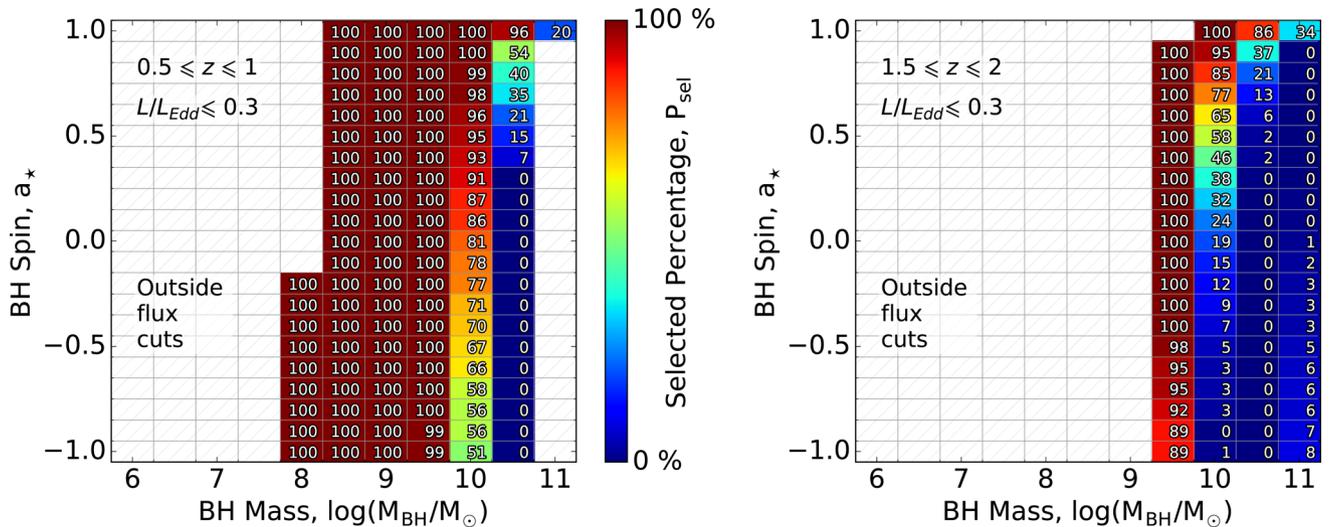
Figure A1. Same as Fig. 4, but only for model SEDs with $L/L_{Edd} \leq 0.3$, but for simplified model SEDs that neglect colour–temperature corrections.

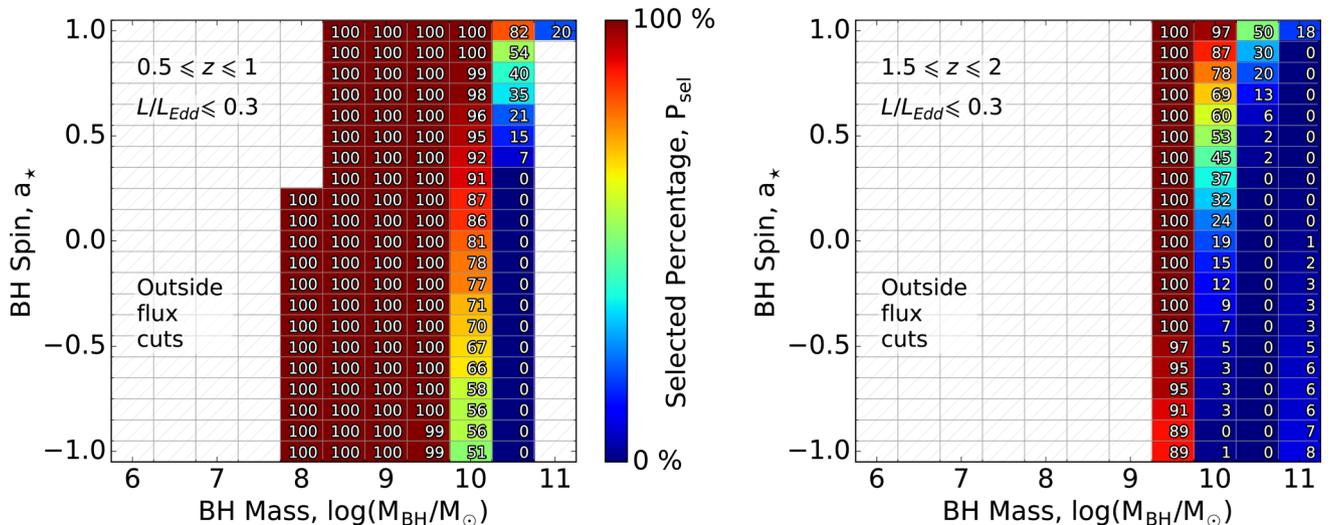**Figure A2.** Same as Fig. 5 (i.e. ignoring temperature colour corrections; $f_{col} = 1$), but only for model SEDs with $L/L_{Edd} \leq 0.3$.

SEDs, we then calculated the synthetic magnitudes of these modified SEDs, and followed the SDSS colour-based target selection algorithm.

Fig. B1 presents the selection percentages of these modified SEDs in the a_*–M_{BH} plane, similarly to what we show in Fig. 4. The general trends with BH spin, at the high-M_{BH} regime, remain essentially identical to what we find for the fiducial set of SEDs. The fraction of SEDs selected by the SDSS algorithm drop with decreasing a_*, reaching $\lesssim 75$ and $\lesssim 10$ per cent for SEDs with $a_* \leq 0$, in the mass bins of $M_{BH} \simeq 10^{10}$ and $> 10^{10} \, M_\odot$, respectively. A closer inspection of the percentages suggests that these are somewhat higher than those derived for the fiducial model SEDs. For example, for maximally retrograde spinning BHs at $z \sim 1.8$ with $M_{BH} \simeq 10^{11} \, M_\odot$, ~ 15 per cent of the modified SEDs would have been selected by the SDSS algorithm, compared with ~ 7 per cent among the fiducial model.

We stress, however, that the SEDs from such high-M_{BH}, low-a_* systems are expected to have extremely weak ionizing radiation fields. Using SEDs very similar to ours, Laor & Davis (2011) estimate that for a non-spinning SMBH with $M_{BH} = 10^{10} \, M_\odot$ only about 2×10^{-6} of the photons would be ionizing (i.e. beyond the Lyman limit). Retrograde spins would result in yet fewer ionizing photons. Therefore, the strong broad high-ionization lines seen in typical quasars (e.g. C IV $\lambda 1549$) are expected to be much weaker than what we considered in our test. This, in turn, would lead to 'redder' colours, and therefore lower completeness levels in the context of the SDSS colour-based selection procedure.

We conclude that the inclusion of emission lines and features of the strengths observed in typical luminous quasars does not significantly alter our main findings (derived from the fiducial SEDs), and that in reality any such effects are most probably expected to be even more limited.
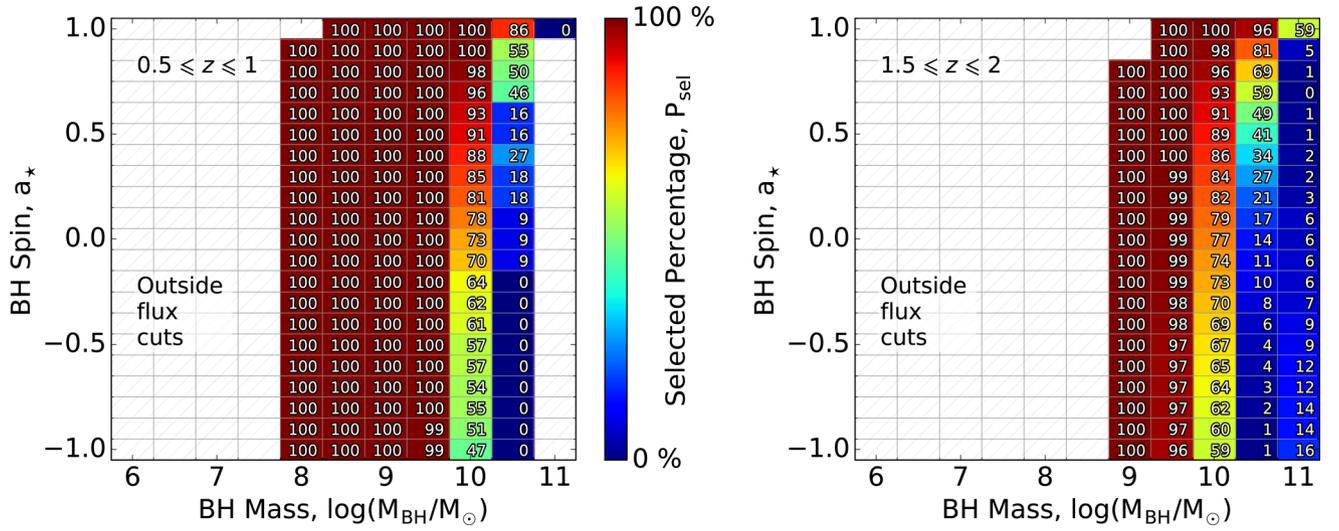
Figure B1. Same as Fig. 4, but for modified model SEDs that include contribution from emission lines and features, typical of SDSS quasars.

This paper has been typeset from a TEX/LATEX file prepared by the author.