Improved redshifts for SDSS quasar spectra

Paul C. Hewett1* and Vivienne Wild2

1Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA
2Institut d’Astrophysique de Paris, 98bis Boulevard Arago, 75014 Paris, France

ABSTRACT

A systematic investigation of the relationship between different redshift estimation schemes for more than 91 000 quasars in the Sloan Digital Sky Survey (SDSS) Data Release 6 is presented. The publicly available SDSS quasar redshifts are shown to possess systematic biases of $\Delta z/(1 + z) \gsim 0.002$ (600 km s$^{-1}$) over both small ($\Delta z \sim 0.1$) and large ($\Delta z \sim 1$) redshift intervals. Empirical relationships between redshifts based on (i) Ca II H&K absorption, (ii) quasar [O III] $\lambda\lambda$3727, 3729 emission and (iv) cross-correlation (with a master-quasar template) that includes, at increasing quasar redshift, the prominent Mg II $\lambda$2799, C IV $\lambda$1549 emission lines are established as a function of quasar redshift and luminosity. New redshifts in the resulting catalogue possess systematic biases, a factor of $\sim 20$ lower compared to the SDSS redshift values; systematic effects are reduced to the level of $\Delta z/(1 + z) \lsim 10^{-4}$ (30 km s$^{-1}$) per unit redshift or $\lsim 2.5 \times 10^{-5}$ per unit absolute magnitude. Redshift errors, including components due both to internal reproducibility and to the intrinsic quasar-to-quasar variation among the population, are available for all quasars in the catalogue. The improved redshifts and their associated errors have wide applicability in areas such as quasar absorption outflows, quasar clustering, quasar-galaxy clustering and proximity-effect determinations.

Key words: catalogues – surveys – quasars: emission lines – quasars: general.

1 INTRODUCTION

The Sloan Digital Sky Survey (SDSS; York et al. 2000) has produced a revolution in both the volume and quality of spectroscopic data available for quasars. The Data Release 5 (DR5; Adelman-McCarthy et al. 2007) and Legacy DR7 (Abazajian et al. 2009) with their associated quasar catalogues (Schneider et al. 2007, 2010, respectively) provide intermediate resolution ($R \sim 2000$), moderate signal-to-noise ratio (S/N; S/N $\sim 15$ per 69 km s$^{-1}$ pixel) and spectra of unprecedented homogeneity, covering essentially the entire ‘optical’ wavelength region ($\lambda = 3000–9180$ Å).

The quality of the Schneider et al. quasar catalogues is truly impressive, with errors in redshift identification reduced to the 0.01 per cent level, and individual redshift estimates, resulting primarily from the SDSS spectroscopic pipeline (and the SDSS DR7 website;8 Stoughton et al. 2002), are accurate to of order $\Delta z/(1 + z) \sim 0.002$. The publication of even individual quasar redshifts, based on moderate resolution spectra, to such accuracy was a significant achievement prior to the mid-1990s, further highlighting the advance represented by the SDSS.

Notwithstanding the quality of the SDSS quasar spectra and the associated redshift estimates, important scientific investigations, including the clustering of quasars themselves (e.g. Croom et al. 2002; Shen et al. 2007), the cross-correlation of quasars and other object populations (e.g. Padmanabhan et al. 2009), the proximity effect (e.g. Bajtlik, Duncan & Ostriker 1988; Kirkman & Tytler 2008) and the origin and properties of associated absorbers (e.g. Nestor, Hamann & Hidalgo 2008; Wild et al. 2008; Tytler et al. 2009), benefit significantly both from reduced systematics in redshift determinations and from the reliable assignment of redshift uncertainties for individual quasars.

In this paper, we present the determination of new redshifts and associated error estimates for more than 89 500 quasars from the SDSS DR6 (Adelman-McCarthy et al. 2008). Our redshift determinations suffer from much smaller systematic effects compared to the default values from the SDSS spectroscopic pipeline. Specifically, systematics are reduced by more than an order of magnitude to $1.0 \times 10^{-4}$ in $\Delta z/(1 + z)$ per unit redshift or, equivalently, 30 km s$^{-1}$ per unit redshift. A detailed comparison of redshifts derived from Ca II H&K absorption, [O II] $\lambda\lambda$3727, 3729 emission, [O III] $\lambda\lambda$4960, $\lambda$2799, C IV $\lambda$1549 emission lines are established as a function of quasar redshift and luminosity. New redshifts in the resulting catalogue possess systematic biases, a factor of $\sim 20$ lower compared to the SDSS redshift values; systematic effects are reduced to the level of $\Delta z/(1 + z) \lsim 10^{-4}$ (30 km s$^{-1}$) per unit redshift or $\lsim 2.5 \times 10^{-5}$ per unit absolute magnitude. Redshift errors, including components due both to internal reproducibility and to the intrinsic quasar-to-quasar variation among the population, are available for all quasars in the catalogue. The improved redshifts and their associated errors have wide applicability in areas such as quasar absorption outflows, quasar clustering, quasar-galaxy clustering and proximity-effect determinations.

Key words: catalogues – surveys – quasars: emission lines – quasars: general.

*E-mail: phewett@ast.cam.ac.uk

http://www.sdss.org/dr7/algorithms/redshift_type.html

© 2010 The Authors. Journal compilation © 2010 RAS
5008 emission and cross-correlation with a new quasar template spectrum provides greatly improved error estimates for individual quasar redshifts. The error estimates incorporate both the uncertainties resulting from the properties of the SDSS spectra, quantified using the very large number of multiple spectra present in the SDSS, and the intrinsic quasar-to-quasar dispersion. The resulting catalogue will allow significant advances in many studies that rely on the determination of systemic quasar redshifts with small systematics and well-determined uncertainties.

The paper is structured as follows. Section 2 describes the quasar sample, before the features of the quasar redshifts available from the SDSS spectroscopic pipeline are illustrated in Section 3. Section 4 includes a description of the procedures involved in generating a master-quasar template for the cross-correlation redshift estimates. Section 5 then describes the procedures employed to provide the new redshift estimates for the SDSS quasars. An assessment of the consistency of the different redshift estimates is given in Section 6 and redshift estimates, based on different rest-frame wavelength regions, are placed on to the same ‘systemic’ reference system. A critical assessment of the internal and external reliability of the new quasar redshifts is presented at this point. 21 cm radio observations of the majority of the SDSS quasars are available from the Faint Images of the Radio Sky at Twenty centimetres (FIRST; Becker, White & Helfand 1995). Section 7 contains a description showing how the new redshift estimation scheme allows spectral energy distribution (SED) dependent composite spectra [for FIRST-detected (FD) quasars in this case] to be constructed, producing significantly improved redshifts. The resulting redshift catalogue, including well-determined error estimates for each quasar, is described in Section 8. A short discussion, including consideration of the origin of the differences with published redshifts and an independent test of the new redshifts, follows in Section 9. The paper concludes with a brief summary of the conclusions as Section 10. We adopt the same convention as employed in the SDSS and use vacuum wavelengths throughout the paper. Absolute magnitudes are calculated in a cosmology with $H_0 = 70\,\text{km s}^{-1}$, $\Omega_M = 0.3$ and $\Omega_{\Lambda} = 0.7$.

2 QUASAR SAMPLE

The quasar sample consists of 91,665 objects, including 77,392 quasars in the Schneider et al. (2007) DR5 catalogue that are retained in the later DR7 quasar catalogue of Schneider et al. (2010). A further 13,081 objects are quasars, present in the additional DR6 spectroscopic plates, identified by one of us (PCH) using a similar prescription to that employed by Schneider et al. (2007), all of which are present in the Schneider et al. (2010) catalogue. An additional 1192 objects, which do not satisfy one, or both, of the emission-line velocity width or absolute magnitude criterion imposed by Schneider et al. (2007), are also included. While formally failing the ‘quasar’ definition of Schneider et al.’s DR5 and DR7 compilations, the objects are essentially all luminous active galactic nuclei (AGN). None of the results in the paper depends on the exact definition of the ‘quasar’ sample used.

The spectra were all processed through the sky-residual subtraction scheme of Wild & Hewett (2005), resulting in a significantly improved S/N at wavelengths $\lambda > 7200\,\text{Å}$. The S/N improvement allows the important quasar rest-frame wavelength regions containing the Mg $\text{H}$ $\lambda$2796 and C $\text{II}$ $\lambda$1908 emission lines to contribute to the cross-correlation redshift determinations (Section 4.2) to much higher redshifts than is possible using the original SDSS spectra. The SDSS DR6 contains a very large number of objects for which multiple spectra are available. For our quasar sample, there are $\approx 9000$ independent pairs of spectra. The catalogue of spectrum pairs allows the accurate determination of redshift reproducibility as a function of S/N, redshift and cross-correlation amplitude, and extensive use of the spectrum pairs is made to quantify the contribution of the SDSS spectra themselves to the quasar redshift errors.

3 SDSS REDSHIFTS

The SDSS spectroscopic pipeline (SPECTRO1D) incorporates a sophisticated scheme$^1$ for determining both the classification (star, galaxy, quasar, etc.) of the spectra and the redshifts of extragalactic objects. Cross-correlation redshift estimates are determined using the Tonry & Davis (1979) technique and a composite quasar template from Vanden Berk et al. (2001). Emission lines are identified via a wavelet transform technique and an independent redshift estimate is derived using the observed-frame wavelength emission-line locations and reference rest-frame emission-line wavelengths, the latter taken from the Vanden Berk et al. (2001) composite quasar spectrum. The reference wavelengths$^3$ adopted from the composite quasar composite can differ from laboratory values due to the complex, often asymmetric, line profiles and apparent ‘velocity shifts’ of the line centroids (Gaskell 1982; Tytler & Fan 1992; Richards et al. 2002).

The SDSS data base and the individual FITS spectrum files contain extensive quantitative information on the determination and reliability of the different redshift estimates. However, the majority of the researchers utilize the ‘final’-redshift estimate $z$ included in the SDSS SPECObjAll table, the individual FITS spectrum file headers, or from the Schneider et al. catalogues.$^4$

If available, the cross-correlation redshift is adopted as the ‘final’ redshift for the spectrum. Some 88 per cent of the quasars possess redshifts derived from cross-correlation and more than a third of such objects also possess consistent emission-line-based redshifts. A further 7 per cent of quasars, where no reliable cross-correlation redshift is available, possess redshifts derived from the emission lines. The remaining 5 per cent of quasars, including a large fraction of pathological objects and spectra of a low S/N, have redshifts derived via manual inspection of the spectra.

3.1 SDSS Princeton redshifts

Independent spectrum classifications and redshift determinations, based on direct $\chi^2$-fitting of template spectra to the data, have been made at Princeton using the SPECTRO1D code.$^5$ The redshift determination, essentially via cross-correlation, differs from the implementation employed in the SPECTRO1D pipeline but the same composite quasar template from Vanden Berk et al. (2001) was used.

3.2 SDSS redshift intercomparison

Fig. 1 shows a comparison of the SDSS final redshifts and Princeton redshifts as a function of quasar redshift.$^6$ The selection of the

\[^{1}\text{http://www.sdss.org/dr7/dm/flatFiles/spSpec.html describes the SPECTRO1D FITS-file data model and lists the emission-line wavelengths.}\]

\[^{2}\text{In the Schneider et al. DR5 and DR7 quasar catalogues, catastrophic redshift errors are virtually absent but otherwise the catalogued redshifts are the ‘final’-redshift estimates from the pipeline reductions.}\]

\[^{3}\text{http://spectro.princeton.edu/}\]

\[^{4}\text{All figures showing redshift differences between estimates } z_1 \text{ and } z_2 \text{ have } \Delta z/(1 + z) = (z_1 - z_2)/(1 + z) \text{ plotted as the y-axis. The choice of which estimate is used in the denominator is usually irrelevant given the scale of the plots.}\]
sub-sample of more than 70,000 spectra is conservative in that only spectra with high-confidence SDSS redshifts, where there is also no inconsistency between the cross-correlation and emission-line redshift determinations, are used. The data in Fig. 1 should essentially represent an internal consistency check and the large differences between redshifts, extending to $\pm 5 \times 10^{-3}$, or $1500 \text{ km s}^{-1}$, are surprising. Perhaps even more striking is the sequence of apparent discontinuities in the behaviour as a function of redshift.

A second illustration of the extent of redshift-dependent systematics comes from comparing the redshift derived from the location of the Mg $\text{II}\lambda\lambda 2796, 2803$ emission in each quasar spectrum with the SDSS redshift. Fig. 2 presents the data for more than 60,000 spectra with S/N $\geq 10$ Mg $\text{II}$ emission-line locations (from the SDSS spectroscopic pipeline$^7$). The rest-frame location of the Mg $\text{II}$ emission line has been shown by many studies over the decades to be well behaved and there is no reason to expect $\pm 500 \text{ km s}^{-1}$ shifts over small redshift intervals or, indeed, an apparent systematic $2 \times 10^{-3}$ change in the location of the Mg $\text{II}$ emission with increasing redshift of the quasars. The systematic redshift differences show similar patterns over the redshift range common to both Figs 1 and 2. Although somewhat more complex to interpret (Section 4.4), the equivalent plot for the C $\text{III}\lambda\lambda 4606, 5008$ is an artefact resulting from the $\pm 1500 \text{ km s}^{-1}$ velocity interval (about the cross-correlation redshift) within which the SDSS reduction pipeline searches for emission lines. Data for 59,260 spectra are included (2628 lie outside the y-axis range plotted).

**Figure 1.** Redshift differences, $\Delta z/(1+z)$, between SDSS and Princeton pipeline reductions. Spectra plotted possess SDSS final redshifts with high confidence ($\Delta\text{Conf} > 0.9$) derived via cross-correlation ($\Delta\text{status} = 3$ or 4). Large differences between redshifts extend to amplitudes of $\pm 5 \times 10^{-3}$ ($\pm 1500 \text{ km s}^{-1}$). Particularly striking is the sequence of apparent discontinuities in the behaviour of the redshift differences as a function of redshift. Data for 69,915 spectra are included (503 lie outside the y-axis range plotted).

**Figure 2.** Redshift differences, $\Delta z/(1+z)$, between SDSS redshifts and redshifts derived from the SDSS-determined Mg $\text{II}$ emission-line locations. Spectra plotted possess Mg $\text{II}$ emission line S/N $> 10$. The solid red line, calculated using a 2001-point running median of the data points, shows the form of the systematic trends with redshift. Systematic redshift differences of $\pm 500 \text{ km s}^{-1}$ shifts over small redshift intervals are evident and, over a larger redshift interval, a prominent systematic trend of $2 \times 10^{-3}$ ($600 \text{ km s}^{-1}$) can be seen. Data for 60,190 spectra are included (2101 lie outside the y-axis range plotted).

**Figure 3.** Redshift differences, $\Delta z/(1+z)$, between SDSS redshifts and redshifts derived from the SDSS-determined C $\text{III}\lambda\lambda 4606, 5008$ emission-line locations. Spectra plotted possess C $\text{III}\lambda\lambda 4606, 5008$ emission-line S/N $> 10$. The wavelength adopted for the C $\text{III}\lambda\lambda 4606, 5008$ centroid has been chosen to produce, on average, zero-offset for $z < 2.2$. The solid red line, calculated using a 2001-point running median of the data points, shows the form of the systematic trends with redshift. Large systematic changes of up to $3 \times 10^{-3}$ ($900 \text{ km s}^{-1}$) over small redshift intervals can be seen. The prominent horizontal `feature' close to $z=4 \times 10^{-3}$ is an artefact resulting from the $\pm 1500 \text{ km s}^{-1}$ velocity interval (about the cross-correlation redshift) within which the SDSS reduction pipeline searches for emission lines. Data for 59,260 spectra are included (2628 lie outside the y-axis range plotted).

with S/N $\geq 10$ Mg $\text{II}$ emission-line locations (from the SDSS spectroscopic pipeline$^7$). The rest-frame location of the Mg $\text{II}$ emission line has been shown by many studies over the decades to be well behaved and there is no reason to expect $\pm 500 \text{ km s}^{-1}$ shifts over small redshift intervals or, indeed, an apparent systematic $2 \times 10^{-3}$ ($600 \text{ km s}^{-1}$) change in the location of the Mg $\text{II}$ emission with increasing redshift of the quasars. The systematic redshift differences show similar patterns over the redshift range common to both Figs 1 and 2. Although somewhat more complex to interpret (Section 4.4), the equivalent plot for the C $\text{III}\lambda\lambda 4606, 5008$ emission (Fig. 3) also shows strong systematic effects as a function of redshift. The form and substantial amplitude of the systematic and random differences in Figs 1–3 led to the initiation of the investigation presented here.

### 4 MASTER-QUASAR TEMPLATE CONSTRUCTION

The generation of the high-S/N quasar template to be used to calculate cross-correlation redshifts begins with a sample of quasars at low redshifts that possess emission-line-determined redshifts. A somewhat more involved procedure is then necessary to incorporate additional quasars at higher redshifts into the master template. In this section, the recipe for each element of the master template construction is outlined.

#### 4.1 Initial low-redshift quasar template

The narrow forbidden emission lines of $[O \text{III}]\lambda\lambda 4960, 5008$ are prominent in many quasar spectra with redshifts $z < 0.8$, and a composite spectrum based on the combination of quasars with

$^7$The S/N constraints applied to the use of emission lines refer to the significance of the emission-line detection by the SDSS spectroscopic pipeline.
A `continuum' is defined using a median filter of 21 pixels and [O iii] λλ4960, 5008 emission then identified using a matched-filter detection scheme applied to a continuum-subtracted version of each spectrum (e.g. Hewett et al. 1985). [O iii] emission is often broad and frequently exhibits strong asymmetries (Heckman et al. 1981); the small filter-scale adopted for the [O iii] detection is chosen with the aim of isolating narrow, well-defined, peaks that may be present. The filter template consists of two Gaussian components of the same width, centred at 4960.30 and 5008.24 Å, respectively, with a flux ratio of 1:3.

Emission features can be identified reliably via detections with a relatively low S/N, particularly given the restricted wavelength range searched in each spectrum. However, given the importance of establishing accurate redshifts as the first step in the construction of the composite quasar, the 8542 spectra possessing [O iii] detections with S/N ≥ 8σ form the starting point for the template construction.

The recipe used to combine spectra with specified redshifts into a composite is as follows.

(i) Pixels falling within 6.0 Å of the strong night-sky lines at 5578.5 and 6301.7 Å are flagged.
(ii) Pixels without valid SDSS data, determined from the SDSS noise array provided for each spectrum, are flagged.
(iii) Spectra are shifted to the rest frame, with the native 69 km s\(^{-1}\) `pixels' of the original SDSS spectra retained. The signal from each spectrum is placed on to the master rest-frame wavelength array using a `nearest pixel' scheme, thereby avoiding the need for any rebinning or interpolation.
(iv) Spectra are normalized using a wavelength interval common to all spectra.
(v) Spectra are median-filtered with a window of 61 pixels to define a `continuum'. Spectrum pixels falling more than 4.5σ below the continuum, along with a grow radius of 2 pixels, are flagged, effectively removing wavelengths affected by strong narrow absorption.
(vi) The median value of all non-flagged pixels at each rest-frame wavelength is calculated (a minimum of 100 spectra must contribute).

At this point a very high S/N composite quasar spectrum is available, extending down to rest-frame wavelength λ ≃ 2300 Å. The [O iii] emission moves beyond the red limit of the SDSS spectra at z > 0.8 and it is necessary to use a cross-correlation scheme, employing a much greater wavelength range of the quasar spectrum, to allow the construction of the master template further into the rest-frame ultraviolet.

4.2 Cross-correlation redshift algorithm

The cross-correlation algorithm is based on a straightforward spatial cross-correlation between an individual quasar spectrum and a high-S/N template spectrum. The key elements of the cross-correlation calculation are (i) a conservative choice of the portions of the quasar spectrum to employ in the calculation, avoiding strong emission lines close to the edges of the observed spectrum, and (ii) application of an essentially identical `window' to both the individual quasar and template spectra prior to the cross-correlation calculation.

For each quasar spectrum, with its companion error array, pixels are excluded from the cross-correlation calculation according to a sequence of rules/tests. The first and last valid pixels, where the SDSS spectrum error array is not set to 0, define the limits of the accessible wavelength range. In the observed frame,

(i) the first 25 pixels at each end are excluded;
(ii) pixels within 6 Å of each of the strong night-sky emission lines at 5578.5 and 6301.5 Å are excluded;
(iii) narrow absorption features are identified by examining a continuum-subtracted spectrum. The continuum is defined using a 61 pixel median filter. Pixels that fall more than 4.5σ below the continuum are flagged and a grow radius of 2 pixels then applied. Thus, a single pixel exceeding the threshold results in the exclusion of 5 pixels.

The quasar spectrum is then transformed to the rest frame using the specified redshift estimate, z\(_{\text{init}}\)\(^{-1}\). In the rest frame,

(i) pixels with λ > 7000 Å are excluded;
(ii) for z\(_{\text{init}}\) > 0.38, pixels with λ > 6400 Å, i.e. the Hα region, are excluded;
(iii) for z\(_{\text{init}}\) < 0.45, pixels with λ < 2900 Å, i.e. the Mg ii region, are excluded;
(iv) for z\(_{\text{init}}\) < 1.10, pixels with λ < 1975 Å, i.e. the C iii] region, are excluded;
(v) for z\(_{\text{init}}\) < 4.00, pixels with λ < 1675 Å, i.e. the C iv region, are excluded;
(vi) pixels with λ < 1275 Å, i.e. the N v and Lyman α lines and the Lyman α forest, are always excluded.

Following the definition of the restricted wavelength interval over which the quasar spectrum is retained, continua, estimated using a large-scale, 601 pixel, median filter, are subtracted from the quasar and the template spectra. Exactly the same wavelength interval is used to estimate the continuum subtracted from the quasar and the template spectra.

With continuum-subtracted quasar, Q\(_{\text{cs}}\), and template, T\(_{\text{cs}}\), spectra available, the cross-correlation, for lag 'l', is performed:

\[
cc(l) = \frac{\sum_i Q_i T_i / \sigma_i^2}{\sqrt{\sum_i (Q_i / \sigma_i)^2 \sum_i (T_i / \sigma_i)^2}},
\]

where \(\sigma_i\) is the noise, as provided in the SDSS FITS files, and the `cs' subscripts have been omitted for clarity.

A quadratic fit is then made to the array of cc(l) values over the interval l = ±npix, with npix = 100. The fit is then refined, performing quadratic fits to narrower pixel intervals centred on the peak of the previous quadratic fit, with the final fit determined over an interval of l = ±npix/5. The output consists of a redshift estimate, z\(_{\text{fin}}\), and a cross-correlation amplitude, cc\(_{\text{max}}\), in the range −1 ≤ cc\(_{\text{max}}\) ≤ 1, which parametrizes the degree of similarity between the two spectra. Extensive experimentation demonstrates that

---

8The additional `jitter' that the simple nearest pixel scheme introduces is small, with a maximum error of 34.5 km s\(^{-1}\) and an increased dispersion of σ = 20 km s\(^{-1}\), less than a third of a pixel, in the extent of features in the resulting composite spectra.

9The SDSS-derived redshift is used to determine the value of z\(_{\text{init}}\) initially, but all the cross-correlation estimates are recalculated using an updated value of z\(_{\text{init}}\) from the cross-correlation calculation itself. The cross-correlation estimates converge to the 10\(^{-3}\) level with just one iteration.
P. C. Hewett and V. Wild

Table 1. Quasar cross-correlation template definition parameters.

| Redshift range | Median $M_I$ | Number | FD number | Wavelength coverage (Å) | Redshift method | Wavelength contribution (Å) |
|----------------|--------------|--------|-----------|--------------------------|-----------------|----------------------------|
| 0.0–0.4        | $-22.50$     | 3958   | 3958      | 2732–8004                | [O iii]         | 2732–8004                  |
| 0.4–0.8        | $-23.86$     | 4584   | 4584      | 2136–6550                | [O iii]         | 2136–6550                  |
| 0.8–1.0        | $-24.89$     | 4071   | 4071      | 1908–5099                | Mg $\scriptstyle{\alpha}_{cc}$ (> 1975 Å) | 1908–5099                |
| 1.0–1.2        | $-25.38$     | 4762   | 393       | 1732–4950                | Mg $\scriptstyle{\alpha}_{cc}$ (> 1975 Å) | 1732–4950                |
| 1.2–1.4        | $-25.77$     | 5118   | 375       | 1589–4176                | Mg $\scriptstyle{\alpha}_{cc}$ (> 1975 Å) | 1589–4176                |
| 1.4–1.6        | $-26.15$     | 5087   | 341       | 1466–3819                | Mg $\scriptstyle{\alpha}_{cc}$ (> 1975 Å) | 1466–3819                |
| 1.6–1.8        | $-26.45$     | 3882   | 262       | 1363–3534                | Mg $\scriptstyle{\alpha}_{cc}$ (> 1675 Å) | 1363–2000                |
| 1.8–2.0        | $-26.71$     | 2797   | 169       | 1275–3275                | Mg $\scriptstyle{\alpha}_{cc}$ (> 1675 Å) | 1275–2000                |

Figure 4. The master-quasar template, plotted as rest-wavelength versus $\lambda F(\lambda)$ (in arbitrary units). The prominent emission-line features of H$\alpha$, [O iii], H$\beta$, [O ii], Mg $\scriptstyle{\beta}$, C $\scriptstyle{\beta}$ i + Si $\scriptstyle{\beta}$ i + Al $\scriptstyle{\beta}$ and C $\scriptstyle{\beta}$ iv are indicated.

4.3 Quasar template extension for redshifts $0.8 < z \leq 1.6$

With the cross-correlation redshift determination procedure in place, it is possible to utilize quasars with redshift $z > 0.8$ to extend the master template. To ensure that the template construction is not adversely affected by the inclusion of spectra with a poor S/N, or the presence of broad absorption line (BAL) troughs, the full quasar sample (Section 2) was restricted to those objects satisfying the following criteria:

(i) SDSS spectrum spectroscopic S/N, SN$_R$+SN$_I$ $\geq 18.0$;
(ii) quasar not identified as BAL quasars by Gibson et al. (2009) or from our own BAL catalogue (Allen et al. in preparation).

Application of the criteria reduces the sample by approximately a half to $\approx 44500$ spectra. Cross-correlation redshifts are then calculated for 4071 spectra with $0.8 < z \leq 1.0$ according to the prescription of Section 4.2. All spectra with $c_{cc_{\max}} \geq 0.2$ and redshifts, $0.8 < z \leq 1.0$, are combined to produce a composite. Then, the original and new composites are combined by taking the average, weighted by the relative number of spectra contributing at each wavelength. The effect is to determine redshifts for quasars using only the wavelength range where the initial (lower redshift) composite is of a high S/N. The procedure is then repeated for intervals of $\Delta z = 0.2$ up to redshift $z = 1.6$. Table 1 summarizes the number of spectra, median absolute magnitudes and wavelength coverage for all of the composites used to generate the final master template spectrum.

The key elements of the scheme are the use of wavelength regions of $\approx 1975$ Å for the calculation of redshifts of quasars up to $z = 1.6$, thereby excluding the C $\scriptstyle{\beta}$ ii and C $\scriptstyle{\beta}$ iv emission lines. The rest-frame ultraviolet region of interest is shown in Fig. 4.

4.4 Luminosity-dependent emission-line shifts

Quasar luminosity-dependent systematic effects related to the rest-frame locations of Ca $\scriptstyle{\beta}$ ii absorption, [O ii], [O iii] and Mg $\scriptstyle{\beta}$ emission are at or below the level of $30 \text{ km s}^{-1}$ (Section 6). However, the same is not true when considering the location of the Mg $\scriptstyle{\beta}$ emission line and the next prominent emission-line complex of C $\scriptstyle{\beta}$ ii $\lambda 1908$, Si $\scriptstyle{\beta}$ ii $\lambda 1892$ and Al $\scriptstyle{\beta}$ i $\lambda 1857$ as one moves further into the ultraviolet. Fig. 5 shows the ratio of the observed-frame centroids for Mg $\scriptstyle{\beta}$ to Mg $\scriptstyle{\beta}$ for quasars in the redshift interval $1.1 < z < 2.2$, where both lines are present in the SDSS spectra. The systematic trend of $\approx 2 \times 10^{-3}$ in the wavelength ratio as a function of quasar luminosity translates directly into a systematic in $\Delta z/(1+z)$ where $z$.

The line centroids are generated as part of the SDSS spectro1d pipeline.
Redshift determinations, in decreasing order of accuracy and increasing quasar redshift, can be obtained using [O III] emission lines, [O II] emission lines, cross-correlation including the Mg II emission line ([Mg II]_cc), cross-correlation including the C IV emission-line complex (C IV]_cc) and, finally, cross-correlation including the C IV emission line (C IV]_cc). For the cross-correlation results, empirical comparisons of redshifts derived using different rest-frame wavelength regions allow conversion relations to be derived as a function of quasar absolute magnitude and redshift. The goal is to build a redshift ‘ladder’ for quasars of increasing redshift that allows the redshift estimates to be placed on the same underlying systemic reference system. The subsections below consider in turn each step in the ladder.

5.1 [O II] and [O III] narrow emission-line redshifts

Redshifts for 13 291 quasars with redshifts \( z < 0.84 \) are available via the detection of [O II] \( \lambda \lambda 4960, 5008 \) emission with an S/N \( \geq 6.0\sigma \) (Section 4.1). Similarly, detection of the [O II] \( \lambda \lambda 3727, 3729 \) emission doublet at an S/N > 6.0\( \sigma \) provides redshift determinations for an additional 3844 quasars, with redshifts \( z < 1.31 \).

In the case of [O II] detections, a single Gaussian, centred at 3728.60 Å, is used. The [O II] doublet consists of two components centred at 3727.09 and 3729.88 Å, respectively. The observed component ratio varies from quasar to quasar but is normally in the range of 0.8:1–0.9:1, leading to the effective wavelength of 3728.60 Å adopted.

5.2 Cross-correlation redshifts including Mg II

Mg II]_cc redshifts are available for a further 12 289 quasars with redshifts \( z \leq 1.10 \). The minimum rest-frame wavelength involved in the cross-correlation is 1975 Å and systematic offsets relative to the emission-line redshifts generated in the previous subsection are not predicted or detectable.

In the interval \( 1.1 < z < 2.1 \), the Mg II]_cc redshifts involve rest-frame wavelengths below 1800 Å and the signal is increasingly dominated by the C III] emission as redshift increases and the Mg II emission line shifts into the far red of the SDSS spectra. Additionally, there is also the luminosity-dependent variation in the location of the C III] emission complex to take into account. The amplitude of the systematic redshift bias is small, only \( \lesssim 2 \times 10^{-4} \) too large at the highest redshift \( z = 2.1 \).

\[1^{11}\]

Absolute magnitudes, \( M_i \), are calculated using the prescription of Schneider et al. (2007).
Taking the Mg\textsubscript{II}\_cc redshifts and the corresponding Mg\textsubscript{II} emission-line centroid determinations from the SDSS pipeline allows the dependence of the cross-correlation bias on redshift and absolute magnitude to be quantified. Treating the systematic as separable in redshift and luminosity, a sample of \( \approx 42,000 \) quasars, in the redshift interval \( 1.1 < z < 2.1 \), shows that linear corrections to the raw cross-correlation redshifts, with slopes of \( 1.61 \times 10^{-4} \) per unit redshift and \( 7.2 \times 10^{-5} \) per unit magnitude (reducing the raw redshift estimates as redshift and luminosity increase), bring any residual systematic trends down to the \( \approx 1 \times 10^{-5} \) level (Fig. 6). Note that the sense and amplitude of the difference between the raw Mg\textsubscript{II}\_cc redshifts and the Mg\textsubscript{II} emission-line centroids are entirely consistent with the existence of the luminosity-dependent emission-line shifts and the way that the master-quasar template spectrum is constructed (Section 4.4). Corrected Mg\textsubscript{II}\_cc redshifts are available for 43,728 quasars in the interval \( 1.1 < z < 2.1 \).

5.3 Cross-correlation redshifts including C\textsubscript{III}] At redshifts \( z > 2.1 \), the Mg\textsubscript{II} line no longer contributes to the cross-correlation redshifts and the full effect of the systematic variation in the rest-frame locations of the Mg\textsubscript{II} and C\textsubscript{III]} emission lines must be taken into account. Fortunately, an empirical determination of the systematic differences between the corrected, unbiased, redshifts derived above and cross-correlation redshifts using only the rest-wavelength region \( 1675 < \lambda < 2650 \) Å, termed C\textsubscript{III]}\_cc redshifts, is straightforward to make. The differences between the corrected Mg\textsubscript{II}\_cc redshifts and raw C\textsubscript{III]}\_cc redshifts, derived using a maximum rest-frame wavelength of \( \lambda = 2650 \) Å, i.e. excluding the Mg\textsubscript{II} emission line, are available for \( \approx 35,000 \) quasars with \( 1.1 < z < 2.0 \). The difference as a function of quasar absolute magnitude is systematic and well represented by a linear trend; a linear fit has a slope of \( 1.67 \times 10^{-4} \), increasing the raw redshifts for increasing bright quasars. Application of the correction removes any detectable systematic effects as a function of absolute magnitude (Fig. 7) or redshift (Fig. 8). The same correction is then applied to C\textsubscript{III]}\_cc redshifts for 13,859 quasars with \( z > 2.1 \).

5.4 Cross-correlation redshifts including C\textsubscript{IV} The intention throughout is to avoid the use of the rest-frame wavelength region including the C\textsubscript{IV} emission line, which is known to show large asymmetric variations in shape, and hence of the line centroid. However, for 3274 quasars in the redshift interval \( 2.1 < z < 4.5 \), the cross-correlation signal from the rest-frame \( \lambda > 1675 \) Å region is too low to produce a reliable C\textsubscript{III]}\_cc redshift. For these 3274 quasars, a cross-correlation redshift determination employing the rest-frame wavelength interval \( \lambda > 1275 \) Å (C\textsubscript{IV}\_cc) is made.

There is a strong luminosity-dependent bias present due to the systematic variation in the location of the rest-frame C\textsubscript{IV} emission-line centroid. The situation is complicated by the presence, in a large number of quasars, of significant absorption bluewards of the C\textsubscript{IV} emission centroid, which biases the line centroid to the...
To decouple the effects on the C IV emission line of quasar luminosity and the presence of absorption, a sub-sample of ∼25 000 quasars with essentially undetectable absorption bluewards of the C IV emission-line centroid is defined.12

An empirical correction for the systematic luminosity-dependent redshift bias (Fig. 9) can then be made in exactly the same way that the C III]_cc redshifts were referenced to the unbiased system. A two-part linear fit to the absolute magnitude dependence, with a slope of 6.67 × 10⁻⁴ for Mₜ < −27.0 and a slope of 3.90 × 10⁻⁴ for Mₜ ≥ −27.0, provides an excellent fit to the systematic trend. The raw C IV_cc redshift estimates are increased for more luminous quasars, reducing systematic redshift differences to undetectable levels. The amplitude of the systematic bias in the raw C IV_cc redshifts is large, a factor of 4 greater than the C III] dependence, and the correction results in a substantial reduction in the dispersion between the C IV_cc and unbiased redshift determinations.

Objects with significant absorption bluewards of the C IV emission line, many of which are BAL quasars, show an extended tail of redshift deviations to high values. A second systematic correction, as a function of the absorption equivalent width (AEW), is then made, using the differences between the corrected C III]_cc redshifts and the absolute-magnitude-corrected C IV_cc values for ≥38 000 quasars. The actual correction applied is based on the empirically determined median Δz/(1 + z) versus AEW relation, but the amplitude of the well-determined correction is closely reproduced by a linear fit with a slope of −2.5 × 10⁻⁵ over the 0–200 range of AEW used. The additional quasar-to-quasar dispersion in redshift at large AEW is significant, but only 574 quasars with AEW > 20 in the final catalogue possess C IV_cc redshifts.

Table 2 summarizes the redshift and absolute-magnitude-dependent corrections made to the redshifts from the different estimation schemes in the ladder.

![Figure 9](https://academic.oup.com/mnras/article/405/4/2302/1045471)

**Figure 9.** Redshift differences, Δz/(1 + z), between the corrected HW cross-correlation redshifts and uncorrected C IV_cc redshifts derived including the C iv emission line, as a function of absolute magnitude. The solid red line is a 2001-point running median of the data points. Data for 34 956 spectra are included (119 lie outside the y-axis range plotted).

5.5 Additional redshifts

Redshifts for an additional 124 quasars are available, although one of the strict criteria described above is not satisfied, e.g. S/N < 6.0σ for an emission-line detection or cc max < 0.2. These redshifts are included in the catalogue but are highlighted by the inclusion of a special status flag.

Finally, the emission-line detection and cross-correlation schemes fail to provide reliable redshift estimates for 1256 quasars. These objects consist primarily of a mix of pathological spectra, including extreme BAL quasars, and spectra of a very low S/N. The objects are included in the redshift catalogue for completeness, with redshifts and redshift errors taken from the Schneider et al. catalogues (Schneider et al. 2007, 2010) and the SPECTROD pipeline for DR6. Again, the source of the redshifts is indicated in the catalogue via a status flag.

6 SYSTEMIC REDSHIFTS AND REDSHIFT UNCERTAINTIES

Section 5 describes the scheme adopted to obtain redshifts using the most reliable estimation procedure for each quasar. Redshifts are referenced to the zero-point provided by the location of the [O III] λλ4960, 5008 emission lines. The goal is to reduce systematic errors in Δz/(1+z) to the level of ≤1 × 10⁻⁴ (30 km s⁻¹) per unit redshift and ≤2.5 × 10⁻⁵ (8 km s⁻¹) per unit absolute magnitude. In this section, the question of referencing the [O III] emission-line redshifts to the systemic system defined by the quasar host galaxies is considered. Starting with the comparison of absorption line and [O III] emission redshifts, the uncertainties in redshift estimates arising from both the intrinsic quasar-to-quasar variation and the reproducibility of the determinations for each technique in the ladder are quantified.

6.1 [O III] and host galaxy systemic redshifts

Redshift estimates based on the detection of photospheric absorption from stars in the spatially averaged spectrum of the quasar host galaxy might be expected to provide a close to ‘ideal’ systemic redshift. Given the nature of the SDSS spectra, coupled with the large luminosity of the quasars at rest-frame optical and near-ultraviolet wavelengths, detection of photospheric absorption is not possible for the majority of objects. However, a direct comparison of redshifts derived from photospheric Ca II λλ3934.8, 3969.6 and [O III] emission is possible for a sample of objects with redshifts z < 0.4. Generating a catalogue of Ca II absorption detections with S/N > 6σ, matched to quasars with [O III] detections in the redshift interval 0.2 ≤ z ≤ 0.4, produces 825 quasars. Restricting the sample to objects that satisfy the luminosity criterion for inclusion in the Schneider et al. SDSS quasar compilations results in 615 quasars with absolute magnitudes covering the full range −24.0 < Mₜ < −22.0. Composite spectra with median Mₜ = −22.2 and −22.6 possess both Ca II absorption and [O III] emission at a high S/N.

Measuring the centroid of the strong Ca II K line at 3934.8 Å (Ca II H is blended with He absorption, producing a shift to longer wavelengths) and the centroid of the [O III] λ4960.30, 5008.24 emission, measured above the 50 per cent peak-height level, shows that the [O III] emission is shifted by 45 ± 5 km s⁻¹ to the blue, with no detectable dependence on luminosity. The offset determined from the composite spectra is in good agreement with the results of Boroson (2005) and the distribution of individual Ca II and [O III] redshift

---

12The absorption strength is parametrized using an integrated absorption equivalent width (AEW), calculated over the velocity range from −29 000 to 0 km s⁻¹, relative to the predicted C iv λ1549 location.
Table 2. Systematic redshift corrections.

| Redshift method | Number | Redshift interval | Redshift correction (|Δz/(1+z)| per unit z) | M_i correction (|Δz/(1+z)| per unit M_i) |
|-----------------|--------|------------------|-----------------|-----------------|
| [O III]         | 13291  | <0.84            |                  |                 |
| [O II]          | 3844   | <1.35            |                  |                 |
| Mg ii_cc        | 12289  | <1.1             |                  |                 |
| Mg II_cc        | 43728  | 1.1–2.1          | 1.61×10^{-4}     | 7.2×10^{-5}     |
| C iii_cc        | 13859  | 1.1–4.1          |                  | 1.67×10^{-4}    |
| C iv_cc         | 3274   | 1.5–5.5          |                  | 6.67×10^{-4} ≤ −27.0 |
| C iv_cc         |        |                  |                  | 3.90×10^{-4} ≥ −27.0 |

6.2 [O III] and [O II] emission-line redshifts

A similar comparison can be made between the [O III] and [O II] emission-line-derived redshifts for more than 7500 quasars, redshifts z < 0.8, with both [O III] and [O II] emission-line detections. A small systematic velocity offset is present, with the [O III]-derived redshifts 24 ± 5 km s^{-1} redwards of the [O II]-derived redshifts. Thus, relative to the systemic reference defined by the Ca II K absorption, the offsets are 21 ± 5 km s^{-1} bluewards for [O II] and 45 ± 5 km s^{-1} bluewards for [O III], in excellent agreement with previous work (e.g. Boroson 2005).

Comparison of 1378 (1103) spectrum pairs results in median errors of 5 × 10^{-5} and 1.4 × 10^{-4} in the reproducibility of Δz/(1+z) for [O III] and [O II], respectively. The smaller error for [O II] results from the typically higher S/N of the emission compared to the [O III] line.

Systematic trends from linear fits to the redshift differences between [O III] and [O II] redshifts, are Δz/(1+z) = 2.1 × 10^{-5} per magnitude and Δz/(1+z) = 5.9 × 10^{-5} per unit redshift.

The empirically determined quasar-to-quasar rms scatter of the intrinsically derived redshifts about the Ca II K line is σ_{intrinsic} = 3.5 × 10^{-4}.

6.3 [O III] and Mg II cc redshifts

[O III] emission-line redshifts compared to Mg II cc redshifts, calculated excluding the [O III] emission lines from the quasar template, for ≃12 500 quasars with [O III] emission-line redshifts show an undetectable offset, median |Δz/(1+z)| < 10^{-5}. Systematic trends, from linear fits to the redshift differences, in sense [O III]–Mg II cc redshifts, are Δz/(1+z) = +1 × 10^{-5} per magnitude and Δz/(1+z) = −1.1 × 10^{-4} per unit redshift. Both luminosity and redshift parametrizations lead to systematics of at most ±3 × 10^{-5} (10 km s^{-1}), for the dynamic ranges present in the sample, more than an order of magnitude below the uncertainties in the individual [O III] redshifts (Section 6.1). The sense and amplitude of the small systematic are consistent with the results from the [O III] to [O II] emission-line comparison (Section 6.2) and are again almost certainly due to the increasing degree of blue asymmetry present in the [O III] emission lines at increasing quasar luminosity. The comparison shows that the Mg II cc redshifts are tied to the reference [O III] emission-line redshifts to very high accuracy and that any systematics present are at most 10^{-4} in Δz/(1+z) or 30 km s^{-1} in velocity.

The empirically determined quasar-to-quasar scatter of the individual Mg II cc redshifts about the [O III] redshifts is σ_{intrinsic} = 2.5 × 10^{-4} or 75 km s^{-1}, which represents an improvement of a factor of ~3 compared to careful determinations of the Mg II emission-line location (e.g. Nestor et al. 2008).

6.4 Mg II cc and C III] cc redshifts

The luminosity-dependent correction to bring the Mg II cc and C iii] cc redshifts into coincidence is highly successful, as evidenced by Figs 7 and 8.

After allowing for the uncertainty in the determination of the Mg II cc and C iii] cc redshifts due to the limited S/N of the SDSS spectra, there is no evidence for an additional quasar-to-quasar redshift scatter associated with the use of the C iii] emission-line region alone.

6.5 C III] cc and C IV cc redshifts

The diversity of the form of the C IV emission line has been known through many studies going back decades. The removal of the systematic quasar luminosity-dependent behaviour, amounting to ≃650 km s^{-1} (over 4 mag in quasar absolute magnitude), improves C IV cc redshifts considerably. However, even for quasars with small AEW values the empirically determined quasar-to-quasar scatter of the individual C IV cc redshifts about the C iii] cc redshifts is σ_{intrinsic} = 8.0 × 10^{-4}. The situation is much worse for quasars with significant absorption bluewards of the C IV emission-line centroid. The systematic correction applied reaches a full 5 × 10^{-3} for
Figure 10. Redshift differences, $\Delta z/(1 + z)$, between the corrected C\textsc{ii},\textsubscript{cc} redshifts and corrected C\textsc{iv},\textsubscript{cc} redshifts derived using the C\textsc{iv} emission-line complex alone, as a function of redshift. Data for 25 008 quasars with AEW $\leq 20$ are included (60 lie outside the y-axis range plotted). The solid red line is a 2001-point running median of the data points. The offset between the two redshift estimators is minimal except for the systematic ‘dip’ centred at $z \approx 2.8$, coincident with the significant drop in detection efficiency for non-BAL quasars in the SDSS.

6.6 Summary

All redshift estimates, except those taken directly from the SDSS, have been increased by 45 km s$^{-1}$ (Section 6.1) to bring the [O\textsc{iii}] emission-line-based estimates on to the systemic system defined by the Ca\textsc{ii} K absorption.

Based on the large sample of repeat spectra, the internal reproducibility of the new cross-correlation redshifts represents an improvement of more than a factor of 2 over the SDSS redshift values. The reproducibility of the new redshifts is indistinguishable from that for the Princeton redshift values up to redshift $z \approx 1.6$. At higher redshifts the Princeton algorithm employs more information, via inclusion of the C\textsc{iv} emission-line region at $\lambda < 1675$ Å, and results in significantly better reproducibility for redshifts $z > 2.0$ than in the scheme presented here. However, the large quasar-to-quasar variation contributing to the redshift uncertainty (Table 3) means that differences in the internal reproducibility are not a critical factor for the redshifts in the final catalogue.

Table 3 summarizes the uncertainties for the different redshift estimates. Accurate redshift reproducibility estimates are available for all quasars, based on an emission-line S/N or $cc_{\text{max}}$ value, and are incorporated in the redshift errors included in Table 4. To provide an indication of the relative contributions of the internal and quasar-to-quasar uncertainties, Column 4 of Table 3 lists the median internal error for each redshift estimate. Column 5 gives the quasar-to-quasar error. The quasar-to-quasar errors, working from the chosen Ca\textsc{ii} absorption reference, have been added in quadrature to produce the cumulative quasar-to-quasar uncertainties in Column 7. Figure 11 shows the final redshift differences over the full redshift range, 0.05 $\leq z \leq 4.5$. The large amplitudes of the systematic trends with redshift and the differences exhibited by individual quasars are evident.

7 QUASARS WITH DETECTIONS IN FIRST

The procedures described in Sections 5 and 6 reduce systematic redshift errors as a function of redshift and absolute magnitude by more than an order of magnitude compared to the publicly available SDSS redshifts. However, a further significant reduction in the remaining relatively large quasar-to-quasar uncertainties will require a detailed investigation of the SED-dependent changes in the properties of the most prominent emission lines (which dominate the redshift determinations). Such an investigation is beyond the scope of this paper, but it is relatively straightforward to consider systematic redshift differences that correlate with the detection of SDSS quasars in FIRST (Becker et al. 1995).

For redshifts $z < 1.1$, the systematic differences between the populations of FD and not-FD (nFD) quasars with Mg\textsc{ii},\textsubscript{cc}
P. C. Hewett and V. Wild

Table 4. Quasar redshift catalogue. The full table is available in the electronic version of the journal – see Supporting Information.

| Name              | RA J2000 (deg) | Dec. J2000 (deg) | z   | σ_z | FIRST | z Alternate | Plate | MJD | Fibre | Code |
|-------------------|---------------|-----------------|-----|-----|-------|-------------|-------|-----|-------|------|
| SDSS J0000006.53+003055.2 | 0.02723       | 0.51534         | 1.823154 | 0.001025 | -1   | -999.0 | 685   | 52203 | 467   | 3    |
| SDSS J000008.13+001634.6 | 0.03390       | 0.27630         | 1.836332 | 0.000614 | -1   | -999.0 | 685   | 52203 | 470   | 3    |
| SDSS J000009.26+151754.3 | 0.03860       | 15.29848        | 1.197436 | 0.000369 | 0    | 1.20035 | 751   | 52251 | 354   | 2    |
| SDSS J000009.38+135618.4 | 0.03909       | 13.93845        | 2.240486 | 0.001468 | 0    | 2.240486 | 750   | 52235 | 82    | 7    |
| SDSS J000010.92–102751.9 | 0.03927      | -10.46443       | 1.851731 | 0.000966 | -1   | -999.0 | 650   | 52143 | 199   | 3    |
| SDSS J000011.41+145545.6 | 0.04755       | 14.92935        | 0.460127 | 0.000357 | 0    | -999.0  | 750   | 52235 | 499   | 1    |
| SDSS J000011.96+000225.3 | 0.04984       | 0.04036         | 0.478321 | 0.000358 | -1   | -999.0  | 387   | 51791 | 200   | 1    |
| SDSS J000012.25–003220.5 | 0.05108       | -0.53905        | 1.437047 | 0.000697 | -1   | -999.0  | 1091  | 52902 | 129   | 3    |
| SDSS J000013.14+141034.6 | 0.05479       | 14.17630        | 0.949947 | 0.000521 | 0    | -999.0  | 750   | 52235 | 98    | 3    |
| SDSS J000013.80–005446.8 | 0.05751       | -0.91300        | 1.840606 | 0.000789 | -1   | -999.0  | 1091  | 52902 | 108   | 3    |

Figure 11. Redshift differences, Δz/(1+z), between the final HW redshifts and SDSS redshifts, as a function of redshift. Note the large range on the y-axis. The solid red line, calculated using a 2001-point running median of the data points, shows the form of the systematic trends with redshift. Data for 90409 quasars are included (1355 lie outside the y-axis range plotted).

cross-correlation redshifts are at the Δz/(1+z) ≤ 10⁻⁴ level. However, for redshifts z > 1.1, once the C iv+[Si iii]+Al iii emission-line complex contributes to the redshift determination, systematic differences become increasingly evident, reaching an amplitude of nearly Δz/(1+z) = 2 × 10⁻³ (600 km s⁻¹) at redshifts z ≥ 4.

The origin of the redshift differences is primarily a systematic change in the ratio of the C iv] to Si iii] emission lines in the FD and nFD populations. The line ratio change results in a shift in the centroid of the blended line; Si iii] is weaker in the FD-detected quasars and the blended line centroid moves redwards. The Mg ii_cc redshifts for the FD quasars are thus too large.

Based on the prescription of Schneider et al. (2007) for matching SDSS quasars to the FIRST survey, there are 4326 FD quasars with z > 1.1 in the DR6 quasar catalogue. The small fraction (7 per cent) of FD quasars, combined with the low amplitude of the systematic redshift differences, means that the inclusion of FD-detected quasars in the construction of the master template results in changes to cross-correlation redshifts of Δz/(1+z) < 10⁻⁴. However, generation of an individual template for the FD quasars results in a quasar template with significant differences in the form of the C iv]+Si iii]+Al iii emission-line complex (Fig. 12).

Construction of the new FD-quasar template proceeds in an identical fashion to that described in Section 4, but the individual quasars used differ. For redshifts z < 1.0 the same quasars used to generate the master-quasar template are employed, whereas at redshifts z > 1.0, only FD quasars are used, with the minimum number of spectra required to generate a composite in a redshift slice reduced from 100 to 50 (Section 4). The number of quasars contributing at each redshift interval is given in Column 4 of Table 1. As evident from Fig. 12, the form of the C iv]+Si iii]+Al iii emission-line complex differs between the master- and FD-quasar templates. The size of the empirical transformations necessary to bring the FD-quasar redshift estimates on to the reference system, in which the Mg ii emission-line centroid does not vary, with either redshift or absolute magnitude, is significantly reduced compared to the quasar population as a whole.

For redshifts 1.1 ≤ z < 2.1 a reduction of Δz/(1+z) = 1.42×10⁻⁴ in the Mg ii_cc redshifts, independent of redshift and absolute magnitude, is necessary. For C iv]+Si iii]+Al iii redshifts an absolute-magnitude-dependent correction of 4.19×10⁻⁵, in the opposite sense to that for the master template, is required, i.e. for bright quasars the C iv]+Si iii]+Al iii redshifts are reduced. However, given the ≈4 mag dynamic range of the quasar sample, the maximum correction for any quasar is Δz/(1+z) ≤ 10⁻³. For the small number of quasars where it is necessary to employ the C iv emission line, the C iv+cc redshifts require a correction with a slope of 3.4×10⁻⁴, in the same sense as the larger correction necessary for the master template, i.e. for bright quasars the C iv+cc redshifts are increased (Fig. 9).
8 THE REDSHIFT CATALOGUE

Table 4 includes redshifts and error estimates for 91,665 quasars. Column 1 gives the SDSS coordinate object name, taken from the SDSS DR7 Legacy Release whenever available. Columns 2 and 3 give the object J2000 right ascension and declination in decimal degrees. The redshift and redshift error are given in Columns 4 and 5, respectively. Column 6 provides a code specifying the FIRST-detection (FD) status of the quasar (−1: not detected; 0: outside FIRST footprint; 1: detected). Column 7 gives the alternate redshift for quasars with detection codes = 0 (derived using the FD-quasar template) and =1 (derived using the master-quasar template) of Column 6. The alternate redshift is assigned a value of ‘−999.0’ for quasars with detection code = −1 of Column 6. Column 8 specifies the origin of the redshift estimate via a numerical code (1: [O iii], 2: [O ii], 3: Mg ii_cc, 4: C iii]_cc, 5: C iv_cc, 6: extra_cc, 7: SDSS). The SDSS spectrum from which the redshift estimate is derived is specified via the spectroscopic plate number, modified Julian Date of observation and fibre number in Columns 9–11, respectively. The redshifts are given to six decimal places but, as evident from the size of the associated errors, the accuracy for individual objects is two orders of magnitude larger. The high level of precision is retained to avoid quantization when comparing different redshift estimates specified to only four decimal places.

The provision of alternate redshifts for FD quasars and quasars whose FIRST detection status is unclear allows the use of an appropriate redshift by researchers with particular definitions of ‘radio’-quasar sub-samples and/or additional radio observations for quasars outside the current FIRST footprint. The primary (Column 4) and alternate (Column 7) redshifts differ only when the primary redshift is derived from cross-correlation (with one of the two quasar templates) and has a value z > 1.1.

Two quasars, SDSS J134415.75+331719.1 and SDSS J142507.32+323137.4, exhibit distinctive double-peak narrow emission. In both cases, the redshift corresponding to the higher velocity system is included in the table.

The majority of researchers will be interested in the combined redshift error (Column 5) arising from the limited S/N of the SDSS spectra and intrinsic variation from quasar to quasar. However, the internal contribution can be recovered straightforwardly via use of the amplitude of the quasar-to-quasar errors listed in Table 3.

Table 5 presents the master-quasar templates used to estimate the cross-correlation redshifts. Column 1 lists the rest-frame wavelength (Å). Columns 2 and 3 include the relative flux (per unit wavelength) and the number of spectra contributing for the master template, respectively, while Columns 4 and 5 provide the same information for the FD-quasar template. The FD-quasar template does not extend quite as far to the blue and the flux column contains entries of ‘−999.0’ for wavelengths λ < 1296.2 Å.

While the spectra should prove of use in the context of redshift estimation, the templates are not suitable for studies of quasar SEDs, where care must be taken in defining the large-scale shape of such composite spectra.

9 DISCUSSION

The approach adopted in this paper to the question of deriving redshifts with a common zero-point over an extended dynamic range in redshift, and hence involving disjoint spectral wavelength coverage, differs from that normally employed. The majority of studies to date have focused on the parametrization of the rest-frame centroid differences between the strongest emission lines present in the quasar spectra (see e.g. appendix A of Shen et al. 2007, for a recent example). Use of the cross-correlation redshifts directly bypasses many of the difficulties associated in providing reliable, reproducible, parametrizations of low S/N, asymmetric, often blended, emission lines present on ‘continua’ that also show significant variation from quasar to quasar. The resultant quasar-to-quasar dispersion and the internal reproducibility of the new HW redshifts represent significant improvements over even the most careful studies utilizing individual emission features.

Systematic, luminosity-dependent relative emission-line shifts have not featured in many previous studies of the quasar population. In part, the lack of such work may reflect the difficulty of performing such studies prior to the availability of the more recent SDSS Data Releases. An exception is the work of Richards et al. (2002) who find a clear relationship between emission-line centroid shifts, of exactly the type discussed here, and emission-line equivalent width. Richards et al. (2002) note that the line equivalent width is directly related to quasar absolute magnitude via the Baldwin effect (Baldwin 1977).

9.1 Comparison with Princeton redshifts and the Vanden Berk et al. (2001) quasar template

Operationally, it is found that straightforward systematic corrections to quasar redshift estimates, as a function of quasar absolute magnitude, reduce the systematic trends as a function of redshift.
moves beyond the red limit of the SDSS spectra (at emission line above $\lambda\lambda$ and [O\textsc{iii}$]$ and [O\textsc{iv}$]$ at $z\lambda\lambda$ and 
$\lambda\lambda$ and [O\textsc{iii}$]$ or [O\textsc{iv}$]$.

2302–2316
600 km s$^{-1}$

Panel (a) shows the prominent emission line of the new quasar composite compared to that of Vanden Berk et al. (2001). The line centroid moves redwards as increasingly large fractions of the line wings are included but at the half peak-height level the HW-template line centroid is only 0.4 ± 0.1 Å (±45 km s$^{-1}$) redwards of the rest-frame reference value of 2798.75 Å, derived from the Mg\textsc{ii} components in the ratio of 2:1.

The strongest ‘jump’ in the relation between the HW and Princeton redshifts in Fig. 13, at $z \simeq 2.2$, derives fundamentally from the large systematic trends in the ratio of Mg\textsc{ii} to ‘C\textsc{iii}’ emission-line locations as a function of quasar absolute magnitude (Fig. 5). The origin of the effect is primarily the change in the ratio of C\textsc{iii}\textsc{λ}1908 to Si\textsc{iii}\textsc{λ}1892 (see Fig. 12 for illustration in the context of FD quasars). A direct comparison of the HW template and Vanden Berk et al. (2001) template is somewhat misleading because the HW redshifts are derived only following the significant absolute-magnitude-dependent corrections. However, using any sensible definition of the emission line, the C\textsc{iii}$\lambda$ Si\textsc{iii}$\lambda$ blend is significantly bluer in the HW template than in the Vanden Berk et al. (2001) composite, producing the increase in the HW redshifts at $z \simeq 2.2$. The Princeton redshifts then become progressively closer to the HW redshifts as the C\textsc{iv} emission line (with its well-established increasing blue asymmetry at increasing quasar luminosity) dominates the Princeton determinations at higher redshifts. Recall though that the

**Figure 13.** Redshift differences, $\Delta z/(1+z)$, between the final HW redshifts and Princeton redshifts, as a function of redshift. Note the large range on the y-axis. Data for 90 979 quasars are included (1306 lie outside the y-axis range plotted).

(absolute magnitude) to $<30$ km s$^{-1}$ per unit redshift ($<10$ km s$^{-1}$ per magnitude). Internal reproducibility represents a factor of >2 improvement over the SDSS redshift determinations. The results presented above, combined with the form of the differences between the Princeton and SDSS redshifts (Section 3), show that the origin of a significant proportion of the improvements achieved is due to differences in the cross-correlation procedure/algorithme employed. However, comparison of the new HW redshifts with the Princeton determinations (Fig. 13) still shows large ($\simeq 600$ km s$^{-1}$ at $z \simeq 2.2$) systematic differences.

The two evident ‘jumps’ in the relationship between the HW and SDSS estimates occur as cross-correlation redshifts including the Mg\textsc{ii}\textsc{λ}2799 emission line become important (at $z \simeq 0.8$) and where Mg\textsc{ii} moves beyond the red limit of the SDSS spectra (at $z \simeq 2.1$). The behaviour can be traced directly to differences in the new quasar template and that of Vanden Berk et al. (2001). Fig. 14(b) shows the excellent agreement between the composites at optical wavelengths where emission lines, including H$\beta$ and [O\textsc{iii}$]\lambda\lambda 4960, 5008$, dominate the redshift determinations (either directly, via emission-line locations or through the contribution of emission lines to the cross-correlation signal). In contrast, Fig. 14(a) illustrates the significant difference in the location of the Mg\textsc{ii}\textsc{λ}2799 emission line in the two composites. At the accuracy levels of interest, absolute wavelength ‘centroids’ of broad emission lines in quasar spectra are dominated by the particular scheme used to define the associated ‘continuum’ and the height above the continuum used to define the ‘line’. However, the centroid of the portion of the Mg\textsc{ii} emission line above half the peak height is 1.2 ± 0.1 Å bluer in the HW template compared to the Vanden Berk et al. (2001) template. The line centroid moves redwards as increasingly large fractions of the line wings are included but at the half peak-height level theHW-template line centroid is only 0.4 ± 0.1 Å (±45 km s$^{-1}$) redwards of the rest-frame reference value of 2798.75 Å, derived from the Mg\textsc{ii} components in the ratio of 2:1.

**Figure 14.** Rest-frame spectra of the new quasar composite compared to that of Vanden Berk et al. (2001). Panel (a) shows the prominent emission line of Mg\textsc{ii}\textsc{λ}2796.35, 2803.53, with the new quasar composite (solid line) and the Vanden Berk et al. (2001) composite (dashed line). The lower plot shows the ratio Vanden Berk et al. (2001)/new. The vertical dotted line indicates the wavelength 2798.75 Å, derived from the Mg\textsc{ii} components in the ratio of 2:1. The significantly bluer location of Mg\textsc{ii} emission in the new composite is evident. Panel (b) shows the same information for the rest-frame wavelength region including H$\beta$ and [O\textsc{iii}$]\lambda\lambda 4960, 5008$. The vertical dotted lines indicate the rest-frame wavelengths of 4862.68, 4960.30 and 5008.24 Å for H$\beta$ and [O\textsc{iii}$]$. While the new composite possesses slightly stronger [O\textsc{iii}$]$ emission, there is no evidence for any detectable offset in the emission-line locations of H$\beta$ or [O\textsc{iii}$]$.
Observed frequency distribution of redshift differences, $\beta = v/c$, for $z_{\text{abs}} > z_{\text{qso}}$. The centroid of the $\beta = 0$ component for the HW redshifts shows no detectable shift over the entire redshift range of the quasars, $1.55 < z < 3.5$.

The observed distributions are shown. No attempt has been made to calculate an absorber density by incorporating the redshift path accessible as a function of $\beta$.

9.2 Associated C IV and Mg II absorbers as quasar redshift diagnostics

The availability of the large SDSS quasar catalogues has stimulated new investigations into the physical origin of associated absorbers, particularly those evident through the presence of C IV and Mg II absorption (e.g. Nestor et al. 2008; Wild et al. 2008; Vanden Berk et al. 2008). A pre-requisite for such investigations is an estimate of the systemic quasar redshifts. Given the large intrinsic variation in the properties of the C IV emission line and the relative invariance of the Mg II emission-line centroid, redshifts based on the location of the Mg II emission line are often employed in studies of both associated C IV and Mg II absorbers in quasars with redshifts $z \leq 2.1$.

Both Mg II and C IV absorber catalogues are available from our investigation of absorber populations in SDSS quasars (e.g. Wild, Hewett & Pettini 2006). Strong narrow absorbers are flagged and ‘removed’ from the quasar spectra prior to the calculation of the cross-correlation redshift determinations (Section 4). The new HW-quasar redshifts are thus essentially independent of the presence of individual absorbers and a comparison of associated absorber velocity distributions, using both SDSS and HW redshifts, provides a powerful test of the redshift accuracy in an astrophysical context of considerable current interest. Fig. 15 shows the observed distribution of redshift differences, $\beta = v/c$, for $z_{\text{abs}} > z_{\text{qso}}$. The centroid of the $\beta = 0$ component for the HW redshifts shows no detectable shift over the entire redshift range of the quasars, $1.55 < z < 3.5$.

10 CONCLUSIONS

A systematic investigation of the relationship between different redshift estimation schemes for more than 91 000 quasars in the SDSS DR6 is presented. Empirical relationships between redshifts based on (i) Ca II H&K host galaxy absorption, (ii) quasar [O II] $\lambda\lambda 3728, 3729$, (iii) [O III] $\lambda\lambda 4960, 5008$ emission and (iv) cross-correlation (with a master-quasar template) that includes, at increasing quasar redshift, the prominent MgII $\lambda\lambda 2799, 2802$, CIII] $\lambda 1908$ and C IV $\lambda\lambda 1549$ emission lines are established as a function of quasar redshift and luminosity. New redshifts in the resulting catalogue possess systematic biases a factor of $\approx 20$ lower compared to the SDSS redshift values; systematic effects are reduced to the level of $\Delta z/(1+z) \lesssim 10^{-4}$ (30 km s$^{-1}$) per unit redshift or $\lesssim 2.5 \times 10^{-5}$ per unit absolute magnitude.

It is important to realize that there will be systematic redshift trends present as a function of the quasar SEDs and the specific example of FD quasars (Section 7) provides an example, related to the
radio properties of the quasar SEDs. One of the primary motivations of this work is to facilitate further studies of SED-dependent systematic emission-line properties, working from redshift estimates whose properties as a function of redshift and absolute magnitude are well understood.

Equally important as the new redshift determinations, well-determined empirical estimates of the quasar-to-quasar dispersion in redshifts are available for each method of redshift estimation and a combined internal+population uncertainty is provided for every quasar in the catalogue.

The improved redshifts and their associated errors have wide applicability in areas such as quasar absorption outflows, quasar clustering, quasar-galaxy clustering and proximity-effect determinations.

ACKNOWLEDGMENTS

We thank the referee, Don Schneider, for providing a very careful and constructive reading of the original draft. We are grateful to James Allen, Bob Carswell and Gordon Richards for encouragement, insights and helpful conversations. PCH acknowledges support from the STFC-funded Galaxy Formation and Evolution programme at the Institute of Astronomy. VW is supported by a Marie Curie Intra-European Fellowship.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The participating institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543

Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634

Adelman-McCarthy J. K. et al., 2008, ApJS, 175, 297

Bajtlik S., Duncan R. C., Ostriker J. P., 1988, ApJ, 327, 570

Baldwin J. A., 1977, ApJ, 214, 679

Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559

Boroson T., 2005, AJ, 130, 381

Croom S. M., Boyle B. J., Loaring N. S., Miller L., Outram P. J., Shanks T., Smith R. J., 2002, MNRAS, 335, 459

Gaskell C. M., 1982, ApJ, 263, 79

Gibson R. R. et al., 2009, ApJ, 692, 758

Heckman T. M., Miley G. K., van Breugel W. J. M., Butcher H. R., 1981, ApJ, 247, 403

Hewett P. C., Irwin M. J., Buncle P., Bridgeland M. T., Kibblewhite E. J., He X. T., Smith M. G., 1985, MNRAS, 213, 971

Kirkman D., Tytler D., 2008, MNRAS, 391, 1457

Nestor D., Hamann F., Hidalgo P. R., 2008, MNRAS, 386, 2055

Padmanabhan N., White M., Norberg P., Porciani C., 2009, MNRAS, 397, 1862

Richards G. T., Vanden Berk D. E., Reichard T. A., Hall P. B., Schneider D. P., SubbaRao M., Thakar A. R., York D. G., 2002, AJ, 124, 1

Schneider D. P. et al., 2007, AJ, 134, 102

Schneider D. P. et al., 2010, AJ, in press

Shen Y. et al., 2007, AJ, 133, 2222

Stoughton C. et al., 2002, AJ, 123, 485

Tonry J., Davis M., 1979, AJ, 84, 1511

Tytler D., Fan X.-M., 1992, ApJS, 79, 1

Tytler D. et al., 2009, MNRAS, 392, 1539

Vanden Berk D. E. et al., 2001, AJ, 122, 549

Vanden Berk D. E. et al., 2008, ApJ, 679, 239

Wild V., 2009, preprint (arXiv:0907.5221)

Wild V., Hewett P. C., 2005, MNRAS, 358, 1083

Wild V., Hewett P. C., Pettini M., 2006, MNRAS, 367, 211

Wild V. et al., 2008, MNRAS, 388, 227

York D. G. et al., 2000, AJ, 120, 1579

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table 4. Quasar redshift catalogue.

Table 5. Quasar cross-correlation templates.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.