The rest-frame optical colours of 99 000 Sloan Digital Sky Survey galaxies

V. Smolčić,[1,2,3]★ Ž. Ivezić,[1,4] M. Gaćeša,[2,5] K. Rakos,[6] K. Pavlovski,[2] S. Ilijić,[7]
M. Obrić,[4,8] R. H. Lupton,[1] D. Schlegel,[1] G. Kauffmann,[9] C. Tremonti,[10]
J. Brinchmann,[9] S. Charlot,[9] T. M. Heckman,[11] G. R. Knapp,[1] J. E. Gunn,[1]
J. Brinkmann,[12] I. Csabai,[13] M. Fukugita[14] and J. Loveday[15]

[1]*Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001, USA*
[2]*Physics Department, University of Zagreb, Bijenička cesta 32, 10000 Zagreb, Croatia*
[3]*Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg D-69117, Germany*
[4]*Department of Astronomy, University of Washington, PO Box 351580, Seattle, WA 98195-1580, USA*
[5]*Physics Department, University of Connecticut, 2152 Hillside Road, Storrs, CT 06269-3046, USA*
[6]*Institut für Astronomie, Universität Wien, Türkenschanstrasse 17, A-1180 Wien, Austria*
[7]*Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia*
[8]*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700AV Groningen, the Netherlands*
[9]*Max-Planck-Institute für Astrophysik, D-85748 Garching, Germany*
[10]*Hubble Fellow, University of Arizona, Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[11]*Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*
[12]*Apache Point Observatory, 2001 Apache Point Road, PO Box 59, Sunspot, NM 88349-0059, USA*
[13]*Department of Physics, Eötvös University, Postfach 32, Budapest H-1518, Hungary*
[14]*Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa, Kashiwa City, Chiba 277-8582, Japan*
[15]*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QJ*

ABSTRACT

We discuss the colours of 99 088 galaxies selected from the Sloan Digital Sky Survey (SDSS) Data Release 1 'main' spectroscopic sample (a flux-limited sample, $r_{Pet} < 17.77$, for 1360 deg^2) in the rest-frame Strömgren system (uz, vz, bz, yz). This narrow-band ($\sim$200 Å) photometric system, first designed for the determination of effective temperature, metallicity and gravity of stars, measures the continuum spectral slope of galaxies in the rest-frame 3200–5800 Å wavelength range. We synthesize rest-frame Strömgren magnitudes from SDSS spectra, and find that galaxies form a remarkably narrow locus ($\sim$0.03 mag) in the resulting colour–colour diagram. Using the Bruzual & Charlot population-synthesis models, we demonstrate that the position of a galaxy along this locus is controlled by metallicity and age of the dominant stellar population. The distribution of galaxies along the locus is bimodal, with the local minimum corresponding to an $\sim$1-Gyr-old single stellar population. The position of a galaxy perpendicular to the locus is independent of metallicity and age, and reflects the galaxy's dust content, as implied by both the models and the statistics of *IRAS* detections.

Comparison of the galaxy locus in the rest-frame Strömgren colour–colour diagram with the galaxy locus in the Hδ_A–D_n(4000) diagram, utilized by Kauffmann et al. to estimate stellar masses, reveals a tight correlation, although the two analysed spectral ranges barely overlap. Furthermore, the rest-frame $r - i$ colour (5500–8500 Å wavelength range) can be 'predicted' with an rms of 0.05 mag using the rest-frame Strömgren colours. *Overall, the galaxy spectral energy distribution in the entire ultraviolet to near-infrared range can be described as a single-parameter family with an accuracy of 0.1 mag, or better.* This nearly one-dimensional distribution of galaxies in the multidimensional space of measured parameters strongly supports the conclusion of Yip et al., based on a principal component analysis, that SDSS galaxy spectra can be described by a small number of eigenspectra. Furthermore, the rest-frame Strömgren colours correlate tightly with the classification scheme proposed by Yip et al. based on the first three eigenspectra. Apparently, the contributions of stellar populations that

★E-mail: smolcic@mpia.de

dominate the optical emission from galaxies are combined in a simple and well-defined way. We also find a remarkably tight correlation between the rest-frame Strömgren colours of emission-line galaxies and their position in the Baldwin–Phillips–Terlevich diagram. These correlations between colours and various spectroscopic diagnostic parameters support earlier suggestions that rest-frame Strömgren photometry offers an efficient tool to study faint cluster galaxies and low surface brightness objects without performing time-consuming spectral observations.

Key words: methods: statistical – surveys – galaxies: fundamental parameters – galaxies: Seyfert – galaxies: statistics.

1 INTRODUCTION

The studies of galaxies have been recently invigorated due to the advent of modern sensitive large-area surveys across a wide wavelength range. The Sloan Digital Sky Survey (SDSS, York et al. 2000; Stoughton et al. 2002; Abazajian et al. 2003) stands out among these surveys because it has already provided ultraviolet (UV) to near-infrared (near-IR) five-colour imaging data, and high-quality spectra for over 100 000 galaxies. The spectroscopic galaxy sample is defined by a simple flux limit (Strauss et al. 2002), and will eventually include close to 1 000 000 galaxies.

A number of detailed galaxy studies based on SDSS data have already been published. Strateva et al. (2001) and Shimasaku et al. (2001) demonstrated a tight correlation between the $u - r$ colour, concentration of the galaxy's light profile as measured by the SDSS photometric pipeline PHOTO (Lupton et al. 2002), and morphology. Blanton et al. (2001) presented the SDSS galaxy luminosity function, and Kauffmann et al. (2003a,b) determined and analysed stellar masses and star formation histories for 100 000 SDSS galaxies. A large number of other, published and ongoing, studies are based on the rich data set provided by the SDSS.

SDSS galaxy spectra have high quality ($R \sim 1800$, 3800–9200 Å wavelength range, spectrophotometric calibration better than 10 per cent; for more details, see Stoughton et al. 2002), and have been used in a number of detailed investigations, including studies of elliptical (Bernardi et al. 2003a,b,c,d; Eisenstein et al. 2003), star-forming (Hopkins et al. 2003), and active galaxies (Kauffmann et al. 2003c; Zakamska et al. 2003, 2004; Hao et al. 2005a,b). Yip et al. (2004) analysed the SDSS spectra of 170 000 galaxies using principal component analysis [Karhunen–Loéve (KL) transform]. They demonstrated that SDSS galaxy spectra can be described by a small number of eigenspectra: more than 99 per cent of the galaxies are found on a two-dimensional locus in the space spanned by the ratios of the first three eigencoefficients. Even the most extreme emission-line spectra can be described to within the measurement noise with only eight eigenspectra. An efficient *single-parameter* classification scheme is proposed, and it suggests three major galaxy classes. Another remarkable result of their analysis is a strong correlation between various spectral lines, encoded in, and readily discernible from the eigenspectra.

It is not yet known what the results of Yip et al. imply for the detailed distribution of galaxies in colour space, but it is obvious that a high degree of structure should exist. Also, a strong correlation is expected between colours and various spectroscopic parameters, such as the strength of the $H\delta_A$ line and the 4000-Å break, that were recently utilized by Kauffmann et al. (2003a) to estimate stellar masses and the dust content for SDSS galaxies.

The main goal of this paper is to investigate whether the intrinsic simplicity of galaxy spectra implied by the results of Yip et al. can be reproduced in rest-frame colour–colour diagrams, and whether the position of a galaxy in these diagrams is related to parameters discussed by Kauffmann et al. (2003a). We present an analysis of correlations between colours and various spectroscopic diagnostic parameters, including model-dependent estimates of stellar masses and dust content. We utilize the 'rest-frame' Strömgren photometric system that measures the continuum spectral slope of galaxies in the rest-frame 3200–5800 Å wavelength range. In principle, any set of bandpasses could be used to synthesize the colours. An additional motivation for synthesizing galaxy colours in the rest-frame Strömgren photometric system is a suggestion that it is an efficient tool for studying faint cluster galaxies and low surface brightness objects without performing time-consuming spectral observations [1] (Fiala, Rakos & Stockton 1986). The unprecedented number of high-quality SDSS spectra can be used to analyse in detail the advantages and limitations of that method, and relate galaxy colours in the Strömgren system to parameters commonly measured using SDSS data.

We describe the synthesis of rest-frame Strömgren magnitudes from SDSS spectra in Section 2. In Section 3, we analyse the distribution of galaxies in the resulting colour–colour diagram, and interpret it using the Bruzual & Charlot (1993, 2003) population-synthesis models. In Section 4, we demonstrate and study in detail correlations between rest-frame Strömgren colours and various diagnostic spectroscopic parameters. In Section 5, we correlate the rest-frame Strömgren colours with the spectral eigencoefficients. In Section 6, we study the properties of the galaxies at sufficiently large redshifts ($\gtrsim 0.18$) to have reliable synthesized uz magnitudes. We discuss and summarize our results in Section 7.

2 THE SYNTHESIS OF REST-FRAME STRÖMGREN PHOTOMETRY

2.1 The Strömgren photometric system

The Strömgren *uvby* narrow-band (~ 200 Å) filter system was designed for measuring the temperature, chemical composition and surface gravity of stars, without resorting to spectroscopy (Strömgren 1966; for a recent compilation of over 100 000 measurements for $\sim 63 000$ stars, see Hauck & Mermilliod 1998). The bandpasses bracket the 4000-Å break and cover three regions in the

---

[1] Observations in the rest-frame Strömgren photometric system use filters adjusted (in hardware) to the known cluster redshift (say, from the brightest cluster galaxy).
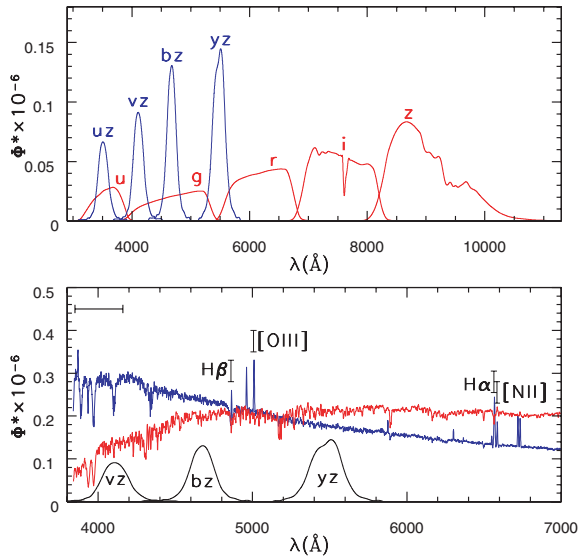
Figure 1. The top panel shows renormalized (see equation 3) filter transmission curves for the SDSS photometric system (*ugriz*), and for the Strömgren photometric system. The bottom panel emphasizes the 3800–7000 Å region. The two spectra are typical for blue and red galaxies, and the four labelled emission lines are used to separate star-forming from AGN galaxies. The horizontal bar in the top left-hand corner marks the wavelength region used in the analysis by Kauffmann et al. (2003).

3200–5800 Å spectral region, which makes them a powerful tool for investigating stellar populations in star clusters, or composite systems such as galaxies.

The filter system used here (*uz, vz, bz, yz*; see the next section for details) is a somewhat modified Strömgren system, such that the filters are slightly narrower, and the filter response curves are more symmetric, than in the original system (Odell, Schombert & Rakos 2002). The system characteristics are (see Fig. 1) as follows.

(i) The *uz* filter ($\lambda_{\mathrm{eff}} = 3500$ Å) is shortwards from the Balmer discontinuity (4000-Å break), and provides a measure of hot stars due to recent star formation.

(ii) The *vz* filter ($\lambda_{\mathrm{eff}} = 4100$ Å) extends bluewards from 4600 Å, but redwards from the Balmer discontinuity. This wavelength region is strongly influenced by metal absorption lines, particularly for spectral classes F and M which typically dominate the galaxy light. The *uz* − *vz* colour is a good measure of the 4000-Å break.

(iii) The *bz* ($\lambda_{\mathrm{eff}} = 4675$ Å) and *yz* ($\lambda_{\mathrm{eff}} = 5500$ Å) filters extend redwards from 4600 Å, where the influence of absorption lines is small. The *bz* − *yz* colour is thus a good measure of the temperature–colour index, practically free of metallicity and surface gravity effects. For old stellar populations, the *vz* − *yz* and *bz* − *yz* colour indices presumably serve as age and metallicity indicators.

2.2 The rest-frame Strömgren photometric system

The rest-frame Strömgren photometric system utilizes filters whose bandpasses are redshifted and stretched to correspond to the Strömgren bands in the rest frame. This method, introduced by Fiala et al. (1986), alleviates the need for *K*-corrections, and thus allows robust colour comparison for galaxies at different redshifts. For example, Rakos & Schombert (1995) studied the colour evolution of 17 clusters of galaxies spanning the redshift range from 0.2 to 0.9. They demonstrated that the colours of the red population are con-

sistent with passive evolution models for a single stellar population which formed around $z = 5$. On the other hand, the fraction of blue cluster galaxies dramatically increases from 20 per cent at $z = 0.4$ to 80 per cent at $z = 0.9$. The rest-frame Strömgren photometric system also provides a good method for detecting cluster members, and enables an efficient spectrophotometric classification of galaxies (Rakos, Dominis & Steindling 2001; Rakos & Schombert 2004, and references therein).

2.3 The synthetic Strömgren photometry from SDSS spectra

The synthetic Strömgren magnitudes are synthesized on the AB system[2] (Oke & Gunn 1983). An AB magnitude, *m*, in a bandpass $S(\nu)$, is defined by

$$m = -2.5 \log_{10} \frac{\int \mathrm{d}(\log_{10} \nu) f_\nu(\nu) S(\nu)}{\int \mathrm{d}(\log_{10} \nu) S(\nu)} - 48.6, \tag{1}$$

where f_ν is the specific flux per unit frequency (for more details see Fukugita et al. 1996). The zero-point (48.6) is chosen such that an object with a specific flux of 3631 Jy has $m = 0$ (i.e. an object with f_ν = constant has an AB magnitude equal to the Johnson *V* magnitude at all wavelengths). The function $S(\nu)$ includes the response of the detector, the transmissivity of the filter, the telescope, and the atmosphere of the Earth at some representative air-mass. It is convenient to express the above equation in terms of the specific flux per unit wavelength (the form used to report SDSS spectra, Stoughton et al. 2002) as

$$m = -2.5 \log_{10} \int \phi^*(\lambda) f_\lambda(\lambda) \, \mathrm{d}\lambda, \tag{2}$$

where we introduced a quantity, hereafter called the *normalized filter*:

$$\phi^*(\lambda) = \frac{10^{19.44} \lambda S(\lambda)}{c \int \lambda^{-1} S(\lambda) \, \mathrm{d}\lambda}. \tag{3}$$

Computing *m* is, therefore, reduced to a simple integral of the SDSS spectrum (with the spectrum given in 10^{-17} erg cm^{-2} s^{-1} Å^{-1}, and λ in Å) with a *precomputed* normalized filter. The normalized filter curves are shown in Fig. 1.

SDSS spectra are obtained using a plate with 640 spectroscopic fibres (diameter 3 arcsec, corresponding to ~6 kpc at the redshift of 0.1), which are fed into two independently calibrated spectrographs (Schlegel et al., in preparation). Spectra are corrected for the Galactic interstellar dust reddening during calibration. As shown by Vanden Berk et al. (2004), the accuracy of synthesized magnitudes for point sources can be improved by requiring that the mean offset between the imaging (*g, r, i*) and synthesized magnitudes is zero for each spectrograph (this information is not available when calibrating spectra). We confirmed their result for point sources, and found that it also improves the synthesized photometry for galaxies (as inferred from the width of the galaxy colour locus, discussed below). The synthesized photometry for galaxies can be further improved by using galaxies (instead of point sources) to correct for the mean offsets between the imaging (we use model magnitudes, see Stoughton et al. 2002 for details) and synthesized magnitudes. Hereafter, we adopt the latter method. When correcting synthesized Strömgren magnitudes, we linearly interpolate (as a function of wavelength) the *g*- and *r*-band corrections.

[2] Note that all the published work by Rakos and collaborators, who pioneered the redshifted Strömgren photometric system, is expressed on the Vega system.
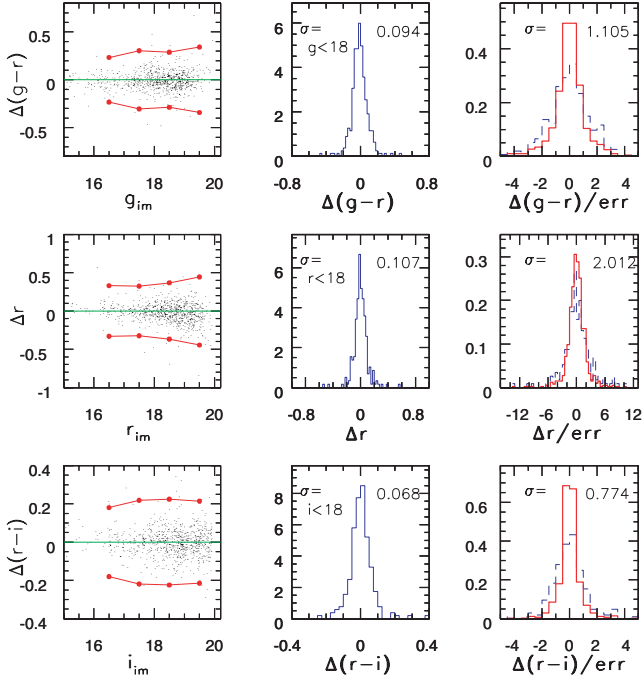
Figure 2. The comparison of imaging and synthesized SDSS magnitudes (colours) for stars, with the latter tied to imaging point spread function magnitudes. The first column displays the colour or magnitude difference as a function of magnitude; the large symbols connected by lines show the $\pm 3\sigma$ envelope. The distribution of colour or magnitude differences at the bright end is displayed in the middle column. The last column shows the distribution of colour or magnitude differences normalized by the expected errors at the bright end (solid line), and for the full sample (dashed line).

2.4 The tests of synthesized magnitudes

We synthesized SDSS magnitudes (g, r, i) and rest-frame Strömgren magnitudes (uz, vz, bz, yz) [3] for 99 088 unique 'main' ($r_{\rm Pet} < 17.77$) galaxies from SDSS Data Release 1 (Abazajian et al. 2003), as well as for stars from the same spectroscopic plates. The accuracy of synthesized magnitudes (relative to imaging magnitudes, which are accurate to 0.02 mag, Ivezić et al. 2003) for a subsample of stars is summarized in Fig. 2. The colours are reproduced with the expected accuracy ($\sim$0.08 mag), while there are systematic errors in the overall magnitude scale at the level of 0.10 mag (rms scatter). These results are in agreement with those of Vanden Berk et al. (2004).

Fig. 3 summarizes the obtained accuracy for galaxies. Again, the colours are reproduced with the expected accuracy, while systematic errors in the overall magnitude scale are $\sim$0.2 mag. Note that using galaxies (instead of point sources) to correct for the offset between imaging and synthesized magnitudes introduces a systematic trend in the Δr distribution at the faint-magnitude end. None the less, as evident from the top and bottom panels, this does not significantly affect the synthetic colours as the overall trend nearly cancels out. In addition, the distribution of colour-differences normalized by the expected errors (last column in Fig. 3, top and bottom panels) shows that the errors are estimated correctly (the width of the distributions is $\sim$1). Therefore, we do not expect this to affect the synthetic photometry significantly. The errors in rest-frame Strömgren colours are expected to be smaller than those in $g - r$ and $r - i$ colours

[3] We discuss the implications of the uz magnitude separately in Section 6 since it can only be synthesized for galaxies with redshifts beyond $\sim$0.18.
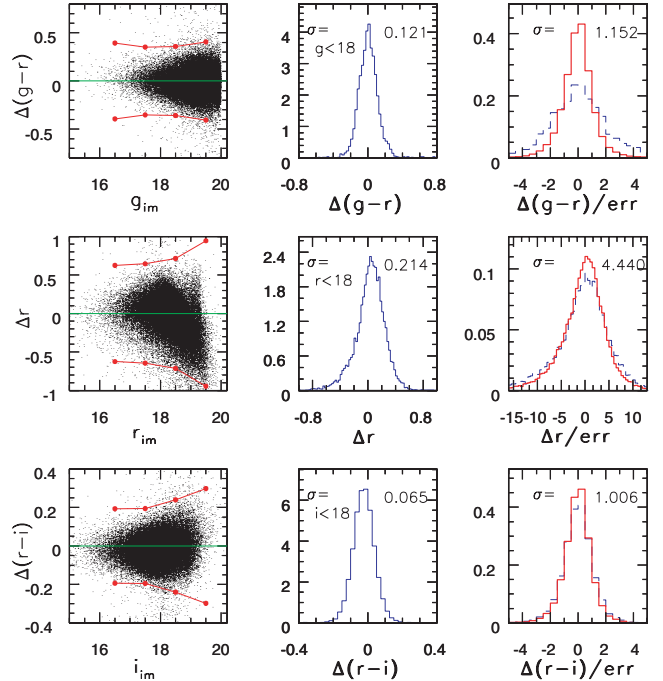
Figure 3. Analogous to Fig. 2, except that the results are shown for galaxies, and the spectrophotometric magnitudes are tied to imaging model magnitudes.

because the relevant wavelength range is shorter (and thus less sensitive to large-scale errors in the spectrophotometric calibration of SDSS spectra). We estimated errors in $vz - yz$ and $bz - yz$ colours by comparing measurements for 1253 galaxies observed twice: the scatter (rms determined from the interquartile range) for the $vz - yz$ colour is 0.09 mag, and for $bz - yz$ colour 0.05 mag. We tested for plate-to-plate systematics by analysing the position of the red peak in the colour distribution of galaxies (see the next section). This position shows a scatter of $\sim$0.02 mag for $vz - yz$ colour, and $\sim$0.01 mag for $bz - yz$ colour. We conclude that the synthesized colours are sufficiently accurate for a detailed investigation of rest-frame colours. This is the first time that rest-frame optical colours with an accuracy of <0.1 mag are available for a sample of $\sim$100 000 galaxies.

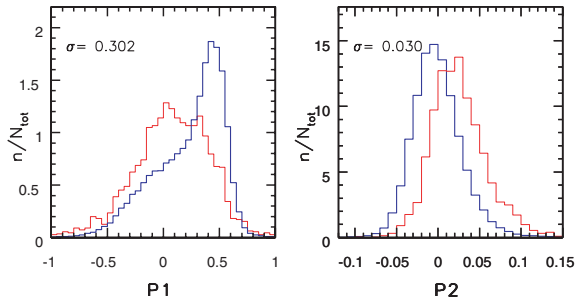# 3 THE COLOURS OF GALAXIES IN THE REST-FRAME STRÖMGREN SYSTEM

In this section, we analyse the distribution of galaxies in the rest-frame Strömgren colour–colour diagram, discuss possible systematic errors and interpret the colours using model spectra.

3.1 The galaxy locus and principal axes

The distribution of 99 088 galaxies in the rest-frame Strömgren colour–colour diagram is shown in Fig. 4. The distribution resembles a remarkably narrow locus. To quantify its width, we define a set of principal axes ($P1, P2$) where $P1$ measures the position along the locus, and $P2$ perpendicular to it. Since the locus is not perfectly straight, we separately fit the blue ($vz - yz < 0.6$) and red ($vz - yz > 0.7$) ends. As the boundary between the two ($P1, P2$) definitions, we choose $vz - yz = 0.646$, where the two fits intersect. We obtain
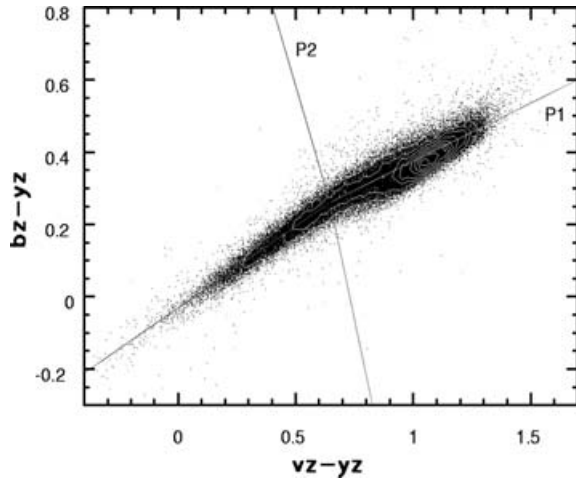
Figure 4. The distribution of 99 088 SDSS 'main' galaxies in the rest-frame Strömgren colour–colour diagram. Each galaxy is represented by a dot, and the overall distribution is outlined by linearly spaced isopleths. The principal axes *P1* and *P2* are defined for the two regions separated by $vz - yz = 0.646$.

Figure 5. The colour probability distributions of the principal rest-frame Strömgren colours for SDSS 'main' galaxies (note different scales for x-axis). The thick lines show the distributions for the whole sample, and the thin lines show the distributions for the subsample of galaxies detected by *IRAS*. The σ values shown in each panel are rms scatter (determined from the interquartile range) for the whole sample.

for the blue side

$$P1 = 0.911 \, (c_1 - 0.646) + 0.412 \, (c_2 - 0.261) \tag{4}$$

$$P2 = -0.412 \, (c_1 - 0.646) + 0.911 \, (c_2 - 0.261), \tag{5}$$

and

$$P1 = 0.952 \, (c_1 - 0.646) + 0.307 \, (c_2 - 0.261) \tag{6}$$

$$P2 = -0.307 \, (c_1 - 0.646) + 0.952 \, (c_2 - 0.261), \tag{7}$$

for the red side, with $c_1 = vz - yz$ and $c_2 = bz - yz$. The coefficients are normalized such that the errors in *P1* and *P2* are the same as the errors in $vz - yz$ and $bz - yz$, in the limit when the latter two are the same. It is noteworthy that the *P2* axis is nearly parallel to the colour index $m_1 = (vz - yz) - 2(bz - yz)$, used as a metallicity estimator for single stars (e.g. Twarog 1980).

The distributions of *P1* and *P2* colours are shown in Fig. 5 by thick lines. The width of the locus in the *P2* direction is surprisingly small[4] (0.03 mag, and 0.32 mag in the *P1* direction). Indeed, the

---

[4] It may be surprising that the width of the *P2* colour distribution is smaller than the errors of the colours used to synthesize it. This is due to covariances in the $vz - yz$ and $bz - yz$ errors.
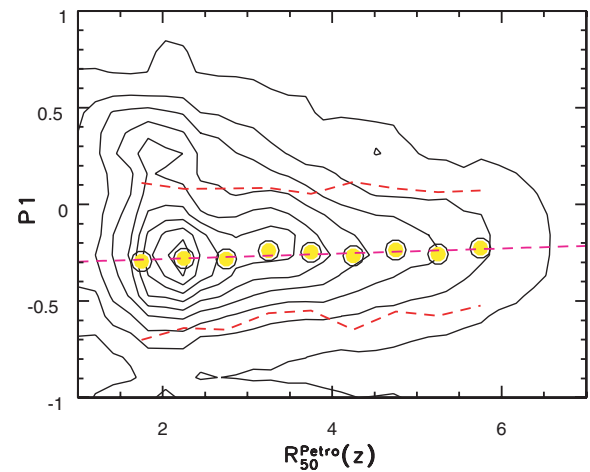
Figure 6. A test for aperture effects, due to 3 arcsec SDSS spectroscopic fibre diameter, on the *P1* colour. The linearly spaced contours show the distribution of galaxies with absolute Petrosian magnitude in the *r* band between -19 and -18 in the *P1* colour versus Petrosian 50 per cent radius in the *z*-band (an estimator of the galaxy's size) diagram. The aperture effect is expected to be the strongest for blue galaxies with $P1 < 0$. The large circles show median *P1* colours for bins of the Petrosian radius in the *z* band, computed for galaxies with $P1 < 0$. The middle line is a best straight line fit to these medians, and the other two lines illustrate 2σ deviation around the median. The slope of the best-fitting line is 0.015 mag arcsec^{-1}.

rest-frame colours for a large number of galaxies have not been measured to date with such an accuracy. This small value testifies both that the errors in synthesized colours are small, and that *the slope of galaxy spectra in the 4000–5800* Å *wavelength range is a nearly one-parameter family*.

### 3.2 Fibre aperture effects on *P1* and *P2*

SDSS spectra are obtained with a fixed 3-arcsec-diameter fibre positioned as close as possible to the centre of the galaxy. Since the fibre aperture may not include the entire galaxy, it is possible that colours derived from SDSS spectra are biased red by the bulge contribution. This bias should be the largest for intermediate- and late-type galaxies where the fibre aperture may miss a significant fraction of the outer disc and its star formation. While this effect was discussed in detail elsewhere (e.g. Kauffmann et al. 2003a; Kewley, Jansen & Geller 2005), here we explore its impact on the *P1* and *P2* colours.

The fibre aperture effect will tend to make the *P1* colour redder for galaxies with large angular size because bulges are generally redder than discs. The expected effect for the *P2* colour is not so obvious. We argue below (see Section 3.4) that the *P2* colour is related to dust, and if this is true, the expected effect, for the *P2* colour, is to become bluer for galaxies with large angular size.

Figs 6 and 7 show the dependence of the *P1* and *P2* colours on the Petrosian 50 per cent radius in the *z* band[5] for galaxies with the *r*-band absolute magnitudes in the range -19 to -18. This magnitude slice is selected because it is dominated by blue galaxies where the expected effect is the strongest (see Fig. 21). As the Petrosian radius increases, the bulge contribution to the flux captured by the fibre also

---

[5] For more details about the Petrosian radii, see Stoughton et al. (2002) and Strauss et al. (2002).
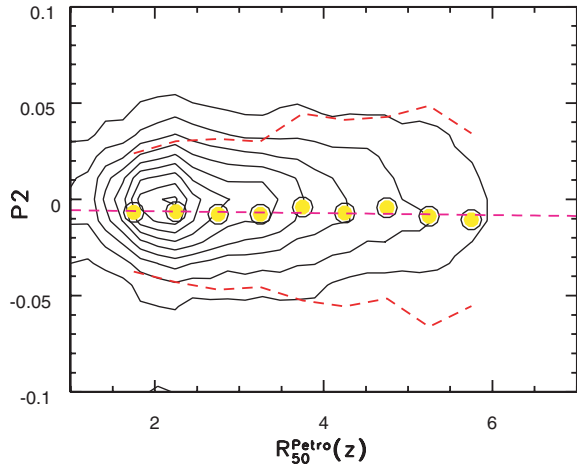
Figure 7. Analogous to Fig. 6, except for the *P2* colour and only for galaxies with *P1 < 0*. The slope of the best-fitting line is -0.0005 mag arcsec^{-1}.

increases. As evident, small but measurable dependence of the *P1* colour on the Petrosian radius is indeed present. Given the observed range of the Petrosian radius, the expected bias in the *P1* colour is $\lesssim 0.05$ mag, with the rms scatter contribution to the *P1* measurement error of ~ 0.02 mag. Compared to the observed *P1* range, this is a small effect. The fibre aperture effect on the *P2* colour is entirely negligible.

3.3 The stellar population model colours

The principal axes (*P1*, *P2*) are defined by the *observed* morphology of the distribution of galaxies in the rest-frame Strömgren colour–colour diagram. Here, we relate these axes to the stellar age and metallicity using the Bruzual & Charlot (1993, 2003) population-synthesis models.

Using the same method as for observed spectra, we synthesize rest-frame Strömgren colours from model spectra. Fig. 8 shows the model colours as a function of the age of a single stellar population, and for three representative metallicities. A single stellar population corresponds to a single epoch of star formation, while real galaxies likely have more complex star formation histories. In addition, for single stellar population models, the star formation rate history, which is an important discriminating parameter (e.g. Kennicutt 1998), is unavailable. Nevertheless, the colour range spanned by a single stellar population provides a robust constraint on the possible colour distribution for any star formation history (because it brackets the possible colour range).

The observed range of the *P1* colour can be explained by stellar populations with ages from 100 Myr to 10 Gyr. Note that *P1 > 0* corresponds to populations older than ~ 1 Gyr, but the age estimate is very uncertain due to the unknown metallicity.

Models provide a simple explanation for the observed small locus width: the *P2* colour is a colour projection which is nearly independent of both age *and* metallicity.[6] At any age, $P2 \sim 0$ to within

---

[6] The *P2* colour, while much less sensitive to metallicity than the *P1* colour, does become redder with increasing metallicity for populations older than 1 Gyr. However, the effect is exceedingly small (0.03 mag dex^{-1}). For individual stars (*P2* is nearly parallel to m_1, cf. Section 3.1), this is large enough because their effective temperature is typically constrained much better than the star formation history of a galaxy.
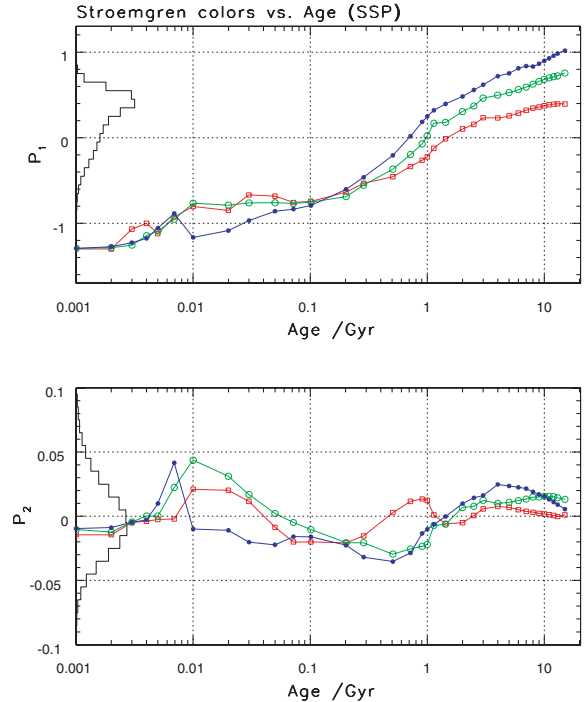
Figure 8. The Bruzual & Charlot model predictions for the dependence of the principal rest-frame Strömgren colours on the age and metallicity ([Z] = 0.004, squares; 0.02, circles; 0.05, dots) of single stellar populations. Note the different scales for the two colours. The histograms on the *y*-axes show the observed colour distributions.

0.02–0.03 mag, and for populations older than ~ 1 Gyr, $P2 \sim 0.01$ to within 0.01 mag. That is, whatever is age or metallicity, *P2* is always 0 to within a few hundreths of a magnitude. This is very different from *P1* which varies by 2 mag as a function of age, and up to 0.6 mag as a function of metallicity. *The ability of models to explain the mean value and the very narrow distribution of the P2 colour is a considerable success because the principal axes P1 and P2 are defined solely by the morphology of the data distribution in the* bz − yz *versus* vz − yz *colour–colour diagram.*

The observed width of the *P2* distribution is wider (0.03 mag) than the model prediction (~ 0.01 mag). While this difference could be easily attributed to the measurement scatter or to the simplicity of the models, we find that the *P2* colour is correlated with the estimate of the galaxy dust content, A_z (the effective dust extinction in the SDSS *z* band), determined by Kauffmann et al. (2003a) and further discussed in Section 4.2. Hence, the larger scatter observed in the *P2* colour than that predicted by the model is, at least partially, due to dust reddening (which is not included in the models).

It could be argued that the $P2$–A_z correlation is simply due to the fact that both quantities are determined from SDSS spectra. However, they are determined using *different*, barely overlapping, spectral ranges. While Kauffmann et al. analysis utilized the 3850–4160 Å spectral range to constrain their models (and then compared imaging $g − r$ and $r − i$ colours to model-predicted colours to estimate dust reddening), *P2* is determined from the 4000–5800 Å wavelength range (see the bottom panel in Fig. 1).

Another reasonable objection is that the A_z estimates are, of course, model-dependent and consequently may not be a good measure of the dust content. However, Obrić et al. (2006) demonstrated that galaxies detected by *IRAS* have systematically higher values of

A_z. This finding strongly suggests that A_z does provide an estimate of the dust content.

3.4 The colours of *IRAS*-detected galaxies

Since the $P2$–A_z correlation implies that $P2$ is also a measure of the galaxy dust content, galaxies detected by *IRAS* should have systematically higher values of $P2$. Using the same SDSS–*IRAS* galaxy sample discussed by Obrić et al. (∼2200 probable matches for the sample of 99 088 galaxies discussed here), we compared the $P1$ and $P2$ distributions for *IRAS*-detected galaxies to those for the full sample. As evident from Fig. 5, galaxies detected by *IRAS* have systematically higher $P2$ values.

The linear definition of the principal axes $P1$ and $P2$ (equations 4–7) is not a perfect description of the galaxy locus shown in Fig. 4 because the locus is curved. Hence, the $P2$ distribution slightly changes along the locus. Since *IRAS*-detected galaxies are biased towards blue galaxies, it could be that the difference in $P2$ distributions for *IRAS*-detected and all SDSS galaxies shown in Fig. 5 is caused by the variation of the $P2$ distribution along the locus. Fig. 9 demonstrates that this is not the case: for each segment of the locus defined by a narrow bin of $P1$, *IRAS*-detected galaxies have systematically higher $P2$ colours. This offset strongly suggests that $P2$ is a measure of the galaxy dust content.

The shift of the $P2$ distributions between SDSS–*IRAS* and all SDSS galaxies is about the same as the $P2$ distribution width, so $P2$
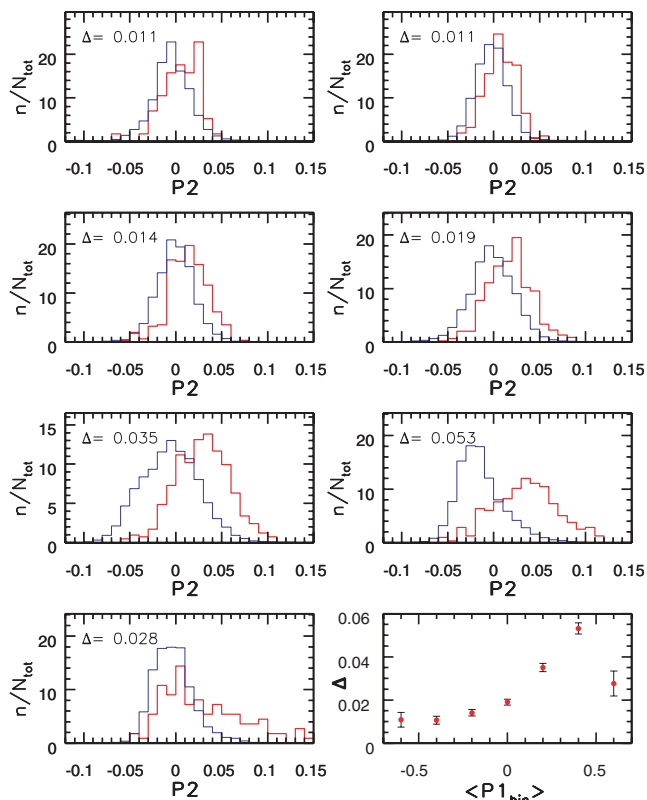
Figure 9. The comparison of $P2$ distributions for all galaxies and for the subsample detected by *IRAS*, for 0.2-mag-wide bins in $P1$, in the range -0.7–0.7 (row by row, from top left-hand to bottom right-hand panel). The thin lines show the distributions for the whole sample, and the thick lines for the *IRAS*-detected subsample. The Δ values shown in each panel are the difference between the medians for the two distributions, and are also shown as a function of $P1$ in the bottom right-hand panel.

is a rather noisy measurement of the dust content. Nevertheless, we demonstrate below that it is sufficiently efficient to study differences in the dust content between various galaxy populations separated using emission lines.

4 THE CORRELATIONS BETWEEN THE REST-FRAME STRÖMGREN COLOURS AND SPECTRAL PARAMETERS

In the previous section, we have shown the connection between the two principal rest-frame Strömgren colours and galaxy physical parameters: $P1$ depends, in a degenerate way, on stellar age and metallicity, while $P2$ is correlated with the dust content, and independent of age and metallicity. In this section, we further analyse correlations between Strömgren colours and various spectral parameters. We show that the position of an emission-line galaxy in a diagram commonly used to separate star-forming from active galactic nucleus (AGN) galaxies is well correlated with the $P1$ colour, that $P1$ and $P2$ colours are correlated with the parameters discussed by Kauffmann et al. (2003a,b), and that the rest-frame $r - i$ colour can be predicted to better than 0.05 mag from the $P1$ colour. All these correlations suggest that galaxy spectra are simple and well-defined superpositions of stellar spectra.

4.1 The Baldwin–Phillips–Terlevich diagram

The Baldwin–Phillips–Terlevich (BPT) diagram (Baldwin, Phillips & Terlevich 1981) is a standard method for classifying emission-line galaxies as star-forming and AGN galaxies. With the advent of SDSS data, it is now possible to construct the BPT diagram for an unprecedented number of galaxies (e.g. Ivezić et al. 2002; Kauffmann et al. 2003c; Heckman et al. 2004; Hao et al. 2005a,b). Obrić et al. (2006) have shown that the presence of emission lines in a galaxy's spectrum is well correlated with the $u - r$ colour and the concentration index determined *solely* from the imaging data. Galaxies without emission lines tend to have larger concentration indices and redder $u - r$ colours than galaxies with emission lines. Furthermore, the distribution of emission-line galaxies in the BPT diagram is also correlated with the $u - r$ colour and concentration index. Galaxies classified as star forming have predominantly blue $u - r$ colours and small concentration indices, while AGN galaxies have redder $u - r$ colours and large concentration indices.

Here, we investigate the correlation between the strength of the emission lines used to construct the BPT diagram, and rest-frame Strömgren colours. To classify a galaxy as an emission-line galaxy, we require a 3σ significant detection of Hα, Hβ, [N II] and [O III 5007] lines (the line strengths are determined as described by Kauffmann et al. 2003c). The bottom panel in Fig. 10 shows the distribution of 43 281 emission-line galaxies, from the sample discussed here, in the BPT diagram. To visualize the correlation with $P1$ and $P2$ colours, the dots are two-dimensionally colour coded according to their position in the $P1$–$P2$ colour–colour diagram shown in the top panel. *There is a strong correspondence between the position of a galaxy in the BPT diagram and its $P1$ colour.* Galaxies in the 'star-forming branch' with small [N II]/Hα ratios, for a given [O III 5007]/Hβ ratio, have predominantly blue $P1$ colours, while AGN galaxies have redder $P1$ colours. This correlation is further illustrated in Fig. 11. The correlation with the $P2$ colour is not discernible.

In the subsequent analysis, we separate emission-line galaxies in three groups according the their position in the BPT diagram: AGN, star forming, and 'unknown'. The adopted separation boundaries
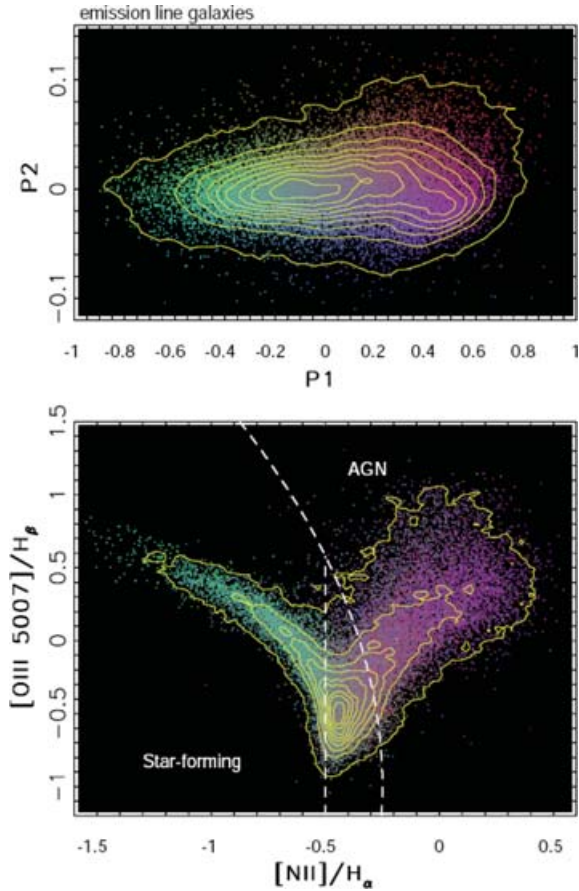
Figure 10. The top panel shows the distribution of SDSS 'main' galaxies in a diagram constructed with the principal rest-frame Strömgren colours. Each dot corresponds to one galaxy; their distribution is outlined by linearly spaced isopleths. The colour code is determined by the position in this diagram, and is used in subsequent figures to visualize correlations of other quantities with *P1* and *P2*. The bottom panel shows the distribution of galaxies in the BPT diagram, constructed with emission-line strength ratios. The dashed lines separate the regions populated by star-forming and AGN galaxies. They are motivated by the modelling results from Kewley et al. (2005) and are fine-tuned to isolate the peak of the observed distribution. For galaxies in the wedge-shaped regions, the classification is uncertain. The dots are coloured using *P1* and *P2*; note the strong correlation between the *P1* colour and [N II]/Hα ratio, although these parameters are determined in different wavelength ranges, and measure the continuum slope and line strengths, respectively.

are shown by the dashed lines. The last category is found at the joint of the two branches, and it is not obvious from the displayed data to which class these galaxies belong. Fig. 12 compares the *P1* distributions of these classes. The majority of galaxies without emission lines and AGN galaxies have *P1* > 0, while star-forming galaxies have *P1* < 0. More than nearly two-thirds of galaxies with *P1* > 0.3 are galaxies without emission lines, and more than nearly two-thirds of galaxies with *P1* < −0.4 are star-forming galaxies. At least 20 per cent of the galaxies with *P1* > 0 harbour an AGN. The fraction of AGN galaxies drops to essentially zero by *P1* ∼ −0.3 (or by *P1* ∼ −0.5 if *all* unclassified galaxies are AGN). This dramatic absence of AGN among the bluest galaxies appears not to be a selection effect, as discussed by Kauffmann et al. (2003c, section 4.1).

The comparison of *P2* distributions for the three subclasses of galaxies is shown in Fig. 13. As expected, AGN and star-forming
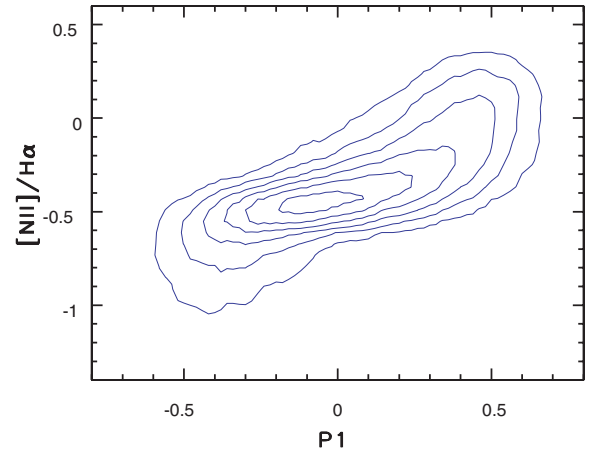
Figure 11. The correlation between the first principal axis in the rest-frame Strömgren colour–colour diagram, and the [N II]/Hα line strength ratio. The distribution of all emission-line galaxies in the sample is shown by linearly spaced isopleths.
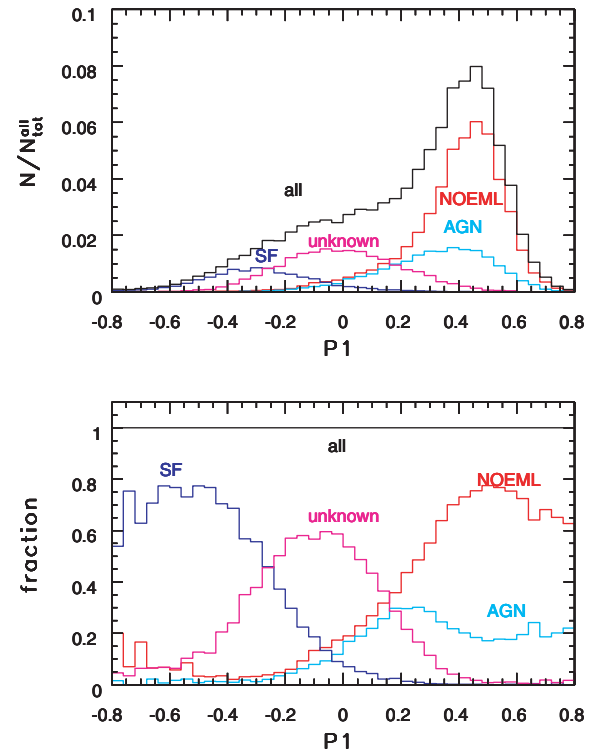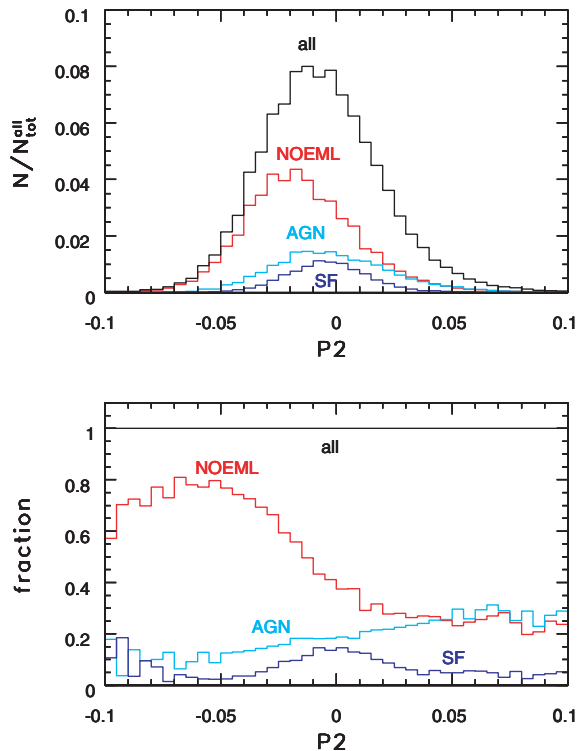

Figure 12. The distributions of *P1*, the first principal axis in the rest-frame Strömgren colour–colour diagram, for subsamples of galaxies separated using emission-line strengths: galaxies without emission lines (NOEML), star-forming galaxies (SF), AGN, and emission-line galaxies which cannot be reliably classified (unknown). The top panel displays counts, and the bottom panel displays the fraction of each subsample in the whole sample.

galaxies have redder *P2* distributions than galaxies without emission lines.

4.2 The Hδ_A – D_n (4000) locus

Kauffmann et al. (2003a) used the distribution of galaxies in the plane spanned by the strength of the Hδ_A line and the 4000-Å break
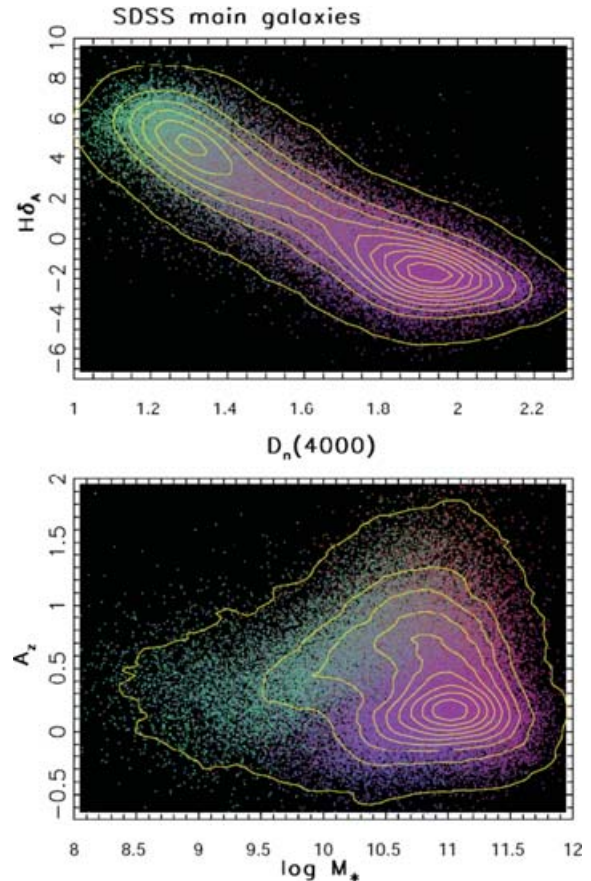
Figure 13. The distributions of *P2*, the second principal axis in the rest-frame Strömgren colour–colour diagram, for subsamples of galaxies separated using emission-line strengths: galaxies without emission lines (NOEML), star-forming galaxies (SF) and AGN. The top panel displays counts, and the bottom panel displays the fraction of each subsample in the whole sample (emission-line galaxies which cannot be reliably classified make up the missing fraction).

Figure 14. The top panel shows the distribution of SDSS galaxies in the plane spanned by the strength of the Hδ_A line and the 4000-Å break [$D_n(4000)$]. Each galaxy is represented by a point colour-coded according to its position in the *P1–P2* colour–colour diagram, shown in the top panel in Fig. 10. Note the strong colour gradient indicating a correlation between the position of a galaxy along this locus and its *P1* colour. The overall, strongly bimodal, distribution is outlined by linearly spaced isopleths. The bottom panel shows the distribution of the same sample in the dust content (A_z) versus stellar mass (M_*) diagram. A_z and M_* are model-dependent estimates derived from the measured Hδ_A, $D_n(4000)$ and imaging $g - r$ and $r - i$ colours. The colour gradients indicate good correlations between the positions in the A_z–M_* and *P1–P2* planes.

[$D_n(4000)$] to obtain model-dependent estimates of stellar masses and the dust content for SDSS galaxies. Given the position of a galaxy in the Hδ_A–$D_n(4000)$ plane, the most probable mass-to-light ratio is drawn from a model library. With the measured luminosity, this ratio then yields the stellar mass. The observed luminosity is corrected for the dust extinction determined by comparing observed imaging $g - r$ and $r - i$ colours to model-predicted colours (the latter do not include the effects of dust reddening). The reddening correction needed to make models agree with data is interpreted as an effective optical depth, A_z, due to the galaxy's interstellar dust. As discussed in Section 3.3, Obrić et al. (2006) demonstrated that galaxies detected by *IRAS* have systematically higher values of A_z, supporting the notion that A_z provides an estimate of the dust content.

The analysis by Kauffmann et al. is based on the 3850–4160 Å spectral range, marked by the horizontal bar in the bottom panel in Fig. 1, while the rest-frame Strömgren colours discussed here are determined using the 4000–5800 Å range. That is, the two studies are based on practically independent spectral ranges (the *vz* filter does have overlap with the Hδ_A line). Here, we demonstrate that despite their different wavelength ranges, and different techniques (colour versus spectral line analysis), the derived inferences about galaxies are similar.

The top panel in Fig. 14 shows the distribution of galaxies in the Hδ_A –$D_n(4000)$ plane, with the dots coloured according to the galaxy's position in the *P1–P2* colour–colour diagram (see the top panel in Fig. 10). As discussed by Kauffmann et al. (2003a), galaxies form a well-defined locus in this plane. The strong colour gradient

along the locus indicates a correlation with the *P1* colour. The bottom panel in Fig. 14 shows the distribution of galaxies in the plane spanned by the derived parameters: the dust content and the stellar mass. The colour gradients indicate a good correlation between the positions in the A_z–M_* and *P1–P2* planes. In particular, galaxies with small masses [$\log(M_*) \lesssim 10$] are dominated by blue galaxies with *P1* < 0, while a correlation between A_z and *P2* is discernible (gradient from purple to red as A_z increases) for high-mass galaxies [$\log(M_*) \sim 11$]. Since emission-line properties, discussed in the previous Section, correlate with *P1* and *P2* colours, it is expected that galaxies without emission lines, star-forming, and AGN galaxies display different distributions in these two diagrams. This is demonstrated in Figs 15, 16 and 17. For example, the detailed morphology of the galaxy distribution shown in the bottom panel in Fig. 14 can be understood as a superposition of three dominant populations with distinctive positions, as seen in the bottom panels in Figs 15, 16 and 17. It is noteworthy that the A_z–M_* correlation for star-forming galaxies shown in the bottom panel in Fig. 16 is
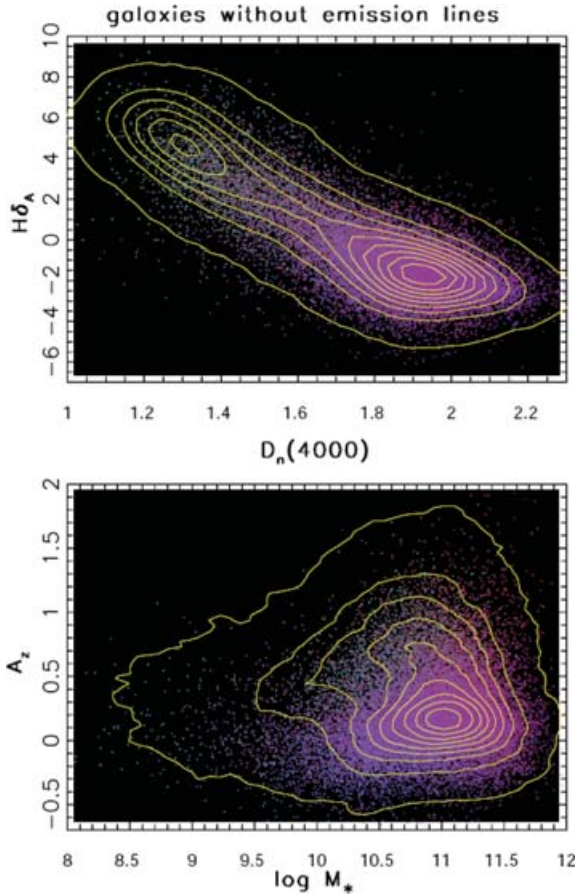
Figure 15. Analogous to Fig. 14, except that only galaxies without emission lines are shown (contours correspond to the full sample).

Figure 16. Analogous to Fig. 14, except that only star-forming galaxies are shown (contours correspond to the full sample).

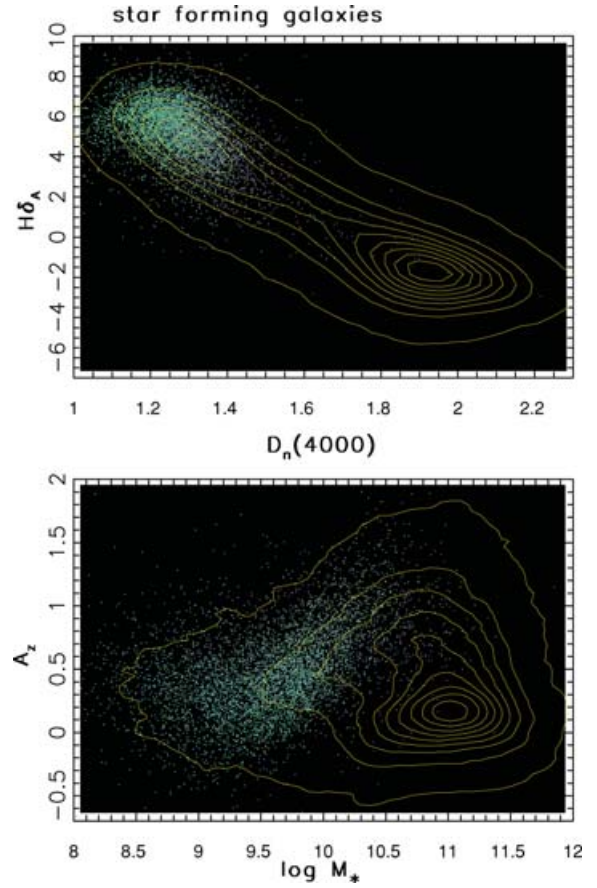consistent with a constant dust-to-stellar mass ratio for all galaxies.

In order to quantify the correlations between the positions in the $H\delta_A$–$D_n(4000)$ and $P1$–$P2$ planes, we define a parameter η which measures the position along the locus in the $H\delta_A$–$D_n(4000)$ plane:

$$\eta = 0.0995\,[D_n(4000) - 1.75] - 0.995\,H\delta_A. \qquad (8)$$

Note that $\eta = 0$ corresponds to $H\delta_A = 0$. Fig. 18 compares the η–$P1$ and A_z–$P2$ distributions for galaxies with and without emission lines. As evident, there is good correspondence between the two pairs of parameters, although they are measured using barely overlapping spectral ranges. In particular, note the good correlation between A_z and $P2$ for emission-line galaxies. Analogous diagrams comparing star-forming and AGN galaxies are shown in Fig. 19. It appears that the A_z–$P2$ correlation is stronger for AGN than for star-forming galaxies. This could be explained as a consequence of the more homogeneous AGN population because star-forming galaxies span a broad range in stellar mass, and hence in metallicity and surface mass density. It also seems that the slope of the A_z–$P2$ correlation is different for the two types of galaxies, but the data are too noisy to derive a robust conclusion.

4.3 The correlations of A_z and $P2$ with the $H\alpha/H\beta$ ratio

An often used estimator of the galaxy dust content is the $H\alpha/H\beta$ ratio. Therefore, at least some degree of correlation should exist between this parameter and the A_z and $P2$ parameters. Fig. 20 shows the

dependence of $\log(H\alpha/H\beta)$ on $P2$ and A_z for star-forming and AGN galaxies. Similarly to A_z versus $P2$ dependence shown in Fig. 19, it appears that different weak correlations exist for these two types of galaxies. The difference in slopes, similar to the difference in slopes for the A_z versus $P2$ correlation, is possibly due to different dust geometries, or perhaps different dust properties, between AGN and star-forming galaxies (note that this difference also exists in the $H\alpha/H\beta$ versus A_z diagram, that is, even when $P2$ is not used). A detailed analysis and interpretation of these correlations is beyond the scope of this work. We conclude that the similarity of the top and bottom panels in Fig. 20 suggests that $P2$ is a good, though somewhat noisier, proxy for A_z.

4.4 The colour–absolute magnitude distribution

The colour distribution of galaxies depends on their absolute magnitude: the fraction of red galaxies increases with luminosity. Furthermore, the median colour of red galaxies becomes even redder with luminosity (Baum 1959; Faber 1973). These long-known results are spectacularly confirmed by the accurate and voluminous SDSS data (e.g. Blanton et al. 2003; Baldry et al. 2004, and references therein). Hence, the distribution of the $P1$ principal colour introduced here is expected to be highly correlated with luminosity.

The top panel in Fig. 21 shows the distribution of 1704 galaxies selected from a narrow redshift range in the absolute magnitude versus $P1$ colour plane. We use the Petrosian r band magnitudes without K-correction and *Wilkinson Microwave Anisotropy Probe*
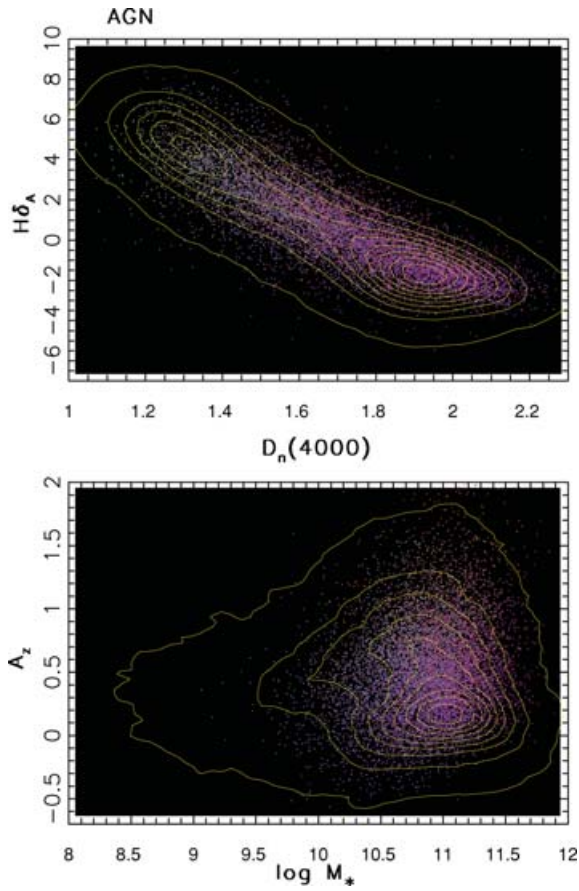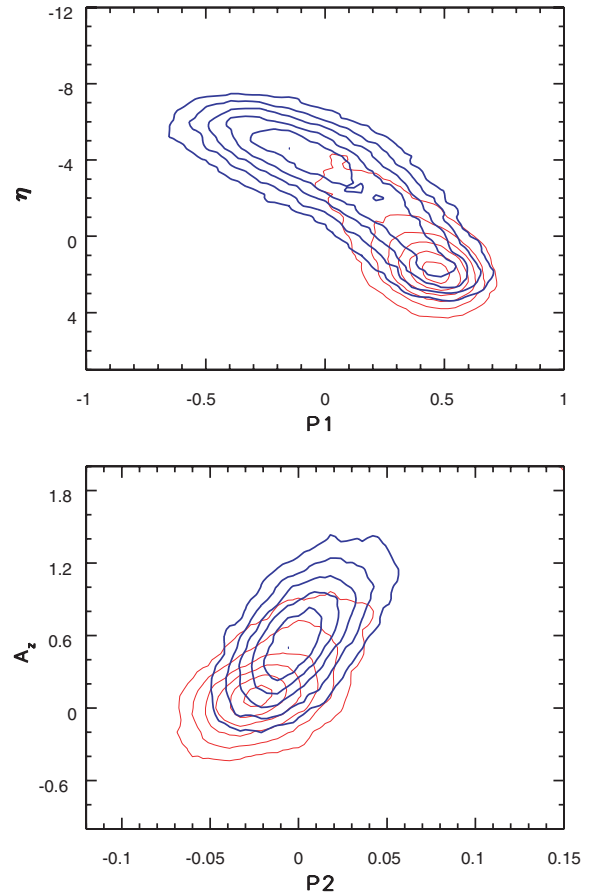
Figure 17. Analogous to Fig. 14, except that only AGN galaxies are shown (contours correspond to the full sample).

Figure 18. The top panel shows the correlation between the position of a galaxy on the locus in the $H\delta_A$ –$D_n(4000)$ plane (η) and the *P1* colour. The thin contours correspond to galaxies without emission lines, and the thick contours correspond to emission-line galaxies. The bottom panel shows the A_z–*P2* correlation.

(*WMAP*) cosmology (Spergel et al. 2003) to compute M_{rPet} (for a description of Petrosian magnitudes, see Stoughton et al. 2002; Strauss et al. 2002). The strong colour–magnitude relation for red galaxies is evident, as well as the dramatic increase in the fraction of red galaxies for $M_{rPet} < -20$. Indeed, red galaxies with $P1 \sim 0.5$ show a well-defined local maximum in their luminosity function (at $M_{rPet} \sim -20.7$).

This figure suggests that $P1 = 0.25$ provides a good separation between blue and red galaxies ($u - r = 2.22$ separator from Strateva et al. 2001 corresponds to median $P1 = 0.18$). The separation of less-luminous blue galaxies and more-luminous red galaxies is much more pronounced in this diagram based on the *P1* colour, than when using other colours such as *K*-corrected broad-band SDSS colours (see e.g. fig. 7 in Blanton et al. 2003). This is presumably because *P1* is a principal axis in a rest-frame colour–colour diagram constructed with narrow-band filters (the galaxy locus in the $r - i$ versus $g - r$ rest-frame colour–colour diagram, discussed in Section 4.7, is twice as broad as the galaxy locus in the $bz - yz$ versus $vz - yz$ plane).

Since the *P1* colour is correlated with numerous other parameters, the subsamples separated using those parameters should have distinct distributions in the colour–absolute magnitude diagram. As an example, we choose the separation into star forming, AGN and galaxies without emission lines. The bottom three panels in Fig. 21 compare the colour–absolute magnitude distribution for each subsample to that for the full sample. As evident, AGN and galaxies without emission lines are predominantly red and luminous, while star-forming galaxies are blue and faint. The faint blue galaxies

without emission lines, visible in the upper middle panel, are galaxies with weak emission lines that did not fulfill the condition of the 3σ detection for all four relevant emission lines.

4.5 The colour–mass distribution

Kauffmann et al. (2003a) estimated stellar masses for SDSS galaxies by multiplying observed luminosities (corrected for extinction using A_z) with the stellar mass-to-light ratio inferred from the position of a galaxy in the $H\delta_A$–$D_n(4000)$ plane. The correlation of the mass-to-light ratio and parameter η, which measures the position of a galaxy along the $H\delta_A$–$D_n(4000)$ locus, is shown in the top panel in Fig. 22. The dashed line is a best-fitting

$$\log(M/L_z) = 0.000630\eta^3 - 0.00174\eta^2 + 0.01396\eta + 0.231, \quad (9)$$

which 'predicts' $\log(M/L_z)$ with an rms of 0.13 [the overall distribution of $\log(M/L_z)$ has an rms of 0.17, i.e. a factor of 1.5]. The upper middle panel shows the mass-to-light ratio as a function of the *P1* colour. The best fit,

$$\log(M/L_z) = 0.273\,P1^3 - 0.328\,P1^2 + 0.516\,P1 + 0.091, \quad (10)$$

reproduces $\log(M/L_z)$ with an rms of 0.10, equal to the median error in estimating $\log(M/L_z)$ using the Kauffmann et al. (2003a)
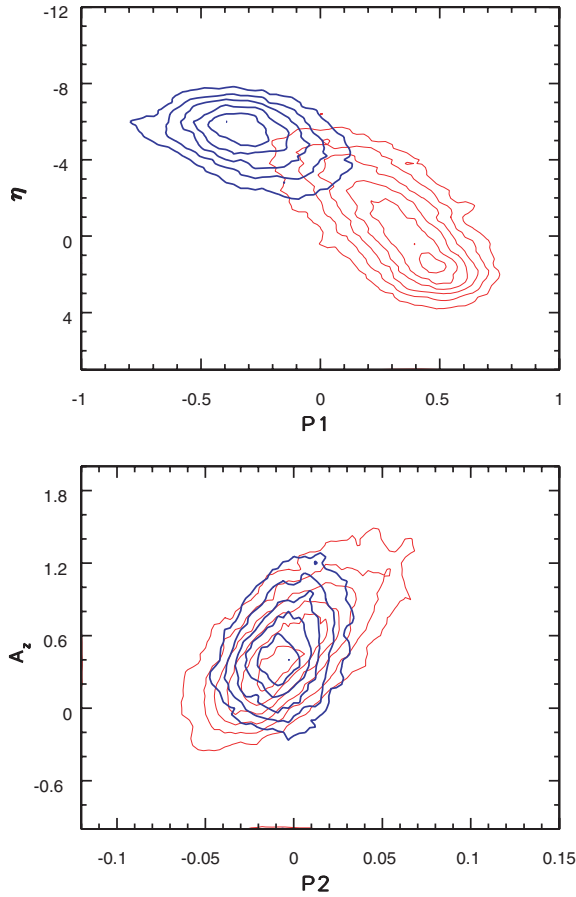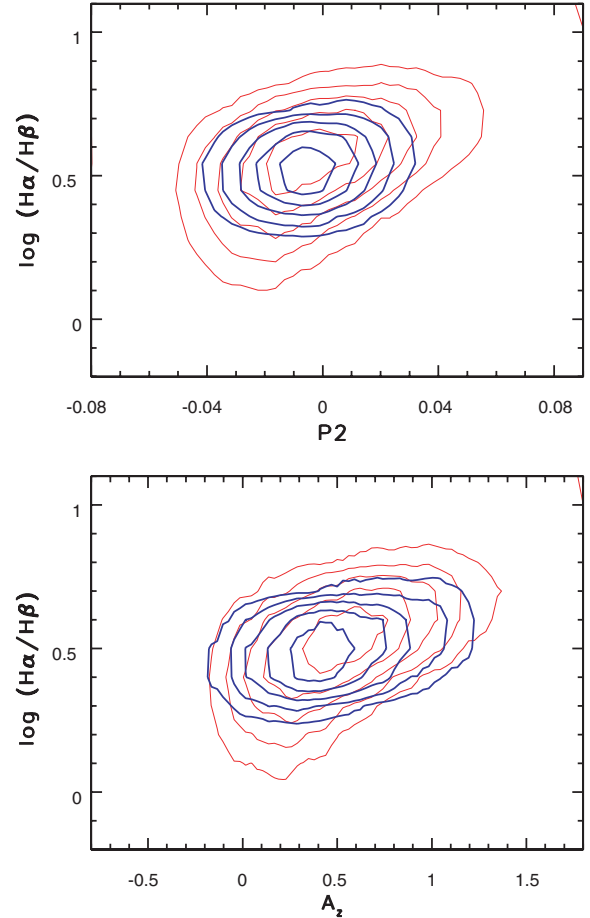
Figure 19. Analogous to Fig. 18, except that star-forming (thick contours) and AGN (thin contours) galaxies are shown.

Figure 20. The comparison of $\log(H\alpha/H\beta)$ versus $P2$ (top panel) and $\log(H\alpha/H\beta)$ versus A_z (bottom panel) correlations for star-forming (thick contours) and AGN (thin contours) galaxies.

method. Hence, *it is possible to use P1 to convert the observed luminosities into stellar masses* with practically the same accuracy as the $H_{\delta_A}-D_n(4000)$ method employed by Kauffmann et al. (2003a). The bottom two panels compare the mass–colour distributions when the mass is determined using *P1*, and when the mass is taken from Kauffmann et al. (2003a); their similarity confirms this conclusion. We note that the overall appearance of the mass–colour distribution, in particular, the strong correlation of mass and colour, essentially reflects the distribution of galaxies in the colour–luminosity diagram (because the mass-to-light ratio has a very small dynamic range).

Since *P1* and luminosity, L_z, are correlated, as described in the previous section, and M/L_z is a function of *P1*, M/L_z and L_z should also be correlated to at least some extent. For example, Bernardi et al. (2003b) determined that for elliptical galaxies $M/L \propto L^{0.14}$. We find that a best power-law fit for *all* galaxies is $M/L \propto L^{0.4}$. It reproduces $\log(M/L_z)$ with an rms of 0.20. Since this scatter is larger than that produced when determining *M/L* from *P1*, we speculate that galaxy colours are a better indicator of the mass-to-light ratio, and that its correlation with galaxy luminosity is probably a consequence of the colour–magnitude relation.

4.6 The colour–velocity dispersion distribution

Stellar velocity dispersion is automatically measured for each SDSS galaxy by fitting the observed spectrum with a linear combination of galaxy template spectra broadened by a Gaussian kernel (Schlegel et al., in preparation). Detailed analysis of velocity dispersions for

elliptical galaxies, as measured by SDSS, is discussed by Bernardi et al. (2003a) and Sheth et al. (2003).

We find a good correlation between the velocity dispersion and the *P1* colour, shown in the top panel in Fig. 23. The correlation is especially strong for velocity dispersions larger than the instrumental resolution of the SDSS spectra [60–70 km s^{-1}, $\log(v_{\rm disp}/100\,{\rm km\,s^{-1}})$ ~ -0.2]. A similar correlation between the velocity dispersion and the $g - r$ colour for elliptical galaxies is discussed by Bernardi et al. (2003d).

The correlation between the velocity dispersion and the *P1* colour appears to be a consequence of the good correlation between the stellar mass and the velocity dispersion, shown in the bottom panel in Fig. 23. We find that the function

$$\log(M/M_\odot) = 10.7 + 2 \log\left(\frac{v_{\rm disp}}{100\,{\rm km\,s^{-1}}}\right). \tag{11}$$

provides a good description for *both* blue and red galaxies. For a constant *M*, in the range $10.5 < \log(M/M_\odot) < 12$, the rms scatter in the velocity dispersion is 40 km s^{-1}, independent of *M*. For a constant velocity dispersion, the rms scatter in $\log(M/M_\odot)$ decreases from 0.33 for $v_{\rm disp} = 100$ km s^{-1} to 0.18 for $v_{\rm disp} = 250$ km s^{-1}.

The similar $\log(M/M_\odot)$ versus $v_{\rm disp}$ behaviour for blue and red galaxies can be considered as circumstantial evidence that the model-dependent mass-to-light ratios must be correct to some extent. The reason is that blue ($P1 < 0.25$) and red galaxies ($P1 > 0.25$)
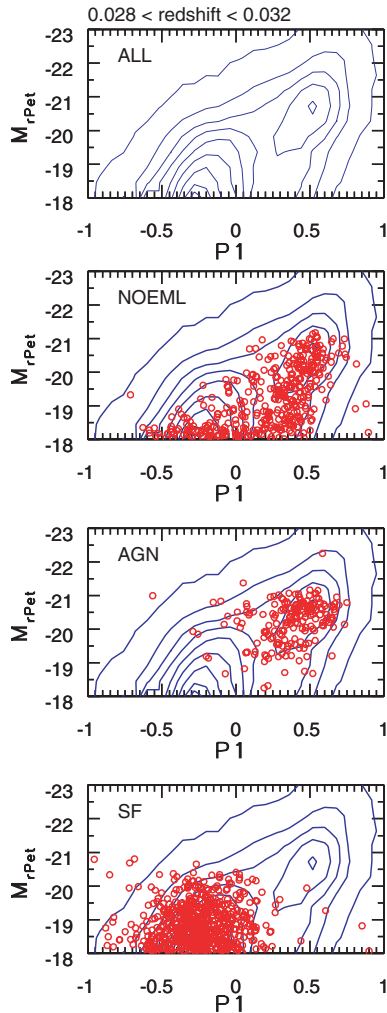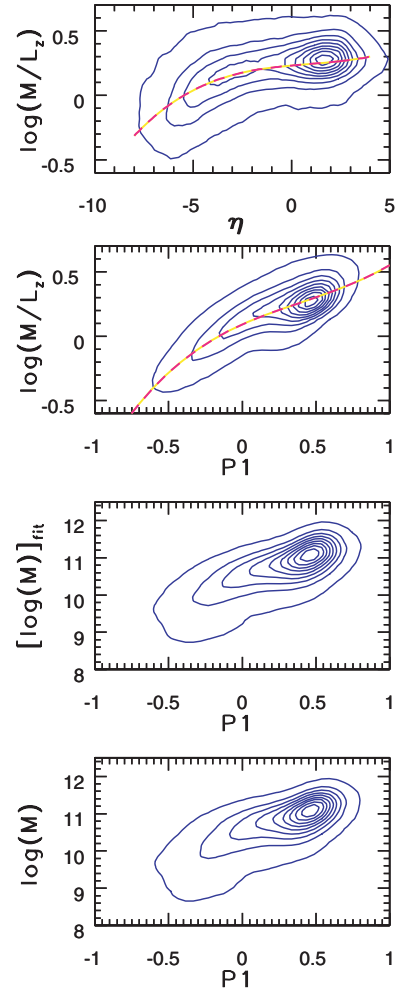
Figure 21. The colour–absolute magnitude diagram for SDSS main galaxies. The top panel shows the absolute *r*-band magnitude as a function of the *P1* rest-frame Strömgren colour for galaxies with $0.028 < z < 0.032$. The remaining panels compare the distribution of all galaxies (contours) to the distributions of three subsamples (symbols) selected using emission lines: galaxies without emission lines in the upper middle panel, AGN galaxies in the lower middle panel, and star-forming galaxies in the bottom panel.
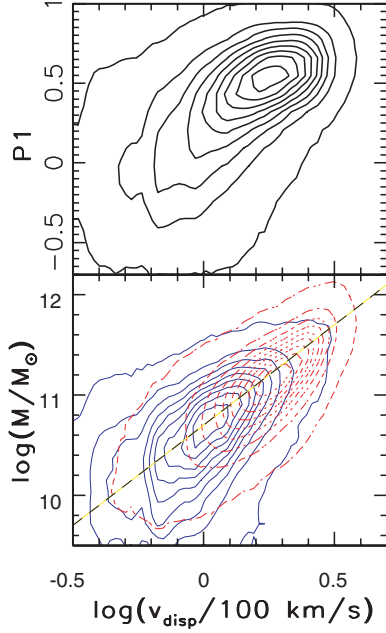

Figure 22. The top panel shows the correlation between the mass-to-light ratio (in solar units) and the position of a galaxy on the $H\delta_A$–$D_n(4000)$ locus, where the model-dependent mass-to-light ratio is taken from Kauffmann et al. (2003). The dashed line is a best-fitting third-order polynomial. The upper middle panel shows the correlation between the mass-to-light ratio and the position of a galaxy on the locus in the rest-frame Strömgren colour–colour diagram. The dashed line is a best-fitting third-order polynomial. The lower middle panel shows the mass–colour diagram using the best-fitting mass-to-light ratio determined from the *P1* colour, and the bottom panel shows analogous distribution using the mass-to-light ratio from Kauffmann et al. (2003). Note the similarity of the two distributions.

have different observed luminosities (median z-band luminosity for red galaxies is larger by a factor of 2.6 than that for blue galaxies), and different velocity dispersions, and only the 'correct' mass-to-light ratios (medians are 1.2 and 1.9 for blue and red galaxies, respectively) will place the inferred stellar masses on the same M–v_{disp} relation. If the same mass-to-light ratio is used for all galaxies, the inferred mass for red galaxies would be smaller by a factor of 1.6 than that for blue galaxies with the same velocity dispersion. That is, the two distributions shown in the bottom panel in Fig. 23 would be vertically offset by $\sim$0.2. The data constrain the offset between the two distributions to $\lesssim$0.1.

4.7 The rest-frame Strömgren colours versus red colours

We have demonstrated in the preceding sections a good correlation between rest-frame Strömgren colours and parameters determined from the blue part of the spectrum [$H\delta_A$ and $D_n(4000)$]. In this section, we analyse the correlation between rest-frame Strömgren

colours and parameters determined from the red part of spectrum. As shown in Section 3.3, the age–metallicity degeneracy cannot be broken using rest-frame Strömgren colours employed here. It is possible that additional information contained at the redder wavelengths could provide a method to separate the age and metallicity effects. If the properties of the red part of the spectrum can be fully predicted using rest-frame Strömgren colours, then such a possibility can be ruled out. Furthermore, if the blue and red parts of the spectrum are dominated by uncorrelated stellar populations, it should *not* be possible to accurately predict $r - i$ colour using blue Strömgren colours.

To characterize the red part of the spectrum, we synthesize the *rest-frame r − i* colour, using the same method as that used for synthesizing rest-frame Strömgren colours. The top panel in Fig. 24 shows the rest-frame $r - i$ colour as a function of redshift.

Figure 23. The top panel shows the correlation between the *P1* rest-frame Strömgren colour and the velocity dispersion. The bottom panel shows the correlation between the galaxy's stellar mass and the velocity dispersion, separately for galaxies with *P1* < 0.25 (solid lines) and *P1* > 0.25 (dashed lines). The diagonal straight line, $\log(M/M_\odot) = 10.7 + 2\log(v_{\rm disp}/100\,{\rm km\,s}^{-1})$, is added for illustration.

Reassuringly, there is no strong dependence on redshift – the upper limit on the change of the median colour with redshift is ∼0.02 mag (in the 0.02–0.12 redshift range). The correlation between the rest-frame $r - i$ colour and the *P1* colour is shown in the upper middle panel. The dashed line is a best-fitting

$$r - i = 0.109\,P1^3 - 0.221\,P1^2 + 0.288\,P1 + 0.238, \qquad (12)$$

which predicts the rest-frame $r - i$ colour with an rms of only ∼0.05 mag. The distribution of residuals is shown in the lower middle panel, and is well described by a Gaussian distribution.

Since *P2* is a measure of dust reddening in a galaxy, some correlation between the $r - i$ residuals and *P2* could be expected, although the errors in the determination of *P1*, *P2* and $r - i$, which are comparable to the magnitude of $r - i$ residuals, may easily mask it. The dependence of residuals on the *P2* colour, shown in the bottom panel, gives a hint of such a correlation. For example, the median $r - i$ residual is 0.03 mag for *P2* = 0.05, while it is −0.02 mag for *P2* = −0.05.

This one-dimensional parametrization of galaxy spectral energy distributions (SEDs) extends to even longer wavelengths. Obrić et al. (2004, 2006) found that SDSS imaging magnitudes in the *u* and *r* bands, together with redshift information, can be used to predict Two-Micron All-Sky Survey (2MASS) *K*-band (2.2-μm) magnitudes with an rms scatter of only ∼0.2 mag, with a significant fraction of this scatter contributed by measurement errors (they estimated that astrophysical scatter is much smaller than 0.1 mag). In summary, *the UV to IR SEDs of galaxies can be described as a single-parameter family with an accuracy of 0.1 mag, or better; in addition, the more detailed spectral diagnostic parameters, such as line strengths, appear well correlated with the overall SED.*
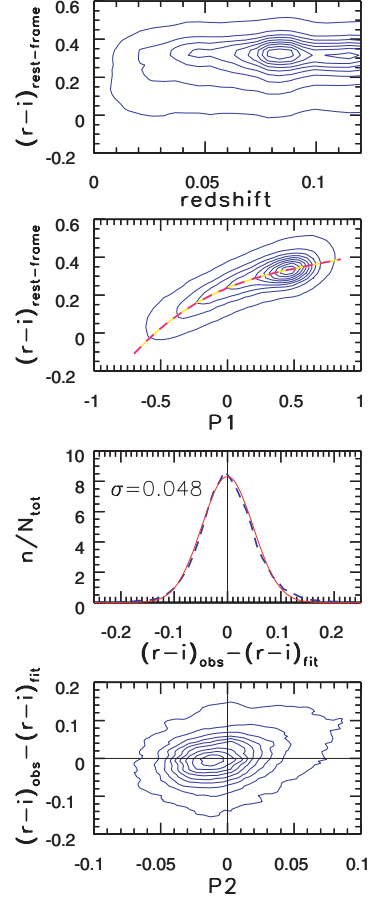

Figure 24. The top panel shows the rest-frame $r - i$ colour as a function of redshift. The upper middle panel shows the correlation between the rest-frame $r - i$ colour and the *P1* rest-frame Strömgren colour. The dashed line is a best-fitting third-order polynomial (see the text). The distribution of residuals between the fit and measured values is shown in the lower middle panel by the thick dashed line. The thin solid line is a Gaussian with the width of 0.048 mag. The bottom panel shows the residuals as a function of the *P2* rest-frame Strömgren colour.

5 THE CORRELATION BETWEEN THE REST-FRAME STRÖMGREN COLOURS AND SPECTRAL EIGENCOEFFICIENTS

In the previous section, we discussed correlations between the rest-frame Strömgren colours and various other spectral parameters. Here, we study the correspondence between a classification scheme based on spectral eigencoefficients proposed by Yip et al. (2004) and the principal colours *P1* and *P2*.

Yip et al. used the KL transform to classify 170 000 SDSS galaxy spectra. These spectra are classified in a plane spanned by the mixing angles between the first three eigencoefficients, $\Phi_{\rm KL}$ and $\Theta_{\rm KL}$. This classification discriminates between early-, late-type and extreme emission-line galaxies (see their section 5 for details). Given correlations between various spectral parameters discussed here, it is expected that the distribution of galaxies in the $\Phi_{\rm KL}$–$\Theta_{\rm KL}$ plane is, at least to some extent, correlated with the principal colours *P1* and *P2*.

Fig. 25 shows this distribution, colour coded according to the position of each galaxy in the *P1*–*P2* plane. The correlation between the *P1* colour and the proposed classification is evident. As discussed
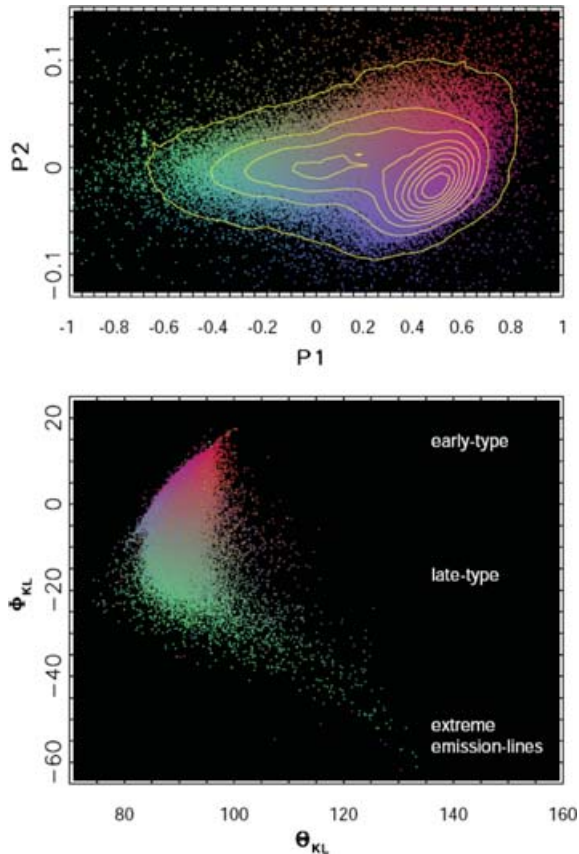
Figure 25. The bottom panel shows a spectral classification plane for galaxies proposed by Yip et al. (2004). Each galaxy is represented by a dot colour-coded according to its position in the *P1–P2* plane, shown in the top panel. The sequence from early-type to extreme emission-line galaxies adopted from Yip et al. (2004) is shown on the right-hand side in the bottom panel. Note the good correspondence between *P1* and *P2* colours and the Yip et al. classification.

in detail by Yip et al., the mean spectra for galaxies selected from small (Φ_{KL}, Θ_{KL}) regions correspond to different spectral types. This suggests that the *P1* colour is a robust measure of galaxy type, and confirms that most of the variance in galaxy SEDs is captured by a small number of eigencomponents, as found by Yip et al.

The correlation with *P2* is not as strong as for *P1*, but appears discernible for early- to late-type galaxies. There is a slight gradient from purple ($P2 < 0$) to red-orange ($P2 > 0$) in the $80 < \Theta_{KL} < 100$ range.

6 THE REST-FRAME UV COLOURS OF $z \gtrsim 0.18$ GALAXIES

Our analysis of the rest-frame Strömgren colours in the preceding sections did not utilize the *uz* magnitudes. Since the SDSS spectra cover only wavelengths longer than 3800 Å, the *uz* magnitudes can be synthesized only for galaxies at redshifts greater than $z \sim 0.18$ (hereafter 'high-z' sample). The 'high-z' sample includes only ~9 per cent of the 99 088 galaxies from the 'main' galaxy sample, and is strongly biased towards red galaxies because of the severe luminosity cut (see Fig. 21). Nevertheless, it includes a sufficient number of blue galaxies to demonstrate a correlation between the $uz - vz$ colour and the *P1* colour, shown in Fig. 26.
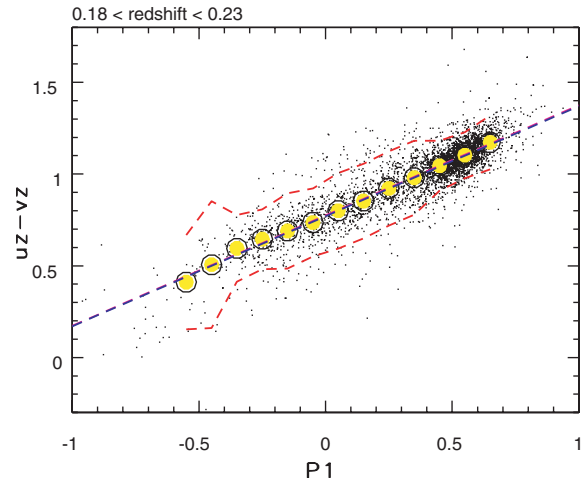
Figure 26. The $uz - vz$ versus *P1* colour–colour diagram for galaxies (dots) with synthesized *uz* magnitudes ($0.18 < z < 0.23$). The large circles show medians for *P1* bins. The middle (straight) dashed line is the best-fit to these medians, and the other two dashed lines mark the 2σ envelope around the medians. The best-fitting relation can be used to 'predict' the $uz - vz$ colour from *P1* with a residual scatter of only 0.08 mag.
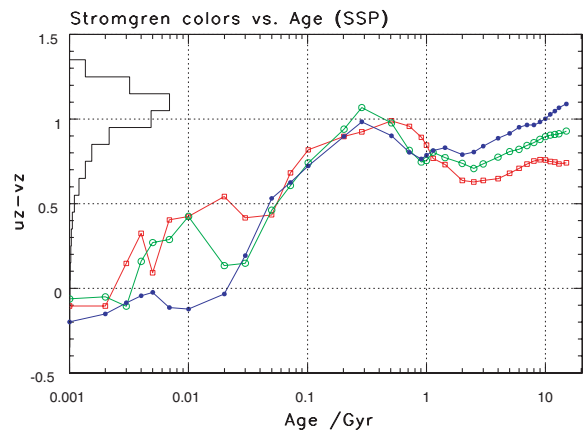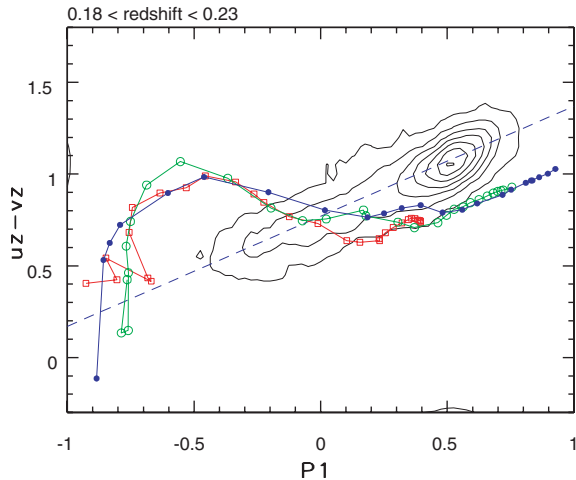

Figure 27. The Bruzual & Charlot model predictions for the dependence of the $uz - vz$ colour on age and metallicity ($[Z] = 0.004$, squares; 0.02, circles; 0.05, dots) of single stellar populations. Note that for small ages ($\lesssim 30$ Myr) metallicity can have a strong impact on the $uz - vz$ colour, while for intermediate ages (30 Myr $\lesssim$ age $\lesssim 2$ Gyr) it plays only a minor role. For old stellar populations ($\gtrsim 2$ Gyr), the $uz - vz$ colour becomes again increasingly sensitive to metallicity. However, not even for the oldest galaxies the model $uz - vz$ colour is sufficiently red to explain the observations.

The best-fitting relation shown in Fig. 26,

$$uz - vz = 0.60\,P1 + 0.77, \qquad (13)$$

can be used to 'predict' the $uz - vz$ colour with a residual scatter of only 0.08 mag! It is noteworthy that the $uz - vz$ residuals are not correlated with the *P2* colour. This small scatter means that either the dependence of the $uz - vz$ colour on metallicity is not as strong as assumed a priori, or that age and metallicity for galaxies are combined in such a degenerate way that the distribution of galaxies in the $uz - vz$ versus *P1* plane is essentially one-dimensional (within this 0.08 mag scatter), similarly to their distribution in the *P1* versus *P2* plane.

In order to distinguish between these two possibilities, we resort again to the Bruzual & Charlot population-synthesis models (see

Figure 28. The $uz - vz$ versus *P1* colour–colour diagram. The observed galaxy distribution is shown by linearly spaced contours (the same distribution is shown by dots in Fig. 26). The Bruzual & Charlot model predictions are shown by lines and symbols, analogously to Fig. 27. The position along each model track is controlled by the age of a single stellar population, while different tracks correspond to different metallicities. Note that the effects of age and metallicity are largely degenerate such that different model tracks do not deviate from each other by more than ∼0.1 mag. Note also the discrepancy between the model tracks and the observed distribution.

Section 3.3). Fig. 27 shows model predictions for the dependence of the $uz - vz$ colour on age and metallicity. At least for the smallest ($\lesssim$30 Myr) and largest ($\gtrsim$2 Gyr) ages, the $uz - vz$ colour appears sensitive to metallicity. However, this 'high-*z*' sample does not include extremely blue galaxies with implied ages below 30 Myr, or *P1* < −0.5 (see Fig. 8), so model predictions are not at odds with the correlation displayed in Fig. 26. The situation is less clear for old red ($\gtrsim$2 Gyr, or *P1* $\gtrsim$ 0.2) galaxies. Model tracks for 10-Gyr-old galaxies (*P1* > 0.4) predict a >0.2–0.3 mag wide range of $uz - vz$ colours, while the observed distribution has the rms scatter of 0.08 mag.

The resolution of this discrepancy is the degenerate dependence of the *P1* and $uz - vz$ colours on age and metallicity. For galaxies older than ∼2 Gyr, *both* colours become redder as age and metallicity increase, by about the same amount. As shown in Fig. 28, for *P1* > 0.4 the model tracks converge and become a single track irrespective of metallicity. Taken at face value, this convergence suggests that the $uz - vz$ colour scatter of elliptical galaxies (at a given *P1*) is due to age spread, and not due to a finite width of their metallicity distribution.

This interpretation has to be taken with caution because models never produce sufficiently red $uz - vz$ colours to explain the observations of the reddest galaxies. While one is tempted to invoke the dust (which is not incorporated in these models) as an explanation of this discrepancy, these galaxies are the reddest and most luminous galaxies which are not supposed to have significant amounts of dust. Furthermore, the models show significantly different behaviour for blue galaxies from that displayed by the data distribution. It is quite plausible that model spectra in the rest-frame UV wavelength range require significant improvements. For example, Fig. 28 indicates that a model spectrum fit to λ > 4000 Å spectral range for a blue galaxy (say, *P1* = −0.5), will have the flux in the 3200–4000 Å wavelength range underestimated by ∼50 per cent (0.5 mag). Note that this discrepancy cannot be easily attributed to dust effects (which

are not accounted for in model spectra) because model spectra are redder, and not bluer, than the observed spectra.

7 DISCUSSION AND CONCLUSIONS

We analysed the synthesized rest-frame Strömgren colours for ∼100 000 SDSS galaxies. This narrow-band photometric system is well suited for studying galaxy colours in the 4000–5800 Å wavelength range. Galaxies form a narrow locus in the rest-frame Strömgren $bz - yz$ versus $vz - yz$ colour–colour diagram, with a width of only 0.03 mag. This small width, in addition to demonstrating that the errors in synthesized colours are small, shows that the slope of galaxy spectra in the 3200–5800 Å wavelength range is practically a one-parameter family. This conclusion is independent of the details of the employed photometric system.

Using the Bruzual & Charlot population-synthesis models, we find that the dependence on age and metallicity is *fully absorbed* in the first principal colour axis, *P1*. This colour is essentially a measure of the Hubble type (in sense of early- versus late-type galaxies, rather than detailed morphological classification). It is not possible to break the age–metallicity degeneracy with any of the observables discussed here. This may not be possible even when using standard methods such as Lick indices, at least in the case of elliptical galaxies (Eisenstein et al. 2003). As demonstrated by Eisenstein et al., the strengths of numerous metal lines, such as Mg b, can be accurately predicted from the strength of the Hδ_A line (see their fig. 5); hence, no additional information can be extracted from the measurements of those lines. On the other hand, Tremonti, Heckman & Kauffmann (2004) report a very tight mass–metallicity relation (0.1-dex scatter in metallicity at a given mass) for ∼50 000 SDSS star-forming galaxies. Thus, it may be that age and metallicity are strongly correlated, too, which then would provide a trivial solution to the problem of breaking the age–metallicity degeneracy. Indeed, such a relation has been claimed for disc stars in our Galaxy (Twarog 1980; however, see also Feltzing, Holmberg & Hurley 2001).

The second principal colour axis, *P2*, is independent of stellar age and metallicity, at least according to the Bruzual & Charlot models. We stress that the definition of this colour is purely *observational*, and the models explain both its value and narrow distribution width. The *P2* distribution width (0.03 mag) measured using SDSS data is smaller than typical observational errors in rest-frame colours using other data available to date ($\gtrsim$0.05 mag). The *P2* colour appears most sensitive to the galaxy dust content, as demonstrated by the analysis of model-dependent estimates of the dust content by Kauffmann et al. (2003a), and also by the redder *P2* colours for galaxies detected by the *IRAS* survey. Of course, *P2* could be *indirectly* related to metallicity, if the dust content depends on metallicity (that is, the observed *P2* colour could be correlated with metallicity as a result of radiative transfer through interstellar dust, rather than due to the dependence of the source function on metallicity).

The principal colours are well correlated with numerous other spectral parameters determined from bluer and redder spectral ranges, as well as with the position of a galaxy in the BPT diagram constructed with emission-line strength ratios. These good correlations suggest that not only are the overall UV–IR SEDs a nearly one-parameter family, but that the same parameter also controls the properties of emission lines. Most notably, there is a good correspondence between the stellar mass-to-light ratio estimates by Kauffmann et al. (2003a) and the principal colour *P1*.

The low dimensionality of galaxy spectra implied by the numerous correlations among the continuum and spectral parameters shown here is in good agreement with the conclusions by Yip

et al. (2004), who applied the principal component analysis to SDSS galaxy spectra. The same conclusion was reached in a detailed study of elliptical galaxies by Eisenstein et al. (2003). These conclusions are valid for the overwhelming majority of all galaxies; however, we emphasize that the analysis presented here is based on low-order statistical moments (such as rms scatter) and cannot rule out a possibility that as many as 5–10 per cent of galaxies behave differently.

Our results support the suggestion by Rakos and collaborators (Fiala et al. 1986) that the rest-frame Strömgren photometry provides an efficient tool for studying faint cluster galaxies and low surface brightness objects, without the need for time-consuming spectroscopy. For example, the stellar mass-to-light ratio can be determined from *P1* with an rms scatter of only 0.10. Since the brightest cluster galaxies can be selected from SDSS data to redshifts of ~0.8 (e.g. Ivezić et al. 2002), this method can be used to quantify the rise in star formation activity with cosmic epoch, and the Butcher–Oemler effect, with an unprecedented accuracy. We are currently undertaking such an observational programme.

ACKNOWLEDGMENTS

REFERENCES

Abazajian K. et al., 2003, AJ, 126, 2081
Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681
Baldwin J., Phillips M., Terlevich R., 1981, PASP, 93, 5
Baum W. A., 1959, PASP, 71, 106
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJ, 585, L117
Bernardi M. et al., 2003a, AJ, 125, 1817
Bernardi M. et al., 2003b, AJ, 125, 1849
Bernardi M. et al., 2003c, AJ, 125, 1866
Bernardi M. et al., 2003d, AJ, 125, 1882
Blanton M. R. et al., 2001, AJ, 121, 2358
Blanton M. R. et al., 2003, ApJ, 592, 819
Blanton M. R. et al., 2003, ApJ, 594, 186
Bruzual A. G., Charlot S., 1993, ApJ, 405, 538
Bruzual A. G., Charlot S., 2003, MNRAS, 344, 1000
Eisenstein D. J. et al., 2003, ApJ, 585, 694
Faber S. M., 1973, ApJ, 179, 731
Feltzing S., Holmberg J., Hurley J. R., 2001, A&A, 377, 911
Fiala N., Rakos K. D., Stockton A., 1986, PASP, 98, 70
Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
Hauck B., Mermilliod M., 1998, A&AS, 129, 431
Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109
Hao L. et al., 2005a, AJ, 129, 1783
Hao L. et al., 2005b, AJ, 129, 1795
Hopkins A. M. et al., 2003, ApJ, 599, 971
Ivezić Ž. et al., 2002, AJ, 124, 2364
Ivezić Ž. et al., 2003, Mem. Soc. Ast. It., 74, 978
Kauffmann G. et al., 2003a, MNRAS, 341, 33
Kauffmann G. et al., 2003b, MNRAS, 341, 54
Kauffmann G. et al., 2003c, MNRAS, 346, 1055
Kennicutt R. C., 1998, ApJ, 498, 541
Kewley L., Jansen R. A., Geller M. J., 2005, PASP 117, 227
Lupton R. H., Ivezić, Ž., Gunn J. E., Knapp G. R., Strauss M. A., Yasuda N., 2002, in Tyson J. A., Wolff S., eds, Proc. SPIE Vol. 4836, Survey and Other Telescope Technologies and Discoveries. SPIE, Bellingham, p. 350
Obrić M. et al., 2006, MNRAS, in press (doi: 10.1111/j.1365-2966.2006.10675.x)
Obrić M. et al., 2004, IAUS, 222, 533
Odell A. P., Schombert J. M., Rakos K. D., 2002, AJ, 124, 3061
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Petrosian V., 1976, AJ, 209, L1
Rakos K. D., Dominis D., Steindling S., 2001, A&A, 369, 750
Rakos K. D., Schombert J. M., 1995, ApJ, 439, 47
Rakos K. D., Schombert J. M., 2004, AJ, 127, 1502
Sheth R. K. et al., 2003, ApJ, 594, 225
Shimasaku K. et al., 2001, AJ, 122, 1238
Spergel D. N. et al., 2003, ApJS, 148, 175
Stoughton C. et al., 2002, AJ, 123, 485
Strateva I. et al., 2001, AJ, 122, 1861
Strauss M. A. et al., 2002, AJ, 124, 1810
Strömgren B., 1966, ARA&A, 4, 433
Tremonti C., Heckman T. M., Kauffmann G., 2004, ApJ, 613, 898
Twarog B. A., 1980, ApJ, 242, 242
Vanden Berk D. E. et al., 2004, ApJ, 601, 692
Yip C. W. et al., 2004, AJ, 128, 585
York D. G. et al., 2000, AJ, 120, 1579
Zakamska N. L. et al., 2003, AJ, 126, 2125
Zakamska N. L., Strauss M. A., Heckman T. M., Ivezić Ž., Krolik J. H., 2004, AJ, 128, 1002

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