DESI complete calibration of the colour–redshift relation (DC3R2): results from early DESI data

J. McCullough [1,2,3]★ D. Gruen [3,4] A. Amon,[5,6] A. Roodman,[1,2] D. Masters,[7] A. Raichoor [8]
D. Schlegel,[8] R. Canning,[9] F. J. Castander [10,11] J. DeRose,[8] R. Miquel,[12,13] J. Myles [1,2] J.
A. Newman,[14] A. Slosar,[15] J. Speagle,[16] M. J. Wilson,[17] J. Aguilar,[8] S. Ahlen,[18] S. Bailey,[8] D. Brooks,[19]
T. Claybaugh,[8] S. Cole,[17] K. Dawson,[20] A. de la Macorra,[21] P. Doel,[19] J. E. Forero-Romero [22]
S. Gontcho A Gontcho [8] J. Guy,[8] R. Kehoe,[23] A. Kremin,[8] M. Landriau,[8] L. Le Guillou,[24] M. Levi,[8]
M. Manera,[13] P. Martini,[25] A. Meisner [26] J. Moustakas,[27] J. Nie,[28] W. J. Percival [29] C. Poppett,[8,30]
F. Prada,[31] M. Rezaie [32] G. Rossi,[33] E. Sanchez [34] H. Seo [35] G. Tarlé,[36] B. A. Weaver,[26] Z. Zhou,[37]
and H. Zou[37] DESI Collaboration

Affiliations are listed at the end of the paper

ABSTRACT

We present initial results from the Dark Energy Spectroscopic Instrument (DESI) complete calibration of the colour–redshift relation (DC3R2) secondary target survey. Our analysis uses 230 k galaxies that overlap with KiDS-VIKING *ugriZYJHK*$_s$ photometry to calibrate the colour–redshift relation and to inform photometric redshift (photo-*z*) inference methods of future weak lensing surveys. Together with emission line galaxies (ELGs), luminous red galaxies (LRGs), and the Bright Galaxy Survey (BGS) that provide samples of complementary colour, the DC3R2 targets help DESI to span 56 per cent of the colour space visible to *Euclid* and LSST with high confidence spectroscopic redshifts. The effects of spectroscopic completeness and quality are explored, as well as systematic uncertainties introduced with the use of common Self-Organizing Maps trained on different photometry than the analysis sample. We further examine the dependence of redshift on magnitude at fixed colour, important for the use of bright galaxy spectra to calibrate redshifts in a fainter photometric galaxy sample. We find that noise in the KiDS-VIKING photometry introduces a dominant, apparent magnitude dependence of redshift at fixed colour, which indicates a need for carefully chosen deep drilling fields, and survey simulation to model this effect for future weak lensing surveys.

Key words: gravitational lensing: weak – techniques: spectroscopic – surveys – galaxies: distances and redshifts.

1 INTRODUCTION

Modern cosmology relies on our ability to observe galaxies as tracers of the structure formation and expansion of the Universe. To do this we must map their on-sky positions and, crucially, their positions in all three dimensions. Measuring the redshift, *z*, for galaxies outside of our local group is a good proxy for this third dimension, because the expansion of the Universe reddens the light from a galaxy in a way that is monotonic with distance. The most accurate way to measure a galaxy's redshift is via the detection of prominent emission or absorption features with sufficient signal-to-noise ratio in the spectral energy distribution (SED). With this, the observed wavelengths of these features are compared to their known rest-frame wavelengths to provide a redshift measurement. Spectroscopic surveys obtain these many-wavelength observations very successfully via slit masks on

the telescope (e.g. on Keck; Oke et al. 1995), and more recently with integral field units (IFUs) (e.g. the Hobby Eberly Telescope Dark Energy Experiment; Gebhardt et al. 2021), and massively multiplexed instruments with independent optical fibres capable of taking many spectra at once (e.g. DESI; Flaugher & Bebek 2014; DESI Collaboration 2022). The majority of galaxies in our Universe are relatively faint and distant, and low-throughput galaxies are more feasibly observed photometrically – in optical, near-infrared, and infrared filters – rather than spectroscopically, due to limits in exposure time. Imaging surveys estimate the distances to galaxies from their colours, measured as the ratio of photometric fluxes in different bandpass filters. Although photometric redshifts are more readily attainable, these estimates are less accurate than spectroscopic redshifts.

Cosmological measurements from wide imaging surveys, like weak gravitational lensing or galaxy clustering, rely on geometric information. The most recent imaging surveys, such as the Dark Energy Survey (DES; Sevilla-Noarbe et al. 2021), Subaru Hyper

★ E-mail: jmccull@stanford.edu

Suprime-Cam (HSC; Aihara et al. 2022), and the Kilo-Degree Survey (KiDS; Kuijken et al. 2019), span a significant fraction of the sky. Accurate estimates of the redshifts of these galaxy samples based on limited information (e.g. photometry rather than spectroscopy) are required to obtain unbiased cosmological constraints. Indeed, one of the foremost difficulties facing imaging surveys for cosmology lies in calibrating the redshift probability distribution (Myles et al. 2021). Typically, redshift distributions are estimated and calibrated for an ensemble of galaxies, $n(z)$, and the uncertainty is modelled as an error on the mean redshift of the distribution in the cosmological analysis – as cosmological parameters are most sensitive to shifts in the mean-z (see e.g. Amon et al. 2022; van den Busch et al. 2022; Dalal et al. 2023; Li et al. 2023a). Calibrating the redshift distribution for an entire ensemble has unique difficulties compared to doing the same for individual galaxies, though individual redshifts have a multitude of other science applications. As ensemble redshift distributions are of the foremost interest to weak lensing cosmology, ensemble calibration is the focus of this paper.

With observations of fluxes in only a few broad-bands, the underlying challenge for determining an accurate redshift is a degeneracy between a galaxy's spectral phenotype and redshift (see Newman & Gruen 2022 for a review). A variety of approaches have historically been used to determine galaxy redshifts. Quiescent elliptical galaxies have a consistent drop in light emission at 4000 Å, which enables accurate photo-zs with this so-called red sequence of galaxies, which allows the bounding of the redshift between different band passes and is particularly useful for identifying cluster members and lensing galaxy samples (e.g. the RedMaPPer algorithm in Rykoff et al. 2014, 2016). While early-type, passive galaxies benefit from having very similar SEDs, this is not necessarily true for many other galaxy types. Template fitting methods can rely on either spectroscopically informed templates or semi-analytic models that are typically constructed with stellar populations (e.g. Brammer, van Dokkum & Coppi 2008). An empirical variation of this is a Principal Component Analysis (PCA), but in essence both methods fit the data to a linear combination of templates (see Salvato, Ilbert & Hoyle 2018 for a review). Regardless of the method for redshift estimation, redshift-type degeneracy is an irreducible problem when working with photometry (see discussion of methods that break age-mass-redshift degeneracy in e.g. Wang et al. 2023).

A correct model for the galaxy population (i.e. the mix of templates and their luminosity functions at each redshift, or a large, fully representative reference sample of galaxies with known spectroscopic redshift) would be required to determine correct $n(z)$ for photometric samples despite redshift-type degeneracies. Both of these solutions at present appear unfeasible, though gains in forward modelling the distribution have been made (e.g. Alsing et al. 2023). The issue can be greatly reduced by observing in additional bands that break degeneracies, which motivates a filter set that goes beyond the standard optical broad-bands (Buchs et al. 2019; Wright et al. 2019). With this increased wavelength coverage, however, the colour space that the photometric observations occupy becomes high-dimensional. The challenge is to associate like-spectroscopic galaxies with photometric galaxies efficiently across that high-dimensional space.

Self-Organizing Maps (SOMs; Kohonen 2004) can serve as a useful tool to subdivide the high-dimensional colour space into a set of SOM cells efficiently, tracing density and coherent galaxy types, as demonstrated in Masters et al. (2015) and utilized in Hildebrandt et al. (2020) and Myles et al. (2021) among others. The use of SOMs for redshift calibration demand spectroscopic galaxy samples to completely populate this space, thereby providing

accurate galaxy redshifts for any combination of colours. Of critical note is that spectroscopic redshifts are typically obtained only for a specific and limited selection of galaxies and must be weighted to become representative of the photometric sample (Hartley et al. 2020). The resulting calibration problem is that without a complete spectroscopic sample that fully populates the photometric colour-space derived redshift distributions are subject to bias and uncertainty given incomplete or undersampled spectroscopic observations.

Some 5000 spectroscopic redshifts have recently been determined by the complete calibration of the colour–redshift relation (C3R2) project in order to populate this colour space and thereby calibrate the colour–redshift relation (Masters et al. 2017, 2019; Stanford et al. 2021). Beyond fully populating each SOM cell, a larger multiplicity of spectroscopic galaxies per SOM cell will be needed to meet requirements for future deep imaging surveys like *Euclid* (Amendola et al. 2013) and Rubin Observatory (Collaboration et al. 2009; Ivezić et al. 2019). This can be achieved via statistical characterization of broad or bimodal redshift distributions in SOM cells, though a well constructed SOM minimizes these features where possible. Additionally, as upcoming imaging surveys are deeper than most spectroscopic samples, it is essential that the magnitude dependence of the redshift at a fixed colour, dz/dm, is understood. Previous examinations of this measurement have shown a small dz/dm trend at fixed colour (Masters et al. 2019).

In this paper, we present the DESI complete calibration of the colour–redshift relation (DC3R2) secondary target survey. This survey supplements the existing spectroscopic samples used for photometric redshift calibration by both populating SOM cells that were previously unfilled, and by increasing the multiplicity per cell. We use DC3R2 to revisit the magnitude dependence of the redshift at a fixed colour with improved statistics, including apparently brighter galaxies, and study trends in dz/dm as a function of colour. Section 2 introduces the survey data used. In Section 3 we describe the construction of the DC3R2 sample. We explore how it calibrates the colour–redshift relationship alongside DESI main survey targets in Section 4. Finally, using this new resource, we examine the magnitude dependence of redshift at fixed colour in the presence of observational effects like photometric scatter (Section 5.2).

2 DATA

DC3R2 is a secondary target programme on the Dark Energy Spectroscopic Instrument (DESI), Section 2.1, that obtained spectroscopic redshifts for galaxies that were targeted with KiDS-VIKING (KV) photometry, Section 2.2. We also analyse DESI main survey targets that fall within the KiDS-VIKING footprint, but were targeted on photometry from the DESI Legacy Imaging Surveys, Section 2.3, including the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019). Additional spectroscopic and imaging surveys were used to validate this analysis, as listed in Section 2.4. For reference, all magnitudes in this paper are absolutely calibrated in the AB system (see Oke & Gunn 1983).

2.1 DESI

The DESI is a ground-based spectroscopic experiment installed at the 4m Mayall telescope (Collaboration et al. 2016a; DESI Collaboration 2022). The DESI instrument is sensitive from 360–980 nm, with 5000 robotically actuated fibres that are capable of taking spectra simultaneously. Over five years, it aims to measure spectra of 40 million galaxies and quasars that will aid in examination of baryon

Table 1. Generalized target selections for each survey contributing to this analysis and their complementary redshift sensitivities and the resulting observed spectra counts before quality cuts are applied. All magnitude cuts are reported in the optical Legacy Survey bands, *grz*, made use of for targeting. Note that the BGS fibre magnitude cuts are colour dependent and simplified here, and that objects may be shared between different target classes. See Fig. A2 for a breakdown of the interplay of each sample in colour–magnitude space.

Sample	No. Spec	Redshift range	Magnitude cut	Colour selection
DC3R2	51 177	$0 < z < 1.6$	$z_{\mathrm{fibre}} < 22.10$	SOM, KV *ugriZYJHK$_s$*
LRG	46 771	$0.4 < z \lesssim 1.0$	$z_{\mathrm{fibre}} < 21.60$	$grz + W1$
ELG	125 027	$0.6 < z < 1.6$	$g > 20, g_{\mathrm{fibre}} < 24.1$	*grz*
BGS bright	156 895	$0 < z \lesssim 0.6$	$r < 19.5, r_{\mathrm{fibre}} \lesssim 22.9$	magnitude limited
BGS faint	–	$0 < z \lesssim 0.6$	$19.5 < r < 20.175, r_{\mathrm{fibre}} \lesssim 22.9$	$grz, + W1$

acoustic oscillations (BAO), the growth of structure through redshift-space distortions and dark energy (Collaboration et al. 2016b). While these are DESI's primary goals, the survey is uniquely capable of providing a multitude of spectroscopic redshifts that have far-reaching uses. Here we exploit its ability to be used to calibrate photometric redshifts, in line with the need of weak gravitational lensing experiments.

The DESI main survey targets relevant to this paper are divided into three galaxy types: luminous red galaxies (LRGs; Zhou et al. 2023), emission line galaxies (ELGs; Raichoor et al. 2023), and the Bright Galaxy Survey (BGS; Hahn et al. 2022). Notably, LRGs are magnitude selected on a bright *z*-band fibre-magnitude alongside *grz* and NIR photometry from *WISE* W1 to target red galaxies from 0.4 < z < 1.0 with high signal-to-noise (Zhou et al. 2023). ELGs, on the other hand, are selected on a *g*-band magnitude alongside *grz* colour cuts to target star forming galaxies from 0.6 < z < 1.6 (Raichoor et al. 2023). Lastly, BGS targets an *r*-band magnitude limited sample (split into a bright and faint subset) that is significantly brighter than the other main surveys, $z \lesssim 0.6$, capable of observation during sub-optimal conditions (Hahn et al. 2022). Each of these main survey samples targets in a complementary part of the colour–redshift space, though their union does not span the full census of galaxy populations DESI is capable of measuring. For precise photometric cuts for each sample, we direct the reader to the respective target selection papers, which use the photometry of the DESI Legacy Imaging Surveys (Dey et al. 2019). For a brief overview of these selections, see Table 1.

The methodology for fitting models to the obtained spectra are explored in Bailey et al. (in preparation), and the validation of these techniques is performed in Lan et al. (2023). The observations that produced data for this analysis come from 2020 December 14 through 2021 July 9, which span a combination of the 'One-Percent Survey' (OPS), Survey Validation (SV), and the beginning of the main survey operations (Y1) – the internal *Fuji* and *Guadalupe* data releases, respectively (DESI Collaboration 2023). SV data have already been made available in the DESI Early Data Release (EDR) (DESI Collaboration 2023). Selection changes for main targets were subject to minute changes between SV1 and the OPS as well as Y1 operations, detailed in the respective paper for each sample. The selection footprint for DC3R2 was modified after the OPS and before Y1 to make use of newly released photometry. Additionally, on 2021 May 12 and 13, a small dedicated tile programme was run for DC3R2 targets with high priority. The selection for DC3R2 targets is outlined in greater detail in Section 3.1, and the optimization for dedicated tile fibres is discussed in Appendix C.

2.2 KiDS-VIKING

The Kilo Degree Survey (KiDS) is a large scale optical *ugri* imaging survey with OmegaCAM on the VLT Survey Telescope (VST)

at the ESO Paranal Observatory (Arnaboldi et al. 2000; Kuijken et al. 2015). Its footprint is overlapped by a near-infrared *ZYJHK$_s$* VIRCAM photometric survey with the 4 m Visible and Infrared Survey Telescope for Astronomy (VISTA), the VISTA Kilo-degree Infrared Galaxy Public Survey (VIKING). The two surveys both span more than a thousand square degrees in the *ugriZYJHK$_s$* bands, to a depth of $r \leq 25$. Their complementary wavelength coverage has been processed jointly to create the 9-band KiDS-VIKING (KV) survey (Wright et al. 2019). This data set provides dereddened, multiband colour information for an overlapping patch of the DESI SV and Y1 footprint, allowing us to crucially associate DESI spectroscopic redshifts with a high-dimensional colour space.

In this analysis, specifically, we make use of the KiDS-450 data release, with observations spanning the Galaxy And Mass Assembly (GAMA) fields (Driver et al. 2011), depicted as G09, G12, and G15 and the shaded green regions in Fig. 1 (de Jong et al. 2017). Through the beginning of main survey operations, we also used the KiDS-1000 release, shown in the blue footprint of Fig. 1, which jointly with KiDS-450 expands the overall footprint and provides a superset of optical and near-infrared photometry (Kannawadi et al. 2019; Wright et al. 2019; Hildebrandt et al. 2020). The only selections we apply to this calibrated photometry is the provided photometric quality flags (e.g. FLAG_GAAP_u) which conservatively mask out (for all nine bands) pixels affected by diffraction spikes, oversaturation, neighbouring bright stars, or other effects in the final coadded images. We require all bands to perform our analysis, so any galaxy with flagged poor photometry in a single band is eliminated from our targeting.

2.3 DESI legacy imaging surveys

In order to target the main survey samples for DESI, *grz* optical photometry was collected on across three instruments to produce the DESI Legacy Imaging Surveys (Dey et al. 2019): the Mayall *z*-band Legacy Survey (MzLS) and the Beijing-ARizona Sky Survey (BASS; Zou et al. 2017) at Kitt Peak National Observatory, and the Dark Energy Camera Legacy Survey (DECaLS) at the Cerro Tololo International Observatory. DECaLS extended the photometric catalogue of the Dark Energy Survey (DES; e.g. see Sevilla-Noarbe et al. 2021 for photometry or Dark Energy Survey Collaboration 2016 for an overview), while MzLS and BASS supplied the same filters at complementary high declination ($\delta > +34$). The photometric catalogue spans approximately 14 000 deg^2 in the Northern hemisphere, relevant for DESI and this analysis. Additionally, the optical reach of these surveys was augmented by the *Wide-field Infrared Survey Explorer* (*WISE*) bandpasses (3.4–22 μm, W1, W2, W3, W4) (Wright et al. 2010; Cutri et al. 2013). More extensive information on the
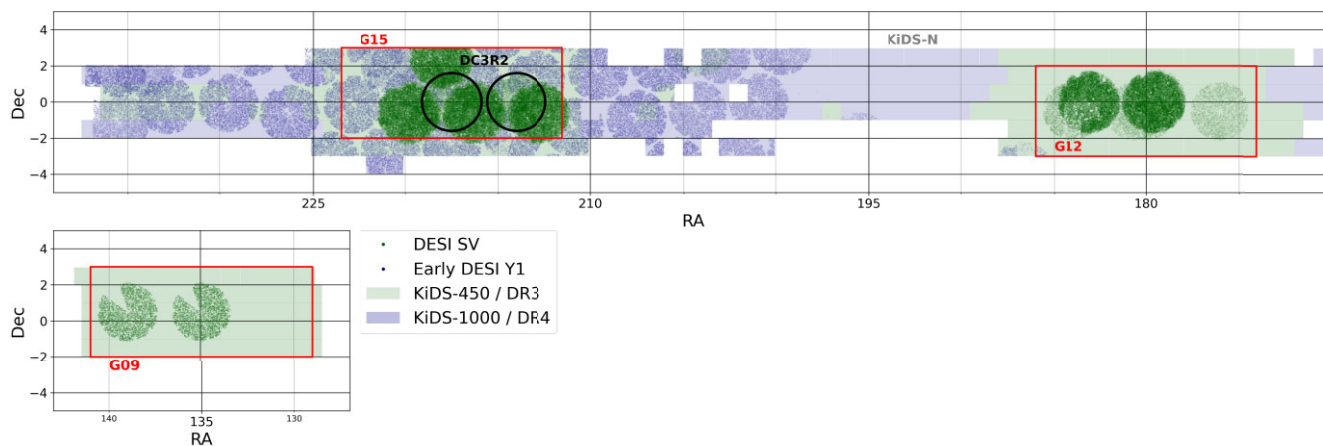
Figure 1. Footprint of spectroscopic redshifts used in this analysis, including main survey targets, depicting objects observed in SV (green points) and through the first 56 d of Y1 operations (blue points). The combined bright and faint footprint for the DC3R2 dedicated tiles are outlined in circles. The broad KiDS-VIKING-N equatorial field provides the *ugriZYJHK$_s$* photometry to match to DESI data, which is inclusive of KV DR4 (shaded) and DR3 (light shaded) that spans the overlap of KiDS-1000 with the DESI footprint (Kuijken et al. 2019). Particularly relevant for SV, prior to the KiDS-1000 release, the majority of our targets lie in the GAMA fields (boxed).

depths and calibration of the Legacy Surveys can be found on the data release page.[1]

2.4 Other spectroscopic surveys

A multitude of spectroscopic surveys have been undertaken in the COSMOS field. Several of these were utilized in this analysis for validation. Among these surveys are the original C3R2 effort (Masters et al. 2017, 2019; Stanford et al. 2021) and the master spectroscopic catalogue from the COSMOS collaboration (M. Salvato, in preparation). The latter includes observations from a variety of wavelength regimes and spectral resolutions across many instruments (VLT VIMOS, VUDS, Keck MOSFIRE, DEIMOS, Magellan IMACS, Subaru FMOS, and many others (Lilly et al. 2007; Trump et al. 2007; Kartaltepe et al. 2010; Balogh et al. 2014; Kriek et al. 2015; Le Fèvre et al. 2015; Silverman et al. 2015; Casey et al. 2017; Hasinger et al. 2018). For the use of this project, these samples were limited to only confident redshifts. Furthermore, they were matched to the Masters et al. (2017) photometry and assigned to the original C3R2 SOM in order to validate our colour–redshift relation.

3 METHODS AND OBSERVATIONS

From 2020 December through 2021 July, DESI observed 328 k main survey targets (ELGs, LRGs, BGS) in the KV footprint, 51 177 of which were also selected as DC3R2 and 1216 that were exclusively DC3R2 targets. The following sections will further break down these numbers into the DC3R2 secondary targets during SV (3.1), Y1 (3.1.2), and our dedicated tiles (3.1.1), and selected main survey targets that overlap. After completeness cuts we find that we have a sample of 230.7k galaxies that occupy the colour space from 0.0 < z < 1.55. The aim of this programme is to fill as much of the high-dimensional colour–magnitude space for future weak lensing surveys as possible, with high multiplicity and as much depth as is feasible when limited to secondary targets. The sample reported here spans the breadth of 56 per cent of the population anticipated for Rubin and Euclid analyses, effectively most of the galaxies anticipated at z <

1.55. Future work can expand to higher redshift populations on NIR and IR instruments. Crucially DESI is well suited to extending this analysis and examining deeper samples in this lower redshift regime with dedicated programmes.

In the following sections, we describe our procedure for target selection (3.1) through the three major phases of the DC3R2 programme: the dedicated tiles (3.1.1), and SV, Y1 spare fibres (3.1.2). Additionally we detail the observations (3.2), redshift completeness (3.2.2), and the weighting schema to provide a representative sample of redshifts for calibration purposes (3.2.3).

3.1 Target selection

The primary DC3R2 targets span across the GAMA-9h, 12h, and 15h equatorial fields (Driver et al. 2011) for a total of 300 deg^2, where sufficiently deep *ugriZYJHK$_s$* colour information was available. We have matched the GAMA fields as reported by KiDS-VIKING (KV) DR3 (Wright et al. 2019) to the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019) Data Release 9 photometry in order to constrain colour alongside DESI z-band fibre fluxes, using the closest match within 1 arcsec. The fibre fluxes from DECaLS were used to ensure our programme could be completed at low exposure cost. In order to calibrate the complete colour redshift relation, DC3R2 aimed to occupy with multiplicity as much as reasonably attainable to DESI of the colour-space described by the Masters et al. (2017) SOM, which is a useful map trained on narrow band COSMOS photometry that spans the approximate depth and breadth of colour for future weak lensing surveys like *Euclid* and LSST. Appropriately it is the subject of several redshift search programmes that also aim to span this colour-space with spectroscopy (Masters et al. 2015; Saglia et al. 2022), which further encourages our use of it. This SOM is then corrected to KiDS-VIKING colours according to the procedure described in Section 4.1, to allow for assignment of our alternate photometric bands. The abundances of galaxies observed across the colour-space for these KiDS-450 fields (G09, G12, G15) can be found in Fig. 2(a). In the neighbouring panel, Fig. 2(b) depicts the distribution of galaxies observed from this catalogue with successfully measured redshifts.

Galaxies are selected by colour, determined by their assignment to cells in this map, as well as by fibre magnitude. We select these
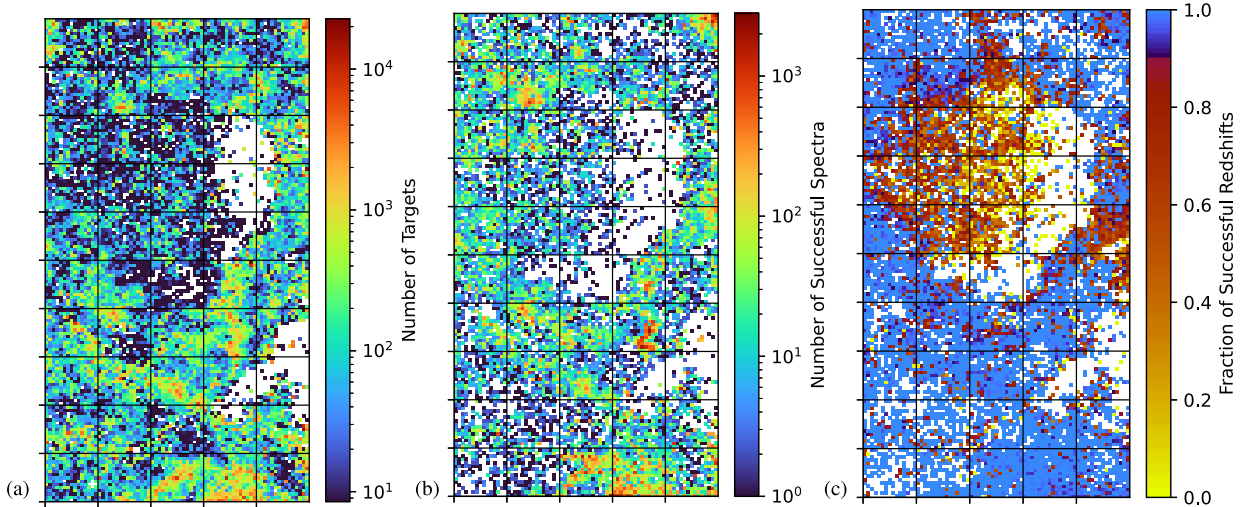
Figure 2. The distributions of targeting, spectroscopy, and redshift completeness across the colour-space. (a) Target abundance by *ugriZYJHK_s* colour self-organizing map cell in the relevant KiDS-VIKING 450 fields (G09, G12, G15), from which spectroscopic targets for DC3R2 were chosen for SV and the dedicated tiles. Objects included in this count meet either DC3R2 selections or any main survey (ELG, LRG, BGS) magnitude cuts in *grz*. (b) Number of successful spectra taken per cell, across DC3R2 as well as all DESI main surveys (ELG/LRG/BGS) that pass redshift quality cuts. (c) Fraction of spectra that meet our redshift quality selections at a given colour, across all surveys, out of all objects observed. If the fraction of spectra in the cell with certain redshifts amounts to less than 90 per cent we consider the spectra that are measured incomplete and potentially biased and exclude the cell from the fiducial analysis.

colours on mean cell magnitude, to take into account visibility of the targets to the fibre, and on a high probability of redshift < 1.6 (i.e. the [O II] feature lies within the DESI wavelength window, discussed further in the following paragraphs). This allows us to target 3692 cells from the C3R2 SOM (≈33 per cent) with a minimum of three galaxies per cell, spread in magnitude. These targets were chosen to enable a first quantification and rejection of faulty photometry and redshift and examine the trend of redshift with magnitude at fixed colour, sorely required for accurate calibration efforts in the future (Masters et al. 2019). From 1800 targets per sq. deg. DC3R2 required only a small random subset (36 per sq. deg., or 2 per cent for our initial request of three galaxies per cell, making it an ideal candidate for a secondary target programmes with very flexible fibre assignment across a high density of targets. While ideally, DC3R2 would sample the full colour space available to DESI, i.e. no selection on z < 1.6, to fully probe the edge of the redshift range without a photometric redshift prior, these galaxies are realistically too faint on average for a spare fibre programme to address in a few 20 min exposures. We benefit from main survey target classes (ELGs especially) probing this redshift regime with higher fidelity. We can summarize the diverse range of target selections in this paper in Table 1, where we find DC3R2 is the broadest in terms of redshift, but augmented by deeper main survey samples.

Explicit choices for the DC3R2 target selection procedure follows from the joint KiDS-VIKING-DECaLS matched catalogue,

(i) To boost redshift completeness of our targets, individual galaxies that would conservatively be expected to take four visits or fewer were selected from the catalogue as MAG_FIBER_Z < 22.10 – too acheive a balmer break SNR of approximately one in the optical, where passive galaxies are the conservative upper bound, see the discussion in Zhou et al. 2023 (their section 4.4) for more information.

(ii) Potential targets were assigned to the Masters et al. (2017) SOM, after transforming the latter to the KiDS-VIKING *ugriZYJHK_s* colour-space (see Section 4.1).

(iii) An envelope of SOM cells (not contiguous) that contain galaxies bright enough to achieve redshift success with two visits (mean(MAG_FIBER_Z) < 21.88 and >95 per cent probability of redshift < 1.6) were selected. The probability of redshift < 1.6 for a given cell was obtained from the COSMOS15 (Laigle et al. 2016) photometric redshifts for each cell as in Masters et al. (2015).

(iv) The full target catalogue was defined as the concatenation of all cells within the above envelope that each had at minimum three viable detections (believed to have observable redshifts in four or fewer DESI visits). From this we prioritized the brightest in each cell (as a lever arm for the measurement in Section 5.2) and multiplicity as described for each component of our survey in the respective section below.

The initial spare fibre catalogue for the SV stage had approximately 970k targets available that met the above criteria. This target list was modified for Y1 to include the larger KiDS-1000 footprint (Kuijken et al. 2019), see 3.1.2 for more information. Our ability to prioritize targets is severely limited by the secondary nature of our programme, and where we do have ability to prioritize is discussed in more depth in Appendix C for the dedicated tiles. We are agnostic to pre-existing spectroscopic coverage in the field and simply cover any available target by chance target-fibre vacancies.

3.1.1 Dedicated tiles

DC3R2 has observed two dedicated tiles, i.e. instrument pointings, at the end of the One Percent Survey within the survey validation period. We chose the two pointings to be centred at 217.5 and 221.0 deg RA on the celestial equator, within the GAMA 15h field (Driver et al. 2011) and also within the Hyper Suprime-Cam Subaru Strategic Program (HSC) survey area (Aihara et al. 2018). Each tile was observed with two different fibre configurations, one for a single 30 min exposure aiming at targets down to z_{fibre} < 21.5, and one for a total of 90 min of exposure time with targets down to z_{fibre} < 22.10. A description of how targets were chosen for each of these pointings can be found in Appendix C, but in summary we optimize priority

for bright-faint pairs (a large span of MAG_GAAP_Z) in a given cell to maximize our lever arm on measuring dz/dm. The dedicated tiles are identified in Fig. 1 as black circles.

3.1.2 Spare fibres

The DC3R2 targets observed outside of the dedicated tiles occurred in two phases within the SV and Y1 periods. During SV, DC3R2 took observations for 44 272 targets. In contrast, during the onset of Y1 6905 DC3R2 objects were observed.

The initial spare fibre, *secondary target*, observations are described by the selections made in 3.1 on the KV DR3, KiDS-450, photometry released at the time.

The spare fibre target selection and prioritization for the Y1 observing period was modified from the initial SV phase to draw targets from the then publicly available larger area KiDS-1000 (DR4) photometric catalogue (Kuijken et al. 2019), as seen in the shaded blue regions of Fig. 1. The strategy was altered to prioritize the brightest galaxy in each cell alongside a randomly drawn target within a Z-magnitude of 21.88. This lower magnitude cut ensures a redshift is obtainable from two visits, which differs from the maximum of four visits in SV, seldom available for spare targets. The random draw traces the magnitude distribution of the cells overall. This pair-wise selection per cell has the benefit of extending the magnitude leverage on dz/dm, as more bright galaxies are available in the larger footprint, and achieving high redshift success in survey spare fibre mode. The same criterion as in Section 3.1 allowed us to select from 5074 cells in the C3R2 SOM for the Y1 fibre proposal. The blue objects in Fig. 1 are successful matches back to this wider catalogue during Y1, sometimes retroactively from main survey targets in the SV phase. The DC3R2 spare fibre programme extended into Y1, and while some of that data are analysed here, we expect an additional 315k targets within our footprint to become available with the Y1 data release, with around 28k of those objects being DC3R2 exclusive targets and the remainder coming from overlap with DESI main survey target classes.

3.2 Observations

The positions of all observed DC3R2 targets on the sky, falling on the equatorial where nine-band photometry is available, are depicted in Fig. 1, including the dedicated tiles in black. An individual DESI pointing can capture up to 5k targets across 8 deg^2, which is visible as the overlapping circular rosettes in this figure.

3.2.1 Redshift determination

Following pixel level calibration and extraction of a 1D spectrum, the DESI redshifting pipeline REDROCK forward models the observed data from a basis consisting of a set of template SEDs for each target class (Bailey et al. in preparation; Guy et al. 2023). This decomposition is done at each point in a fine grid of redshift values z and the χ^2 of the difference of observed data and the best linear combination of templates is determined. A minimum in the $\chi^2(z)$ indicates a potentially optimal redshift-template solution. The z that globally minimizes χ^2 is taken as the redshift of the observed object.

3.2.2 Completeness

The primary metric of redshift confidence in the DESI survey is $\Delta\chi^2$, the difference between the χ^2 of the best-fitting redshift and template combination and the second lowest local minimum of χ^2

as a function of redshift. The larger the $\Delta\chi^2$, the more confident one can be that the first redshift is correct (Bailey et al., in preparation). As each main survey sample selection (Hahn et al. 2022, Zhou et al. 2023, and Raichoor et al. 2023) has different characteristic SED features, the ability of REDROCK to fit a given spectrum depends on the type of galaxy. The DESI visual inspection efforts have determined minimum $\Delta\chi^2$ selections and additional quality cuts that maximize purity and completeness for each main survey sample (Guy et al. 2023; Lan et al. 2023).

For this analysis, DC3R2 galaxies that we consider to have high confidence redshifts must pass the same completeness criteria as the BGS sample, i.e. $\Delta\chi^2 > 40$. This is the strictest among the $\Delta\chi^2$ cuts for DESI main survey samples and was chosen to account for the broad variation of SED types in our target selection, informed by the visual inspection done for the BGS sample in Lan et al. (2023) that optimizes good redshift purity and inclusion for our comparably bright targets. ELGs and LRGs have significantly lower $\Delta\chi^2$ requirements in the DESI main survey, which combined with further quality cuts give high fractions of confident redshifts, as REDROCK and its template sets have been optimized to obtain good fits to these SEDs. Galaxies with different SEDs may need more conservative metrics to minimize outliers, hence our adoption of the $\Delta\chi^2 > 40$ cut. Our analysis indicates this is more than sufficient in most regions of colour-space, but it ought to be kept in mind that $\Delta\chi^2$ is only one avenue of selecting robust redshifts and the threshold required depends strongly on the SED type of the galaxy. Other features, e.g. the signal-to-noise in strong features like [O II], may also be strong predictors of successful redshift measurements in REDROCK.

We report the distribution of successful redshifts in the SOM from all target classes based upon their individual success criteria in Fig. 2(b). While a cell may have successful redshift measurements, those measurements may be biased in certain regions of colour-space where obtaining good redshifts is more difficult. How we account for this bias is described in the following section. The redshifts measured from these successful spectra are visible in Fig. 3(a) after accounting for these incomplete regions. We can see in Fig. 3(b) the general agreement with previous spectroscopy in the COSMOS field. The colour-redshift calibration is further discussed in Section 4.2.

For the sake of future survey efforts to calibrate the colour-redshift relation, in Fig. 4 we report the exposure times required by DESI for targets of eight-band colours in the SOM to achieve $\Delta\chi^2 > 40$ for objects scaled to a magnitude of MAG_GAAP_Z = 21.0, as following from the discussion in Appendix D.

For the fibre configurations with 90-min exposure time over the DC3R2 dedicated tiles, 98.1 per cent of targets achieved the DC3R2 criterion for success with no flagged warnings (ZWARN = 0). For the galaxies in the two single 30-min exposures, 93.4 per cent meet this redshift success criterion. The overall redshift success rate for the dedicated tile targets is 95.8 per cent. For the entire DC3R2 ancillary programme that also made use of spare fibres (footprint depicted in Fig. 1), the success rate is 93.4 per cent for unique DC3R2 targets, and 94.7 per cent for shared targets, with a total of 13 270 targets observed during the dedicated programme. While many of these overlap with main survey targets, the higher DC3R2 priority in fibre selection ensures that the selection of these objects are less biased towards the main survey selections in colour than similar overlaps in the spare fibre fields.

3.2.3 Sample weighting scheme

One of the primary methods of photometric redshift calibration in weak lensing is to appropriately reweight a sample of known
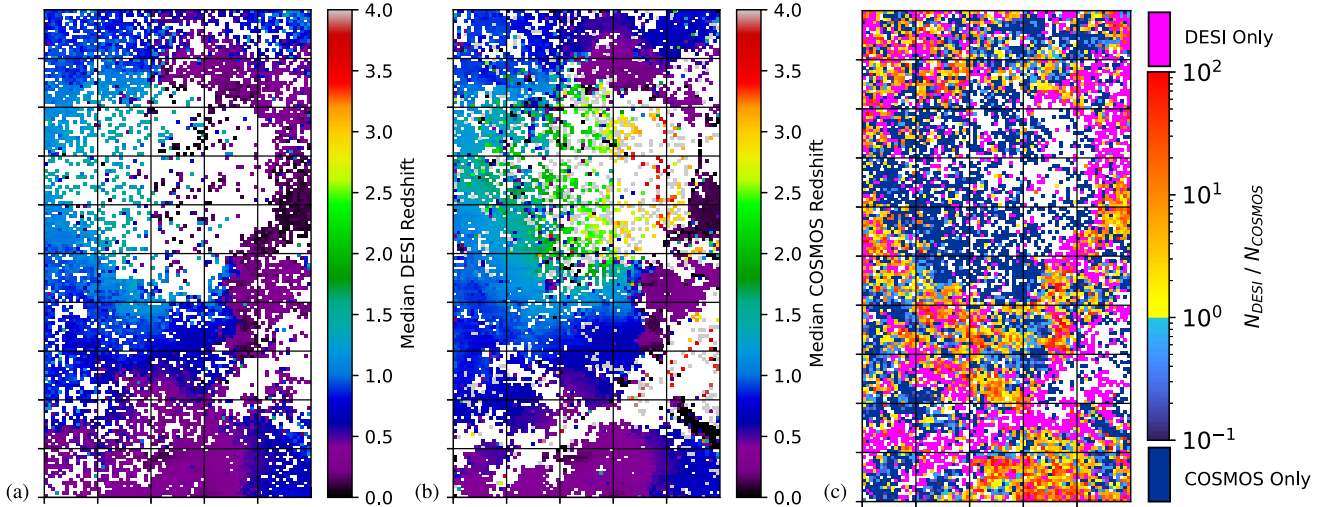
Figure 3. Median DESI redshifts compared to previously measured and curated spectroscopy in the COSMOS field reveals substantial gains in the low redshift regions and overall multiplicity. Note that these colour-spaces are not identical and this comparison is a generalized one, as per the discussions in Section 4.1, Appendix B. (a) Median spectroscopic redshifts for a given cell for all DC3R2 and main survey targets that meet our quality cuts. (b) Median spectroscopic redshifts for a given cell for the COSMOS field (as discussed in 2.4), from prior, non-DESI spectroscopy. (c) The ratio of spectroscopic redshift counts between initial DESI efforts and prior COSMOS data for the same map, where solid values depict spaces exclusively added by DESI compared to COSMOS. The transition occurs where DESI doubled existing spec-z counts.
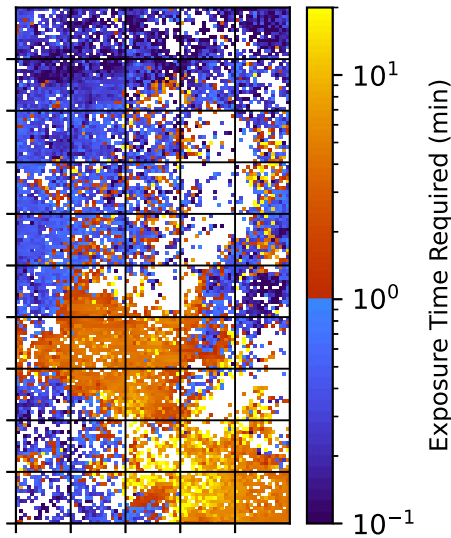
Figure 4. Median DESI exposure times (in minutes) necessary for each SOM cell, for targets with a DECam fibre magnitude of $z_{fiber} = 21$, to reach a high redshift confidence $\Delta\chi^2 = 40$, as per Appendix D. Full exposure time quantiles and median colours for each cell are reported in the attached data products. BGS targets are excluded from this plot, due to their typically much larger sky brightness. The split colour bar is chosen to approximately separate quiescent, passive galaxies (lighter orange) from emission line galaxies (blue), which occurs for these magnitudes near one minute of exposure.

redshifts to accurately approximate the redshift distribution of a much larger source galaxy sample. This requires a full understanding of the selection acting on both samples. For the true redshift sample, this selection can be intentional in the definition of a spectroscopic target catalogue or unintentional due to incomplete redshift recovery among the targeted galaxies. Only with the selection accounted for an in addition with the weak lensing source galaxy sample selection

applied can the reweighted spectral sample be representative and the estimated redshift distribution therefore be unbiased.

For this survey, we must take into account the overlap of DC3R2 targets with DESI main survey targets. The selection of the latter is not based on KiDS-VIKING colours but on DESI Legacy Survey g, r, z, $W1$, $W2$ photometry (Myers et al. 2023). The observation prioritization of the DESI main survey targets may affect relative abundances of certain redshifts within a given SOM cell. This is accounted for by reweighting according to different subselections of our targets, namely:

(1) **Dedicated tiles**: Our dedicated tile sampling method over a small area will be unaffected by DESI main survey oversampling. It is our *fiducial* sample for this reason.

(2) **DC3R2 exclusive targets**: These are targets observed during regular DESI operations that are only observed as targets of our dedicated programme, i.e. that do not meet DESI main survey target selection criteria. This sample will be potentially biased *away* from the types and redshifts of galaxies observed by DESI's main surveys in SOM cells that contain a mix of both.

(3) **DESI main survey targets**: DESI main survey targets that overlap with our selection will be more likely to be observed due to their high priority in comparison with DC3R2 exclusive targets in SOM cells that contain a mix of both. Each class of main survey targets has different magnitude and colour cuts than the DC3R2 targets. SV and Y1 main survey targets have different colour selections that have evolved over the course of the survey and are weighted accordingly.

We can re-weight these subselections to leverage the full DESI survey beyond our fiducial sample, (1). The basic ruleset for usage and weighting is described here, with the aim to be providing a reliable and unbiased sample of spectroscopic galaxies over as many cells as possible with a large magnitude span within available cells. We want each cell to have a collection of representative galaxies that are corrected for overabundances caused by preferential observations

of galaxies of a particular SED type. For all DESI-observed galaxies that have KiDS-VIKING photometry in a given cell:

(i) We split the sample in the cell into categories of (2) and (3) by observation flags DESI_TARGET and SCND_TARGET, available bitmasks in the DESI catalogue that indicate main survey and secondary programmes, respectively.

(ii) For those in (2), we apply a cut on redshift confidence and warning flags (DELTACHI2 > 40, ZWARN = 0).

(iii) For those in (3), we check how complete the *targeting* is for main survey targets in colour space by comparing the full main survey target catalogue occupation for the cell (with only colour cuts applied) to the full KiDS-VIKING occupation of that cell. If the targeting completeness (fraction of photometric targets that are also main survey targets) is below 90 per cent for a given colour, we exclude the main targets from this cell and skip the following steps. This effectively ensures that the colour cut on the main survey sample does not significantly bisect a SOM cell. As the period covered by the observations spanned several iterations of LRG/ELG cuts (see Myers et al. 2023), only cells where both the stricter Y1 cut and the looser SV cuts did not bisect the cell were retained.

(iv) For those in (1), (2), and (3), we check the redshift completeness of the sample type in the given cell. This is done on the subset of galaxies in a cell that pass the specific survey target selection (e.g. the DECaLS based magnitude and colour cuts, or DC3R2 magnitude selection). If among this subset within a cell the redshift completeness is less than 90 per cent, we exclude all of these for both DC3R2 and main survey galaxies from the weighting scheme due to the risk of a significant redshift-dependent selection bias, and do not include low confidence targets in the next step. See a simple visual description of this in Fig. 2(c). The large number of incomplete spectra seen in the centre top left of this figure do not lack sufficient exposure time, but spectral features within the DESI sensitivity range. The galaxies likely to be $z > 1.6$ predominantly lack spectral features in the optical.

(v) We weight the samples of (2) and (3) high confidence redshifts so that both samples contribute proportionally to their abundances in the targeting catalogue, within a given cell and bin of z_{fibre}. If only one sample type exists in a given magnitude/cell bin, no such re-weighting is done.

We see a linear relationship be written explicitly as follows. For a given magnitude/colour bin b within the g, r, z magnitude cube of a given SOM cell c let N^a denote the number of actual (or *observed*) and N^p the number of potential targets. Further, denote the unique sample (ELG, LRG, BGS (B), BGS (F), DC3R2 spare fibres, DC3R2 dedicated tiles) of a target by the index s. Where each class of galaxy lives in the SOM is delineated by colour in Fig. 5, where we can see each of the samples are highly complementary. The weight of the redshift of an object towards the estimated redshift distribution of a cell c depends on s and b and can be factorized as

$$w(b, c, s) = w_b(b, c) \times w_s(b, c, s), \qquad (1)$$

where $w_b(b, c)$ is the weight of the magnitude bin towards the full sample in the cell and $w_s(b, c, s)$ is the weight of the sample that the galaxy belongs to for the given cell and magnitude bin. Individual samples in magnitude–colour bins require these weights to properly reproduce relevant target abundances in our final reported redshift distributions. An illustration of how samples populate a choice of magnitude bins is visible for the SV3 selections in Appendix A, Fig. A2 where a more thorough description of magnitude bins can be found. For objects that are in more than one sample (typically only true for some DC3R2 targets with main survey samples), the largest

of its sample weights is used.

$$w_b(b, c) = \frac{N_b^p / N_c^p}{N_b^a / N_c^a} \qquad (2)$$

Here $N_c^{a,p}$ and $N_b^{a,p}$ refer to the number of actual or potential targets in cell c or in magnitude bin b of that cell, respectively. Similarly, the reweighting of samples among that bin and cell is given by

$$w_s(b, c, s) = \frac{N_{s,b}^a / N_b^a}{N_{s,b}(\text{spec}) / N_b(\text{spec})} \qquad (3)$$

No reweighting is done for objects in (1) aside from the redshift completeness check, as these targets are not expected to be biased within a cell. The distribution of weights normalized for each SOM cell, $\Sigma_{s \in c} \Sigma_{b \in c} w_s(b, c, s) = N_c(\text{spec})$, are depicted in Fig. 6 for the main survey samples. The purpose of these weights, to construct less biased $p(z|c)$ distributions, can be demonstrated in Section 5.1, where they produce noticeable shifts in the inferred redshift distributions.

4 THE COLOUR–REDSHIFT RELATION

In this section, we describe the colour-space used in this analysis (Section 4.1) and discuss the characterization of redshift distributions in that space given DESI observations (Section 4.2).

4.1 Description of colour-space

This analysis makes use of the SOM first developed in Masters et al. (2015), which includes galaxies to Rubin and *Euclid*-like depth, $i \lesssim 24.5$ or approximately 98 per cent complete at $i = 25.3$ from Laigle et al. 2016, in *ugrizYJH*, with the modifications and added K_s-band from Masters et al. (2017). Our map exists in analogous nine-band KiDS-VIKING photometric colours to the original training photometry. As at the time of this paper KiDS-VIKING-like photometry is not available across the same fields, our transfer function to shift the map is evaluated by measuring flux in the KiDS-VIKING bandpasses for the best model template fit to the narrow band photometry in the original SOM cell. The nature of the template fitting and the templates themselves to produce these SEDs are elaborated upon in Laigle et al. (2016). While KV is not the photometry that Masters et al. (2015) was trained on, though making use of a transformed version of this SOM has the advantage that individual cells will be populated by galaxy populations of approximately the same true colours. Thus the deep samples collected by C3R2 and the DC3R2 survey will have similar redshifts for similar cells and can help inform future redshift searches. We can conclude that our map, though photometrically different, will span a similar colour space to Masters et al. (2017) and therefore *Euclid* and LSST by construction. For more elaboration on the differences between the SOM in this analysis and Masters et al. (2017), see Appendix B, and for the exact colours of our map see Appendix A.

4.2 Characterization of colour-space using DESI

With DESI targets mapped to *ugriZYJHK*$_s$ colour-space, we have a powerful statistical sample that constrains the high-dimensional colour–redshift relation for future cosmological analyses. With more than 230k galaxies covering 56 per cent of the map, we present redshifts that can span $0 < z < 1.6$ with high multiplicity. DESI provides both secure spectroscopic redshifts alongside REDROCK SED models fitted to each galaxy, which allows us to explore the evolution of galaxy-type across the map, as seen in Fig. 5. The colours of the SOM, broadly have smooth transitions, as do the redshifts as
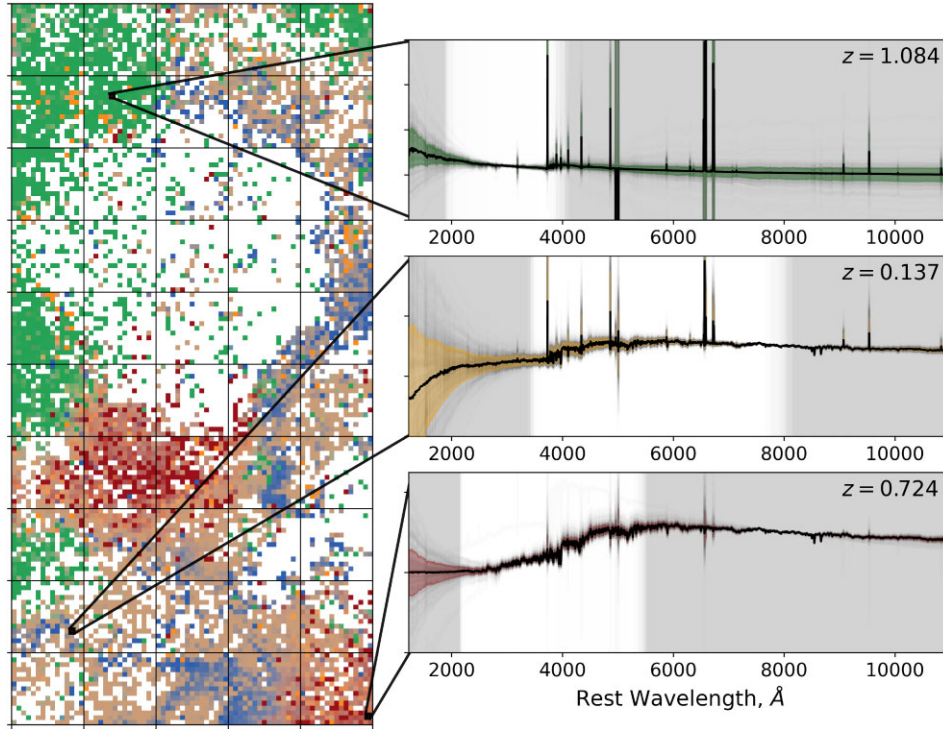
Figure 5. Depiction of the SOM with four complementary samples, post completeness cuts: emission line galaxies (green, top), luminous red galaxies (dark red, bottom), the Bright Galaxy Survey (dark blue), and DC3R2 selected targets (orange, middle), where the strength of each colour channel is directly proportional to the fraction of galaxies in that cell that come from the sample noted. The distributions of normalized REDROCK SED fits are shown for three choices of SOM cell that span star forming (top), star burst (middle), and passive (bottom) galaxy types, with a coloured envelope for the 68 per cent quantile region about the median of all templates shown in black. The shaded regions denote rest-frame wavelengths not observed by DESI according to the best-fitting redshift and individual galaxy spectrum fits are drawn in grey. Normalization is done at a point of featureless continuum, 3000 Angstroms for star forming galaxies and at 8000 Angstroms for mostly quiescent galaxies.
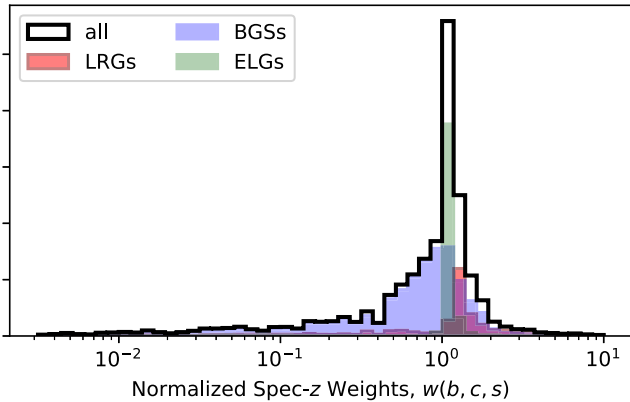
Figure 6. Distribution of the weights for main target classes. Overrepresented spectroscopic targets will have lower weight by construction, in order to generate a more representative $p(z)$.

seen in Fig. 3(a) which depicts the median redshift across this map. Galaxies with alike colours tend to have alike redshifts, except for where sharp delineating features express an innate degeneracy – where small shifts in colour can have large consequences for redshift inference.

The gains that have been made in constraining the $p(z|c)$ can be seen in the other panels of Fig. 3, where we examine the additional colour coverage of bright, low-redshift cells that were underobserved in the COSMOS field in the past. The red and magenta regions in Fig. 3(c) depict regions where DESI dominates the redshift information in the map by spectroscopic count. Cosmic variance and simple undersampling can distort the broader colour-redshift relation we measure by biasing populations of galaxies within a given SOM cell (see an extended discussion in Appendix B). DC3R2 and DESI together provide much higher confidence in the distributions of redshifts at these magnitudes and colours. This is exemplified in Fig. 7, which depicts the number of cells in the map at each redshift with spectroscopic coverage. We see a dominant DESI + KV contribution in the BGS and ELG redshift regimes (the two peaks), which sharply drops off at $z = 1.55$ (becoming subdominant to COSMOS at $z > 1.375$), where the [O II] line passes out of the optical range. DESI jointly with KiDS-VIKING calibrates 87 per cent of the LSST/*Euclid* colour-space (as a fraction of cells) at $z < 0.35$, 84 per cent from $0.8 < z < 1.2$, and 77 per cent overall for $z < 1.65$. While this is the overarching calibration extent of early DESI data in the KV fields, we can see that the higher coverage where DESI main survey targets supply galaxy spectra reveals that future efforts to span deeper magnitudes, and DC3R2-like efforts to bridge the spaces between main target categories could make real gains to fill more of the space approximately between $0.4 < z < 0.8$. For the full colour space independent of redshift, the DESI-KV sample, inclusive of DC3R2, calibrates 56 per cent of the space.

Similarly, DESI will produce better constraints on the $p(z|c)$ in the future, jointly with photometry of improved depth and with the full extent of the five year survey. With currently available KiDS-
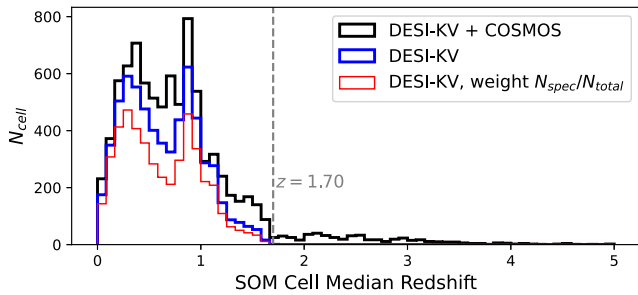
Figure 7. Existing spectroscopic coverage of the Rubin/Euclid colour space, shown as a histogram of SOM cell counts against their median redshifts. Contributions from this work (DESI + KV photometry, middle, blue), are combined with data taken in the COSMOS field for reference (top, black). The contribution of DC3R2 populated cells, weighted by the count of DESI spectra per combined spectroscopic count (lowest, red), demonstrate that jointly DC3R2 alongside DESI now provides a majority of spectroscopic calibration at redshifts below $z \approx 1.2$. The dashed line at $z = 1.7$ provides a coarse upper bound for the detection limit of the [O II] feature on silicon based detectors, and thus the upper redshift limit anticipated by these samples.
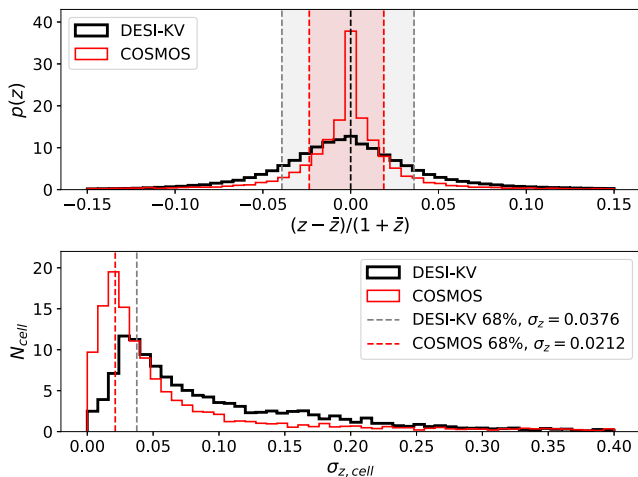
Figure 8. (Top) The distribution of redshifts across all cells as separation from mean redshift in a given cell and their respective 1σ region for DESI-KV and (for comparison) COSMOS. (Bottom) Distribution of the unbiased estimator for σ_z cell widths across the SOM. The broader peak observed in this work is likely due to the shallower KiDS-VIKING wide field photometry that DESI spectroscopic redshifts are matched to. Noted in the legend are the shaded regions from the top panel.

VIKING, we can see in Fig. 8 that redshifts for a given colour are more broad than in COSMOS – by a factor of $\sim$1.7 – likely due to photometric scatter affecting cell assignment. The 68 per cent-region for the span of redshifts in a cell in aggregate, $(z_c - \bar{z}_c)/(1 + \bar{z}_c)$, with the DESI-KV sample is [-0.0392, 0.0361], yielding an approximate $\sigma_z = 0.0376$. In the bottom panel we can see how the distributions vary in individual cells, with the median of cells falling at approximately this aggregate (marked by the dashed lines). DESI provides a powerful data set, and if similar photometry to the COSMOS field were supplied for our spectra, we might anticipate our estimate of the variance in our overall redshift distributions statistically decreasing by a factor of 0.317 over our current reported, unweighted uncertainty (i.e. $\sigma_{z, \text{COSMOS}}^2 / \sigma_{z, \text{DESI–KV}}^2$). With an already substantial increase in number of spectra, deeper photometry would provide strong advantages. While the depth of the sample can be increased, the bounds on redshift range are fixed on the current instrument.

5 CHARACTERIZATION OF SPECTROSCOPIC SELECTION BIASES ON REDSHIFT CALIBRATION

Ideal photometric redshift calibration for weak lensing surveys has to account not just for the dependence of redshift on observed colour. Even at fixed colour, magnitude and explicit and implicit selections of the spectroscopic or the weak lensing source galaxy sample can have an impact on the resulting redshift distribution (e.g. Gruen & Brimioulle 2017; Newman & Gruen 2022). Testing or accounting for this requires a spectroscopic sample that simultaneously spans the colour-space and the depth of the wider shape catalogue for weak lensing, and a full understanding of the selection process. Accounting for selection effects is difficult, but may be achieved via increasingly realistic simulations of the survey transfer function (e.g. Myles et al. 2021; Everett et al. 2022 for weak lensing source galaxy selection accounted for in redshift calibration). For future surveys, spanning the depth will no longer be possible across all colours on account of the required spectroscopic exposure times. A subcell calibration that accounts for magnitude variation may become necessary at this stage, and the accuracy of this calibration will heavily depend on the depth and multiplicity of the available spectroscopic redshifts.

Here we investigate the impact of these systematics on redshift calibration with the DC3R2 spectroscopic sample. Section 5.1 tests the effects of magnitude cuts and spectroscopic selections on the redshift distributions of realistic Stage-III lensing survey redshift bins. Sections 5.2 and 5.3 explicitly check for trends in redshift as a function of magnitude at fixed colour, and the impact of noise in observed colours on those. Finally, Sections 5.4 and 5.5 compare the systematic effects found to the requirements on future surveys.

5.1 Impact on calibration of redshift bins

Selection effects in spectroscopic samples will distort crucial estimates for weak lensing surveys, like that of the mean redshift per tomographic bin. To explore these selection effects we use the KiDS-450 galaxy abundances in our SOM cells to infer redshift distributions for five KiDS-like tomographic bins. This procedure follows where,

$$p(z|b, \text{sel}) = \sum_{\text{cell} \in b} p(z|c) \, p(c|\text{sel}) \qquad (4)$$

for a given selection, *sel*, and bins, *b*, comprised of cells, *c*. The first term amounts to the distribution of spectroscopic redshifts in a given cell and the second amounts to a weighting factor that relies on the abundances of galaxies in the calibrated sample (here, KiDS-450). The resulting $n(z)$s are depicted in Fig. 9(a). We create these bins by sorting cells covered by DESI spectroscopy by their median redshift. The lowest redshift cells are assigned to the first tomographic bin, until the number of KV galaxies with shapes occupying those cells reaches one fifth of the overall sample. We continue in this fashion, associating roughly equal numbers of calibrated galaxies with each bin, *b*. Note, that we can see in Fig. 9(a) that this binning does not place equal numbers of *calibrating* galaxies, i.e. spectroscopic redshifts, in each bin.

We then perform a set of tests on the impact of selection biases by applying further selections to the spectroscopic sample and determining how that changes the estimated $n(z)$s. Table 2 summarizes the resulting change in mean redshift for the tomographic bins.

Every comparison ensures that the colour selection for each bin does not change as the spectroscopic selection effect is applied, i.e. is made for bins defined by the same cell envelope both before and
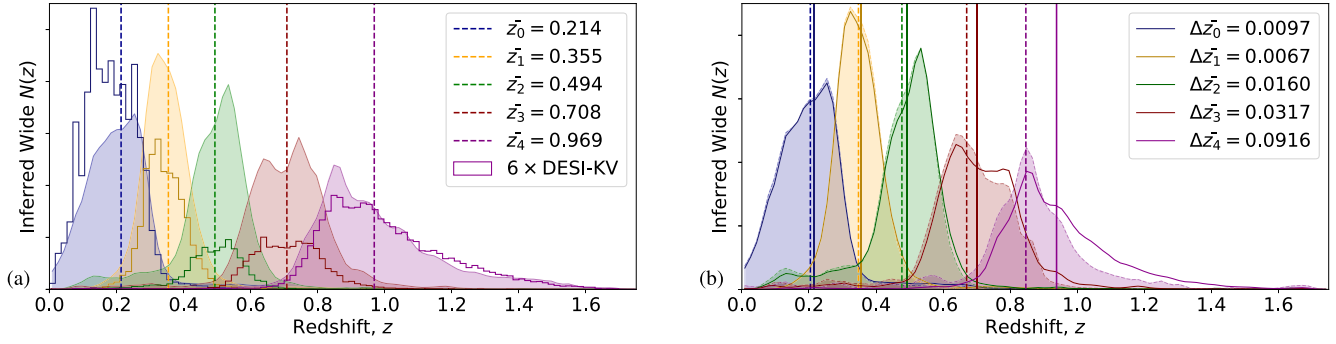
Figure 9. A simple redshift calibration scheme for the G09, G12, G15 fields of KiDS-450 using DESI + DC3R2 spectroscopy reveals that applying selection effects on the spectroscopic catalogue can significantly underestimate the mean redshift of a photometric sample of the same observed colour. (a) DESI spectroscopic redshifts (histogram outlines) and redshift distributions of photometric bins from re-weighting these by KiDS-450 colour-space abundance (shaded). Shown are five tomographic bins of equal galaxy count selected by KV colour after weighting (i.e. the bins from inferred redshift have equal size). The dashed lines represent the mean of the weighted bin, depicted in the legend. Estimated redshift distributions are smoothed by a Gaussian kernel for visualization purposes. (b) Inferred redshift distributions of KiDS-450 using a representative spectroscopic calibration sample (unfilled distributions) with the mean redshift of each bin indicated by the solid vertical line. When using a calibration sample with an additional selection of MAG_GAAP_Z < 21.0, the inferred distributions (shaded) are altered and the mean redshift of each bin (dashed vertical line) is biased to low redshift. The bias in mean redshift for each bin is noted in the legend, showing the substantial effect on the inferred redshifts of a photometric sample if the spectroscopic calibration sample is biased to brighter galaxies.

Table 2. Consolidation of the shifts in mean redshift inferred for a tomographic bin with KiDS-VIKING-like cell abundances if different selections are made on the spectroscopic sample. $\Delta \bar{z}_i$ is always the mean redshift of the new selection subtracted from the fiducial mean. The final row is a repeat of the first magnitude cut with the weights applied both before and after the selection. Error bars are produced by bootstrapping cells with replacement in the selection.

Selection	$\Delta \bar{z}_0$	$\Delta \bar{z}_1$	$\Delta \bar{z}_2$	$\Delta \bar{z}_3$	$\Delta \bar{z}_4$
$Z < 21.0$	0.0097 ± 0.0057	0.0067 ± 0.0027	0.0160 ± 0.0041	0.0317 ± 0.0058	0.0916 ± 0.0068
$\Delta \chi^2 > 40$	0.0007 ± 0.0057	0.0010 ± 0.0027	0.0005 ± 0.0040	0.0015 ± 0.0058	0.0045 ± 0.0119
$dz < 0.113$	0.0182 ± 0.0064	0.0086 ± 0.0030	-0.0115 ± 0.0040	-0.0036 ± 0.0057	0.0034 ± 0.0124
$w(b, c, s)$ scheme	-0.0041 ± 0.0055	-0.0062 ± 0.0031	-0.0152 ± 0.0042	-0.0155 ± 0.0055	-0.0033 ± 0.0130
$Z < 21.0$, w	0.0016 ± 0.0054	0.0014 ± 0.0031	0.0031 ± 0.0042	0.0323 ± 0.0055	0.0642 ± 0.0071

after the additional spectroscopic selection. This means that if the selection eliminates certain SOM cells from spectroscopic coverage, we use the estimated mean redshift of the same reduced set of cells even for the fiducial sample in order to isolate the effect of selection bias at fixed colour.

The spectroscopic selection effects we examine are:

(i) **Magnitude – (MAG_GAAP_Z < 21.0)** For $dz/dm = 0$ and a well populated colour space, a magnitude cut would introduce no change in mean redshift. In the presence of significant $dz/dm > 0$, however, a magnitude cut would bias the inferred redshift towards a lower mean value. This operation and the result is demonstrated in Fig. 9(b). We indeed find a small ($|\Delta z| < 0.01$) impact of the magnitude cut on the mean redshift inferred for the two lowest redshift bins, but an impact that exceeds current calibration requirements on the mean redshift inferred for the higher redshift bins, up to almost $\Delta z = 0.1$. This cut changes the mean magnitude in each bin, and the spectroscopic redshift counts according to Table 3. As this cut was chosen to be moderately bright for demonstration, we also explore how less severe selections affect this metric in Fig. 10 as half magnitude steps down from the KV limiting magnitude in the i-band. This demonstrates the importance of representative spectroscopic redshifts in our toy model, which is constructed in a way that fundamentally differs from DES, KiDS, and HSC as our SOM is not trained on a magnitude in addition to its colours.

(ii) **Quality Flags – ($\Delta \chi^2 > 40$)** As a consequence of enforcing a more severe confidence in the fitted model, this will remove spectra in an SED-dependent way from the ELG and LRG samples where,

Table 3. Changes in properties of a given tomographic bin when their calibrating spectroscopic sample is subject to MAG_GAAP_Z < 21, describing the decrease in mean magnitude of the calibrating galaxies, and the fraction of spectroscopic calibrating galaxies cut from each bin. Higher redshift bins are dominated by fainter spectra and suffer the largest population shifts.

Bin	0	1	2	3	4
$\left(\overline{Z_{\mathrm{mag}}} \right)_{\mathrm{fid}}$	19.54	19.80	20.20	20.78	21.45
$\left(\Delta \overline{Z_{\mathrm{mag}}} \right)_{\mathrm{cut}}$	-0.163	-0.197	-0.308	-0.499	-0.949
$\Delta N_{\mathrm{spec}} / N_{\mathrm{spec,\,fid}}$	0.020	0.033	0.113	0.273	0.821

by default, some objects with smaller $\Delta \chi^2$ are considered confident redshifts. When implemented, less than 2 per cent of spectroscopic galaxies are removed. The impact of this stricter selection is small ($|\Delta z| < 0.005$) but coherent across all redshift bins, i.e. biases all redshift distributions towards a lower mean by preferentially removing higher redshift galaxies from the sample within a given cell. In the highest redshift bin, where the fraction of ELG targets contributing to the calibration is large, the effect exceeds the redshift calibration requirements of future weak lensing experiments (cf. also Hartley et al. 2020 for the impact of selection based on conventional redshift confidence flags).

(iii) **Removal of cell outliers:** Galaxies with large deviation from the median redshift of an ensemble of similar colour are more likely to be outliers of various types, e.g. due to blending, AGN light, or redshift determination errors. We define a sample of outliers based
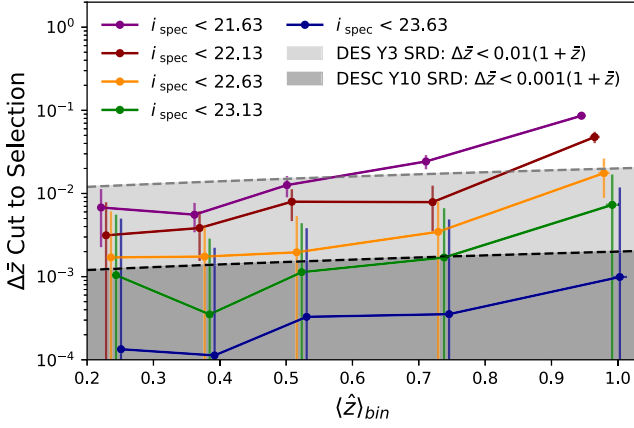
Figure 10. Lack of representative spectra for an inference biases the mean redshift of a bin for our sample matched to KV. Depicted are the shifts in mean redshift for our tomographic bins should the calibrating spectroscopic galaxies be preferentially brighter than the source galaxies. We show several choices of magnitude cut incrementally brighter than the KiDS limiting magnitude of $i = 23.63$, and compare that to the Rubin Science Requirements Document (SRD) (LSST Dark Energy Science Collaboration 2018). Small offsets have been applied to the mean redshift per bin, $\langle\hat{z}\rangle_{\rm bin}$, for visualization.

on the criterion

$$dz = \frac{|z - \mathrm{median(z)_{cell}}|}{(1 + \mathrm{median(z)_{cell}})} > 3\sigma_{z,\mathrm{all}} \,, \qquad (5)$$

where $\sigma_{z,\mathrm{all}}$ is the aggregate standard deviation in cells found in Section 4.2, from which $3\sigma_{z,\mathrm{all}} = 0.113$. The selection corresponds to the central 99.7 per cent of redshifts if these were to follow Gaussian distributions of colour-independent width at any fixed colour and thus would reject only the most egregious outliers. In practice, these distributions are not Gaussian and this selection cuts 5.1 per cent of objects. The selection is relative to the median rather than the mean to robustly deal with cells that are undersampled or heavily affected in their mean redshift by outliers. The impact of removing the outlier population defined this way is maximally of the order of $\Delta z \approx 0.02$, and has reduced impact in the highest redshift bins. While a number of these outliers can be attributed to photometric scatter across degenerate regions in the SOM (neighbouring cells with large separation in redshift), others may be real examples of broad or bimodal cell distributions and ought to be examined with visual inspection.

(iv) **Applying weights, w:** Spectroscopic redshifts are weighted according to the scheme described in Section 3.2.3, to account for prioritization of spectroscopy for targets of certain morphologies, colours, and SED-types that are not necessarily representative. As the weights can be somewhat noisy in undersampled regions of colour space (in targeting), additionally this comparison is restricted to cells where the resulting shift in $\bar{z}_c$ from applying the weights is small. While this will result in an underestimation of the true weighting effect for all galaxies, these are the regions where weights can be applied confidently due to the spectroscopic and targeting counts. We apply a cell envelope selection where $|\bar{z}_c - \bar{z}_{c,w}| < 0.08$, which retains > 90 per cent of cells and eliminates the sparsely occupied, and noisiest cells. Note that the effect of this scheme in Table 2 is always to lower the mean redshift in a given bin. The impact is maximally $\Delta z \approx 0.01$ in the most sparsely populated spectroscopic bins. Additionally, if these weights are applied in the same way to the magnitude selection ($Z < 21$), we see in Table 2 that they mitigate

the bright spectroscopic bias in high bins, but do not eliminate it.

5.2 Magnitude dependence of redshift at fixed colour

Past analyses on photometry and spectroscopy in the COSMOS field have shown the magnitude dependence of redshift at fixed colour cell to be small and well described by a linear behaviour with $dz/dm \approx 3 \times 10^{-3}$ (Masters et al. 2017). The equivalent measurement made in the KV data with DESI redshifts is depicted in Fig. 11(b). Here we examine the relation of differences in VIKING Z-band magnitude and redshift between pairs of galaxies that occupy the same colour-cell in the SOM. A cell with n galaxies contributes $n(n-1)/2$ data points to Fig. 11(b). We see a linear relationship, with a large amount of scatter. The raw slope measured, $dz/dm = 0.0250 \pm 0.0009$, is an order of magnitude larger than that reported in previous studies (see Appendix E for methodology). This ought not to be taken at face value, and Section 5.3 (Fig. 11a) explores photometric scatter as the largest source of bias in this measurement, as well as the chief difference between the data set used in this study and that of previous ones. Furthermore, comparisons of dz/dm will be affected to lesser degrees by the other second order effects discussed in Section 4.1.

5.3 Systematic uncertainty due to photometric scatter

With the depth of our SOM resembling that of future deep surveys, it could be problematic that we use a relatively shallow photometry in KiDS-VIKING to assign our spectroscopic galaxies. Photometric scatter enters our calculation through perturbations in the cell assignment, which comprises our definition of fixed colour. Galaxies with larger flux uncertainties tend to be fainter, which is also correlated with higher redshifts. While the direction of scatter will be impacted by the colour topography of our individual SOM, we can broadly expect faint, high redshift galaxies to scatter more often into brighter, lower redshift cells and induce a strongly positive dz/dm. We can imagine photometric scatter as an asymmetric smearing of the redshift distributions in a given cell.

A potential systematic effect of this photometric scatter on the measured slope can be explored with a test described here. We can perform a direct measurement on the slope introduced by photometric noise by (1) removing any existing dz/dm from our data by randomly shuffling all spectroscopic redshifts in a given SOM cell, thus nulling dz/dm while preserving the $p(z|c)$, (2) applying a random Gaussian draw of the width of the flux error reported in the catalogue for each band, thus perturbing measured colours. (3) Reassigning galaxies to the SOM based on these perturbed fluxes, we repeat our measurement of dz/dm. Additionally, we can (4) iteratively reshuffle and perturb many times to reduce statistical uncertainty in the measurement from the limited sample size. The only contributor to this final, measured dz/dm will be photometric scatter, making it a measurement on this systematic that can be compared across photometric surveys of different depth. Using (4) to quintuple the number of points available to us for this measurement, we observe in Fig. 11(a), that $(dz/dm)_{sc} = 0.0132 \pm 0.0007$ for the reported KV errors associated with our DESI targets in this SOM. The slope induced this way can be measured across colour space, i.e. across the SOM. This colour-dependent effect is subtracted off in the second panel of Fig. 12, leaving an estimate of an intrinsic dz/dm that is corrected at first order, at least under the assumption that the reported flux measurement errors are accurate. While individual cells carry significant large intrinsic $(dz/dm)_c$ (order of ≈ 0.01), the average across all cells is much lower at $(dz/dm)_{c,\mathrm{avg.}} = -7 \times 10^{-4}$.
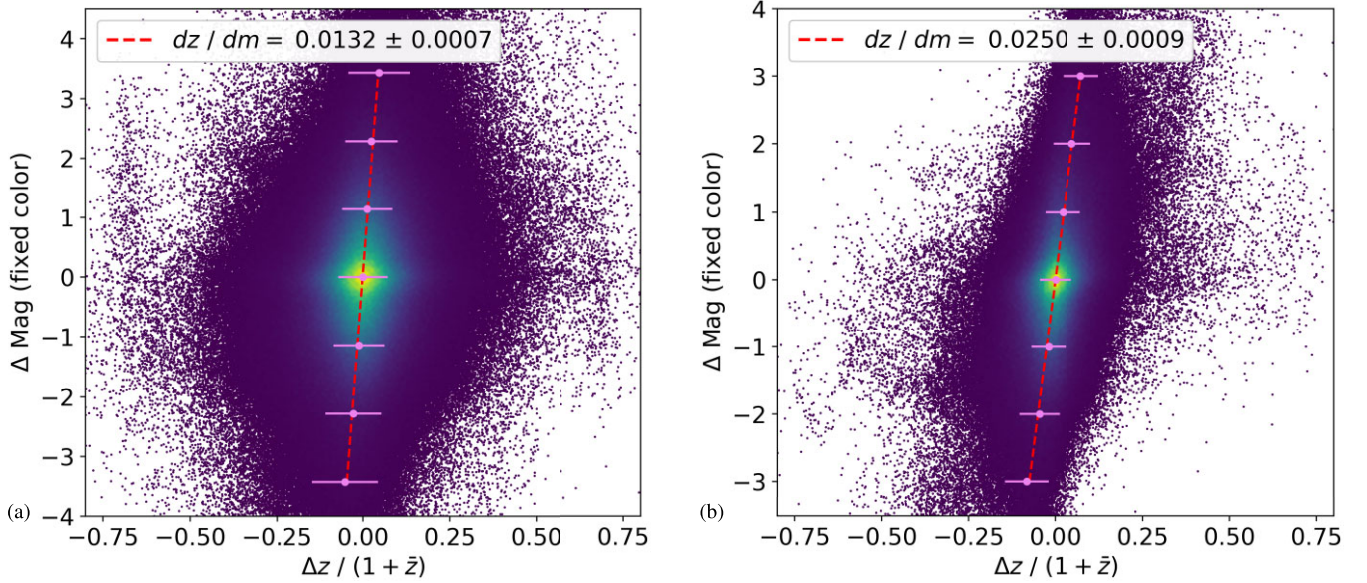
Figure 11. Plots depicting the systematic from photometric scatter induced in dz/dm (left) and the raw measurement from the joint DC3R2-KV sample (right), where each data point is the difference between a unique pair of galaxies occupying the same SOM colour-cell in magnitude and redshift. The best-fitting slope (dashed red) is fit from a collection of medians and σ (pink) in magnitude slices. If reported photometric errors are accurate, this systematic dominates our measurement. (a) Systematic induced in dz/dm by KiDS-VIKING photometric errors in conjunction with our given SOM resolution for DC3R2 spectroscopic redshifts at fixed colour, with change in magnitude in KiDS-VIKING MAG_GAAP_Z. Section 5.3 describes the procedure of scattering galaxies by their reported errors to perform this measurement. (b) Total measured change in redshift with magnitude (Z-band) at fixed colour as observed in the shallow KiDS-VIKING photometry is dramatically affected by photometric scatter in contrast to previous studies on deeper COSMOS field photometry.
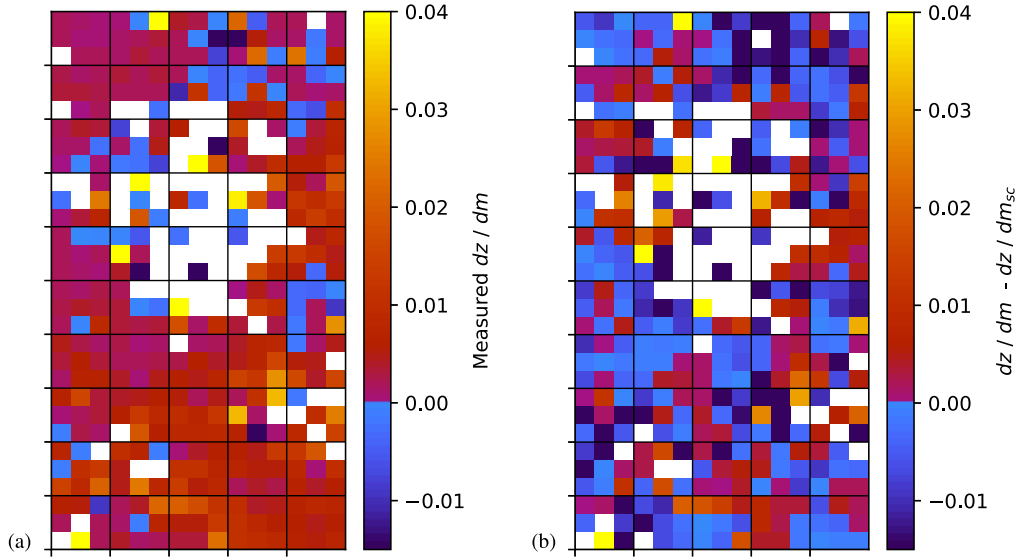
Figure 12. Colour dependence of dz/dm in the raw data (left) and corrected for the contribution of photometric scatter to the overall slope (right). (a) Raw measurement of change in redshift with magnitude at fixed colour, dz/dm across the SOM, deresolved for statistical power into 5 x 5 super cells. (b) Intrinsic change in redshift with magnitude at fixed colour, dz/dm as observed across the SOM, corrected to first order by subtracting off the contribution due to photometric scatter (i.e. dz/dm_{raw} − dz/dm_{sc}).

5.3.1 Impact of photometric noise levels

We can see in Fig. 11(a) that the systematic measured for KiDS-VIKING-like error in a SOM of *Euclid*-like resolution is significant, $(\mathrm{d}z/\mathrm{d}m)_{\mathrm{sc}} \approx 0.013$, and the likely dominant contributor to our overall measured slope. With this effect in mind, we revisit the original measurement performed in Masters et al. (2019) using the same field and similar photometric filters. Repeating the test of Section 5.3

used to measure the impact of photometric scatter, but with the lower noise levels of the COSMOS photometry, we find that the original measurement had a contributing noise bias of $(\mathrm{d}z/\mathrm{d}m)_{\mathrm{sc}} \approx 0.003 \pm 0.0004$. This is similar to the measured slope in Masters et al. (2019), and means that the intrinsic dz/dm in that sample is roughly consistent with zero. With the limitations of the coarse photometric scatter test we have applied we do not claim to have a measurement of the *intrinsic* dz/dm in COSMOS to better than 0.003 as a result.

Also worthy of note is that doubling the photometric error for KV associated DESI redshifts produces a $(\mathrm{d}z/\mathrm{d}m)_{\mathrm{sc},2} = 0.0256 \pm 0.0010$, which is larger than the observed raw slope in Section 5.3. Thus, in case the photometric error in the KV catalogues should be underestimated at levels that have been reported in other studies doing source injection into survey images (Everett et al. 2022, e.g.), it could be that the average intrinsic slope of the DC3R2 sample is indeed consistent (across the full sample, not simply cell-by-cell) with zero as well.

As noise in the KiDS-VIKING photometry introduces a large systematic effect on our desired measurement, it is worth exploring what photometry is needed to drive down the error and improve constraints on $\mathrm{d}z/\mathrm{d}m$. Given the uncertainty on our estimation of the systematic effect of photometric noise, the hypothesis that $\mathrm{d}z/\mathrm{d}m \approx 0$ seems to be consistent with current data. Yet improved photometry such as that from HSC (Aihara et al. 2018) could allow us to use DESI spectra to better effect. With a Gaussian-draw, background limited error model based upon the effective exposure times and limiting magnitudes of Aihara et al. 2018 (see their table 8) we generate HSC-like flux errors for *grizy* to augment KiDS-VIKING *uJHKs*. We find that $(\mathrm{d}z/\mathrm{d}m)_{\mathrm{sc}} \approx 0.008$, which halves the effect, though does not eliminate it. In the limit of noiseless optical (*ugriz*), we still find $(\mathrm{d}z/\mathrm{d}m)_{\mathrm{sc}} \approx 0.004$. We can imagine the LSST scatter with these objects would be comparable to this value, without improved NIR or IR follow up. Note that this is still larger than the slope found with the COSMOS photometry that includes the very deep UltraVISTA NIR data, demonstrating both that such small systematic errors are achievable in principle, and that they rely on deep NIR that is difficult to achieve with currently operating instruments.

5.4 Future survey requirements

The tests presented above have shown limitations to how well we can constrain the trend of mean redshift with magnitude at fixed colour in the presence of noisy photometry. Here, we connect this to requirements for how well do we need to know $\mathrm{d}z/\mathrm{d}m$ for future, stage IV weak lensing surveys. We can approximate the redshift calibration error due to imperfectly known magnitude dependence as

$$\frac{\Delta z}{(1+\bar{z})} \approx \Delta(\mathrm{d}z/\mathrm{d}m) \times (\bar{m}_{\mathrm{wide}} - \bar{m}_{\mathrm{spec}}), \qquad (6)$$

where $\Delta(\mathrm{d}z/\mathrm{d}m)$ is our error in the magnitude dependence of redshift at fixed colour, $\bar{z}$ is the mean redshift of the sample being calibrated, and $(\bar{m}_{\mathrm{wide}} - \bar{m}_{\mathrm{spec}})$ is the offset between the mean magnitudes of wide field observed galaxies and the spectroscopic sample used to calibrate them.

The requirement for final Rubin analyses is expected to be $|\Delta z|/(1+\bar{z}) \approx 0.001$ (LSST Dark Energy Science Collaboration 2018). With the limiting *i*-band magnitude of final Rubin data and the mean magnitude of the current DESI sample for $\bar{m}_{\mathrm{wide}} = 26.3$ and $\bar{m}_{\mathrm{spec}} = 21.1$, respectively, we find that we need to determine $|\Delta(\mathrm{d}z/\mathrm{d}m)| \leq 0.0002$. If we assume the $\Delta(\mathrm{d}z/\mathrm{d}m) \approx 0.004$ that was determined in Section 5.3.1 to be the limit obtainable with current NIR photometry and deep Rubin observations, our needs and capabilities are at odds.

If instead of considering the faintest galaxies used by Rubin, we estimate the sample's mean magnitude, using the slope of the luminosity function from Gruen & Brimioulle (2017), observed to the limiting *i*-band magnitude for LSST, we obtain $\bar{m}_{\mathrm{wide}} = \bar{i}_{\mathrm{Rubin}} \approx 24.90$. This provides a more realistic, less conservative requirement for $\Delta(\mathrm{d}z/\mathrm{d}m) \leq 0.0006$, which is still stricter than all our estimates

for the effect of photometric scatter on this slope by almost an order of magnitude.

5.5 Interpretation

The previous sections studied how selections on the spectroscopic calibration sample provided here may impact the redshift distributions estimated using that sample. We performed these tests using a simple redshift calibration scheme that assumes that at a fixed observed colour, the distribution of true redshifts of a photometric and a spectroscopically selected galaxy sample are identical. We note that none of the recent, Stage-III, analyses have relied on a redshift calibration scheme that was quite this simple, and hence the biases we identify are not expected to be present at the level found here in recent analyses. For example, Fig. 10 depicts a case where the calibrating spectra are up to two magnitudes brighter than the limiting magnitude of the sample, whereas the formal KiDS analysis made use of a variety of spectroscopic sources that spanned the depth of their photometric sample (Busch et al. 2022). Similarly DES did not strictly apply a spectroscopic selection and made use of narrowband photometric redshifts where there were no representative spectroscopic redshifts (Myles et al. 2021). Despite this, extrapolation of redshift calibration samples to photometric objects of similar colour but fainter magnitude is likely required for future, deeper photometric surveys, and hence it is useful to study the potential pathways for bias in such an application of our sample. Both the DES and KiDS SOMs (Wright et al. 2019; Myles et al. 2021) are trained on colours and a magnitude (or luptitude) – and DESC will be unable to do this without throwing away a substantial part of their sample. Hence the SOM for this analysis is matching spectroscopic *z*s to galaxies based on their colours alone.

A necessary condition for a non-zero bias in such a redshift calibration scheme is that the expectation value of redshift depends on properties relevant for the spectroscopic target selection and measurement success other than just a galaxy's observed colour. The most salient such effect we identify is via a trend of mean redshift with observed magnitude at given observed colour, $\mathrm{d}z/\mathrm{d}m$. We find this slope to be much steeper than reported in previous studies. The cause for this steepness is an effect of the larger photometric noise in the colours measured for our target sample, rather than a large dependence of redshift on magnitude at fixed *true* colour. Most of the imaginable selection effects are related to a galaxy's magnitude, and hence the non-zero $\mathrm{d}z/\mathrm{d}m$ propagates into several of the bias tests we perform and merits further interpretation.

The bias imposed by photometric scatter on $\mathrm{d}z/\mathrm{d}m$ varies with colour, as seen in Fig. 12, and potentially also with other selection choices. The scatter itself is asymmetric as it does not shift objects into neighbouring cells isotropically in the map, and will induce a larger $\mathrm{d}z/\mathrm{d}m$ in areas with large colour-redshift degeneracies (i.e. a small error in colour could lead to a large offset in the mean local redshift). Our ability to understand the intrinsic $\mathrm{d}z/\mathrm{d}m$ present in the data depends heavily on the quality of our photometry and/or our ability to correctly model photometric measurement errors. As seen in the right panel of the same figure, the intrinsic $\mathrm{d}z/\mathrm{d}m$ is similarly colour dependent, but just as likely to be negative as positive (indicating that it is likely a noisy measurement).

Our method to constrain this bias relies on the reported photometric errors. However, if these errors were to be misestimated by an order of unity, potentially in a colour-dependent way, the entire measured slope could be the result of this systematic, as discussed in Section 5.3.1. Literature analyses have demonstrated that the deviation between observed photometry and the true fluxes of a galaxy requires

realistic image simulations processed by photometric pipelines to estimate, and reported errors are frequently an underestimate at levels that indeed reach factors of two (e.g. Huang et al. 2017; Everett et al. 2022). No literature study exists on the accuracy of reported KV GAAP flux errors specifically, but the relative linearity of the algorithm and preliminary studies on image simulations (Li et al. 2023b) imply that the misestimation of the error in GAAP is non-zero but not a factor of two.

Consequently, Section 5.1 demonstrates that selection effects in magnitude suffer from dz/dm dependence significantly in higher redshift bins, where objects tend to be fainter, and this emphasizes the role that deep photometry, in contrast to merely multiband coverage, plays when making maximal use of spectroscopic redshifts. We can see with the σ_z cut, on the narrowness of the redshift distribution for a given cell, that potential outliers in a given cell that arise either from photometric scatter or a misattributed redshift can have dramatic effects on the redshift distribution. Since this effect is largest in the lowest redshift bin – where objects are brightest – we might suspect that these outliers are dominated by cell-to-cell photometric scatter. Cutting these cells does further limit the colour space of a given weak lensing analysis, and ought to be avoided where possible. For this reason, visual inspection works like Lan et al. (2023) and careful selection of the redshift sample are crucial to future analyses.

It has been established (a) that LSST will require a very accurate measurement of dz/dm if calibrated by DESI alone (Section 5.4), and (b) that even with perfect optical photometry the photometric scatter contribution to dz/dm with existing DESI spectra is larger than the acceptable error by almost an order of magnitude (Section 5.3.1). The strategies to account for this in future weak lensing redshift calibration efforts can be threefold:

(i) **Deeper photometry**: Despite (b), specifically improving NIR/IR measurements will decrease the systematic contributor to dz/dm and potentially allow for a measurement that meets photo-z requirements.

(ii) **Deeper spectroscopy**: Deeper spectra will reduce a given survey calibration's dependence on dz/dm, mitigating how well known this slope has to be in (a). After DESI's current survey, it will be uniquely situated to push for deeper spectra.

(iii) **Modelling**: Perhaps the most viable approach, one can take into account the effect of photometric scatter with appropriate modelling of the data. Future analyses could constrain the bias via forward-modelling from a colour space of true photometry. In this approach, the intrinsic and observed dz/dm are folded into the inference alongside the systematic in a methodologically appropriate way, and the size of the systematic effect is no longer a limiting factor.

6 CONCLUSIONS

Photometric redshift calibration – estimating the redshifts for galaxies where we only observe them through a collection of filters – requires a thorough understanding of the colour–redshift relation, which is very non-linear across the full galaxy population. This paper presents a catalogue of spectra, the DESI complete calibration of the colour–redshift relation (DC3R2) secondary target survey, designed to aid in the redshift calibration for a large fraction of the weak lensing source galaxy samples of future photometric surveys. The data includes associated weights that turn the survey samples (ELGs, LRGs, BGS) into one that is representative at a given $ugrizYJHK_s$ and limiting magnitude of 23.68 (5σ, i-band), using KiDS-VIKING as test-bed. We chose to select targets on this nine-band photometry in order to break redshift-type degeneracies, and

KiDS-VIKING provided the most constraining data over this area. With this unprecedented quantity of DESI spectroscopic redshifts, we examine how it benefits future surveys and allows us to examine spectroscopic selection effects.

(i) The DESI sample reported here calibrates the redshift distribution of roughly 56 per cent of the galaxies in COSMOS via 230k spectroscopic redshifts. For the photometric colour space that will be visible to DESC and *Euclid* (approx. 98 per cent complete at $i = 25.3$), this sample corresponds to coverage of 6248 cells out of 11 250. Approximately 41 per cent of the full COSMOS colour-space and galaxy population is calibrated from spectroscopic galaxies that are classed as DC3R2 targets (inclusive of overlap with main survey targets), and 4 per cent from uniquely DC3R2 objects.

(ii) However, even though this sample provides an incredible quantity of high-quality spectra, we find that the combination of uncertain photometry and a variety of spectroscopic selection effects can produce substantial biases on the redshift inference for a given lensing redshift bin. We demonstrate that introducing a preference for brighter magnitude calibration spec-zs in the presence of photometric scatter effects or intrinsically high dz/dm biases the mean redshift on the bin of the order of $\Delta z \approx 0.01$, especially for higher redshift bins (see Fig. 10 for a breakdown). This effect is present even for the shallower test-bed data used here and will be exacerbated for fainter samples, as it enters the inference as an induced magnitude dependence of redshift at fixed colour. Fewer photometric bands further worsens the effect, as breaking degeneracies becomes more difficult (e.g. the uncertainty in mean redshift for the year 3 HSC, Rau et al. 2023, *grizy* analysis was larger than that of KiDS-450, Wright et al. 2020 $ugriZYJHK_s$ by a factor of 1.25, even with KiDS having an additional bin).

(iii) Results for this work are expressed in a colour-space that is similar to past work on the Masters et al. (2017) SOM, as our map is a transformation of this space into KiDS-VIKING colours. As demonstrated in Fig. B1, we recover very similar galaxy SEDs and redshifts per cell, but comparisons like that of Fig. 3(c) ought to be taken with the knowledge that the colours and photometric noise levels between different surveys are not identical.

(iv) Our analysis reveals a general agreement in the colour-redshift relation with previous spectroscopic surveys that have explored this colour-space. When accounting for effects induced by photometric noise, we also find agreement in the magnitude dependence of redshift at fixed colour with the result of Masters et al. (2017).

(v) Photometry quality has an important role in redshift calibration. This study has found that photometric errors need to be well understood for modelling the colour–redshift relation and especially magnitude-dependent effects (dz/dm). In order to constrain this slope for future surveys we require either better photometry than expected, deeper spectroscopy, or improved analysis methodology (discussed in Section 5.5). We find that the use of KV photometry in this map does not supply as strong a constraint as past data sets do on the slope dz/dm. However, close examination of the systematic dz/dm induced by photometric scatter in a high resolution SOM has strengthened the case for the null hypothesis in this work and in past survey data sets. While dz/dm ≈ 0 is potentially consistent with our data, deeper photometry or survey simulation will be needed to constrain the slope sufficiently for future weak lensing efforts.

The spectroscopic redshifts measured by the complete 5-yr DESI survey will provide unparalleled support for future redshift calibration in weak lensing surveys. To fully leverage the powerful quantity of data discussed in this paper, more accurate photometry and deeper spectroscopic redshifts will be necessary to constrain the magnitude

dependence of redshift at fixed colour. Exploration into the capability of massively multiplexed spectroscopic instruments, like DESI, to attain redshifts of fainter sources than currently targeted is important. Looking ahead, future surveys will provide additional wavelength and sky area coverage that can improve photometric redshift calibration. Among these is a follow-on programme to DC3R2 using the 4-m Multi-Object Spectroscopic Telescope (4MOST; de Jong et al. 2019) that will observe targets across the same SOM used in this work, though more uniformly to redshift ∼1.55 (Gruen & McCullough 2023). Together with deeper spectroscopic campaigns (e.g. DESI-II; Schlegel et al. 2022) and campaigns including high-quality infrared spectroscopy, these data will form the basis for constraining a model of the galaxy population seen by deep photometric surveys, including its redshift distribution.

DATA AVAILABILITY

The data and code used to generate the figures in this paper will be made available with its publication.[2] The redshifts for objects taken during SV are available in the DESI Early Data Release[3] (see DESI Collaboration 2023). Visualization tools exist to explore the data across the colour-space in browser and are available via github.[4]

REFERENCES

Aihara H. et al., 2018, PASJ, 70, S4
Aihara H. et al., 2022, PASJ, 74, 247
Alsing J., Peiris H., Mortlock D., Leja J., Leistedt B., 2023, ApJS, 264, 29
Amendola L. et al., 2013, Living Rev. Relat., 16, 6
Amon A. et al., 2022, Phys. Rev. D, 105, 023514
Arnaboldi M., Capaccioli M., Mancini D., Scaramella R., Sedmak G., Kurz R., 2000, in Bergeron J., Renzini A., eds, From Extrasolar Planets to Cosmology: The VLT Opening Symposium. Springer, Berlin, p. 204
Balogh M. L. et al., 2014, MNRAS, 443, 2679
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Buchs R. et al., 2019, MNRAS, 489, 820
van den Busch J. L. et al., 2022, A&A, 664, A170
Casey C. M. et al., 2017, ApJ, 840, 101
Collaboration L. S. et al., 2009, LSST Science Book, Version 2.0. preprint (arXiv:0912.0201)
Collaboration D. et al., 2016a, The DESI Experiment Part II: Instrument Design. preprint (arXiv:1611.00037)
Collaboration D. et al., 2016b, The DESI Experiment Part I: Science,Targeting, and Survey Design. preprint (arXiv:1611.00036)
Cutri R. M. et al., 2013, in Cutri R. M.et al., eds, Explanatory Supplement to the AllWISE Data Release Products, Explanatory Supplement to

---

[2] https://zenodo.org/record/8328495
[3] https://data.desi.lbl.gov/doc/releases/edr
[4] https://jmccull.github.io/DC3R2_Overview

the AllWISE Data Release Products. Infrared Processing and Analysis Center, California Institute of Technology Pasadena

DESI Collaboration, 2023, AJ, 167, 33

DESI Collaboration 2022, AJ, 164, 207

DESI Collaboration 2023, preprint (arXiv:2306.06308)

Dalal R. et al., 2023, Phys. Rev. D, 108, 123519

Dark Energy Survey Collaboration 2016, MNRAS, 460, 1270

Dey A. et al., 2019, AJ, 157, 168

Driver S. P. et al., 2011, MNRAS, 413, 971

Everett S. et al., 2022, ApJS, 258, 15

Flaugher B., Bebek C., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 282

Gebhardt K. et al., 2021, ApJ, 923, 217

Gruen D., Brimioulle F., 2017, MNRAS, 468, 769

Gruen D., McCullough J., 2023, The Messenger, 190, 28

Guy J. et al., 2023, AJ, 165, 144

Hahn C. et al., 2023, AJ, 165, 24

Hartley W. G. et al., 2020, MNRAS, 496, 4769

Hasinger G. et al., 2018, ApJ, 858, 77

Hildebrandt H. et al., 2020, A&A, 633, A69

Huang S. et al., 2017, PASJ, 70, S5

Ivezić Ž. et al., 2019, ApJ, 873, 111

de Jong J. T. A. et al., 2017, A&A, 604, A134

de Jong R. et al., 2019, The Messenger, 175, 3

Kannawadi A. et al., 2019, A&A, 624, A92

Kartaltepe J. S. et al., 2010, ApJ, 709, 572

Kohonen T., 2004, Biol. Cybern., 43, 59

Kriek M. et al., 2015, ApJS, 218, 15

Kuijken K. et al., 2019, A&A, 625, A2

Kuijken K. et al., 2015, MNRAS, 454, 3500

LSST Dark Energy Science Collaboration 2018, The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document, preprint (arXiv:1809.01669)

Laigle C. et al., 2016, ApJS, 224, 24

Lan T.-W. et al., 2023, ApJ, 943, 68

Le Fèvre O. et al., 2015, A&A, 576, A79

Li X. et al., 2023a, Phys. Rev. D., 108, 123518

Li S.-S. et al., 2023b, A&A, 670, A100

Lilly S. J. et al., 2007, ApJS, 172, 70

Masters D. et al., 2015, ApJ, 813, 53

Masters D. C., Stern D. K., Cohen J. G., Capak P. L., Rhodes J. D., Castander F. J., Paltani S., 2017, ApJ, 841, 111

Masters D. C. et al., 2019, ApJ, 877, 81

Myers A. D. et al., 2023, AJ, 165, 50

Myles J. et al., 2021, MNRAS, 505, 4249

Newman J. A., Gruen D., 2022, ARA&A, 60, 363

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Oke J. B. et al., 1995, PASP, 107, 375

Raichoor A. et al., 2023, ApJ, 165, 126

Rau M. M. et al., 2023, MNRAS, 524, 5109

Rykoff E. S. et al., 2014, ApJ, 785, 104

Rykoff E. S. et al., 2016, ApJS, 224, 1

Saglia R. et al., 2022, A&A, 664, A196

Salvato M., Ilbert O., Hoyle B., 2018, Nature Astronomy, 3, 212

Schlegel D. J. et al., 2022, A Spectroscopic Road Map for Cosmic Frontier: DESI, DESI-II, Stage-5. preprint (arXiv:2209.03585)

Sevilla-Noarbe I. et al., 2021, ApJS, 254, 24

Silverman J. D. et al., 2015, ApJS, 220, 12

Stanford S. A. et al., 2021, ApJS, 256, 9

Trump J. R. et al., 2007, ApJS, 172, 383

Wang B. et al., 2023, APJ, 944, L58

Wright E. L. et al., 2010, AJ, 140, 1868

Wright A. H. et al., 2019, A&A, 632, A34

Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., Joachimi B., Kannawadi A., Kuijken K., 2020, A&A, 640, L14

Zhou R. et al., 2023, AJ, 165, 58

Zou H. et al., 2017, PASP, 129, 064101

APPENDIX A: THE SOM AND COLOUR-SPACE

The SOM described in Section 4.1 is fully described in Fig. A1 in the KiDS-VIKING *ugriZYJHKs* colours, transformed from the original Masters et al. (2017) SOM using COSMOS narrow-band best-fitting SEDs re-evaluated in the KiDS-VIKING filters.

Additionally, the complete set of magnitude bins chosen for the weighting schema in Section 3.2.3 are depicted in Fig. A2, with the bin edges tuned to encapsulate sample selection boundaries in each band. The hatch patterns specify the SV3 selection boundaries for each sample, while the main and SV1 selections have mild differences. The occupancy of these magnitude colour bins heavily depends on the quantity of each main survey target.
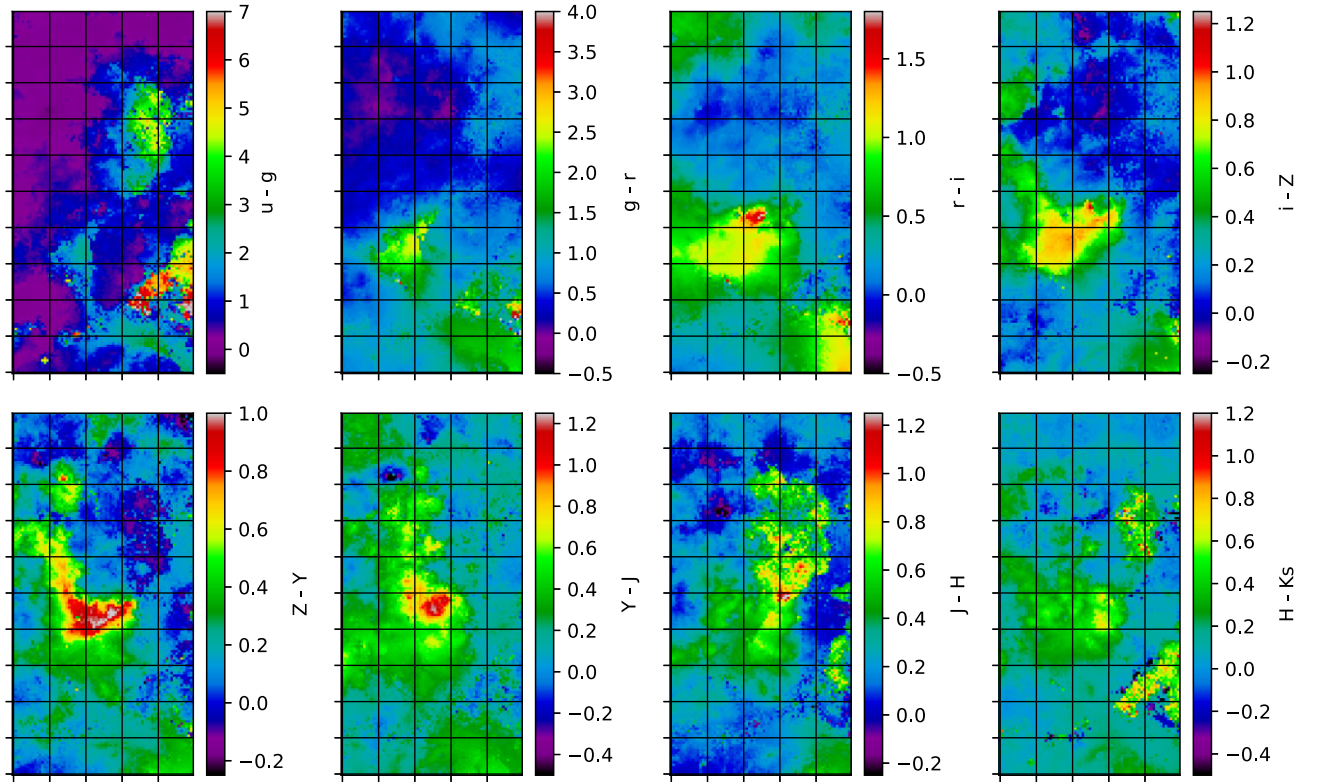
Figure A1. Depiction of the Masters et al. (2017) SOM corrected to KiDS-VIKING *ugriZYJHK_s* colours, as used for this analysis.
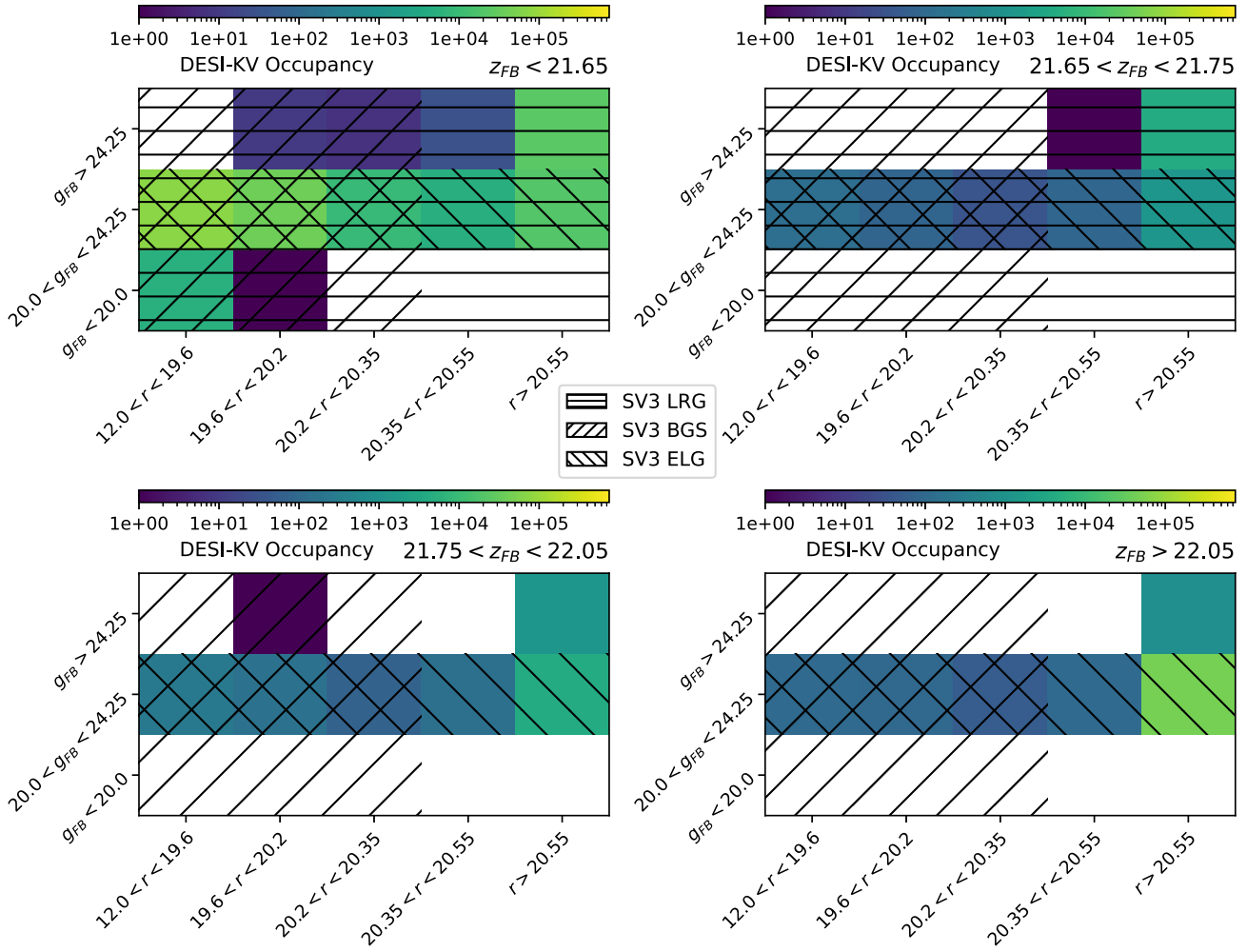
Figure A2. Depiction of the occupancy of the entire DESI-KV sample in the *grz* magnitude bins (3 x 6 x 4) used to produce the weight scheme in Section 3.2.3 that renders cell-level redshift distributions more representative.

APPENDIX B: SOM-C3R2 AGREEMENT

The $p(z|c)$ distributions in a given cell may differ in the original C3R2 SOM (Masters et al. 2017) to that from the map used for this work. Combining results from differing surveys should come with the following caveats, with systematic factors contributing to changes in redshift distributions:

(i) **Cosmic variance** – The original C3R2 survey observes a different limited volumes of the Universe compared to the the overlap region of DESI and KiDS-VIKING used for DC3R2, with galaxy samples of limited number. This difference in cosmic variance can influence shifts in the redshift distributions seen in each cell. If each cell is well described as $\delta(z)$, this impact of altered abundances is minimal on the calibrated colour–redshift relation. This paper has found that for the photometry used, the redshift distributions are comparably broader, in which case altered abundances may overrepresent certain contributors to the overall cell redshift. A small sample is also liable to miss cell contributors that are rare and may not accurately measure the tails of these SOM cell distributions. However, the DESI/KV overlapping footprint is sufficiently large that cosmic variance will be a minor concern.

(ii) **Photometric noise** – KV photometry has a higher noise level than the COSMOS catalogue that used to train the original the map.
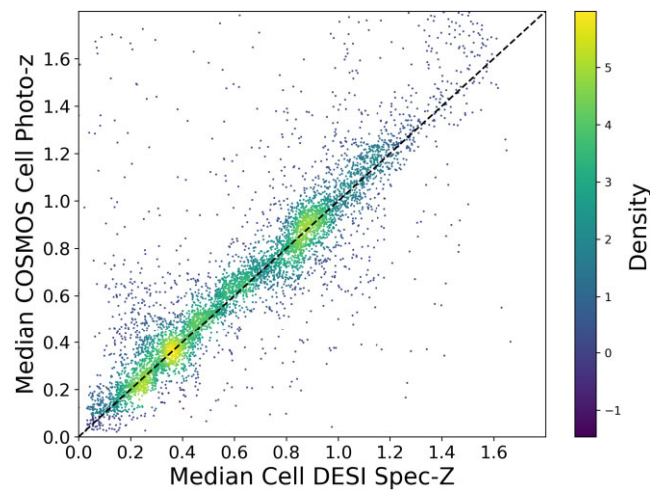

Figure B1. Median DC3R2 redshifts of a SOM cell in the transformed KV mapping compared to the COSMOS photometric redshift of the same SOM cell prior to the transformation. This reveals the effect of both the SOM colour transformation as well as the differences in sample depth. Broadly speaking, there is good recovery between redshifts for objects of the same colours.

The effect of this a scattering of galaxies between neighbouring cells, effectively lowering the resolution of the map. This effect would influence measurable quantities like dz/dm, which is discussed further in Section 5.3, as this noise is asymmetric between cells.

(iii) **The SOM colour transformation –** As the filters covered by KiDS-VIKING are different from those used in Masters et al. (2017), a best fit SED for each cell was defined based on template fits to COSMOS narrow-band galaxies in Laigle et al. (2016) and evaluated in the relevant KiDS-VIKING bands to transform the colour of that cell. The median of all COSMOS galaxies in the original SOM produced the best-fitting SEDs for this evaluation. For a sample with overlap in COSMOS/KiDS, we ought to show that the same galaxies reliably get assigned to the same cell given sufficiently low photometric uncertainty. Before a joint analysis of the samples assigned to cells based on their original COSMOS and based on KV photometry could be made, which we do not attempt here, it would first have to be shown that the colour transformation does not introduce e.g. spurious magnitude dependence of redshift within colour cells. We can see a broad recovery of COSMOS cell photo-z against the DESI spectroscopic median redshift per cell in Fig. B1, which indicates that this transformation produces a colour map with very similar galaxies on a cell-by-cell basis. However, cell boundaries in colour space could have minor distortions. Outliers at high z are likely to result from spectroscopic incompleteness in $z \geq 1.6$. Photometric scatter is likely the largest contributor to disagreement at lower redshifts, which we elaborate on the consequences of more fully in Section 5.2 and care to characterize for future surveys.

(iv) **Differing redshift fitting algorithms in spectral modelling –** Robustness tests on DESI's spectral modelling pipeline, REDROCK (Bailey et al., in preparation), are expanded upon in Guy et al. (2023) and Lan et al. (2023). Additionally, for a small overlap of objects in common fields, we can directly compare the REDROCK modelled redshift to those found by other surveys to test their reliability.

(v) **Spectroscopic selection effects –** Quality flags chief among them, we can cross-check on other spectral surveys to inform our redshift quality cuts that limit outliers. Perhaps even more importantly we can look at our relevant findings as a function of REDROCK's $\Delta \chi^2$ to see if this selection affects quantities like dz/dm or shifts our understanding of the $P(z|c)$. Certain galaxy types are more conducive to confident redshifts. Past studies have found that if colour cuts are implemented on spectroscopic galaxies that are not available to the weak lensing surveys, redshift calibration errors can be as large as $\Delta z \approx 0.04$ (Hartley et al. 2020). This analysis mitigates this effect by removal of detections that do not have sufficiently high completeness in a given cell which accounts for these sample specific colour cuts (see Section 3.2.3).

(vi) **Cell assignment algorithm–** The metric that is minimized for each galaxy to be assigned a cell can dramatically impact the distribution across the map. We attempt to imitate the cell assignment procedure from Masters et al. (2015), which minimizes contributions from flux measurements with larger error and from drop-outs, and when running the assignment on the original COSMOS catalogue recover the identical assignment for more than 99 per cent match.

APPENDIX C: DEDICATED TILE TARGET OPTIMIZATION

The dedicated tiles comprise of two equatorial pointings at 217.5 degrees and 221.0 degrees RA, with a bright (30 min, $z_{\text{fibre}} < 21.5$) and faint (3×30 min, $z_{\text{fibre}} < 22.10$) set of exposures each. A mix

of bright and faint targets were selected according to the procedure described here.

The aim of target choice is to optimize our measurement of dz/dm within a cell, and measure it for as many colour-cells as possible. We devise a metric to select targets that best do this between the bright and faint selection criteria. A function that describes the energy of our targeting procedure can be written below that optimizes our fibre choice when minimized. We chose the following energy function, E:

$$E = \frac{1}{N_{\text{cells}}} \text{Var} \left(\left\langle \frac{\mathrm{d}z}{\mathrm{d}m} \right\rangle \right) = \frac{1}{N_{\text{cells}}(\rangle 1 \text{ target}) \sum_{i \in N} 1/\sigma_i^2}. \quad \text{(C1)}$$

Where the variance of the measurement of the slope, σ^2 of a given cell, i is proportional to

$$\sigma_i^2 \propto \frac{1}{\sum_{j \in \text{cell i}} \left(Z_j - \bar{Z} \right)^2} \quad \text{(C2)}$$

That prioritizes having a large span in ΔZ per cell, as well as maximal coverage in the colour space, N_{cell}, with multiple targets per cell. A simple simulated annealing algorithm was implemented to minimize this function, by varying choice of target on a fibre-by-fibre basis. For certain fibres, only a single target will be visible in both the bright and faint exposure. The optimization becomes useful when multiple targets are visible to each fibre.

APPENDIX D: SCALING OF EXPOSURE TIME MAP

The exposure times in Fig. 4 and corresponding templates were drawn from a diverse sample with varying observation conditions and fluxes across a given colour. The rationale for the reported Z = 21.0 magnitude, is that for a point source in a sky limited background, these exposure times might be scaled accordingly to estimate future spectral survey needs, e.g. 4MOST, (Jong et al. 2019). For a sky limited background,

$$\text{SNR} \propto \sqrt{t_{\text{exp}}}, \quad \text{(D1)}$$

where t_{exp} is the total exposure time for a given object. In order to scale existing exposure time measurements in this regime to those that would be necessary for a Z = 21.0 object, we use the following prescription,

$$t_{\text{exp,sc}} = 10^{0.8(21.0-Z_{\text{o}})} t_{\text{exp,o}}, \quad \text{(D2)}$$

where $t_{\text{exp, sc}}, t_{\text{exp, o}}$ are the scaled and original exposure times, respectively, and Z_{o} is the Z magnitude of the object being scaled. This follows from the sky limited case for a point source and ought to be treated as a rough estimate of the exposure times necessary. Additionally, it is useful to develop a scaling for the quality of the redshift. As the REDROCK quantity of DELTACHI2, $\Delta \chi^2$ is our primary quality cut for redshift completeness, and the most extreme requirement for any galaxy is $\Delta \chi^2 > 40$. We can add another factor of $\frac{40}{\Delta \chi_o^2}$ to account for this, with some scatter. Thus, Fig. 4 also includes projections from low-confidence redshift targets, where our overall analysis does not. These are targets that have been observed at low signal-to-noise, or lack helpful spectral features in the available wavelength range. There may thus be model uncertainty between several potential redshifts or spectral templates, rendering realistic projections for exposure times in cells that are otherwise incomplete.

Our scaling relationship follows as an approximation,

$$t_{\text{exp,sc}} \approx \frac{40}{\Delta \chi_{\text{o}}^2} 10^{0.8(21.0-Z_{\text{o}})} t_{\text{exp,o}} \quad \text{(D3)}$$
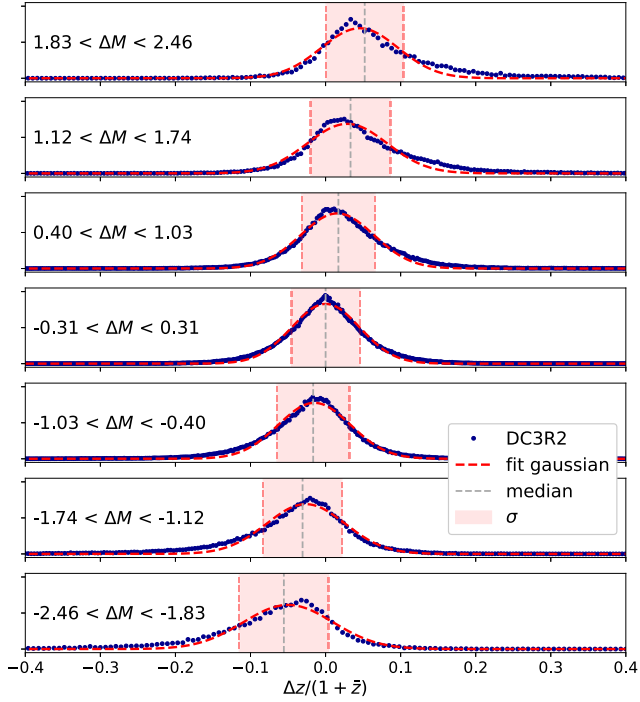
Figure E1. Example set of histograms (points) for each magnitude bin, with median (thin dashed vertical line) and Gaussian fits (bold dashed) to measure dz/dm. While there is some skew in the most extreme magnitude bins, the Gaussian is a good approximation, for which the fit σ describes the width.

APPENDIX E: Dz/Dm MEASUREMENT

The slopes measured in Section 5.2 for dz/dm were done in such a way to be comparable to the procedure for the original measurement in Masters et al. (2017). The data points in Fig. 11 are drawn from unique pairs in a given SOM cell, producing $\binom{n}{2}$ points for a cell with n spectroscopic galaxies. To fit the linear model, we take slices in ΔMAG_GAAP_Z and fit each histogram of points to a normal distribution with a median and width, σ. Tails in the distribution cause deviations from the fitted normal that become larger with increasing $|\Delta M|$, which may lead to slight underestimation of the slope uncertainty. The fitting process is demonstrated visually in Fig. E1. The medians at the bin midpoints are then fit with a linear regression that produces our estimate of dz/dm for a given sample.

For Fig. 12, where there are fewer data points for each cell than for the entire sample, only two slices in magnitude were chosen, where they were chosen to jointly span [min(ΔM_{cell}), max(ΔM_{cell})].

[1]*Kavli Institute for Particle Astrophysics and Cosmology, Department of Physics, Stanford University, Stanford, CA 94305, USA*

[2]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

[3]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstraße 1, D-81679 München, Germany*

[4]*Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany*

[5]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

[6]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

[7]*Infrared Processing and Analysis Center, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA*

[8]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

[9]*Institute of Cosmology and Gravitation, Faculty of Technology, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*

[10]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*

[11]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, E-08193 Barcelona, Spain*

[12]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

[13]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*

[14]*Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA*

[15]*Brookhaven National Laboratory, Upton, NY 11973, USA*

[16]*Dunlap Institute for Astronomy and Astrophysics, University of Toronto, Toronto M5S 3H4, Canada*

[17]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

[18]*Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA*

[19]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

[20]*Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA*

[21]*Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México*

[22]*Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711 Bogotá, Colombia*

[23]*Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA*

[24]*Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), F-75005 Paris, France*

[25]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*

[26]*NSF's NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA*

[27]*Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA*

[28]*National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing 100012, P. R. China*

[29]*Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*

[30]*Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA*

[31]*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain*

[32]*Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, USA*

[33]*Department of Physics, Sejong University, Seoul 143–747, Korea*

[34]*CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*

[35]*Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA*

[36]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*

[37]*National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing 100012, P. R. China*

This paper has been typeset from a T$_E$X/L^{A}T$_E$X file prepared by the author.