The SAMI Galaxy Survey: Data Release One with emission-line physics value-added products

Andrew W. Green,[1][*][†] Scott M. Croom,[2,3][*] Nicholas Scott,[2] Luca Cortese,[4]
Anne M. Medling,[5,6][‡] Francesco D'Eugenio,[3,5] Julia J. Bryant,[1,2,3]
Joss Bland-Hawthorn,[2] J. T. Allen,[2,3] Rob Sharp,[5] I-Ting Ho,[7] Brent Groves,[5]
Michael J. Drinkwater,[3,8] Elizabeth Mannering,[1] Lloyd Harischandra,[1]
Jesse van de Sande,[2] Adam D. Thomas,[5] Simon O'Toole,[1] Richard M. McDermid,[1,9]
Minh Vuong,[1] Katrina Sealey,[1] Amanda E. Bauer,[1] S. Brough,[10] Barbara Catinella,[4]
Gerald Cecil,[11] Matthew Colless,[5] Warrick J. Couch,[1] Simon P. Driver,[4]
Christoph Federrath,[5] Caroline Foster,[1] Michael Goodwin,[1] Elise J. Hampton,[5]
A. M. Hopkins,[1] D. Heath Jones,[12] Iraklis S. Konstantopoulos,[1,13] J. S. Lawrence,[1]
Sergio G. Leon-Saval,[2] Jochen Liske,[14] Ángel R. López-Sánchez,[1,9]
Nuria P. F. Lorente,[1] Jeremy Mould,[15] Danail Obreschkow,[4] Matt S. Owers,[1,9]
Samuel N. Richards,[1,2,3,16] Aaron S. G. Robotham,[17] Adam L. Schaefer,[1,2,3]
Sarah M. Sweet,[15] Dan S. Taranu,[3,4] Edoardo Tescari,[3,18] Chiara Tonini[18]
and T. Zafar[1]

Affiliations are listed at the end of the paper

ABSTRACT

We present the first major release of data from the SAMI Galaxy Survey. This data release focuses on the emission-line physics of galaxies. Data Release One includes data for 772 galaxies, about 20 per cent of the full survey. Galaxies included have the redshift range $0.004 < z < 0.092$, a large mass range ($7.6 < \log M_*/M_\odot < 11.6$), and star formation rates of $\sim 10^{-4}$ to $\sim 10^1 M_\odot \, yr^{-1}$. For each galaxy, we include two spectral cubes and a set of spatially resolved 2D maps: single- and multi-component emission-line fits (with dust-extinction corrections for strong lines), local dust extinction, and star formation rate. Calibration of the fibre throughputs, fluxes, and differential atmospheric refraction has been improved over the Early Data Release. The data have average spatial resolution of 2.16 arcsec (full width at half-maximum) over the 15 arcsec diameter field of view and spectral (kinematic) resolution of $R = 4263$ ($\sigma = 30 \, km \, s^{-1}$) around $H\alpha$. The relative flux calibration is better than 5 per cent, and absolute flux calibration has an rms of 10 per cent. The data are presented online through the Australian Astronomical Observatory's Data Central.

Key words: surveys – galaxies: general.

1 INTRODUCTION

Our textbooks provide a reasonable picture of how the first dark matter structures assembled out of the primordial matter perturbations (Peacock 1999; Mo, van den Bosch & White 2010). But just how gas settled into these structures to form the first stars and galaxies, and how these evolved to provide the rich diversity of galaxies we see around us today, remains an extremely difficult problem to unravel. The outstanding questions regarding galaxy formation and evolution include

(i) What is the physical role of environment in galaxy evolution?
(ii) What is the interplay between gas flows and galaxy evolution?

[*] E-mail: andrew.green@aao.gov.au (AWG); scott.croom@sydney.edu.au (SMC)
[†] How each author contributed to the paper is listed at the end.
[‡] Hubble Fellow.

(iii) How are mass and angular momentum built up in galaxies?

Mass is thought to be the primary discriminant driving the huge variety of galaxies observed, setting their star formation rate (e.g. Peng et al. 2010, 2012), metallicity (e.g. Tremonti et al. 2004), and morphology. However, in addition to mass, the environment of a galaxy also plays a central role in controlling such properties (e.g. Lewis et al. 2002; Blanton & Moustakas 2009, and Dressler 1980; Cappellari et al. 2011b, respectively). Despite the wealth of data at hand, the physical processes that drive environmental differences are still uncertain. The processes are likely to depend on whether a galaxy is the central galaxy or a satellite in its parent halo, the mass of the parent halo, and local galaxy–galaxy interactions (e.g. Davies et al. 2015). We can only test the interplay of these processes observationally with samples that span a large range of environments with many galaxies.

Gas flow (or lack thereof) in and out of a galaxy controls its evolution with time. Inflows have formed discs, fuelled generation upon generation of new stars, and fed supermassive black holes. In current galaxy-formation theory, galactic-scale outflows are required to explain the problem that the theoretical cold-dark-matter mass function is so different from the observed stellar-mass function of galaxies (e.g. Baldry et al. 2012). A process with strong mass dependence is needed to resolve this problem. Outflows offer the most promising solution (e.g. Silk & Mamon 2012), and are clearly detected by combining gaseous emission-line ionization diagnostics with kinematics (e.g. Sharp & Bland-Hawthorn 2010; Fogarty et al. 2012; Ho et al. 2014, 2016a). Gas inflows can be traced using the measurement of misalignment between gas and stellar kinematics (e.g. Davis et al. 2011; Davis & Bureau 2016) and by searching for flattened metallicity gradients (Kewley et al. 2010; Rich et al. 2012).

The mass and angular momentum of a galaxy are most directly probed by its kinematic state. A galaxy's accretion and merger history is central to defining its character, and aspects of this history are encoded in the line-of-sight velocity distributions. By studying the detailed kinematics of galaxies across the mass and environment plane, we unlock a new view of galaxy evolution (Cortese et al. 2016; van de Sande et al. 2017). Recent three-dimensional or integral-field spectroscopy (IFS) observations have defined a new set of morphological classifications in terms of dynamical properties (e.g. Cappellari et al. 2011a; Emsellem et al. 2011), such as the separation into fast rotators (rotation dominated) and slow rotators (dispersion dominated). It is important to understand how these kinematic properties are distributed across the mass–environment plane, and to make direct comparison to simulations that are now becoming available to measure more complex dynamical signatures (e.g. Naab et al. 2014).

Cosmological-scale hydrodynamic simulations can now form thousands of galaxies with realistic properties in $\sim 100\,\mathrm{Mpc}^3$ volumes (e.g. Vogelsberger et al. 2014; Schaye et al. 2015). These simulations allow study of how gas enters galaxies (e.g. Codis et al. 2012) and the impact of feedback (e.g. Genel et al. 2015). Those at higher resolution (e.g. Brooks et al. 2009; Hopkins et al. 2014) are probing details of disc formation, gas flows, and feedback, though not yet within a full cosmological context. Direct, detailed comparison of spatially resolved data to these simulations is required to advance our understanding.

Over the past 20 years, imaging surveys from the *Hubble Space Telescope* (far field) and the Sloan Digital Sky Survey (SDSS; near field) have been particularly effective in identifying evolution of galaxy parameters with cosmic time and with environment across large-scale structure. This has been matched by extensive surveys using multi-object spectroscopy (e.g. York et al. 2000; Colless et al. 2001; Driver et al. 2011). These large samples have produced some key results such as the importance of environment (e.g. Lewis et al. 2002). However, these surveys were usually limited to a single spectrum within a fixed fibre aperture at the centre of each galaxy; spatial information had to be drawn from multi-wavelength broad-band images. It has long been recognized that the complexity of a galaxy cannot be captured with a single average or central spectrum. Three-dimensional imaging spectroscopy, or IFS, is needed to quantify each galaxy.

Driven by pioneering work using Fabry–Perot interferometry (Tully 1974) and lenslet arrays (Courtes et al. 1988), IFS has exploited the plunging costs of large-area detectors to become a common method for studying galaxies today (e.g. Hill 2014). The first generation of IFS surveys, sampling tens to hundreds of galaxies, have only recently been completed. Examples include ATLAS^{3D} (Cappellari et al. 2011a), CALIFA (Sánchez et al. 2012), and SINS (Förster Schreiber et al. 2009). These surveys demonstrated that there is much to learn from both the stellar and gaseous components in data of this kind. However, these surveys all used instruments that target individual galaxies one at a time and are, therefore, not optimal for surveying thousands of galaxies. To move beyond samples of a few hundred requires systems that can obtain IFS observations of multiple galaxies at a time ('multiplexing').

Multiplexed IFS has only recently become possible. The FLAMES instrument on the VLT (Pasquini et al. 2002) was the first, with 15 integral-field units (IFUs) each having 20 spatial resolution elements in a 2×3 arcsec field of view. Most of the multiplexed systems use optical fibres to collect the resolved light from the focal plane. The fibres make possible not only IFUs with large fields of view (such as the PPAK IFU used for CALIFA; Kelz et al. 2006), but also ease deployment of IFUs over large focal planes and allow the spectrograph to be mounted on the floor rather than on the telescope, simplifying design and improving stability. Fibre-based systems are therefore preferred for wide-field, multi-object IFS in the optical bands.

With the aim of carrying out IFS surveys targeting thousands of galaxies, we developed the Sydney/AAO Multi-object Integral-field spectrograph (SAMI; Croom et al. 2012) on the 3.9 m Anglo-Australian Telescope (AAT). SAMI provides a multiplex of $\times 13$ with each 15 arcsec diameter IFU comprising a compact fused fibre bundle with minimized cladding between the fibre cores (hexabundles; Bland-Hawthorn et al. 2011; Bryant et al. 2011, 2014). The MaNGA survey (Bundy et al. 2015), operating on the Apache Point 2.5 m Telescope, has also begun a similar project, with an IFU multiplex of $\times 17$. Meanwhile, the high-redshift KMOS-3D and KROSS surveys (Wisnioski et al. 2015; Magdis et al. 2016) are making spatially resolved observations of high-redshift galaxies.

In this paper, we present Data Release One (DR1) of the SAMI Galaxy Survey, building on our Early Data Release (EDR) in 2014 (see Allen et al. 2015). We provide data cubes for 772 galaxies and value-added products based on detailed emission-line fitting. Future releases will provide more galaxies and products. In Section 2, we review the SAMI Galaxy Survey itself, including the selection, observations, data reduction, and analysis. In Section 3, we describe the core data being released, with discussion of data quality in Section 3.4. The emission-line physics value-added products are described in Section 4. The online data base is introduced in Section 5. We summarize this paper in Section 6. Where required, we assume a cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2 BRIEF REVIEW OF THE SAMI GALAXY SURVEY

The SAMI Galaxy Survey is the first integral-field spectroscopic survey of enough galaxies to characterize the spatially resolved variation in galaxy properties as a function of both mass and environment. Specific details concerning the survey can be found in papers describing the SAMI instrument (Croom et al. 2012; Bryant et al. 2015), the SAMI-GAMA sample target selection (Bryant et al. 2015), the SAMI cluster sample target selection (Owers et al. 2017), data reduction (Sharp et al. 2015), and the EDR (Allen et al. 2015). Below we review key aspects of the survey, but for detailed discussions of these aspects, please refer to the papers above.

2.1 The SAMI instrument

SAMI is mounted at the prime focus of the AAT and has 1 deg diameter field of view. SAMI uses 13 fused optical fibre bundles (hexabundles; Bland-Hawthorn et al. 2011; Bryant et al. 2011, 2014) with a high (75 per cent) fill factor. Each bundle combines 61 optical fibres of 1.6 arcsec diameter to form an IFU of 15 arcsec diameter. The 13 IFUs and 26 sky fibres are inserted into pre-drilled plates using magnetic connectors. Optical fibres from SAMI feed into AAOmega, a bench-mounted double-beam optical spectrograph (Sharp et al. 2006). AAOmega provides a selection of different spectral resolutions and wavelength ranges. For the SAMI Galaxy Survey, we use the 580V grating at 3700–5700 Å and the 1000R grating at 6250–7350 Å. With this setup, SAMI delivers a spectral resolution of $R = 1812$ ($\sigma = 70\,\mathrm{km\,s^{-1}}$) for the blue arm, and $R = 4263$ ($\sigma = 30\,\mathrm{km\,s^{-1}}$) for the red arm at their respective central wavelengths (van de Sande et al. 2017). A dichroic splits the light between the two arms of the spectrograph at 5700 Å.

2.2 Target selection

In order to cover a large dynamic range in galaxy environment, the SAMI Galaxy Survey is drawn from two regions with carefully matched selection criteria. The majority of targets are from the Galaxy And Mass Assembly (GAMA) survey (Driver et al. 2011), and we denote this as the SAMI-GAMA sample. However, the volume of the SAMI-GAMA region does not contain any massive galaxy clusters, so a second set of targets are drawn from specific cluster fields. This we denote as the SAMI cluster sample (Owers et al. 2017).

DR1 includes galaxies only from the SAMI-GAMA sample, and the selection for these targets is described by Bryant et al. (2015). Briefly, the sample is drawn from the 4 × 12 deg fields of the initial GAMA-I survey (Driver et al. 2011), but uses the deeper spectroscopy to $r < 19.8$ of the GAMA-II sample (Liske et al. 2015). The high completeness of the GAMA sample (98.5 per cent) leads to high-reliability group catalogues (Robotham et al. 2011) and environmental metrics (Brough et al. 2013). The GAMA regions also provide broad-band imaging from the ultraviolet to far-infrared (Driver et al. 2016).

The selection limits for the SAMI-GAMA sample, shown in Fig. 1, consist of a set of volume-limited samples with stellar-mass limits stepped with redshift. We select using stellar masses determined from only g- and i-band photometry and redshift, using the relationship given in equation 3 of Bryant et al. (2015). This determination is based on the relationship between mass-to-light ratio and colour derived by Taylor et al. (2011), and assumes a Chabrier (2003) initial mass function.

Bryant et al. (2015) describe the process of allocating target galaxies to fields for observation.

2.3 Observing strategy

Our standard observing sequence consists of a flat-field frame (from the illuminated AAT dome) and arc frame, followed by seven object frames each of 1800 s exposure. A flat-field and arc are taken to end the sequence. The seven object exposures are offset from one another in a hexagonal dither pattern (see Bryant et al. 2015, fig. 16), with the subsequent frames radially offset from the first exposure by 0.7 arcsec in each of six directions 60 deg apart. This offset is applied based on the most central guide star in the field, using an offset in pixels on the guide camera. Variations in atmospheric refraction and dispersion between different exposures cause the effective offsets to differ for different galaxies on the same field plate. However, the high fill factor of SAMI hexabundles minimizes the effect on data quality (see Section 3.4.5 and Appendix A). The change in offset across the field is measured as part of the alignment process during data reduction as described in section 5.2 of Sharp et al. (2015).

Where possible, twilight-sky frames are taken for each field to calibrate fibre throughput. Primary spectrophotometric standards are observed each night that had photometric conditions to provide relative flux calibration (i.e. the relative colour response of the system). For non-photometric nights, the standard observed on the closest available photometric night is used.

2.4 Data reduction

Raw telescope data are reduced to construct spectral cubes and other core data products in two stages that are automated for batch processing using the 'SAMI Manager', part of the SAMI PYTHON package (Allen et al. 2014). The specifics of both stages are detailed in Sharp et al. (2015). Subsequent changes and improvements to the process are described in section 3 of Allen et al. (2015) and in Section 3.2 below.

The first stage of data reduction takes raw 2D detector images to partially calibrated spectra from each fibre of the instrument, including spectral extraction, flat-fielding, wavelength calibration, and sky subtraction. Processing for this stage uses the 2DFDR fibre data reduction package (AAO software Team 2015) provided by the Australian Astronomical Observatory.[1] This stage outputs the individual fibre spectra as an array indexed by fibre number and wavelength, and referred to as 'row-stacked spectra' (RSS).

In the second stage, the RSS are sampled on a regular spatial grid to construct a three-dimensional (two spatial and one spectral) cube. Processing for the second stage is done within the SAMI PYTHON package (Allen et al. 2014). This stage includes telluric correction, flux calibration, dither registration, differential atmospheric refraction correction, and mapping input spectra on to the output spectral cube. The last of these stages uses a drizzle-like algorithm (Fruchter & Hook 2002; Sharp et al. 2015). The spectral cubes simplify most subsequent analysis because the cube can be read easily into various packages and programming languages, and spatial mapping of the data is straightforward. However, in creating the spectral cube, additional covariance between spatial pixels is introduced that must be correctly considered when fitting models and calculating errors (Sharp et al. 2015).

---

[1] Different versions of 2DFDR are available, along with the source code for more recent versions at http://www.aao.gov.au/science/software/2dfdr
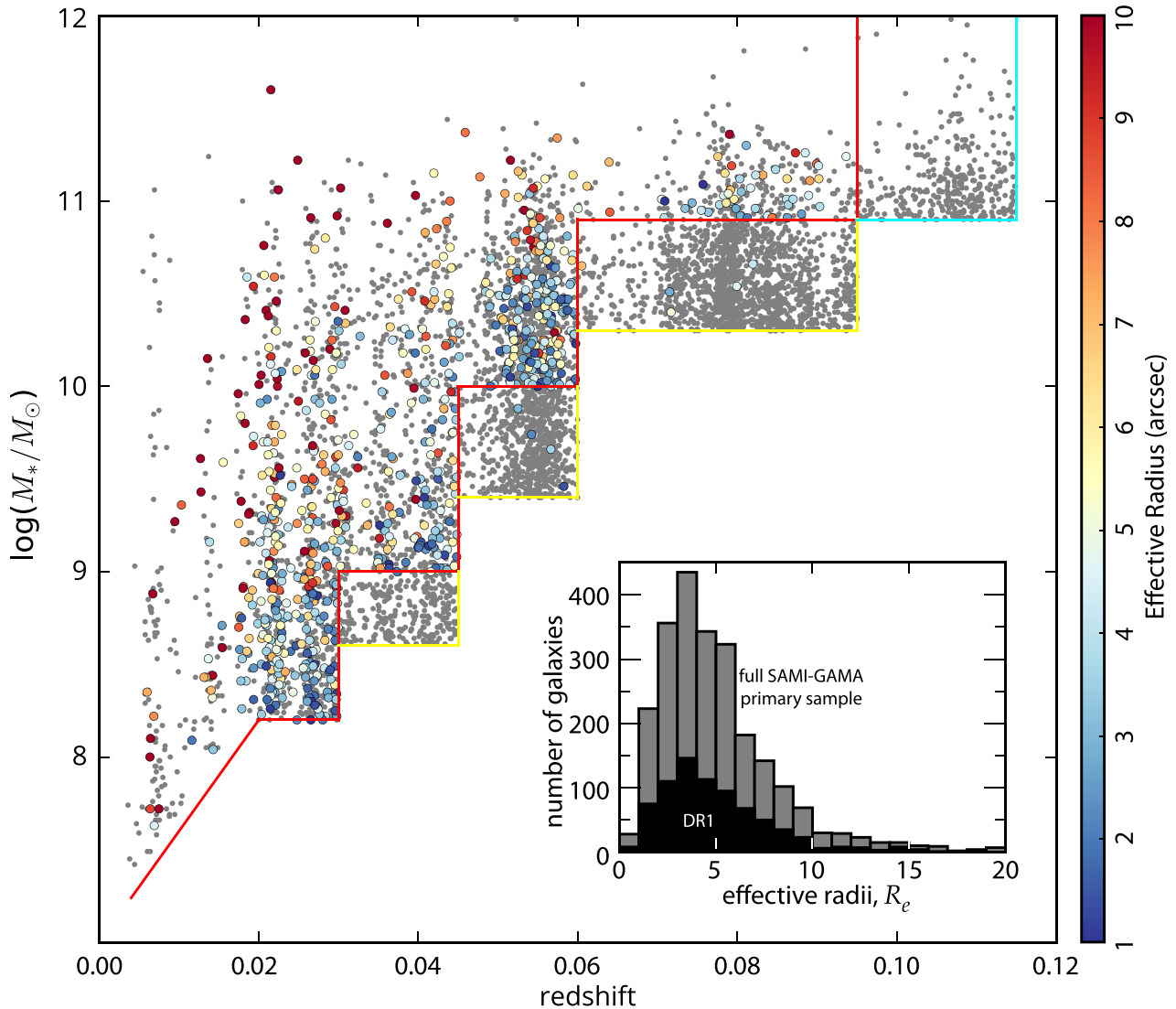
Figure 1. The SAMI-GAMA portion of Galaxy Survey targets in the redshift versus stellar mass plane. The primary targets lie above the red line, and secondary targets lie above the cyan (higher redshift) or yellow (lower mass) line. Light grey points show the full SAMI-GAMA sample, while the targets comprising DR1 are coloured by effective radius (R_e) in arcsec. The inset histogram illustrates that the R_e distribution of the DR1 galaxies (black) is representative of the full primary sample (grey).

2.5 Comparing SAMI with other large IFS surveys

2.5.1 Spatial resolution

The SAMI Galaxy Survey has less spatial resolution elements per galaxy than most first-generation IFS surveys. Fig. 2 shows the distribution of spatial resolutions delivered in this release. By design, the SAMI Galaxy Survey has uniform sampling of galaxies in units of effective radius (R_e). First-generation surveys were based on instruments with a single IFU with a large field of view on the sky and many spatial samples. For example, CALIFA uses the PPAK fibre bundle (Kelz et al. 2006) that contains 331 science fibres and uses this bundle to target a single galaxy at a time. In contrast, SAMI has 793 target fibres, a factor of 2.4 more fibres, but distributes them over 13 targets, with a much smaller field of view per IFU. The ATLAS³D and CALIFA surveys target lower redshift galaxies better matched in size to their larger IFUs, leading to higher spatial resolution – CALIFA's 2.68 arcsec fibres cover 0.28–1.63 kpc

over the redshift range of their survey, while our measured FWHM (full width at half-maximum) covers 0.21–4.7 kpc over our redshift range. Therefore, these first-generation surveys continue to serve as a benchmark for local (<100 Mpc) galaxies, while second-generation surveys will provide much larger samples of slightly more distant galaxies (typically >100 Mpc).

2.5.2 Spectral resolution

In the neighbourhood of the Hα emission line, the SAMI Galaxy Survey has higher spectral resolution than most other first- and second-generation surveys. In the blue arm, the large number of spectral features visible drives the survey design to broad wavelength coverage (3700–5700 Å), leading to a resolution of $R \simeq 1812$. However, in the red arm, by limiting spectral coverage to a ~1100 Å region around the Hα emission line, we can select a higher spectral resolution, $R \simeq 4263$. This selection is distinct
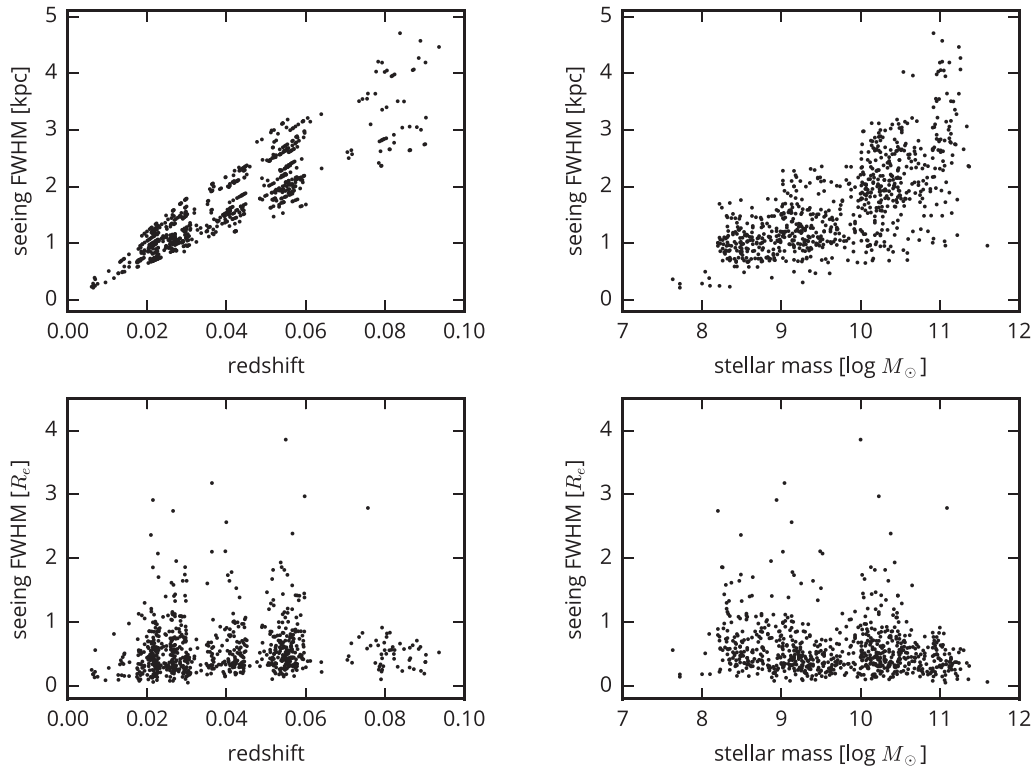
Figure 2. Distribution of spatial resolution achieved for each galaxy in the DR1 catalogue as a function of redshift and mass. Spatial resolution is shown as the FWHM of a Moffat-profile fit to the simultaneously observed star after reconstruction into a spectral cube (described further in Section 3.4.2).

from most other surveys, such as CALIFA and MaNGA, with $R \simeq 850$ and $R \simeq 2000$ respectively around the Hα line. Therefore, analyses based on SAMI data can better separate distinct kinematic components (e.g. in outflows; see Ho et al. 2014, 2016a), can more accurately measure the gas velocity dispersion in galaxy discs (Federrath et al. 2017), and can investigate the kinematics of dwarf galaxies. The trade-off for the higher spectral resolution in the red arm is more limited spectral coverage, which only extends to ~7400 Å, whereas MaNGA reaches to ~1 μm.

2.5.3 Surface brightness sensitivity

We estimate our surface brightness sensitivity in the SDSS g band, as the SAMI blue arm completely contains the filter bandpass. Specifically, we use the final blue cubes and estimate the signal to noise (S/N) in a region 200 Å wide at the centre of the g band (4686 Å). The median S/N per spaxel (0.5 arcsec × 0.5 arcsec) per spectral pixel (1.04 Å) is 3.1 at a surface brightness of 23 AB mag arcsec^{-2} in the g band. To compare this to other surveys, we scale (accounting for covariance) to the same equivalent apertures and spectral sampling as their quoted sensitivities. The CALIFA survey has S/N = 3 at a surface brightness of 23 AB mag arcsec^{-2} in the g band, but within 1 arcsec × 1 arcsec spaxels and a 2.3 Å resolution element (Sánchez et al. 2016). The equivalent S/N for the same aperture in SAMI is 6.6. MaNGA quotes an S/N of 4–8 (Å^{-1} per 2 arcsec fibre) at an r-band limit of 23 AB mag arcsec^{-2}. Within the same aperture, our median S/N is 7.1, but in a different band.

2.5.4 Environment measures

The SAMI Galaxy Survey also benefits from more complete and accurate environmental density metrics than other IFS surveys. The

GAMA survey has much greater depth ($r < 19.8$ versus $r < 17.8$) and spectroscopic completeness (>98 per cent versus ≃ 94) than the SDSS on which the MaNGA survey is based (Driver et al. 2011 and Alam et al. 2015, respectively). Therefore, GAMA provides several improved environmental metrics over SDSS, including group catalogues and local-density estimates (Robotham et al. 2011 and Brough et al. 2013, respectively). For example, 58 per cent of primary survey targets are members of a group identified from GAMA (containing two or more galaxies based on a friends-of-friends approach; see Robotham et al. 2011), but only 15 per cent are members of a group identified from SDSS (Yang et al. 2007).

2.5.5 Range in mass

The SAMI survey provides a broader range in mass of galaxies than MaNGA at the expense of more variability in the radial coverage of galaxies. Our target selection aims to be 90 per cent complete above the stellar-mass limit for each redshift interval targeted while covering a large range in stellar mass ($8 \lesssim \log(M_*/M_\odot) \lesssim 11.5$). This selection results in a more extensive sampling of low-mass galaxies than previous surveys. It also differs from the MaNGA selection, which targets galaxies in a relatively narrow luminosity range at each redshift. The MaNGA selection leads to less variability in the radial extent of the data relative to galaxy size.

3 CORE DATA RELEASE

The galaxies included in DR1 are drawn exclusively from the SAMI-GAMA sample. The included core data products are the regularly gridded flux cubes (spectral cubes). All of the core data included have met minimum quality standards, and the quality of the final data has been measured with care.

Table 1. SAMI-GAMA sample primary and filler targets (see Fig. 1) observed by end of 2016 and their DR1 release.

	Number of targets		
	Primary	Filler	Total
Whole SAMI-GAMA catalogue	2404	2513	4917
Observed to 2015 June	822	10	832
In this release	763	9	772
Observed through 2016	1267	44	1311

3.1 Galaxies included in DR1

Galaxies in DR1 are drawn from all 832 galaxies observed in the SAMI-GAMA sample through 2015 June (AAT semesters 2013A to 2015A). This includes all galaxies in the survey's EDR (but the data for those galaxies have been reprocessed for this release). Table 1 includes the additional 18 months of observing to the end of 2016, and shows how the DR1 galaxy numbers compare to the current progress of the SAMI Galaxy Survey in the GAMA regions. The distribution of these targets in the stellar mass–redshift plane, on the sky, and in the star formation rate–stellar mass plane can be seen in Figs 1, 3 and 4, respectively.

We have not included some observed galaxies in DR1 for quality control reasons. From the 832 galaxies, we removed those with

(i) fewer than six individual exposures meeting the minimum standard of transmission greater than 0.65 and seeing less than 3 arcsec FWHM (48 galaxies removed); and

(ii) individual observations that span more than one month for a single field and have differences in their heliocentric velocity frames of greater than 10 km s^{-1} (12 galaxies removed).[2]

After removing observations that did not meet these data quality requirements, 772 galaxies remain.

Galaxies included in DR1 may have a small bias towards denser regions over the full field sample. The order in which galaxies are observed over the course of the survey is set by the tiling process, which allocates galaxies to individual observing fields. Tiling is based only on the sky distribution of galaxies – not their individual properties. Initial tiles are allocated preferentially to regions with higher sky density to maximize the efficiency of the survey over all. Fig. 3 shows the three GAMA-I fields (G09, G12, and G15) and the sky distribution of galaxies in this data release compared with the overall SAMI field sample.

DR1 galaxies are distributed across the full range of the primary sample in redshift, stellar mass, and effective radius as illustrated in Fig. 1. A Kolmogorov–Smirnov test indicates that the DR1 sample has the same effective radius distribution as the SAMI field sample (*D*-statistic=0.025, *p*-value=0.85). However, there is a difference in the distribution of stellar mass (*D*-statistic=0.08, *p*-value=0.001), such that lower mass galaxies are slightly over-represented in the DR1 sample.

3.2 Changes in data reduction methods since the EDR

For DR1, we use the SAMI PYTHON package snapshot identified as MERCURIAL changeset `0783567f1730`, and 2DFDR version 5.62

with custom modifications. The version of 2DFDR is the same as for our EDR (Allen et al. 2015), and all of the modifications are described by Sharp et al. (2015). These changes have been integrated into subsequent public release versions of 2DFDR. Changes in the SAMI package are described in the rest of this section.

3.2.1 Fibre-throughput calibration

To achieve good flux calibration and uniform image quality, the relative throughput of each of the 819 fibres (including 26 sky fibres) must be normalized to a common value. We have improved the approach for normalizing the fibre throughputs over that used in our EDR.

The fibre-throughput calibration used in our EDR had two shortcomings that limited data quality, particularly from the blue arm of the spectrograph. In our EDR, the relative throughput of individual fibres was primarily determined from the integrated flux in the night-sky lines for long exposures, and from the twilight flat-fields for short exposures. However, the blue data (3700–5700 Å) include only one strong night-sky line, 5577 Å, so are particularly susceptible to two problems. First, sky lines are occasionally impacted by cosmic rays, leading to poor throughput estimates for individual fibres. Secondly, the limited photon counts in the sky line limit the estimates of the relative throughput to $\simeq$ 1–2 per cent. Sky lines are also spatially and temporally variable, but these variations are averaged over our 30 min integrations, such that the residual sky after subtracting using the sky lines throughput is also $\simeq$ 1–2 per cent.

For DR1, the relative fibre throughputs were calibrated from either twilight flat-field frames or from dome flat-field frames for fields where no twilight flat was available. The night-sky spectrum was then subtracted using this calibration. If the residual flux in sky spectra was excessive (mean fractional residuals exceeded 0.025), then the fibre throughputs were remeasured using the integrated flux in the night-sky lines (as in the EDR). If all sky lines in a fibre were affected by bad pixels (typically only an issue for the blue wavelength range, which covers only a single sky line), then the mean fibre-throughput calibration derived from all other frames of the same field was adopted. The sky subtraction was then repeated with the revised throughput values. The method that provided the final throughput calibration is listed with the cubes in the online data base. This approach ensures that, for the calibration options available, the best option is used to calibrate the fibre throughputs.

3.2.2 Flux calibration

The flux calibration process has been improved over our EDR to better account for transparency changes between individual observations of a field and improve overall flux calibration accuracy. In our EDR, the absolute flux calibration was applied after forming all cubes for a field of 12 galaxies and 1 secondary standard star. All objects in the field were scaled by the ratio of the field's secondary standard star observed *g*-band flux to the SDSS photometry *after combining individual observations into cubes* (for full details, see section 4.4 of Sharp et al. 2015).

For DR1, this scaling has also been applied to each individual RSS frame for a given field *before forming cubes*, i.e. the scaling is now applied twice. This additional scaling ensures that differences in transparency between individual observations are removed before the cube is formed, which improves the local flux calibration accuracy and removes spatial 'patchiness' in the data. The accuracy of the overall flux calibration is discussed in Section 3.4.3.

---

[2] 2DFDR does not include a heliocentric velocity correction for the wavelength scale, so mapping to a heliocentric frame would require a second rebinning. This would reduce resolution and complicate covariance, so we choose not to do so. 2DFDR is being updated to include this correction, and future releases of SAMI data will have the spectral cubes corrected for heliocentric velocity.
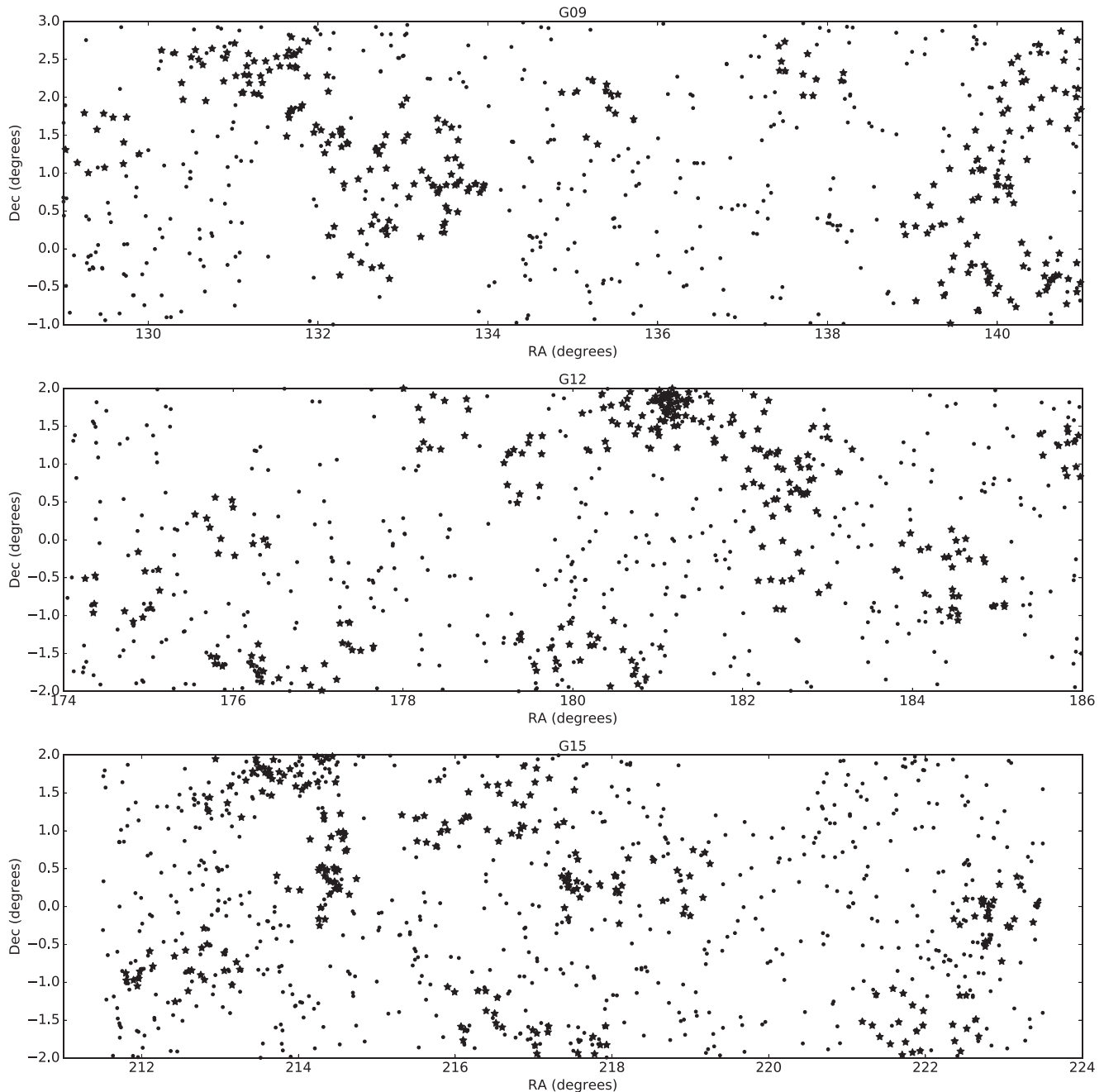
Figure 3. Distribution on the sky of the SAMI-GAMA sample, covering GAMA regions G09, G12, and G15. The primary targets of the complete field sample are shown by the small points, and targets included in this DR1 are shown by stars.

3.2.3 Differential atmospheric refraction correction

For DR1, we have improved the correction for differential atmospheric refraction over that in our EDR. The atmospheric dispersion is corrected by recomputing the drizzle locations of the cube at regular wavelength intervals (see section 5.3 of Sharp et al. 2015). In our EDR, the drizzle locations were recomputed when the accumulated dispersion misalignment reached 1/10th of a spaxel (0.05 arcsec). We found that this frequency caused unphysical 'steps' in the spectra within a spaxel. In DR1, we recalculated the drizzle locations when the accumulated dispersion misalignment reached 1/50th of a spaxel, i.e. five times more often than in the EDR. This significantly reduced the impact of atmospheric dispersion on the local

flux calibration within individual spaxels. As described in section 5.7 of Sharp et al. (2015), we do not recompute at every wavelength because this would dramatically increase the size of the covariance data by a factor of ∼6 to approximately 1 GB for each spectral cube. Appendix A elaborates on how atmospheric dispersion affects the quality of the data.

3.3 Core data products included

Several core data products are included in DR1: flux spectral cubes with supporting information, GAMA catalogue data used for the target selection, and Milky Way extinction spectra.
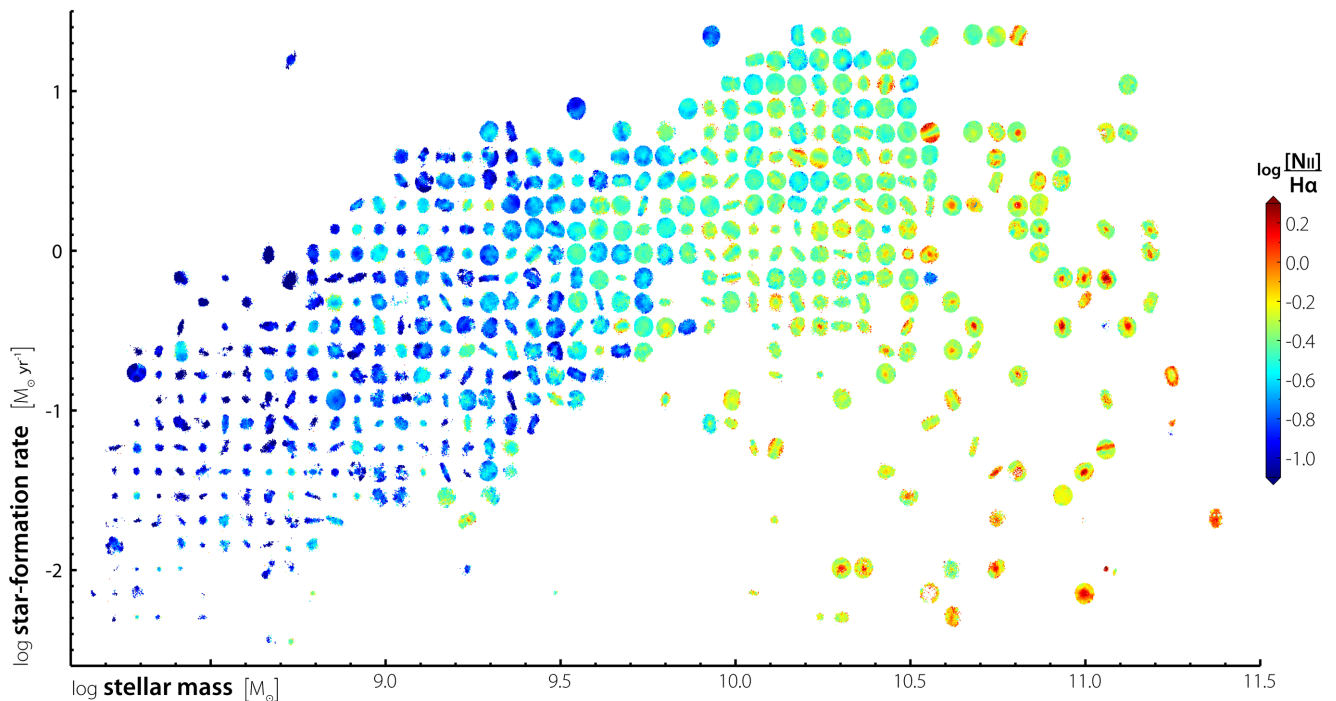
Figure 4. The spatially resolved maps of [Nɪɪ]/Hα within 15 arcsec diameter of DR1 galaxies are arrayed by stellar mass and star formation rate. Not all in DR1 appear because some have insufficient [Nɪɪ] and/or Hα flux for their S/N ratio to exceed 3 across their extent. Some maps have been shifted slightly to avoid overlap, so stellar masses and star formation rates shown are indicative, not exact.

3.3.1 Spectral cubes

The position–velocity spectral flux cubes are the products most users will value. These cubes are presented with the following supporting data, all sampled on the same regular grid.

Variance. The uncertainty of the intensities as a variance, including detector-readout noise and Poisson-sampling noise propagated from the raw data frames.

Spatial covariance. Covariance between adjacent spatial pixels introduced by drizzle mapping on to the regular grid. The covariance and the format of this five-dimensional array are described in section 5.7 of Sharp et al. (2015).

Weights. The effective fractional exposure time of each pixel, accounting for gaps between individual fibres, dithering, etc. These are described in section 5.3 of Sharp et al. (2015).

A world-coordinate system (WCS) for each cube is included. This WCS maps the regular grid on to sky (right ascension and declination) and wavelength coordinates. The origin of the spatial coordinates in the WCS is defined using a 2D Gaussian fit to the emission in the first frame of the observed dither sequence. The wavelength coordinates are defined in the data reduction process from arc-lamp frames. The given coordinates are air wavelengths in the reference frame of the observatory. The accuracy of the spatial coordinates is discussed in Section 3.4.4 and that of the wavelength coordinate in section 5.1.3 of Allen et al. (2015).

Also provided for each spectral cube are estimates of the point spread function (PSF) of the data in the spatial directions. The PSF is measured simultaneously with data collection using the secondary standard star included in each SAMI field. We provide the parameters of a circular-Moffat-profile fit to that star image (i.e. the flux calibrated red and blue star cubes summed over the wavelength

axis). The Moffat profile has form

$$f = \frac{\beta - 1}{\pi \alpha^2} \left(1 + \left(\frac{r}{\alpha} \right)^2 \right)^{-\beta}, \qquad (1)$$

where α and β parametrize the fit and $r^2 = x^2 + y^2$ is the free variable denoting spatial position (Moffat 1969). The reported PSF is the luminosity-weighted average over the full (i.e. red + blue) SAMI wavelength range. With the parameters of the Moffat-profile fit, we also provide the corresponding FWHM, W, as given by

$$W = 2\alpha \sqrt{2^{(1/\beta)} - 1}, \qquad (2)$$

measured in arcseconds. The distribution of measured PSF is discussed in section 5.3.2 of Allen et al. (2015), and is unchanged in DR1.

These data are typically made available as FITS files (Wells, Greisen & Harten 1981), though the archive system may also provide other formats in future.

Finally, for convenience, we include the exact versions of the GAMA data used in the sample selection of the SAMI field sample. Note that in some cases, newer versions of these data are available from the GAMA survey and should be used for scientific analysis.

3.3.2 Milky Way dust-extinction correction

SAMI spectral cubes are not corrected for dust extinction, either internal to the observed galaxy or externally from Milky Way dust. However, we do provide a dust-extinction-correction curve for each galaxy to correct for the latter. Using the right ascension and declination of a galaxy, we determined the interstellar reddening, $E(B - V)$, from the Planck v1.2 reddening maps (Planck Collaboration XI 2014) and the Cardelli, Clayton & Mathis (1989) extinction law to provide a single dust-correction curve for each spectral cube. Note that this curve *has not been applied to the spectral cubes.* To
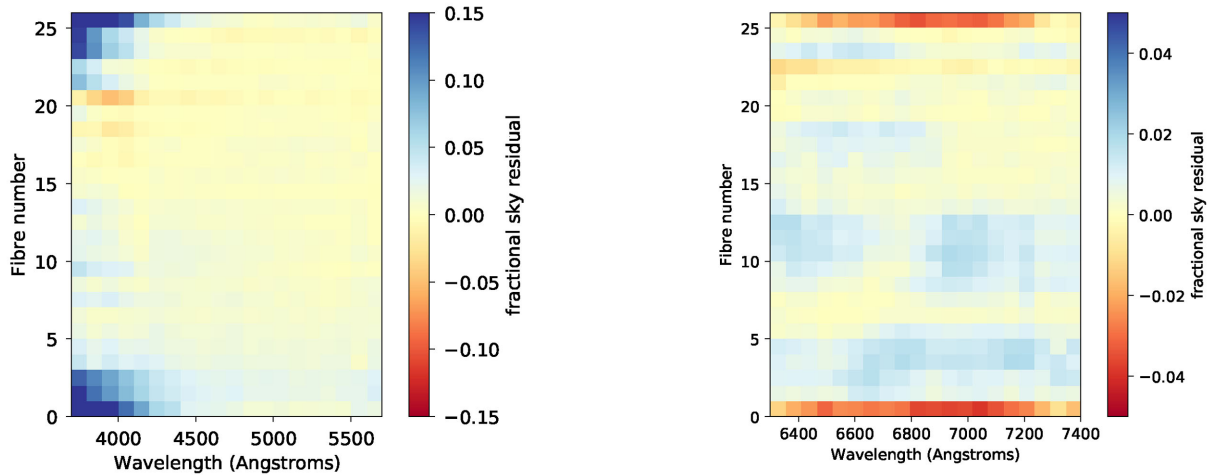
Figure 5. The median fractional sky subtraction residuals as a function of wavelength and fibre number for SAMI sky fibres in the blue (left) and red (right) arms of the spectrograph. The sky fibres are regularly spaced along the SAMI slit, so that sky fibre number also corresponds to approximate location on the AAOmega CCDs. If the sky subtraction was perfect, these residuals would be zero – instead they indicate the likely sky subtraction residuals in science fibres adjacent to these sky fibres. For each sky fibre shown, the spectral direction is sub-divided into 20 uniform bins, and the residual flux is summed in each of these bins, before determining the median residual (across different observed frames). This reduces the impact of shot noise on the residual estimate and allows us to see systematic variations in sky subtraction. A strong increase in the residual in the left-hand corners of the blue CCD is particularly apparent. Note the difference in colour scale between the two images.

correct a SAMI cube for the effects of Milky Way dust, the spectrum of each spaxel must be multiplied by the dust-correction curve.

3.4 Data quality

We now discuss data quality measurements for the core data released. Allen et al. (2015) discuss the quality of the data in our EDR, including fibre cross-talk, wavelength calibration, flat-fielding accuracy, and other metrics. Where data quality does not differ between our EDR and DR1, we have not repeated the discussion of Allen et al. (2015). Instead, we discuss the data quality metrics potentially affected by changes in the data reduction.

3.4.1 Sky subtraction accuracy

The changes to fibre-throughput calibration (see Section 2.4) remove occasional (less than one fibre per frame) catastrophically bad throughputs. It does not change the overall average sky subtraction accuracy, as presented by Allen et al. (2015). The lack of change in sky subtraction precision suggests that fibre throughput and photon counting noise in the blue 5577 Å line are not currently a limiting factor in the precision of sky subtraction.

Systematic residuals after subtracting sky continuum arise from sources such as scattered light in the spectrograph. The residuals are shown as a function of wavelength and sky fibre number in Fig. 5. To clarify the impact of sky subtraction errors, we sum the residual flux in wavelength bins (20 uniform bins per spectrograph arm). The sum reveals sky residuals that would otherwise be dominated by CCD read noise and photon counting errors in a single 0.5–1-Å-wide wavelength channel. Fig. 5 shows that across most of both the blue and red arm CCDs, residuals of the sky-continuum subtraction are ∼1 per cent. However, a strong residual appears at the short-wavelength corners of the blue CCD. This is due to a ghost in the spectrograph caused by a double bounce between the CCD and air–glass surfaces of the AAOmega camera corrector lens (Zhelem, private communication). The ghost results in poor fitting of the fibre profiles, which in turn results in poor extraction and then sky

subtraction. A solution to this using twilight-sky flats to generate fibre profiles has now been developed, but has not been applied to the data in DR1. In the red arm, sky fibres 1 and 26 show an increased sky residual. These fibres exhibit higher residuals because they are at the very edge of the slit/detector and the current scattered light model does not accurately capture the turn-down of the scattering profile at the edge of the slit. A modified scattered light model will be implemented in future releases.

3.4.2 Point spread function

The spatial PSF is measured by fitting a Moffat function to the reconstructed image of the secondary standard star in each SAMI field. SAMI fibres have a diameter of 1.6 arcsec, therefore in seeing ≲ 3 arcsec, the PSF in the individual dithered exposures is under-sampled. Stacking images introduces additional uncertainty from misalignment of the seven frames (fig. 15 of Allen et al. 2015), and from combining exposures with slightly different seeing. Therefore, the PSF of the final spectral cube is degraded from the PSF of the individual frames. To determine the seeing in the input frames, we have fitted a model point source with seeing to the measured intensities in the secondary standard star for each SAMI observation. In Fig. 6, we compare the mean of the seeing measurements for each individual observation (input FWHM) with the FWHM measured from the reconstructed stellar image in the final combined spectral cube (output FWHM). For small input FWHM (≈1 arcsec), the output FWHM increases by 50 per cent. This regime is likely dominated by PSF undersampling. When the input FWHM exceeds ≈1.5 arcsec, the output FWHM is typically 10 per cent larger. No stars have FWHM >3.0 arcsec because such data are excluded by a quality control limit. In summary, DR1 spectral cubes have a mean PSF of 2.16 arcsec (FWHM).

3.4.3 Flux calibration

The relative flux calibration as a function of wavelength in DR1 is consistent with that in the EDR. By comparing SAMI data with
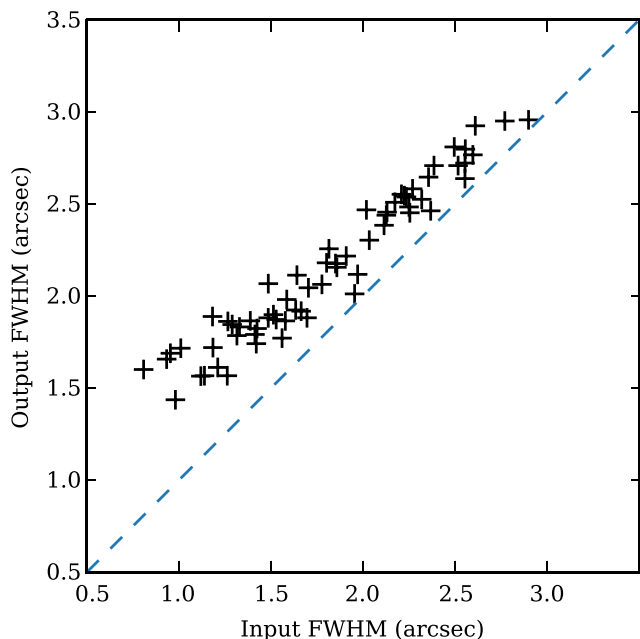
Figure 6. Comparison of the FWHM measured on the final reconstructed secondary standard cubes (output FWHM) versus the mean of the estimated seeing FWHM using model fits to individual dithered exposures (input FWHM). FWHMs are from Moffat-profile fits. The dashed blue line is the 1:1 relation. The output FWHM is typically larger than the input by 10 per cent. The mean FWHM for DR1 is 2.16 arcsec, and the standard deviation is 0.41 arcsec.

SDSS g- and r-band images, Allen et al. (2015) showed that SAMI-derived $g - r$ colours have 4.3 per cent scatter, with a systematic offset of 4.1 per cent, relative to established photometry.

To test the absolute flux calibration, we directly compared SAMI cubes to SDSS g-band images. The blue SAMI cubes were convolved with the SDSS g-band filter and then the integrated g-band flux was calculated within an 8 arcsec diameter circular aperture. An equivalent integrated aperture flux was calculated directly from SDSS images, after first aligning the images (see Section 3.4.4 below) and convolving them to the median SAMI seeing of 2.1 arcsec FWHM (assuming a Gaussian PSF). 137 of the faintest galaxies (with integrated g-band fluxes $< 100 \,\mu$Jy) were not included, to avoid extra scatter from low S/N. A further five galaxies were removed due to regions of bad pixels biasing the integrated fluxes. The distribution of flux ratios is shown in Fig. 7. The median SAMI/SDSS flux ratio is 1.051 ± 0.005 and the flux ratio rms is 0.10, with 95 per cent of objects having flux ratios within ± 0.16 of the median. The agreement between SAMI and SDSS flux calibrations is considerably better than that given by Allen et al. (2015), but most of this improvement derives from a better comparison method, which does not depend on Petrosian or model magnitudes, rather than actual changes to the flux calibration method of SAMI.

3.4.4 WCS and centring of fibre bundles in cubes

The accuracy of the WCS is limited by the stability and accuracy of the single Gaussian fit on the observation chosen as the reference (typically the first frame, see Section 3.3.1 and section 5.2 of Sharp et al. 2015). By fitting to the individual observed galaxies, we lose some robustness. However, we minimize the impact of mechanical errors (plate manufacturing, movement of the connectors within the drilled holes, and uncertainty of the bundle positions) on the WCS
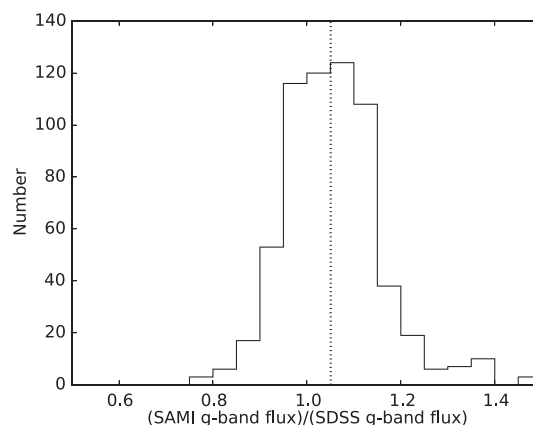

Figure 7. The distribution of the measured g-band flux ratio between SAMI cubes and SDSS images within an 8 arcsec diameter circular aperture. The vertical dotted line shows the median flux ratio of 1.051 ± 0.005.

accuracy. Examining the data, we have identified three possible failure modes of our approach.

(i) The fit may identify a bright star within the field of view of the hexabundle instead of the galaxy of interest. Examples include galaxies 8570 and 91961.

(ii) The catalogue coordinate may not correspond to a peak in the surface brightness of the object, such as one with a very disturbed morphology, or for objects where the catalogue coordinate has been intentionally set to be between two galaxies (galaxies with BAD_CLASS=5 in the target catalogue), see Bryant et al. (2015) for details. Examples include galaxy 91999.

(iii) Finally, the circular Gaussian distribution may not represent the true flux distribution well, leading to some instability or bias in the fit result. Examples include large, extended galaxies such as 514260.

In these cases, the WCS origin may not be very accurate, and the hexabundle field of view may not be well centred in the output spectral cube.

We carry out two tests to characterize uncertainties in the WCS. The first is an internal check that considers offsets at different stages of the alignment process to constrain the expected WCS uncertainties. The second cross-correlates the reconstructed SAMI images with SDSS broad-band images to measure the offset between SAMI and SDSS coordinates. These two tests, which we detail in the following paragraphs, suggest that the WCS accuracy is $\lesssim 0.3$ arcsec for most galaxies, except for the failures noted above.

The internal tests to examine WCS uncertainties use alignment offsets to infer bounds on the typical size of the WCS uncertainties. The first dither pointing of an observation aims to centre each galaxy in its bundle. The dither-alignment transformation aligns the galaxy centroid positions in a dither with the galaxy centroid positions in the first ('reference') frame of an observation. Fig. 8 shows the rms of the residuals for all bundles in a dither after the dither was aligned with the reference frame. The residuals are shown for transformations that are translation-only, translation, and rotation, and using the full transformation of a translation, rotation, and scaling. At least translation is necessary because the dithers are deliberately spatially offset. However, rotation is also important in aligning the dither frames to the centre of the cubes as the SAMI instrument plate holder has a small (~ 0.01 deg) bulk rotation away from its nominal orientation. This rotation suffices to generate offsets from the nominal bundle centres of up to ~ 1 arcsec at the edge of the field
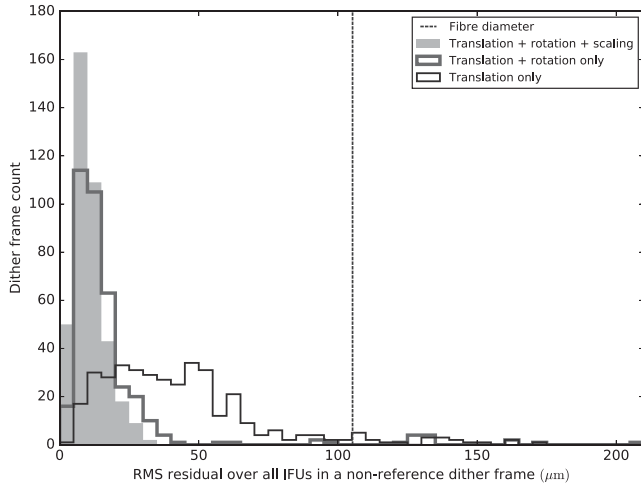
Figure 8. Histograms of the residuals after aligning dither frames. The alignment attempts to simultaneously place bundle centroids of all IFUs in a dither on to the centroids of the 'reference' dither frame. Distributions of residuals are shown for transformations with only a translation, a translation, and a rotation, and the full transformation of a translation, a rotation, and a scaling.

of view. A further improvement is gained using the modification of the plate scale, due to differential atmospheric refraction causing small positional shifts over the course of an observation. The mean rms of $\sim 11\,\mu m$ (0.16 arcsec) for the full transformation reflects how accurately the data are spatially combined for a typical galaxy and hence provides a lower limit to the WCS uncertainty.

The cross-correlation test of the WCS accuracy compares the spatial flux distribution of the final, reconstructed SAMI cubes to SDSS *g*-band images. Each cube is multiplied by the SDSS-*g*-band-filter response and then summed spectrally. The resulting image is then cross-correlated with an SDSS *g*-band image. These SDSS images are centred on the expected coordinates of the galaxy (based on the GAMA input catalogue), are 36×36 arcsec in size, and have been re-sampled to the same 0.5 arcsec pixel scale as the SAMI cubes. The cross-correlation offset (measured using a fit to the peak in the cross-correlation image) is then the difference between the SAMI WCS and the SDSS WCS. These differences are shown in Fig. 9. Outliers in most cases are caused by the cross-correlation centring on bright stars that are present in the SDSS image, but not in the SAMI field of view. Visual checks of outliers also identified five galaxies with gross errors in their SAMI cube WCS, caused by the data reduction centroiding on a bright star in the SAMI field of view rather than the target galaxy (catalogue IDs 8570, 91961, 218717, 228104, and 609396). When outliers are removed using an iterative 5σ clipping of the radial offset (that removes 2.5 per cent of objects), the mean of the remaining differences is -0.077 ± 0.011 arcsec in right ascension and -0.020 ± 0.013 arcsec in declination. Even with clipping, there are residual non-Gaussian tails to the positional offsets. The root-mean-square scatter is 0.31 arcsec in right ascension and 0.36 arcsec in declination. 68 per cent of the objects have a radial offset of less than 0.36 arcsec, and 90 per cent have a radial offset of less than 0.73 arcsec.

Given that the result of the measurement of the WCS uncertainty in the cross-correlation test is consistent with the bounds suggested by the internal tests, we expect that it is representative of the actual uncertainty in our WCS for most targets. The targets subject to one of the failures mentioned above will have a much larger error in their WCS. We have not explicitly corrected for the failures, or shifted
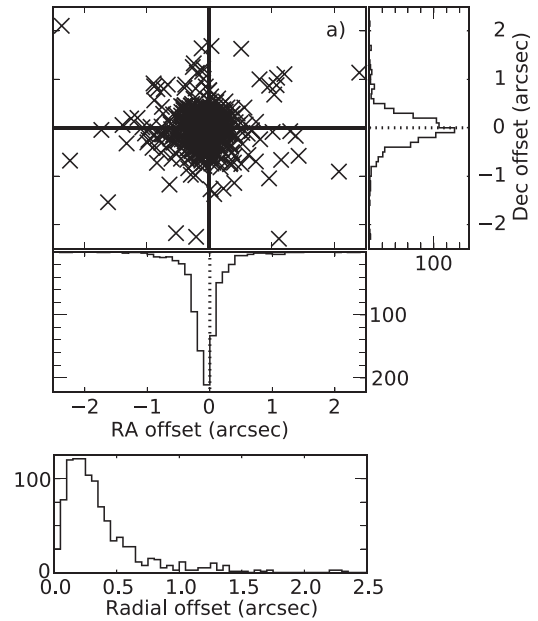

Figure 9. The difference between SAMI and SDSS astrometric solutions based on cross-correlation of images. Top: the distribution of RA and declination differences between SAMI and SDSS, with histograms of the differences in declination and right ascension along the axes. Bottom: histogram of the distribution of differences in radial offset.

the WCS for each galaxy into the SDSS position, as the positional offsets also include the cases where the offset is caused by poor cross-correlation resulting from other bright sources in the SDSS image. Therefore, we do not flag galaxies with poor WCS fits. The only exceptions are the five egregious examples detected 'by eye' and listed above.

3.4.5 Impact of aliasing from sampling and DAR on SAMI data

The combined effects of DAR and limited, incomplete spatial sampling can cause the PSF of IFS data to vary both spatially and spectrally within a spectral cube, an effect we call 'aliasing'. We describe this in Appendix A, but Law et al. (2015) also provide an excellent discussion. Aliasing can cause issues in comparing widely separated parts of the spectrum on spatial scales comparable to, or smaller than, the size of the PSF. Examples are spectral colour and ratios of widely spaced emission lines. We therefore check the impact of aliasing on our data and discuss options for reducing this impact.

To test the impact of aliasing on SAMI data, we check the variation in colour within galaxies expected to have uniform colour across their extent. Uniform colour galaxies are chosen to be passive (no significant emission lines) and to have weak (or flat) stellar-population gradients. The centres of these galaxies also have steep radial profiles (typically a de Vaucouleurs profile). Using only spaxels in the blue SAMI cubes that have a median S/N >15 (per 1.04 Å pixel), we smooth them with a Gaussian kernel in the spectral direction ($\sigma = 15$ Å) to reduce noise, and then sum the flux in two bands at wavelengths 3800–4000 and 5400–5600 Å. These bands are chosen to be narrower than typical broad-band filters, but broad enough to obtain a high total S/N (typically S/N > 100). The separated bands are also more sensitive to the size of the aliasing effects that varies slowly with wavelength (see Appendix A). For each galaxy, we then estimate the rms scatter in the colour formed by the
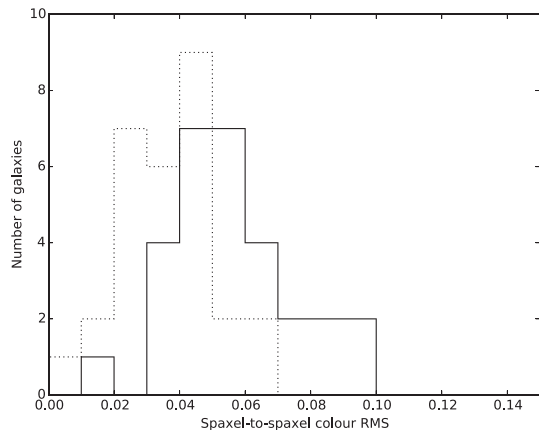
Figure 10. Histogram of rms scatter in colour between spaxels that are 0.5 × 0.5 (solid line) or 1.0 × 1.0 (dotted line) arcsec in size. Galaxies tested are chosen to be passive with uniform colour and the rms is calculated independently for each galaxy.

ratio of the flux in these two bands. Fig. 10 shows the distribution of rms scatter measurements in the spaxel-to-spaxel spectral colour for 29 galaxies. For the default 0.5 × 0.5 arcsec spaxels (solid line in Fig. 10), the median scatter is 0.052 and the 5th–95th percentile range is 0.033–0.093. Summing spaxels 2 × 2 within the cubes so that we have 1.0 × 1.0 arcsec spaxels (dotted line in Fig. 10) leads to a reduced rms with a median value of 0.035 and the 5th–95th percentile range is 0.012–0.061. The reduction in scatter cannot be due to increased S/N due to greater binning, as the summed bands used are too high in S/N. The reduction in scatter when the data are binned to larger spaxels is consistent with the scatter being caused by aliasing in DAR re-sampling.

Aliasing from DAR re-sampling can also affect line ratios. The ratio of the Hα and Hβ emission lines is typically used to estimate dust attenuation. Variations in the PSF at these two wavelengths cause the ratio to reflect not only the true ratio of the two lines, but also the difference in the PSF between the two wavelengths. The latter effect will be most pronounced where there is a sharp change in flux with spatial position in either of the two lines (such as near an unresolved H II region). In such a region, there will be variations pixel to pixel (smaller than the PSF) that are larger than would be indicated by the variance information of the data alone.

One possible method for reducing the impact of aliasing on SAMI data is to smooth it. For example, smoothing the Hα–Hβ line-ratio map by a 2D Gaussian kernel of Gaussian-σ of 0.5 arcsec (one spatial pixel) and truncated to 5 × 5 pixels removes most of the variation caused by aliasing without greatly affecting the output spatial resolution. This smoothing brings the noise properties of the Hα–Hβ line ratio into agreement with Gaussian statistics and significantly reduces variation in the normalized spectra for (point-source) stars. The best choice for the smoothing kernel σ probably ranges between 0.2 and 1 arcsec, depending on the science goal and the level of DAR aliasing associated with the galaxy properties and observational conditions. Smoothing should only be necessary when no other averaging is implicit in the analysis (e.g. smoothing is not necessary for measuring radial gradients).

Aliasing can also appear in another slightly different form that arises from the combination of the incomplete coverage of the focal plane (fill factor) within a bundle. To help address the ~75 per cent fill factor, individual observations are offset or dithered relative to one another by 0.7 arcsec (see Section 2.3). However, differential

dispersion across the 1 deg field changes over the ≈3.5 h observation, causing the dither sizes and directions to vary across the field plate.[3] In rare cases, this can lead to some pixels in the final spectral cubes being poorly covered by input pixels from the original observations (and hence have a low weight, see section 5.1 of Sharp et al. 2015). An example of this is SAMI ID 318936 (shown in Fig. 11), the data for which show a less smooth velocity field because the input fibre footprints have aligned to give poor coverage of bands parallel to the rotation (minor) axis of the galaxy. The choice of dithering strategy is discussed in more detail in section 6 of Sharp et al. (2015).

Alternative data reconstruction schemes may reduce the effects of aliasing from the DAR re-sampling. Smoothing options are discussed further in Medling et al. (2018, submitted) as they pertain to the emission-line value-added products (described briefly in Section 4). In general, only results that depend on the highest possible spatial resolution are likely to be sensitive to aliasing.

4 EMISSION-LINE PHYSICS VALUE-ADDED DATA PRODUCTS

With the core data products described above, our DR1 also includes value-added products based on the ionized gas emission lines in our galaxies. We provide fits for eight emission lines from five ionization species, maps of Balmer extinction, star formation masks, and maps of star formation rate for each galaxy. In this section, we provide an overview of how these data products have been prepared, but full details can be found in Ho et al. (2016b), for emission-line fitting, and Medling et al. (2018, submitted), for emission-line fitting and star formation rate and Balmer-decrement measurements. Examples of these products are shown in Fig. 11 for a selection of galaxies spanning the range of stellar masses in DR1.

4.1 Single- and multi-component emission-line fits

We have fitted the strong emission lines ([O II] 3726,3729, Hβ, [O III] 4959,5007, [O I] 6300, [N II] 6548,6583, Hα, and [S II] 6716,6731) in the spectral cubes with between one and three Gaussian profiles. We fit with the LZIFU software package detailed in Ho et al. (2016b). These fits include corrections for underlying stellar-continuum absorption. LZIFU produces both a single-component fit and a multi-component fit for each spatial pixel of the spectral cube. The latter fits select the optimum number of kinematic components in each spatial pixel.

All lines are fitted simultaneously across both arms of the spectrograph. The blue and red spectral cubes have FWHM spectral resolutions of $2.650^{+0.122}_{-0.088}$ and $1.607^{+0.075}_{-0.052}$ Å, respectively. Assuming that the kinematic profiles are consistent for all lines, the higher resolution in the red helps to constrain the fits in the blue, where individual kinematic components may not be resolved.

LZIFU first fits underlying stellar-continuum absorption using the penalized pixel-fitting routine (PPXF; Cappellari & Emsellem 2004), then uses MPFIT (the Levenberg–Marquardt least-squares method for IDL; Markwardt 2009) to find the best-fitting Gaussian model solution.

Our continuum fits combine template spectra of simple stellar populations from the Medium resolution INT Library of Empirical

---

[3] In aligning the data before combining into the spectral cube, the actual offsets are measured and used instead of the nominal 0.7 arcsec observing pattern, see section 5.2 of Sharp et al. (2015).
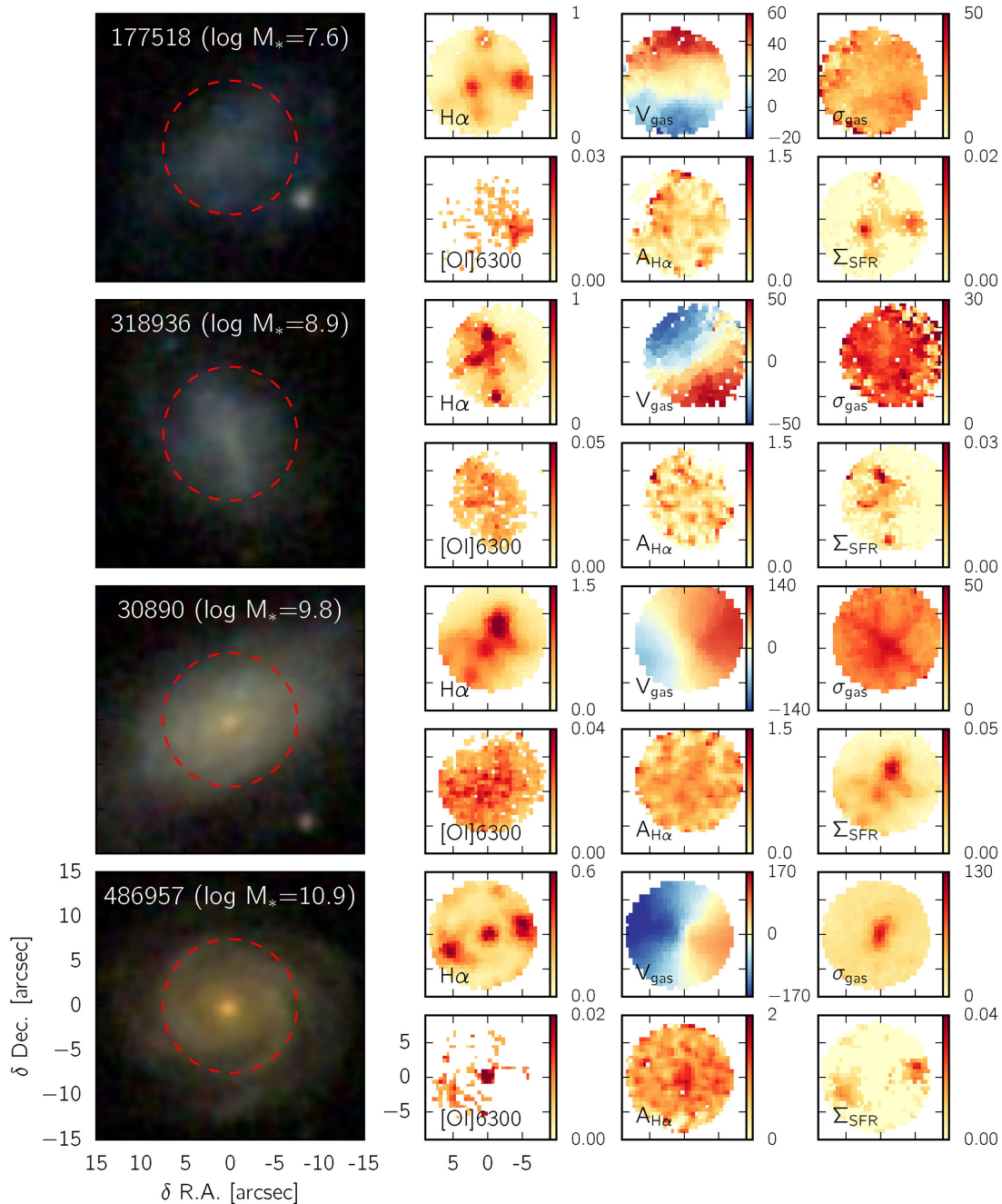
Figure 11. Examples of value-added products for four galaxies spanning the range of stellar masses included in our DR1. The red dashed circles on the SDSS three-colour images (left) indicate the radius of the SAMI fibre bundle. The small panels show the various value-added products: $H\alpha$ flux, gas velocity (v_{gas}), gas velocity dispersion (σ_{gas}), [O I] $\lambda6300$ flux, $H\alpha$ attenuation correction factor ($A_{H\alpha}$), and star-formation-rate-surface-density Σ_{SFR} maps. The units are 10^{-16} erg s^{-1} cm^{-2} spaxel for the flux maps, km s^{-1} for the kinematic maps, magnitude for $A_{H\alpha}$, and $M_\odot$ yr^{-1} kpc^{-2} for Σ_{SFR}.

Spectra (MILES; Vazdekis et al. 2010). These spectra are based on the Padova isochrones (Girardi et al. 2000). The selected templates have four metallicities ($[M/H] = -0.71, -0.40, 0.0, +0.22$) and 13 ages (logarithmically spaced between 63.1 Myr and 15.8 Gyr). In fitting the template spectra to our observed data, Legendre polynomials (orders 2–10) are added (not multiplied) to account for sky subtraction residuals and other possible non-stellar emission within the observed spectral cubes (e.g. scattered light). We expect to improve the scattered light subtraction in the data reduction phase for future releases, and reduce or eliminate the need for these additive polynomials (though they may be necessary for other reasons, see section 3.1 of Cappellari 2017). A reddening curve parametrized by Calzetti et al. (2000) is also applied during fitting. Note that the MILES templates have slightly lower spectral resolution than the red arm of our spectra; therefore, in low-stellar-velocity-dispersion galaxies ($\sigma < 30$ km s^{-1}), the template may underestimate the $H\alpha$ absorption. To account for this and other systematic errors from mismatched templates, we calculate the expected uncertainty in the
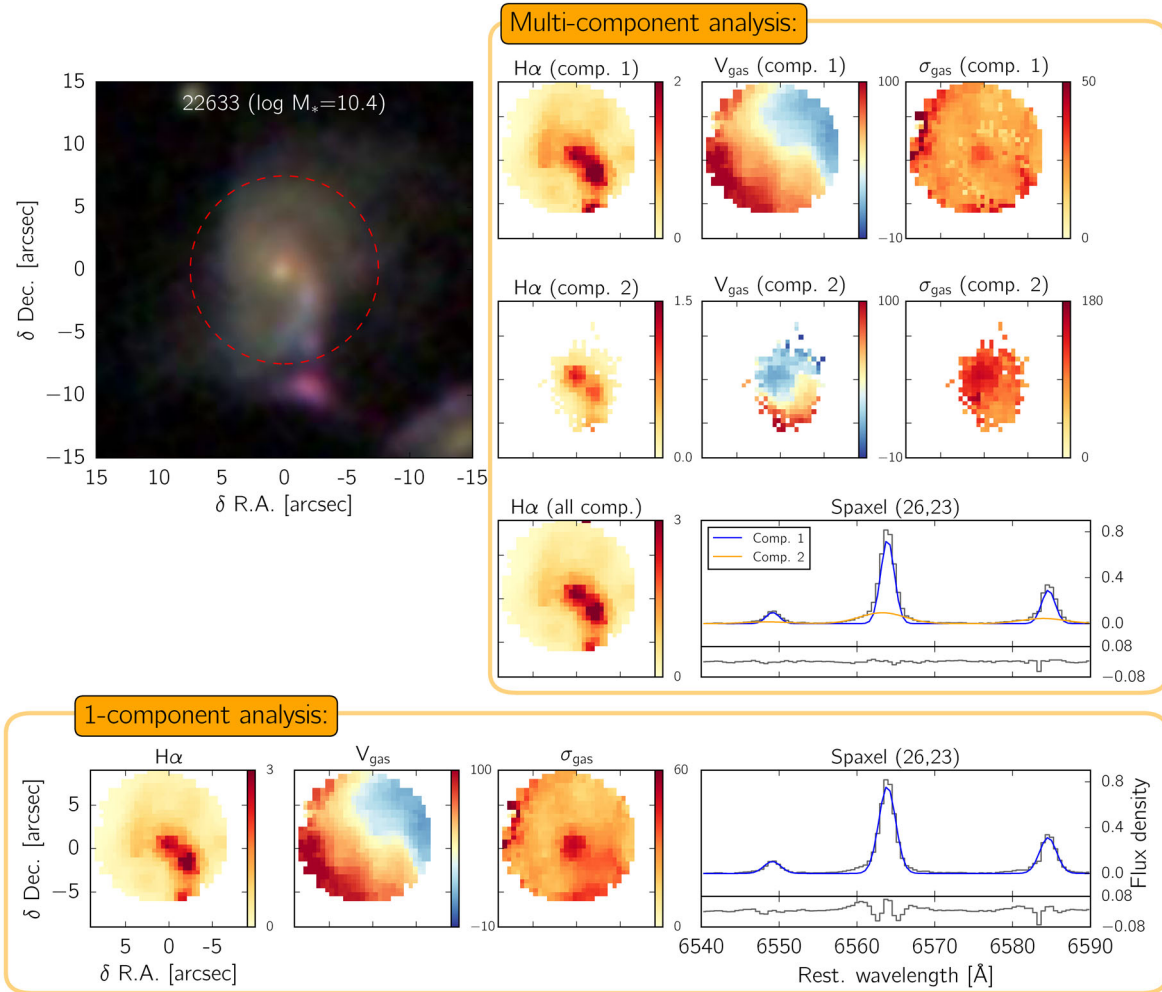
Figure 12. An example comparing multi-component analysis to single-component analysis. The red dashed circle in the SDSS colour image (top left) indicates the SAMI fibre bundle. With the multi-component analysis, we demonstrate that there are two distinct kinematic components in GAMA 22633. The two kinematic components show different velocity dispersions in the spectral fitting panel. The two components also show different H α distribution (H α maps for comp. 1 and 2) and velocity structures (v_{gas} maps for comp. 1 and 2). The nature of the second kinematic component cannot be determined with the one-component analysis that only captures the more dominant narrow kinematic component (bottom row). Including the second kinematic component is necessary to properly model the line profile and reduce the residual.

Balmer absorption from the uncertainty in stellar-population age as measured from the size of the D_n4000 break. This uncertainty is added into the Balmer-emission-flux uncertainty in quadrature.

Each emission line in each spaxel is fitted separately with one, two, and three Gaussian components. In each case, a consistent velocity and velocity dispersion are required for a given component across all lines. Clouds of gas with different ionization mechanisms and kinematics (e.g. AGN versus star-forming clouds) will be recovered as different components given sufficient S/N. For each galaxy, DR1 includes two sets of fits: one that uses a single Gaussian for each line in each spatial pixel ('single component'), another that includes one to three components for each spatial pixel ('recommended components'). Examples of these two fits are shown in Fig. 12. For the fits with recommended components, the number of fits included for each spatial pixel is chosen by an artificial neural network trained by SAMI Team members (LZComp; Hampton et al. 2017). For the recommended components, we also require that each component has S/N ≥ 5 in H α; if this condition is not met, we reduce the number of components until it does.

LZIFU does not consider non-Gaussianity or the effects of 'beam smearing' (described by Davies et al. 2011 and Green et al. 2014). The effects of non-Gaussianity have been explored separately in van de Sande et al. (2017), where variations in the shape of the velocity distribution are generally well below the average velocity dispersions found by LZIFU fits. When expanding the line spread function as a Gauss–Hermite series (van der Marel & Franx 1993), the moments beyond second order (σ) are found to be small with the median value for h_3 (equivalent to skewness) and h_4 (equivalent to kurtosis) being -0.01 in the blue arm and 0.00 in the red arm, with a 1σ spread of 0.016 (van de Sande et al. 2017). However, the choice of functional shape of the velocity distribution is complex, and further investigation may be appropriate depending on the goals of a particular analysis.

The single-component fits include eight maps of line fluxes, and a map each of ionized gas velocity and velocity dispersion. The [O II] 3726,3729 doublet is summed because the blue spectral resolution prevents robust independent measurements of its components. Flux maps of [O III] 4959 and [N II] 6548 are omitted because they are

constrained to be exactly one-third of [O III] 5007 and [N II] 6583, respectively.

The recommended-component fits include maps of the total line fluxes (i.e. the sum of individual components) for each emission line. Additionally, for the Hα line, three maps show fluxes of the individual fit components, and there are three maps each of the velocity and velocity dispersions, which correspond to the individual components of the Hα emission line. The maps showing individual components of Hα flux, velocity, and velocity dispersion are ordered by component width, i.e. first corresponds to the narrowest line and third to the widest. Where there are fewer than three components, higher numbered components are set to the floating point flag NaN, as are all maps without a valid fit.

Fig. 4 illustrates the value of the emission-line fits and the richness of our DR1. It shows how the nature of gas emission changes within galaxies as a function of their stellar mass and star formation rate. At lower stellar masses, emission is driven by star formation, and the gas typically has lower metallicity, which is represented by lower [N II]/Hα ratios (blue). At higher stellar masses, low-star-formation-rate galaxies often host AGN, often resulting in the prominent peak in [N II]/Hα ratio at the centre of the galaxy (red).

Our DR1 includes total-flux model spectral cubes (continuum model plus all fitted emission lines) for direct comparison with the spectral cubes, and maps of quality flags to highlight issues such as bad continuum fits or poor sky subtraction.

4.1.1 Accuracy of GAMA redshifts and systemic velocities from emission-line fits

LZIFU-derived velocities are with reference to the catalogued GAMA redshifts that are listed in the SAMI input catalogue (see Bryant et al. 2015). The GAMA redshifts are on a heliocentric frame and sourced from various surveys such as the main GAMA spectroscopic programme (Hopkins et al. 2013), SDSS (York et al. 2000), and 2dFGRS (Colless et al. 2001). To check the velocity scale of the SAMI cubes, we construct aperture spectra by summing across a 1R_e ellipse. For SAMI cubes that do not extend to 1R_e, we sum over the whole SAMI cube. Each aperture spectrum is then fitted with LZIFU using exactly the same process as the individual cube spaxels.

Fig. 13 shows the velocity difference between the assumed GAMA redshifts and that measured in the aperture spectra. The median difference is $-1.6\,\mathrm{km\,s^{-1}}$ and a robust 1σ range based on the 68 percentile range is 43.9 $\mathrm{km\,s^{-1}}$. The GAMA redshifts used in the SAMI input catalogue were measured using the RUNZ code, and GAMA reports an error on individual RUNZ-derived emission-line redshifts of 33 $\mathrm{km\,s^{-1}}$ from repeated observations (using a robust 68 percentile range; Liske et al. 2015). By subtracting the two in quadrature, we estimate an intrinsic scatter between SAMI DR1 and GAMA of 35 $\mathrm{km\,s^{-1}}$. Reflected in this scatter are (1) differences between the velocity distribution of the light recovered by the small aperture of GAMA survey fibres and our apertures, (2) systematically larger errors on GAMA velocities from previous surveys such as 2dFGRS, and (3) actual velocity errors present in SAMI data. Differences in coverage of the two apertures (1) result in an offset if the small GAMA aperture is not centred on the velocity field. Typical 2dF positioning errors of 0.3–1 arcsec can easily cause velocity differences of >20 $\mathrm{km\,s^{-1}}$ (see section 10.3 of Law et al. 2016). The SAMI aperture spectrum can also be dominated by high Hα flux in the outer parts of galaxies in some cases, e.g. Richards et al. (2014), which would contribute to a velocity difference. Systematically larger velocity errors (2) appear for subsets of
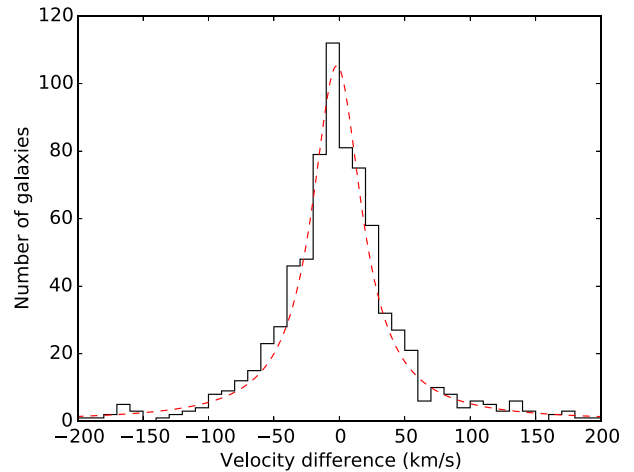

Figure 13. The distribution of rest-frame velocity differences between redshifts catalogued by GAMA and those from LZIFU for SAMI 1R_e aperture spectra, both corrected to the heliocentric reference frame. The red dashed line is a Lorentzian fit to the distribution.

the GAMA survey, such as the 2dFGRS data, which have a velocity resolution of $\sigma \simeq 190\,\mathrm{km\,s^{-1}}$ and rms redshift uncertainty of 85 $\mathrm{km\,s^{-1}}$ (Colless et al. 2001). Actual velocity errors (3) in SAMI data should therefore be small compared to the 35 $\mathrm{km\,s^{-1}}$ upper limit described.

The distribution of velocity differences is well described by a Lorentzian distribution, as found by Liske et al. (2015) for the GAMA velocity uncertainties. The best-fitting Lorentzian is shown by the red dashed line in Fig. 13. The galaxies in the wings of the distribution of velocity differences tend to be those that have lower S/N ratio in the emission-line flux.

4.2 Star formation value-added products

Included with DR1 are value-added products necessary for understanding the spatially resolved star formation. These will be described in detail in a companion paper by Medling et al. (2018, submitted). These are as follows.

(i) Maps of Hα extinction: these are derived by assuming a Balmer decrement (Hα/Hβ ratio); unphysical ratios have extinction corrections set to 1 (no correction). Uncertainties in the extinction correction are also provided.

(ii) Masks classifying each spaxel's total emission-line flux as 'star-forming' or 'other': these are derived using the line-ratio classification scheme of Kewley et al. (2006).

(iii) Maps of star formation rate: these are derived from Hα luminosities and include the extinction and masking above. The conversion factor used is $7.9 \times 10^{-42}/1.53\mathrm{M_\odot}\,\mathrm{yr^{-1}(erg\,s^{-1})^{-1}}$ from Kennicutt (1998). The factor of 1.53 converts from a Salpeter initial mass function (Salpeter 1955) to a Chabrier initial mass function (Chabrier 2003).

5 ONLINE DATABASE

The data of this release are presented via an online data base interface available from the Australian Astronomical Observatory's Data Central.[4] Data Central is a new service of the observatory that

---

[4] Data Central's URL is http://datacentral.aao.gov.au

will ultimately deliver various astronomical data sets of significance to Australian research. Users of the service can find summary tables of the galaxies included in our DR1, browse the data available for individual galaxies, and visualize data interactively online. The service provides for downloading individual and bulk data sets, and a programmatic interface allowing direct access to the data through the HTTP protocol. Also provided are extensive documentation of DR1, the individual data sets within it, and the formatting and structure of the returned data.

Data Central presents data in an object-oriented, hierarchical structure. The primary entities of the data base are astronomical objects, such as stars or galaxies. These entities have various measurements and analysis products associated with them as properties. For example, each galaxy in our DR1 is an entity in the data base, with properties such as red and blue spectral cubes, LZIFU data products, and star formation maps. In the future, these galaxies may also have data from other surveys associated as properties. This structure is designed to provide an intuitive data model readily discoverable by a general astronomer, and is an evolution of the survey's original concept for structuring, archiving, and serving the data: samiDB (Konstantopoulos et al. 2015).

Further development of Data Central is planned. Most relevant to the SAMI Galaxy Survey will be addition of all data products of the GAMA survey, enabling seamless querying of SAMI and GAMA as a single data set. Also planned are more tools for interacting with the data online. As this development progresses, the online user interface is expected to continue to evolve, but the data of DR1 (and their provenance) are stable and in their final form on the Data Central service.

6 SUMMARY AND FUTURE

The SAMI Galaxy Survey is collecting optical IFS for ∼3600 nearby galaxies to characterize the spatially resolved variation in galaxy properties as a function of mass and environment. The survey data are collected with the SAMI instrument on the AAT. Survey targets are selected in two distinct samples: a field sample drawn from the GAMA survey fields and a cluster sample drawn from eight massive clusters.

With this paper, we release spectral cubes for 772 galaxies from the GAMA sample of the survey, one-fifth of the ultimate product. We also release value-added products for the same galaxies, including maps of emission-line fits, star formation rate, and dust extinction. These data are well suited to studies of the emission-line physics of galaxies over a range of masses and rates of star formation. The spectral cubes enable a multitude of science in other areas.

The next public data release of the SAMI Galaxy Survey is planned for mid-2018, and will include further data and value-added products.

ACKNOWLEDGEMENTS

Author contributions. AWG and SMC oversaw DR1 and edited the paper. SMC is the survey's Principal Investigator. JBH and SMC wrote the introduction. JB oversaw the target selection and wrote those parts of the paper. NS wrote sections on the changes to the data reduction, and oversaw the data reduction with JTA and RS. ITH oversaw the emission-line fits and produced Figs 11 and 12, with BG providing additional clarifying text. AMM ran quality control on the emission-line fits, produced the higher order value-added data products, and coordinated ingestion of these to the data base. BG helped coordinate preparation of value-added products for release. MJD and LC oversaw the formatting and preparation of all data for inclusion in the online data base. JvdS prepared the survey overview diagram, Fig. 4. SMC checked the flux calibration accuracy and wrote Section 3.4.3 and prepared Fig. 7. ADT and SMC measured the accuracy of the WCS information and wrote the corresponding Section 3.4.4. RMM provided heliocentric velocity corrections. FDE and JTA created Figs 6 and 7 and contributed to the data reduction software and to the assessment of the data quality, Section 3.4. AWG, EM, LH, SO, MV, KS, and AMH built the online data base serving the data. Remaining authors contributed to overall team operations including target catalogue and observing preparation, instrument maintenance, observing at the telescope, writing data reduction and analysis software, managing various pieces of team infrastructure such as the website and data storage systems, and innumerable other tasks critical to the preparation and presentation of a large data set such as this DR1.

REFERENCES

AAO software Team, 2015, Astrophysics Source Code Library, record ascl:1505.015

Alam S. et al., 2015, ApJS, 219, 12

Allen J. T. et al., 2014, Astrophysics Source Code Library, record ascl:1407.006

Allen J. T. et al., 2015, MNRAS, 446, 1567

Baldry I. K. et al., 2012, MNRAS, 421, 621

Bland-Hawthorn J. et al., 2011, Opt. Express, 19, 2649

Blanton M. R., Moustakas J., 2009, ARA&A, 47, 159

Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396

Brough S. et al., 2013, MNRAS, 435, 2903

Bryant J. J., O'Byrne J. W., Bland-Hawthorn J., Leon-Saval S. G., 2011, MNRAS, 415, 2173

Bryant J. J., Bland-Hawthorn J., Fogarty L. M. R., Lawrence J. S., Croom S. M., 2014, MNRAS, 438, 869

Bryant J. J. et al., 2015, MNRAS, 447, 2857

Bundy K. et al., 2015, ApJ, 798, 7

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Cappellari M., 2017, MNRAS, 466, 798

Cappellari M., Emsellem E., 2004, PASP, 116, 138

Cappellari M. et al., 2011a, MNRAS, 413, 813

Cappellari M. et al., 2011b, MNRAS, 416, 1680

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Chabrier G., 2003, PASP, 115, 763

Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320

Colless M. et al., 2001, MNRAS, 328, 1039

Cortese L. et al., 2016, MNRAS, 463, 170

Courtes G., Georgelin Y., Monnet R. B. G., Boulesteix J., 1988, in Robinson L. B., ed., Instrumentation for Ground-Based Optical Astronomy. Springer-Verlag, New York, p. 266

Croom S. M. et al., 2012, MNRAS, 421, 872

Davies R. et al., 2011, ApJ, 741, 69

Davies L. J. M. et al., 2015, MNRAS, 452, 616

Davis T. A., Bureau M., 2016, MNRAS, 457, 272

Davis T. A. et al., 2011, MNRAS, 417, 882

Dressler A., 1980, ApJ, 236, 351

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

Driver S. P. et al., 2016, MNRAS, 455, 3911

Emsellem E. et al., 2011, MNRAS, 414, 888

Federrath C. et al., 2017, MNRAS, 468, 3965

Fogarty L. M. R. et al., 2012, ApJ, 761, 169

Förster Schreiber N. M. et al., 2009, ApJ, 706, 1364

Fruchter A. S., Hook R. N., 2002, PASP, 114, 144

Genel S., Fall S. M., Hernquist L., Vogelsberger M., Snyder G. F., Rodriguez-Gomez V., Sijacki D., Springel V., 2015, ApJ, 804, L40

Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, A&AS, 141, 371

Green A. W. et al., 2014, MNRAS, 437, 1070

Hampton E. J. et al., 2017, MNRAS, 470, 3395

Hill G. J., 2014, Adv. Opt. Technol., 3, 265

Ho I.-T. et al., 2014, MNRAS, 444, 3894

Ho I.-T. et al., 2016a, MNRAS, 457, 1257

Ho I.-T. et al., 2016b, Ap&SS, 361, 280

Hopkins A. M. et al., 2013, MNRAS, 430, 2047

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Kelz A. et al., 2006, PASP, 118, 129

Kennicutt R. C., Jr, 1998, ApJ, 498, 541

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Kewley L. J., Rupke D., Zahid H. J., Geller M. J., Barton E. J., 2010, ApJ, 721, L48

Konstantopoulos I. S. et al., 2015, Astron. Comput., 13, 58

Law D. R. et al., 2015, AJ, 150, 19

Law D. R. et al., 2016, AJ, 152, 83

Lewis I. et al., 2002, MNRAS, 334, 673

Liske J. et al., 2015, MNRAS, 452, 2087

Magdis G. E. et al., 2016, MNRAS, 456, 4533

Markwardt C. B., 2009, in Bohlender D. A., Durand D., Dowler P., eds, ASP Conf. Ser. Vol. 411, Astronomical Data Analysis Software and Systems XVIII. Astron. Soc. Pac., San Francisco, p. 251

Medling A. M. et al., 2018, MNRAS, submitted

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge Univ. Press, Cambridge

Moffat A. F. J., 1969, A&A, 3, 455

Naab T. et al., 2014, MNRAS, 444, 3357

Owers M. S. et al., 2017, MNRAS, 468, 1824

Pasquini L. et al., 2002, The Messenger, 110, 1

Peacock J. A.ed., 1999, Cosmological Physics. Cambridge Univ. Press, Cambridge, p. 704

Peng Y.-j. et al., 2010, ApJ, 721, 193

Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, ApJ, 757, 4

Planck Collaboration XI, 2014, A&A, 571, A11

Rich J. A., Torrey P., Kewley L. J., Dopita M. A., Rupke D. S. N., 2012, ApJ, 753, 5

Richards S. N. et al., 2014, MNRAS, 445, 1104

Robotham A. S. G. et al., 2011, MNRAS, 416, 2640

Salpeter E. E., 1955, ApJ, 121, 161

Sánchez S. F. et al., 2012, A&A, 538, A8

Sánchez S. F. et al., 2016, A&A, 594, A36

Schaye J. et al., 2015, MNRAS, 446, 521

Sharp R. G., Bland-Hawthorn J., 2010, ApJ, 711, 818

Sharp R. et al., 2006, in McLean I. S., Iye M., eds, Proc. SPIE, Vol. 6269, Ground-based and Airborne Instrumentation for Astronomy. SPIE, Bellingham, p. 62690G

Sharp R. et al., 2015, MNRAS, 446, 1551

Silk J., Mamon G. A., 2012, Res. Astron. Astrophys., 12, 917

Taylor E. N. et al., 2011, MNRAS, 418, 1587

Tremonti C. A. et al., 2004, ApJ, 613, 898

Tully R. B., 1974, ApJS, 27, 415

van de Sande J. et al., 2017, ApJ, 835, 104

van der Marel R. P., Franx M., 1993, ApJ, 407, 525

Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639

Vogelsberger M. et al., 2014, MNRAS, 444, 1518

Wells D. C., Greisen E. W., Harten R. H., 1981, A&AS, 44, 363

Wisnioski E. et al., 2015, ApJ, 799, 209

Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153

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

APPENDIX A: ALIASING CAUSED BY DIFFERENTIAL ATMOSPHERIC REFRACTION CORRECTION AND LIMITED RESOLUTION AND SAMPLING

The effects of differential atmospheric refraction can combine with limited spatial resolution and incomplete sampling to introduce aliasing into the spectra on scales comparable to the PSF. This aliasing is not unique to IFS, though the generally poorer spatial sampling (relative to the seeing) and incomplete spatial coverage (i.e. from gaps between the fibres/spaxels) tend to exacerbate the effect. We will use the much simpler case of a long-slit spectrograph to explain the effect.

To understand the impact of aliasing on spectral data in the presence of differential atmospheric refraction, we consider a simple long-slit image[5] of a white continuum source (i.e. one with a flat spectral energy distribution in wavelength space). The slit has been aligned with the parallactic angle so that atmospheric refraction acts along the length of the slit. For illustrative purposes, we will consider the fairly extreme example of an object observed at a zenith distance of 60 deg. Throughout this section, we assume that the seeing is Gaussian, with 1 arcsec FWHM.

Consider a long-slit image of this object with a spatial scale of 1 arcsec pixel^{-1}. This image is shown (before correction for DAR) on the left of Fig. A1(a). Note that the PSF, even before correction, varies considerably along the wavelength axis due to the poor spatial

---

[5] For our purposes, a long-slit image is an image of a set of simultaneously observed spectra with spatial coordinate along the slit (chosen to be oriented along the parallactic angle for our examples) on the vertical axis and wavelength coordinates along the horizontal axis.
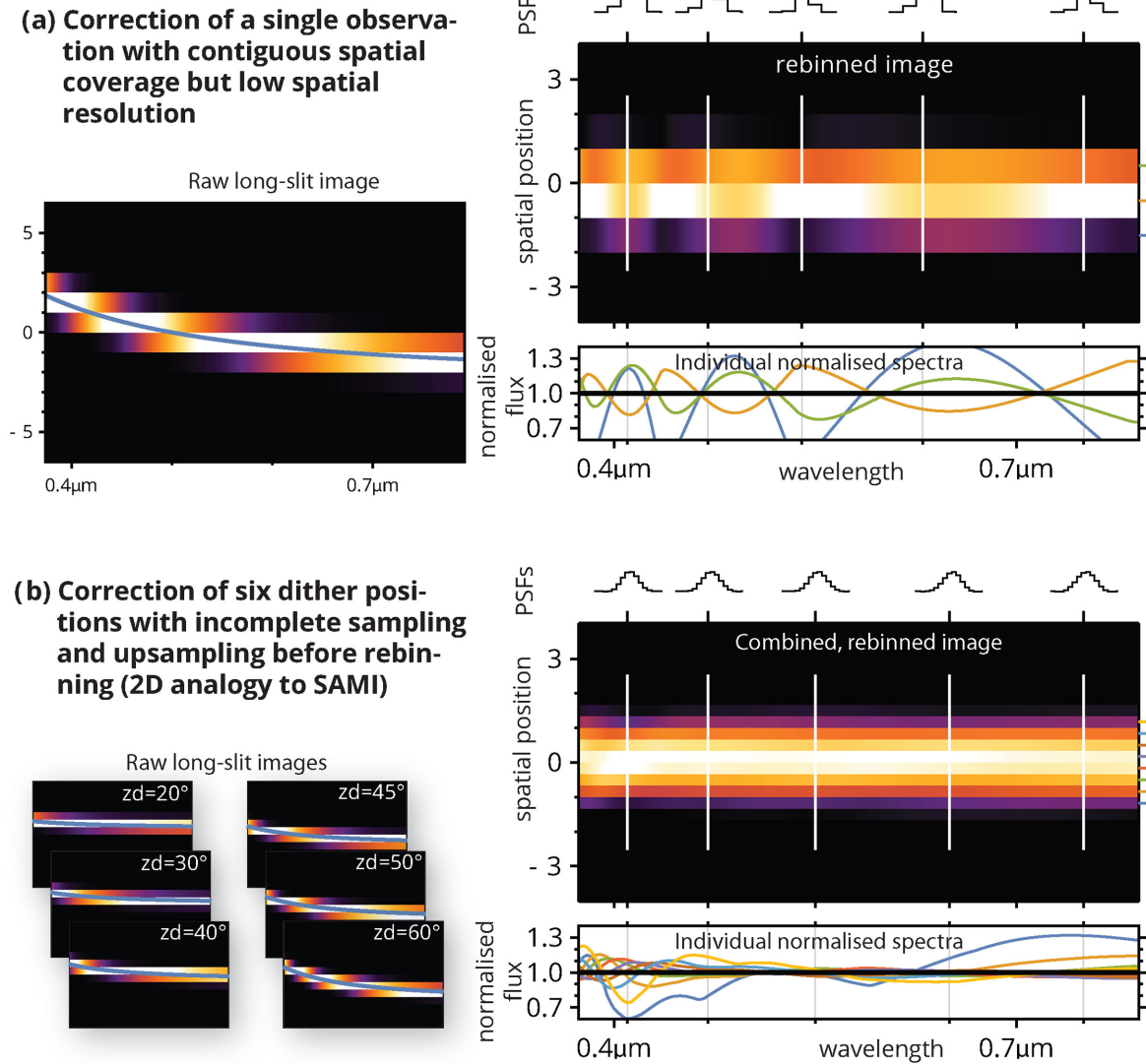
Figure A1. Long-slit observations showing the effects of DAR and results of correcting for it in (a) a simple, single long-slit observation at high airmass, and (b) the combination of several long-slit observations at different airmasses and dither positions – a close 2D analogy with our 3D data. The example shown in part (a) is observed at a zenith distance of 60 deg. The pixels are 1 arcsec, and the underlying Gaussian PSF has an FWHM of 1 arcsec. In part (b), input frames are taken at six airmasses ranging from zenith distance (zd) of 20 to 60 deg. The underlying PSF is the same, and the spatial pixels are also 1 arcsec, but the spatial sampling is incomplete. In the rebinning to correct for DAR, the data are upsampled to 0.33 arcsec pixels before combining. For each panel: the left-hand side shows the raw long-slit image with a line showing the DAR at the centre of the slit overplotted. The right-hand side shows the reconstructed long-slit image after DAR correction, including any rebinning. Where the pixels are brightest (white), the PSF centre falls at the centre of the pixel, while where the two vertically adjacent pixels have intermediate brightness (amber), the PSF centre falls between the two pixels. Vertical white lines mark the location of the spatial PSF shown above the image, with the difference from the mean PSF shown in red (scaled up to show detail). The plot below shows individual spectra from the image. The spatial location of each spectrum along the slit is shown by the corresponding coloured tick on the right of the image. These spectra have been normalized to highlight the relative differences in the spectra, which are entirely the result of the aliasing. *Note*: some PDF renders will attempt to smooth the pixels shown in this figure; we recommend using ACROBAT READER to see the actual, pixelated images as we intend.

sampling of the data. A correction for DAR is applied by shifting the pixels by the amount of the refraction along the spatial direction and rebinning to the original regular grid. After correction, the image of the object no longer shows a position shift with wavelength (shown on the right in Fig. A1a). However, aliasing of the rebinning and sampling are readily visible, causing the individual spectra at each spatial location (shown below the image) to vary within the PSF, and the PSF (shown above the image) to vary with wavelength.

Now, let us extend our example to be a close, 2D analogy with our own 3D spectral cubes. This extended example is shown graphically

in Fig. A1(b). First, we observe the source at several dither positions and airmasses. Secondly, we introduced gaps in the spatial coverage that are smaller than and within the 1 arcsec pixels (and therefore not readily apparent in the individual frames on the left). The dithering ensures that information falling in the gaps in one frame will be picked up in another frame. It also tends to smooth out the aliasing because individual dithers will each have a slightly different aliasing PSF, which will be averaged out in the combination. Finally, to bring our long-slit example closer to the actual process used in SAMI, we add another complication: upsampling. SAMI fibres are 1.6 arcsec,

but we sample the multiple observations on to a 0.5 arcsec output grid. Note that, in combining these six individual frames, it is also necessary to track the weights of the individual output pixels, which account for the gaps in the input data. This extended example has all the same characteristics and similar sampling dimensions of our actual SAMI data, except that we are working with only one spatial dimension instead of two.

Reviewing the resulting combined, DAR-corrected long-slit image shows that, despite its seeming smoothness, the PSF exhibits subtle but important variations with wavelength and spatial position. This long-slit image is shown on the right of Fig. A1(b). The image is fairly smooth because upsampling and several dither positions and airmasses have averaged out some of the aliasing. Yet the subtle differences in the PSF at different wavelengths are still present. These differences are much more apparent in the plot of individual spectra, where the spectrum at each spatial position has been normalized to highlight the relative differences. The spatial location of each of these spectra is shown by the corresponding coloured tick on the right edge of the image. Spectra further from the centre of the PSF (and with lower total flux) tend to have larger relative deviations from the actual spectral shape (this trend matches our analysis of observations of individual stars with SAMI).

Pixelated (discretely sampled) data observed with DAR present show effects of aliasing. These effects are exacerbated by poor spatial resolution and incomplete sampling. Combining observations with many dithers and different airmasses helps to average the aliasing out. Upsampling combined with sub-pixel dithering of the observations can also reduce the severity of the aliasing. Aliasing is not typically seen in long-slit data because the PSF is typically well sampled. However, the tension in IFS between spatial sampling and sensitivity, and the incomplete sampling present in many designs has led to noticeable aliasing in IFS data. Although we have only demonstrated the effect in 2D, long-slit data, DAR is only a 2D effect, so our treatment of aliasing readily extends to 3D IFS data.

The general impact of aliasing is that the PSF varies both with spatial and spectral position within either (2D) long-slit images or (3D) spectral cubes. This effect is subtle, and in many cases can be safely ignored without affecting results. There are, however, two important exceptions. The first exception is the cases where the PSF must be known to very high accuracy. The second is when comparing data that are widely separated in wavelength, for example emission-line ratios or spatially resolved colours. Any analysis that averages over scales larger than the PSF will not be affected by aliasing, such as measures of radial gradients in galaxies and analysis that requires spatial binning to bring out faint signals.

[1]*Australian Astronomical Observatory, 105 Delhi Rd, North Ryde, NSW 2113, Australia*

[2]*Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW 2006, Australia*

[3]*ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO)*

[4]*International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia*

[5]*Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia*

[6]*Cahill Center for Astronomy and Astrophysics, California Institute of Technology, MS 249-17, Pasadena, CA 91125, USA*

[7]*Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*

[8]*School of Mathematics and Physics, University of Queensland, QLD 4072, Australia*

[9]*Department of Physics and Astronomy, Macquarie University, NSW 2109, Australia*

[10]*School of Physics, University of New South Wales, NSW 2052, Australia*

[11]*Department Physics and Astronomy, University of North Carolina, Chapel Hill, NC 27599, USA*

[12]*English Language and Foundation Studies Centre, University of Newcastle, Callaghan NSW 2308, Australia*

[13]*Atlassian, 341 George St Sydney, NSW 2000, Australia*

[14]*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany*

[15]*Swinburne University, Hawthorn, VIC 3122, Australia*

[16]*SOFIA Operations Center, USRA, NASA Armstrong Flight Research Center, 2825 East Avenue P, Palmdale, CA 93550, USA*

[17]*SUPA School of Physics & Astronomy, University of St Andrews, St. Andrews KY16 9SS, UK*

[18]*School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia*

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