ESO Science Archive Facility - Phase 3 Data Release Description
ESO programme The UVES Spectral Quasar Absorption Database, Data Release DR1
provided by Michael Murphy, published on 2019-03-21

# *The UVES Spectral Quasar Absorption Database (SQUAD) DR1*

## Abstract

The UVES Spectral Quasar Absorption Database (SQUAD) Data Release 1 (Murphy et al., 2019, hereafter M19) comprises 467 fully reduced, continuum-fitted high-resolution UVES spectra of quasars with redshifts $z$ = 0–5, with the primary goal of enabling statistical studies of large quasar and absorber samples at high resolution. This document focuses on the selection and basic properties of the quasar sample, and the properties of the 467 spectra. The reader is referred to M19 for the quasar catalogue, scientific applications of the spectra and a catalogue of the damped Lyman-alpha systems (DLAs) in the DR1 spectra. Note that much of the content of this document is repeated directly from M19 with only minor modifications.

## Overview of Observations

This first data release comprises 467 fully reduced, continuum-fitted high-resolution quasar spectra from the Ultraviolet and Visual Echelle Spectrograph (UVES) on the ESO-VLT. It is defined as containing the 475 quasars in the ESO UVES archive whose first exposure (longer than 100 sec) was observed before 30th June 2008. All exposures of these quasars (longer than 100 sec) observed before 17th November 2016 were included in the final, combined spectra in DR1. In total, 3088 exposures were selected and successfully processed, with a total exposure time of $10.09 \times 10^6$ sec (2803 hrs, an average of 5.9 hrs per quasar). The quasars provide continuum-to-noise ratios of 4–342 (median 20) per 2.5 km/s pixel at 5500 Å. All the observed quasars contained in DR1 are catalogued in Table 1 of M19, where we list quasar names, coordinates, redshifts and optical/infra-red photometry sourced from several databases. The table includes properties of individual exposures contributing to each combined spectrum such as the number of UVES exposures, their total duration, the ESO Program IDs, the UVES wavelength settings, slit widths and on-chip binnings used, and the prevailing seeing values reported in the ESO Science Archive (when available, the minimum, median and maximum seeing are reported, and “NA” is reported when seeing information is not available). Finally, the table summarises some important properties of the final spectra.

## Release Content

### Fully Reduced, continuum-fitted quasar spectra

In DR1 we provide 467 UVES SQUAD spectra (final spectra could not be produced for 8 of the 475 quasars selected for DR1). The spectra are fully reproducible from the

raw, archival UVES exposures with open-source software, including our UVES_popler tool for combining multiple extracted echelle exposures which we document in M19. All processing steps described in the Release Notes below are completely transparent and can be improved upon or modified for specific applications using the ancillary data we provide as part of DR1.

## Release Notes

Here we provide a brief overview of the main steps involved in producing the DR1 spectra, from the science and calibration exposure selection, to the data reduction and the post-reduction procedures (e.g. combination, continuum-fitting, etc). The reader should refer to M19 for a complete description of the data calibration and the reduction process.

### Quasar selection

The quasar candidates, satisfying the date criteria outlined in Overview of Observations, were selected by cross-matching the coordinates of all "science" observations in the ESO UVES archive (i.e. with DPR.CATG set to "SCIENCE") with the MILLIQUAS quasar catalogue (Flesch 2015, updated to version 5.2[1]). While this catalogue aims to include all quasars from the literature (up to August 2017), it will not include unpublished quasars. To identify such cases, we checked the ESO proposal titles and observed object names (as labelled by the observers) for all programs that observed any MILLIQUAS quasar with UVES and searched for any objects observed in those programs that may be quasars (and not already reported in MILLIQUAS). This approach identified 9 of the final 475 quasars selected for DR1, and a further 18 objects that, upon data reduction and exposure combination, were clearly not quasars (17 stars and one galaxy). While it is possible that some quasars were not selected by our approach, our manual checking of the proposal titles and object names should ensure this number is very small or zero.

All quasar candidates were identified in the SuperCosmos Sky Survey database (Hambly et al. 2001) to determine a complete set of J2000 coordinates. In cases where a spectrum is available from the Sloan Digital Sky Survey (SDSS DR14; Abolfathi et al. 2018; Pâris et al. 2018), the SDSS coordinates were used in preference. These coordinates were used to name all quasars in DR1 (unique "DR1 Name" field for each quasar in Table 1 of M19). Quasar emission redshifts were taken from a hierarchy of cross-matched databases: SDSS, NED, SIMBAD and, if the quasar appeared in none of these databases, our own approximate measurement from our final spectrum. The latter was required in 13 cases, but in one of these (J031257−563912) no emission line could be identified from which a redshift could be estimated (its redshift is thus set to zero). This provides a nominal, adopted redshift for each quasar, named "$z_{\mathrm{em,Adopt}}$" in Table

[1] See http://quasars.org/milliquas.htm

1 of M19. The sky position and redshift distributions of the quasars are plotted in Figure 1.

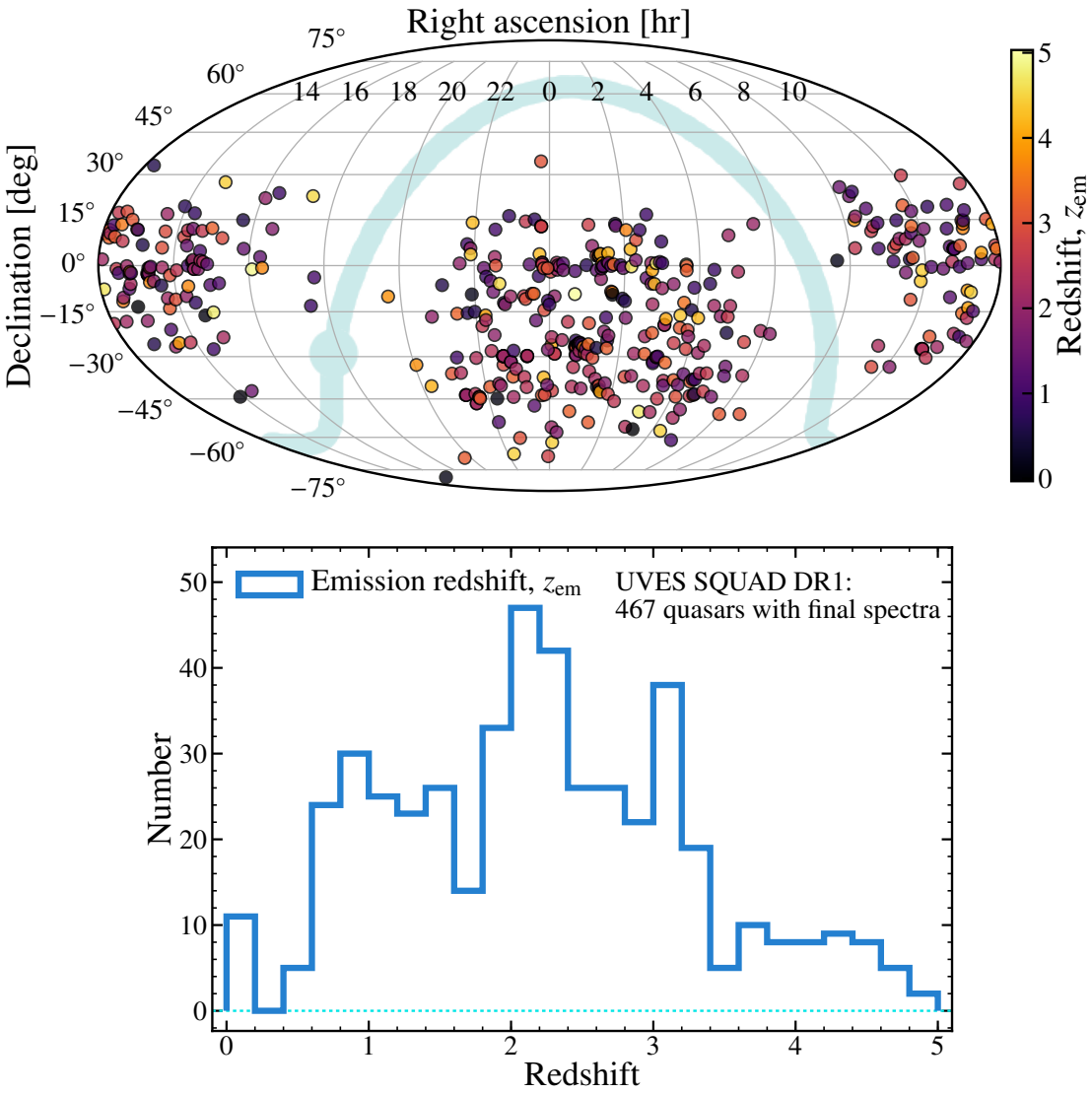


***Figure 1**: Top panel: Sky distribution of the 467 DR1 quasars for which final spectra were produced. The colour/shading of the points encodes the quasar redshift as indicated. The light blue shaded stripe and bulge represent the Galactic plane and centre. Bottom panel: Distribution of emission redshifts ("$z_{em,Adopt}$") for the DR1 quasars with final spectra.*

We note that, for DR1, we only consider UVES scientific exposures taken through an entrance slit; UVES has an image slicer option but we exclude such data for the moment. The slit width and on-chip binning determine the nominal resolving power, i.e. that expected for a fully illuminated slit, as is the case for thorium–argon (ThAr) exposures. However, the quasar exposure's resolving power will be somewhat larger than this, especially if the seeing full-width-at-half-maximum (FWHM) is significantly less

than the slit width. Table 1 in M19 therefore provides the range of slit widths, binnings and seeing during the observations as a guide. We do not include quasar observations made through UVES's iodine absorption cell; these require additional calibration exposures and cannot be combined with non-absorption cell observations of the same quasars. Finally, we exclude exposures taken with the Fibre Large Array Multi Element Spectrograph (FLAMES) mode of UVES.

## Calibration selection

A range of calibration exposures are required to reduce each quasar exposure. To ensure that the best-matching calibrations were selected within a specified "calibration period" before and after each quasar exposure, we used a custom-written code, UVES_headsort (Murphy 2016a). This generally selects the calibration exposure(s) closest in time to the corresponding quasar exposure for five different calibration types:

• **Wavelength calibration:** A single ThAr exposure with the same spectrograph settings (i.e. wavelength setting, on-chip binning and slit-width), was generally selected. Given the UVES operations model, in most cases the ThAr exposure was taken at least several hours after the quasar exposure. Indeed, the median time difference for all 3088 processed DR1 exposures is 5.4 h. However, preference was given to "attached" ThAr exposures, i.e. those taken immediately after quasar exposures without any grating angle changes. An attached ThAr exposure was identified as having the same grating encoder value as the corresponding quasar exposure. In a very small number of cases, particularly for exposures taken before 2001, a slightly different slit width was allowed for the matched ThAr exposure compared to the quasar exposure.

• **Order format and definition:** ThAr and quartz lamp exposures taken through a short slit are used to identify the echelle orders and define a baseline trace across the CCD. A single exposure of each type with the same spectrograph settings (except for the much shorter slit), was selected in all cases.

• **Flat field:** Five quartz lamp exposures with the same spectrograph settings were selected. In a small number of cases, especially for early UVES data (before 2003), some quasar exposures only had 3 or 4 matching flat field exposures; rarely, only a single flat field exposure could be found for quasar exposures taken before 2002.

• **Bias:** The five bias (zero-duration) exposures taken on the same CCD as the quasar exposure were selected in all but rare cases from early UVES operations (before 2002).

## Data Reduction

### Reduction of individual exposures

After determining the best set of calibration exposures for a given quasar exposure, UVES_headsort outputs a reduction script for use with ESO's Common Pipeline Library

(CPL, version 4.7.8[2]) of UVES data reduction routines, specifically via the ESO Recipe Execution Tool (esorex) command-line interface. This provides a highly streamlined data reduction pipeline – typically, a quasar exposure can be matched with calibrations and fully reduced within several minutes – while allowing low-level access to the data reduction parameters for improving the reduction if required. Most of the reduction steps are standard for UVES data and are explained in detail in the UVES CPL pipeline manual[3]. Briefly, these standard steps are:

I. ThAr lines are identified on the format definition frame and used to constrain a physical model of the UVES echellogram. This identifies the diffraction order numbers and spectral setup of the exposure which assists the order definition [step (ii)] and enables the automatic wavelength calibration in step (iv) below.
II. The order definition exposure is used to establish a baseline trace for object light along each echelle order. This acts as an initial guide for extracting the quasar flux.
III. The bias and flat-field exposures are combined to form masters which are used to correct the quasar exposure for bias and dark-current offsets and pixel-to-pixel sensitivity variations in the subsequent steps.
IV. The ThAr flux is extracted along the default trace in the wavelength calibration exposure (corrected for the blaze function using the master flat) and the ThAr lines are automatically, iteratively matched with those in the list carefully selected for UVES in Murphy et al. (2007). This allows a polynomial (air) wavelength solution to be established for the entire CCD (i.e. air wavelength versus pixel position for each echelle order).
V. The quasar flux is optimally extracted, with weights determined by averaging the quasar flux along small spectral sections (normally 32 pixels) and either fitting a Gaussian function to this average profile or using it directly, depending on its signal-to-noise ratio (S/N). The sky flux is extracted simultaneously in this process and is subtracted from the quasar flux in each extracted spectral pixel. The 1-sigma flux uncertainty is also determined from the quasar flux, sky flux and CCD noise characteristics. The flux and uncertainty spectra are corrected for the blaze function using the master flat.

Step (iv) was repeated for the DR1 spectra to improve their wavelength calibration. The optimal extraction weights from step (v) were used to re-extract the ThAr spectra and perform a refined wavelength calibration process. This ensures that the same pixels, with the same statistical weights, are being used to establish the wavelength scale for the quasar spectrum (e.g. it naturally negates the effects of spatially tilted ThAr lines on the CCD). UVES_headsort's reduction scripts also modify CPL's defaults for the wavelength polynomial degree, the number of ThAr lines to search for and select before performing the iterative polynomial fitting, and the tolerance allowed between the fitted and expected wavelength of ThAr lines. Typically, these new defaults simultaneously increase the number of lines used in the wavelength calibration, reduce the residuals around the final wavelength solution, and marginally improve the accuracy of the solution (due to increasing the polynomial degree). In some cases, particularly

[2] http://www.eso.org/observing/dfo/quality/UVES/pipeline/pipe_reduc.html

[3] https://www.eso.org/sci/software/pipelines/uves/uves-pipe-recipes.html

with the very blue wavelength settings (e.g. the standard 346 and 390-nm settings), these new defaults were modified manually to achieve a more robust wavelength solution (i.e. to increase the number of ThAr lines used).

After step (v), the CPL pipeline redisperses the flux and uncertainty arrays onto a linear wavelength grid (i.e. all pixels have the same size in wavelength), merges the spectra from adjacent spectral orders, and corrects the spectral shape using an estimate of the instrument response curve. However, because the resolving power remains reasonably constant across the wavelength range of grating cross-dispersed echelle spectrographs, and UVES covers more than a factor of three in wavelength range (~3050–10500 Å), a constant dispersion in wavelength is inappropriate; it inevitably oversamples the resolution element in the bluest parts of the spectrum and/or undersamples it in the reddest parts. Also, merging adjacent orders should account for small instrument response and/or blaze correction imperfections and variations by scaling their relative flux before averaging, but the accuracy of this is severely limited in a single exposure due to lack of S/N. However, almost all quasars in DR1 were observed in multiple exposures, so there is an opportunity to improve the merging of adjacent orders by considering all exposures together. And, finally, if the spectra from multiple exposures are to be combined, they will have to be redispersed, again, onto a common wavelength grid after correction for heliocentric motions. For these reasons, we use the original (not redispersed), extracted flux and uncertainty arrays of each order (not flux calibrated), from every exposure, to produce each quasar's final spectrum. This was performed using the custom-written code, "UVES_popler" (Murphy 2016b), specifically designed for combining the UVES data reduction pipeline products to produce a final, continuum-fitted spectrum. Below we summarise the overall approach of UVES_POPLER and how it was applied to create the DR1 spectra.

Note that, in the Phase 3 data release of DR1, each combined quasar spectrum is associated with an ancillary ".tar.gz" file that contains all the products from the above reduction steps. These are the files required for reproducing the combined spectrum using the UVES_popler tool, as described below. We describe the ancillary data further in the Data Format section.

### Combining reduced exposures with UVES_popler

UVES_POPLER reads the extracted flux and uncertainty arrays for each echelle order of each quasar exposure and the wavelength calibration polynomials derived from their corresponding ThAr exposures. Operation then proceeds in two phases: the automatic and manual phases. It is important to note that both phases are entirely reproducible and transparent: all parameters of the automatic phase, and relevant details of all manual "actions" subsequently performed in the manual phase, are recorded in a UVES_popler log (UPL) file; any user can understand how a spectrum has been formed and modified, and re-run and modify the entire process themselves. UPL files for all DR1 quasars are provided in this DR1 as part of the ancillary data files for each combined spectrum (see Data Format section below).

The automatic phase attempts to combine the spectra from all orders in all exposures and perform a basic continuum fit. Its main steps are:

I. **Data validation:** Reject pixels whose uncertainty indicates problems in the extraction (e.g. negative or extremely small uncertainties). This normally occurs near the order edges for UVES CPL-reduced spectra.

II. **Residual cosmic ray rejection:** Reject pixels, and their immediate neighbours, whose flux is much larger than the mean flux for their neighbouring 34 pixels. This rejects "cosmic rays" and/or bad pixels not already rejected in the optimal extraction step of UVES CPL-reduced spectra.

III. **Vacuum and heliocentric corrections:** The wavelength scales for the individual exposures are converted from air wavelengths to vacuum, and their correction for heliocentric motion is calculated and applied.

IV. **Redispersion:** A common log-linear, vacuum–heliocentric wavelength scale is established, with a constant velocity dispersion specified by the user, that covers the remaining pixels in the contributing exposures. The flux and uncertainty spectra from all exposures are linearly redispersed onto this common grid.

V. **Order scaling and combination:** The spectra in all echelle orders are combined in an interactive process starting from the highest S/N order. It is combined with the next highest "rank" order: that with the highest combination of S/N and wavelength overlap. The next highest rank order is combined with the previous two, and so on until all orders are combined. The flux (and uncertainty) in each order is optimally scaled to match the combined spectrum from the previous iteration. For each spectral pixel, the combined flux is the weighted mean of the fluxes from the contributing spectra, which is determined through an iterative clipping process to remove discrepant values.

VI. **Continuum fitting:** Each contiguous section of the combined spectrum is broken into "chunks", typically 20000 km/s wide below the quasar Ly-α emission line and 2500 km/s above it, which overlap half of the adjacent chunks. An iterative polynomial fit is performed to each chunk: at each iteration, pixels with flux significantly below (typically >1.4σ) or above (typically >3.0σ) the current fit are rejected for the next iteration. To form a smooth, final continuum, the final fits from adjacent chunks are averaged with a weight that decreases linearly from unity at the chunk's centre to zero at its edge.

The automatic phase of UVES_popler generally produces excellent "quick-look" spectra that are entirely adequate for many scientific goals, particularly those focussing on individual absorption systems whose transitions collectively occupy only a small fraction of the pixels. However, individual UVES exposures nearly always contain some artefacts that inhibit larger, statistical studies (and are often a nuisance to others as well) because, for example, they can mimic real absorption features in blind searches. The automatic continuum fits redwards of the Ly-α forest are generally very reliable, except in the vicinity of absorption features wider than the chunk size or across very narrow quasar emission lines. However, the automatic continuum in the Ly-α forest is not generally useful; reliable automatic continuum placement is a notorious problem in quasar spectroscopy that limits the speed with which high-resolution spectra can

be analysed. Unfortunately, we have not solved that problem here. For these reasons, a manual phase of operation is required.

The manual phase of UVES_popler allows interactive "actions" to be performed on the contributing echelle orders or combined spectrum to improve the quality of the latter and its continuum fit. These actions include:

- Clip (and unclip) pixels from contributing orders or the combined spectrum.
- Manually fit or draw (spline) a new continuum to part of the combined spectrum.
- Automatically fit the continuum for the entire spectrum again.
- Manually fit or draw (spline) a continuum to (part of) a contributing order to re-shape its flux (and uncertainty) array to that of the combined continuum.
- Scale an order's flux and uncertainty array by a constant factor.
- Rerun the automatic order scaling algorithm starting from the highest rank order not manually scaled by the user.

In general, a user will select portions of a spectrum to manually improve using the above actions based on their specific scientific goals. For example, for studying the intergalactic medium, it will be important to remove artefacts and re-fit the continuum in the Ly-α forest region.

In Section 4.2 of M19, we described our approach to improving the DR1 quasar spectra for use towards as many different scientific goals as possible, particularly large statistical studies of DLAs, the intergalactic medium and metal absorption systems. As part of the ancillary files associated with each Phase 3 combined spectrum, we also provide the complete record of parameters used for the automatic phase, and all subsequent manual actions for all DR1 quasars, as UPL files. A key aspect of this approach is that, with the specific treatment of each quasar made completely transparent and reproducible, any user may further improve the DR1 spectrum by using UVES_popler to add manual actions to the UPL file.

It is important to note that we produce and release here a single final spectrum for each quasar. That is, we combine all available CPL-reduced exposures of a quasar regardless of variations in slit width and CCD binning. While a large range of slit widths are available for UVES (0.3–10 arcsec), in practice the range used for a specific quasar is very narrow, presumably because achieving a threshold S/N is most often the immediate observational goal and, for faint (i.e. most) quasars, this is severely affected by the choice of slit width (UVES is a natural-seeing instrument). For example, 385 of the 467 final spectra comprise exposures with a single slit width[4], while only 16 combine exposures with slit widths differing by more than 0.3 arcsec. Nevertheless, different slit widths and CCD binnings will produce individual exposures with different effective resolving powers, thereby affecting the resolution of the final spectrum[5]. If separate combination of exposures of different resolutions is required, UVES_popler can easily be used to construct such "sub-spectra" of a given quasar using its UPL file,

[4] The two arms of UVES can have different slit widths for a dual-arm observation. However, for this example we have ignored cases where the different arms used consistent but different slit widths. That is, the number of spectra for which a single slit width was used at any given wavelength will be somewhat larger than 385.

[5] A nominal, mean resolving power is calculated for each final spectrum in M19.

as discussed in M19. For example, this technique has been employed in the analysis of J051707-441055 by Kotuš et al. (2017).

To make each DR1 quasar spectrum useful for as many scientific goals as possible, our approach was to "clean" it to at least a minimum standard in the manual phase of UVES_popler. Clearly, this cleaning process is the most time-consuming stage, and all authors of M19 contributed to it, so ensuring a strictly uniform standard for all DR1 quasars was not practical. Nevertheless, some more general cleaning steps were taken for each quasar in DR1 with a view to making the final spectrum as useful as possible. These are briefly outlined in the follow.

I. **Artefact and bad data removal:** Very obvious artefacts that are similar, though not identical, in different spectra have been removed manually. A prominent and common example occurs in the bluest 4–5 orders of the red arm spectra due to several bad pixel rows in the corresponding CCD; see Fig. 3 of M19. Another example is "bends": echelle order spectra that have different shapes where they overlap. This can occur for several reasons, e.g. time evolution in the flat-field lamp spectral shape, or poor extractions of the quasar flux near order edges, perhaps due to poorly constrained object traces. When very severe, these affected the final flux spectrum, and so were corrected. Bends in contributing orders were corrected either by removing the bent section or by fitting a continuum to the order (or part thereof) and re-normalising it to match the combined spectrum's continuum shape.

II. **Order rescaling and combination:** In spectral regions with very low S/N, or in echelle orders affected by severe artefacts, the relative scaling between an echelle order's spectrum and the combined spectrum can be very poorly or spuriously determined. This occurs frequently in the bluest orders of the 346 and 390-nm settings. It also occurs if the broad trough of a DLA straddles two echelle orders, and below the Lyman limit in the rest frame of DLAs and Lyman limit systems. In these latter examples, there is simply no flux to allow a relative scaling between adjacent orders; this is certainly a disadvantage of the order-scaling algorithm in UVES_popler. To address this in the DR1 spectra, we manually adjusted the scaling of the highest-ranked order with an obvious scaling problem and re-ran the automatic scaling algorithm starting at that order. This process was repeated for lower-ranked orders to achieve a final spectrum that, visually, appears properly scaled. For the extreme blue orders, where the S/N degrades significantly, the best manual scaling factor to choose is often quite unclear, so there may be significant scaling differences between orders in regions of final spectra with S/N ≤ 5 per pixel.

III. **Continuum fitting:** As already discussed, the continuum fit in UVES_popler's automatic phase is generally not useful in the Ly-alpha forest, near wide absorption features or over narrow emission lines. Thus, for the wide absorption and narrow emission features, we manually fit a new continuum only around the problematic region. See Section 4.2.3 of M19 for details.

## Data Quality

All the authors, and several others, contributed to the manual cleaning and continuum fitting steps outlined in the previous section. Of course, this may lead to varying quality and homogeneity among the final spectra. To reduce this, one author of M19 (MTM) reviewed all DR1 spectra and modified or added manual actions to improve and homogenise them, where necessary. While the purpose of the general cleaning steps above is to ensure a minimum quality and usefulness for all DR1 spectra, some spectra – or, most often, certain aspects of some spectra – have received much more extensive attention, including manual changes to the spectrum not described above. These are generally spectra that have already been published elsewhere. One example is the very detailed study of J051707-441055 to constrain cosmological variations in the fine-structure constant by Kotuš et al. (2017). Beyond the basic cleaning steps outlined above, this study focussed on correcting the individual exposures for known, long-range distortions of the UVES wavelength scale (e.g. Rahmani et al. 2013; Whitmore & Murphy 2015) and velocity shifts between exposures caused by varying alignment of the quasar within the UVES slits. Such improvements are included in the DR1 versions of the spectra when available.

## Known issues

Section 5.2 of M19 describes the details of several remaining artefacts, systematic effects and limitations of the DR1 spectra. These are briefly summarised as follows:

I. **Continuum errors:** The automatic continuum fits test to overestimate the continuum level in DR1 spectra, even in regions of unabsorbed continuum, typically by factors of 0.05–0.25 times the mean, normalized flux uncertainty. In addition to this bias, the manual fitting of continuum sections in the Ly-α forest will tend to underestimate the continuum level at redshifts above ~3. This may be as little as ~3% at $z$~3 but may increase substantially at higher redshifts, with up to ~17% underestimation expected at $z$~4–5. Finally, the incidence of Lyman-limit systems (where there is significant remaining flux bluewards of the limit) and broad absorption line (BAL) features makes the definition of the continuum dependent on the scientific question being addressed. For example, what constitutes the continuum is very different when studying a Lyman limit or the Ly-α forest bluewards of it. Generally, in such cases, we attempted to fit or interpolate a continuum that would be useful for most users. However, we urge those studying Lyman-limit systems and BALs in DR1 spectra to redefine the continuum placement accordingly.

II. **Telluric features:** No attempt was made to remove telluric absorption features from the DR1 spectra, so the usual caution should be used in studying affected wavelength regions (see M19 for details).

III. **Cosmic rays and bad pixels:** Many of these remain in the individual, extracted exposures that contribute to a final DR1 combined spectrum. The combination process removes many of these, especially when 5 or more exposures contribute to the final spectrum. However, M19 describes how this is not the case in regions of low local relative flux; users may wish to remove such features manually according to their scientific goals.

IV. **Unidentified absorption artefacts:** M19 describes how several, relatively weak (typically <10% of the continuum) absorption-like features were noticed during the cleaning of many DR1 spectra and how these tended to occur at similar wavelengths in different spectra. M19 provides a composite spectrum of DR1 spectra that serves to highlight these features and makes detecting weaker ones possible. Users may wish to utilise this composite spectrum to identify or mask spurious spectral features that may affect their absorption line analyses.

V. **Underestimated uncertainties at low flux levels:** M19 demonstrated that the flux uncertainty spectra produced by the CPL pipeline are underestimated at low flux levels, seemingly due to the sky noise contribution being underestimated (see Section 5.2.5 of M19 for details). The factor by which the uncertainty is underestimated is larger for lower S/N exposures, but for high S/N exposures it remains evident in the cores of saturated absorption lines.

VI. **Bad data in individual exposures:** Artefacts of various kinds may still remain in individual exposures, even though they do not appear, by eye, to influence the combined spectrum. These may indeed influence the combined spectrum at a lower level, so users are encouraged to use UVES_popler and the ancillary data provided with each DR1 spectrum released in ESO's Phase 3 to visualize the individual spectra for detailed analyses.

VII. **Blaze function variations and remnants:** Some "bends" in the flux spectra between overlapping wavelength regions of different exposures (or even adjacent echelle orders in the same exposure) remain in the DR1 combined spectra. These have been removed when they had an obvious influence on clear absorption lines, but they may still have a weaker influence when they were less than "obvious" and not removed.

VIII. **Zero level errors:** The zero level in some final DR1 spectra can be significantly different to zero by ~2–4% of the continuum level. This is caused by inaccurate zero level (i.e. sky flux and extraction) estimation in the CPL pipeline and typically affects low S/N spectra most severely.

*IX.* **Wavelength scale shifts and distortions:** The wavelength calibration accuracy of UVES has been the specific focus of many quasar absorption studies. The wavelength scale is set by comparison with a ThAr lamp exposure, and several effects shift and/or distort the true quasar wavelength scale with respect to this: (i) Mechanical drifts and changes in the refractive index of air; (ii) Differences in alignment of the quasar in a slit between exposures and/or between the two different slits of UVES (blue and red arms); (iii) Intra-echelle order wavelength scale distortions; (iv) Long-range wavelength scale distortions. These effects are described in detail in Section 5.2.9 of M19. Most DR1 spectra are not corrected for these effects as they are generally fairly small – typically <20% of a (unbinned) pixel. See M19 for exceptions to this.

# Data Format

## Files Types

The ESO Phase 3 version of each UVES SQUAD DR1 quasar combined spectrum is provided as a FITS file named according to the adopted J2000 coordinates of the quasar in M19 as follows:
`JHHMMSS±DDMMSS.fits`,
where HHMMSS is the hour-minutes-seconds representation of the right-ascension and DDMMSS is the degrees-arc minutes-arc seconds representation of the declination.

The first (and only) extension header data unit (HDU) of the Phase 3 FITS file contains the following data arrays in binary table format:

1. "WAVE": Vacuum–heliocentric wavelength in units of Å.
2. "FLUX": Continuum-normalised flux (no units).
3. "ERR": Continuum-normalised 1-sigma flux uncertainty (no units).
4. "CONTINUUM": The continuum level in arbitrary units.
5. "STATUS": An integer indicating a valid pixel (1) or an error code (negative integers).
6. "NPBCLIP": The number of pixels from exposures that were available for contributing to the combined spectrum pixel, i.e. before the sigma-clipping procedure.
7. "NPACLIP": The number of pixels from the exposures that actually contributed to the combined spectrum pixel, i.e. after the sigma-clipping procedure.
8. "CHBCLIP": The $\chi^2$ of pixel fluxes, around their weighted mean, from exposures that were available for contributing to the combined spectrum pixel.
9. "CHACLIP": The $\chi^2$ of pixel fluxes, around their weighted mean, from exposures that actually contributed to the combined spectrum pixel.

## Ancillary data

The primary header of a ESO Phase 3 FITS file specifies two ancillary data files associated with each quasar using the header keywords "ASSON1" and "ASSON2". These files have the following naming conventions and content:

"ASSON1": This is the UVES SQUAD DR1 final spectrum as released with M19. This is named according to the J2000 coordinates: `JHHMMSS±DDMMSS_Final_Spectrum.tar.gz`. This ".tar.gz" file contains the DR1 final spectrum and the UPL file used to create it from the CPL pipeline products. This file is available via ESO Phase 3 or directly from the UVES SQUAD DR1 web portal: https://github.com/MTMurphy77/UVES_SQUAD_DR1 (Murphy et al. 2018). It contains the same data arrays as the ESO Phase 3 version described above, but with an additional uncertainty array that should better approximate the expected RMS in the flux, given that the redispersion and exposure combination process will introduce correlations between pixel fluxes (and uncertainties) in the combined spectrum. This additional uncertainty array is understood (and used by default) by VPFIT, the widely-used software for fitting absorption profiles in quasar spectra (Carswell & Webb 2014). For a full description of the

data format of this ancillary file, see the "Notes_FITS_Files.txt" file in UVES SQUAD DR1 web portal (Murphy et al. 2018).

"ASSON2": This file contains all the CPL data reduction products associated with a DR1 spectrum. This is named according to the J2000 coordinates: `JHHMMSS±DDMMSS_Reduction.tar.gz`. This file is available via ESO Phase 3 or directly from the UVES SQUAD DR1 web portal: https://github.com/MTMurphy77/UVES_SQUAD_DR1 (Murphy et al. 2018). The files within this "tar.gz" file are required to recreate the final DR1 spectrum of a quasar using the UPL file. Users can also use the UPL file as a starting point for modifying the final spectrum or improving various aspects of the spectrum. The UPL file is included here for convenience; it is the same file as included in the file specified by "ASSON1".

## Acknowledgements

This research has made use of the services of the ESO Science Archive Facility, and this DR is based on data products created from observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere.
Users of data from the UVES SQUAD DR1 should cite Murphy et al. (2019): *Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, MNRAS, 482, 3458.*

## References

Abolfathi B. et al., 2018, ApJS, 235, 42

Carswell R. F., Webb J. K., 2014, VPFIT: Voigt profile fitting program, Astrophysics Source Code Library (ascl:1408.015)

Flesch E. W., 2015, PASA, 32, e010

Hambly N. C. et al., 2001, MNRAS, 326, 1279

Kotuš S. M., Murphy M. T., Carswell R. F., 2017, MNRAS, 464, 3679

Murphy M. T., 2016a, UVES_headsort: VLT/UVES pipeline preparation, doi:10.5281/zenodo.44766

Murphy M. T., 2016b, UVES_popler: POst PipeLine Echelle Reduction software, doi:10.5281/zenodo.44765

Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2018, UVES SQUAD Data Release 1, doi:10.5281/zenodo.1345974, https://github.com/MTMurphy77/UVES_SQUAD_DR1

Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, MNRAS, 482, 3458 (M19)

Murphy M. T., Tzanavaris P., Webb J. K., Lovis C., 2007, MNRAS, 378, 221

Pâris I. et al., 2018, A&A, 613, A51

Rahmani H. et al., 2013, MNRAS, 435, 861

Whitmore J. B., Murphy M. T., 2015, MNRAS, 447, 446