The Third Data Release of the KODIAQ Survey

John M. O'Meara1, Nicolas Lehner2, J. Christopher Howk2, and J. Xavier Prochaska3

1 W.M. Keck Observatory 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA
2 Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA
3 UCO/Lick Observatory, Department of Astronomy & Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

Received 2020 September 16; revised 2020 October 12; accepted 2020 October 14; published 2020 December 23

Abstract

We present and make publicly available the third data release (DR3) of the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) survey. KODIAQ DR3 consists of a fully reduced sample of 727 quasars at 0.1 < \( z_{\text{em}} \) < 6.4 observed with the Echelle Spectrograph and Imager at moderate resolution (4000 \(< R < 10,000\)). DR3 contains 872 spectra available in flux calibrated form, representing a sum total exposure time of \( \sim 2.8 \) megaseconds. These coadded spectra arise from a total of 2753 individual exposures of quasars taken from the Keck Observatory Archive (KOA) in raw form and uniformly processed using a data reduction package made available through the XDIL distribution. DR3 is publicly available to the community, housed as a higher level science product at the KOA and in the igmspec database.

Unified Astronomy Thesaurus concepts: Lyman limit systems (981); Quasars (1319); Damped Lyman-alpha systems (349); Intergalactic medium (813); Circumgalactic medium (1879); Double quasars (406)

Supporting material: machine-readable tables

1. Introduction

Studies of absorption lines toward quasars have impacted a wide range of fields within astrophysics and cosmology. From studies of individual systems to determine the primordial D/H ratio (e.g., Cooke et al. 2018), to studies of tens of high-resolution, high signal-to-noise spectra to constrain the temperature–density relation in the intergalactic medium (IGM; Hiss et al. 2019), to a determination of the mean free path of ionizing radiation from a stack of thousands of lines of sight (Prochaska & Wolfe 2009), quasar lines of sight have become one of the most valuable tools available to understand the universe at cosmological distances.

Large quasar and galaxy surveys such as the Sloan Digital Sky Survey (SDSS; Ross et al. 2020) provide spectra in bulk in a ready-for-analysis state. By contrast, quasar spectra from larger ground-based facilities are often not available to the community in a science-ready form, if at all, and with little information on the provenance of the data. To help remedy the situation, a number of groups have provided the community with pipeline-reduced data from large aperture telescopes (e.g., Murphy 2018). As a part of these endeavors, a primary goal of the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ; Lehner et al. 2014) is to uniformly reduce and make publicly available spectra of quasars obtained with instruments on the Keck telescopes over the last two plus decades. In the first (O’Meara et al. 2015) and second (O’Meara et al. 2017) data releases (DR1 and DR2, respectively), KODIAQ has provided spectra from Keck’s high-resolution echelle spectrograph, HIRES. The DR2 sample encompasses DR1 and adds additional spectra, bringing the sample to a total of 300 unique quasar sightlines. In this work, we significantly increase the number of quasars with the third data release (DR3) from KODIAQ, this time adding 727 quasar spectra obtained with the Echellette Spectrograph and Imager (ESI; Sheinis et al. 2002) instrument on the Keck II telescope. Although lower resolution than HIRES, ESI spectra are of significant resolution to both serve statistical studies of the Ly\( \alpha \) forest as well as metal line absorption, along with studies of individual absorbers such as damped Ly\( \alpha \) (DLA) systems (e.g., Rafelski et al. 2012).

2. The Data

The ESI data in DR3 were obtained by individual PIs between 1999 December (shortly after instrument commissioning) and 2015 January. ESI is an echellette system, with 10 echelle orders spanning the range 3930 \( < \lambda < 10930 \) Å. Cosmic defects in the ESI CCD impact some pixels in the wavelength regions 4460 \( < \lambda < 4565 \) Å. Depending on seeing conditions, slit width, and desired wavelength, ESI has a resolution of 4000 \( < R < 10,000 \). Observations in DR3 are made using the full range of ESI slit widths, ranging between 0.3 and 6.0, with the majority of observations made with the 0.5, 0.75, and 1.0 slits. The observations are summarized in Table 1. The coordinates and quasar emission redshifts presented in Table 1 were obtained by passing the coordinates from the raw data file first through the Simbad databases and then, if not found, the NED database, selecting the quasar closest on the sky. If an object was not found within 30 arcseconds of a quasar in either database, the coordinates were taken directly from the data header. We caution that in the case of close quasar pairs, the coordinates, chosen to correspond to the nearest object, may be inaccurate and refer to the other pair member.

2.1. Data Reduction

Once the data were downloaded and reviewed for calibrations, they were uniformly reduced using the ESIRedux\(^6\) code. ESIRedux is a suite of IDL routines within the XDIL\(^7\) package of astronomical utilities. ESIRedux is functionally very similar to HIRedux, which is described in O’Meara et al. (2015), but
with a few important differences which we outline below. The workflow of ESIredux is as follows. First, the raw data is grouped into setups. Unlike for HIRES, ESI is essentially fixed in wavelength coverage and echelle order placement on the detector, so setups are governed only by slit width and detector binning. Next, for each setup, a bias level and gain is determined, and a flat-field image is constructed from a combination of internal lamps, dome flats, and if present,
Table 2

| ESI DR3 Name       | Alternate HIRES DR2 Name | $z_{\text{em}}$ |
|--------------------|--------------------------|----------------|
| J001602-001225     | J001602-001224           | 2.090          |
| J003501-091817     |                          | 2.423          |
| J004054-091526     |                          | 4.977          |
| J005814-011530     |                          | 2.508          |
| J010311-131617     |                          | 2.681          |
| J013340-040059     |                          | 4.154          |
| J013421-330756     |                          | 4.532          |
| J013515-021349     |                          | 1.820          |
| J015741-010629     |                          | 3.572          |
| J020944+051714     |                          | 4.174          |
| J022554+005451     |                          | 2.980          |
| J024401-013403     |                          | 4.053          |
| J024851+180249     |                          | 4.422          |
| J025518+004847     |                          | 4.015          |
| J073149+285448     |                          | 3.671          |
| J074521+473436     |                          | 3.221          |
| J074711+273903     |                          | 4.170          |
| J074749+443417     |                          | 4.432          |
| J081240+320808     |                          | 2.703          |
| J081435+502946     |                          | 3.885          |
| J081740+135134     |                          | 4.379          |
| J082107+310751     |                          | 2.613          |
| J082540+354414     |                          | 3.846          |
| J082619+314848     |                          | 3.098          |
| J083141+524517     |                          | 3.912          |
| J084424+124546     | J084424+124548           | 2.489          |
| J090033+421547     | J090033+421546           | 3.294          |
| J092708+582319     |                          | 1.915          |
| J092914+282529     |                          | 3.406          |
| J093337+284532     |                          | 3.421          |
| J095500-013006     |                          | 4.434          |
| J100841+362319     |                          | 3.129          |
| J110155-294141     |                          | 2.620          |
| J110216+300140     |                          | 3.128          |
| J110314+544040     |                          | 2.996          |
| J110512+354534     |                          | 4.912          |
| J110576+455553     |                          | 4.137          |
| J110045+112239     |                          | 4.707          |
| J113130+604420     |                          | 2.906          |
| J113508+222715     |                          | 2.880          |
| J115538+053050     |                          | 3.476          |
| J120207+323538     |                          | 5.292          |
| J121117+042222     |                          | 2.548          |
| J121134+090220     |                          | 3.292          |
| J124610+303313     |                          | 2.560          |
| J124924+023339     |                          | 2.121          |
| J130426+120245     |                          | 2.085          |
| J131215+423900     |                          | 2.573          |
| J134002+110630     |                          | 2.926          |
| J134544+262506     | J134544+262506           | 2.030          |
| J134916-033715     |                          | 3.992          |
| J135317+532825     |                          | 2.912          |
| J141906+592312     |                          | 2.329          |
| J142438+225600     |                          | 3.620          |
| J143912+295448     |                          | 3.004          |
| J144331+272436     |                          | 4.420          |
| J144445+291905     |                          | 2.638          |
| J155810-003120     |                          | 2.842          |
| J155814+405337     |                          | 2.638          |
| J16009+472444      |                          | 3.220          |
| J162548+266685     |                          | 2.531          |
| J164656+551444     | J164656+551446           | 4.037          |
| J173352+540030     |                          | 3.435          |
| J175603+574847     | J175603+574848           | 2.110          |
| J175746+755916     |                          | 3.050          |

(This table is available in machine-readable form.)

Figure 2. Redshift distribution of KODIAQ DR3 QSOs.

twilight sky flats. The flat-field images serve the traditional role of removing pixel-to-pixel variations, but also provide the location of the slit edges for each echelle order. Twilight flats are used primarily in the bluest orders, where the internal and dome flats have lower flux. A two-dimensional wavelength image for the detector is then created using a combination of CuAr, HgNe, and Xe lamps. The wavelength solutions are compared to archived databases, and the residuals to the fit are most commonly found to be subpixel. Cosmic-ray rejection is then performed using standard frame comparison techniques. As with HIRedux, sky subtraction and object definition and extraction is performed using the methods outlined in Bochanski et al. (2009) on an order-by-order basis, resulting in a spectrum of each order with counts as a function of wavelength. Finally, for objects with multiple exposures, the individual orders are weighted-mean combined.

A key difference between the ESI data in KODIAQ DR3 and the HIRES data from DR2 is that the data in DR3 have been flux calibrated instead of continuum normalized. For expediency, all exposures in DR3 were flux calibrated using a single archival sensitivity function derived from a median of
sensitivity functions from multiple flux standard stars. Examples of DR3 ESI quasi-stellar object (QSO) spectra are shown in Figure 1.

2.1.1. Issues of Note

A number of issues should be noted before using data from DR3 for precision analysis. First, the reddest echelle order, with a central wavelength of approximately $\lambda = 9800$ Å was often poorly extracted. In these cases, that order has been rejected, and the data trimmed. The rightmost column in Table 1 indicates the wavelength range of the spectra in DR3. Second, the detector defects at $\lambda \approx 4500$ Å are present in the reduced data, and caution should be used near those wavelengths. Finally, some spectra suffer discontinuities at the regions in wavelength where the echelle orders overlap. These discontinuities are the result of order-to-order differences in flux calibration. Users that intend an analysis of the intrinsic quasar spectra should first inspect the data for these occurrences. We note that all intermediate data reduction steps for the data in DR3 are provided in the Keck Observatory Archive (KOA), so that users may adjust specific steps, such as flux calibration, to match their needs. Finally, Table 1 shows that a number of quasars were observed on multiple dates, and often by multiple PIs. We did not attempt to combine the data across these observations, given frequent differences in slit width and potential inaccuracies in flux calibration.

https://koa.ipac.caltech.edu/Datasets/KODIAQ/; doi:10.26135/KOA1
2.2. KODIAQ at the KOA and Igmspec

As with DR2, the data products of DR3 are available for community download from the KOA. As DR3 stems from a different instrument, KOA hosts DR3 as a separate contributed data set from DR2. Nevertheless, a number of quasars have data in both DR3 and DR2. Table 2 lists those quasars. Note that in a few cases, the HIRES DR2 name does not exactly match the ESI DR3 name. Those cases are highlighted in Table 2. As with DR2, within DR3 at KOA, users can search for and download individual quasar data, or the full DR3 data set at once. DR3 will make available all intermediate data reduction products available, grouped by observing run. Spectra from DR3 will also be ingested into the igmspec database.

3. Properties of KODIAQ DR3

The KODIAQ DR3 consists of ESI spectra for 727 quasars. As in previous data releases, quasars are named according to their J2000 R.A./decl. coordinates as resolved by SIMBAD or NED, unless the quasar is not listed in these databases, in which case we use the position on the sky as provided by the raw data headers. In Table 1 we list the quasar redshift as provided by SIMBAD or NED. If the quasar is unlisted, a crude estimate is made for the redshift using the Lyα and other emission lines. The full DR3 is comprised of 2753 individual exposures, grouped into 872 spectral coadds. The aggregate exposure time of the DR3 is ~2.8 megaseconds.

3.1. General Properties

Figure 2 shows the quasar redshift distribution for the DR3. The median redshift for the sample is \( \bar{z} = 3.21 \) and the sample spans the range in redshift of \( 0.10 < z < 6.44 \). The distribution of DR3 quasars on the sky is shown in Figure 3.

In Figure 4, we show median signal-to-noise per pixel, derived from the pipeline optimal extraction values of the flux and \( 1\sigma \) errors, within \( \pm 5 \) Å for the 872 in DR3 at 3 rest wavelengths \( \lambda_r \) where \( \lambda_r = \lambda_{\text{obs}}/(1 + z_{\text{em}}) \) and \( \lambda_{\text{obs}} \) is the observed wavelength in the spectrum. The three values of \( \lambda_r \) are chosen as markers for typical wavelengths of interest for IGM/ circumgalactic medium (CGM) studies, namely \( \lambda_r = 1450 \) Å for studies of heavy element absorption free of Lyα forest contamination, \( \lambda_r = 1170 \) Å for studies of the forest itself, and \( \lambda_r = 1450 \) Å for determination of the HI column density \( (N_{\text{HI}}) \) in Lyman limit systems from the shape of the Lyman break.

3.2. Cosmological Properties

Given the large number of quasars in DR3, and with a significant fraction of the spectra having moderate to high signal-to-noise, the sample can lend itself well to a variety of statistical analysis in cosmology and elsewhere. Here, we highlight a number of statistical properties of DR3 with respect to its ability to be used for studies of neutral hydrogen and heavy element absorption.

In Figure 5 we show the rest wavelength coverage provided by the quasars in DR3 in the quasar rest frame. Vertical lines denote the rest-frame wavelengths of commonly studied lines. The full DR3 illustrates that the number of quasars that can be used for statistical studies of H I and heavy element absorption numbers well into the hundreds. Figure 6 shows this large sample in a different way, namely showing as a function of...
redshift how many quasars can be used for specific ions. We can further subdivide the DR3 sample by exploring how many quasars provide coverage for specific ions as a function of redshift and signal-to-noise. Figure 7 provides this summary for four ions: H I Lyα, C IV, O VI, and Mg II.

Finally, Figure 8 provides the redshift sensitivity function \(g(z)\) for the DR3 sample. The function \(g(z)\) is calculated in the same manner as in O’Meara et al. (2015, 2017). \(g(z)\) is presented in Figure 8 for four values of the signal-to-noise ratio, and for H I Lyα and C IV as representative ions. In all the above metrics, DR3 represents a significant sample for statistical studies of hydrogen and heavy element absorption, even at moderate to high signal-to-noise, which is a requirement for studies of weak absorption.

4. Summary and Future

We present here a summary of and make publicly available the 727 quasar sample in the KODIAQ DR3. As with KODIAQ DR1 and DR2, all spectra and intermediate data reduction products are made available at the KOA. Unlike DR1 and DR2, the data are roughly uniform in wavelength coverage (and nearly always larger per quasar in wavelength coverage than the HIRES spectra in previous releases), but like DR1 and DR2, they sample a large range of quasar redshift, and provide a wide range of signal-to-noise.

When using data products from DR3, we ask that publications provide the standard KOA acknowledgment, acknowledgment of the original PIs that obtained the data, and the following text: “some or all of the data presented in this work were obtained from the Keck Observatory Database of Ionized Absorbers toward QSOs (KODIAQ), which was funded through NASA ADAP grants NNX10AE84G and NNX16AF52G,” along with a citation to this paper.

Support for this research was made by NASA through the Astrophysics Data Analysis Program (ADAP) grants NNX10AE84G and NNX16AF52G, along with NSF grant award number 1516777. This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration. The data presented herein were obtained at the W.M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W.M. Keck Foundation.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. The authors wish to recognize and sincerely appreciate the work of the entire WMKO staff over the last two plus decades, and to the efforts of the team at the NASA Exoplanet Science Institute (NExScI) who are responsible for maintaining the KOA.

Facility: Keck(LRIS); Keck(HIRES).
Software: XIDL; ESIRedux.

ORCID iDs
John M. O’Meara @ https://orcid.org/0000-0002-7893-1054
Nicolas Lehner @ https://orcid.org/0000-0001-9158-0829
J. Christopher Howk @ https://orcid.org/0000-0002-2591-3792
J. Xavier Prochaska @ https://orcid.org/0000-0002-7738-6875

References
Bochanski, J. J., Hennawi, J. F., Simcoe, R. A., et al. 2009, PASP, 121, 1409
Cooke, R. J., Pettini, M., & Steidel, C. C. 2018, ApJ, 855, 102
Hiss, H., Walther, M., Olofse, I., & Hennawi, J. F. 2019, ApJ, 876, 71
Lehner, N., O’Meara, J. M., Fox, A. J., et al. 2014, ApJ, 788, 119
Murphy, M. 2018, Mtmurphy77/Uves_Popler: Uves_Popler: Post-Pipeline
Echelle Reduction Software, v1.00, Zenodo, doi:10.5281/zenodo.1297190
O’Meara, J. M., Lehner, N., Howk, J. C., et al. 2015, AJ, 150, 111
O’Meara, J. M., Lehner, N., Howk, J. C., et al. 2017, AJ, 154, 114
Prochaska, J. X., & Wolfe, A. M. 2009, ApJ, 696, 1543
Rafelski, M., Wolfe, A. M., Prochaska, J. X., Neeleman, M., & Mendez, A. J.
2012, ApJ, 755, 89
Ross, A. J., Bautista, J., Tojeiro, R., et al. 2020, MNRAS, 498, 2354
Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, PASP, 114, 851