The Astronomical Journal, 154:114 (5pp), 2017 September https://doi.org/10.3847/1538-3881/aa82b8

The Second Data Release of the KODIAQ Survey

J. M. O'Meara[1], N. Lehner[2], J. C. Howk[2], J. X. Prochaska[3], A. J. Fox[4], M. S. Peeples[4,5], J. Tumlinson[4,5], and B. W. O'Shea[6]

[1] Department of Chemistry and Physics, Saint Michael's College, One Winooski Park, Colchester, VT 05439, USA
[2] Department of Physics, University of Notre Dame, 225 Nieuwland Science Hall, Notre Dame, IN 46556, USA
[3] University of California/Lick Observatory, Santa Cruz, 1156 High Street, CA 95064, USA
[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[5] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[6] Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA
Received 2017 June 29; revised 2017 July 23; accepted 2017 July 25; published 2017 August 29

Abstract

We present and make publicly available the second data release (DR2) of the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) survey. KODIAQ DR2 consists of a fully reduced sample of 300 quasars at $0.07 < z_{\rm em} < 5.29$ observed with HIRES at high resolution ($36{,}000 \leqslant R \leqslant 103{,}000$). DR2 contains 831 spectra available in continuum normalized form, representing a sum total exposure time of ~4.9 megaseconds on source. These co-added spectra arise from a total of 1577 individual exposures of quasars taken from the Keck Observatory Archive (KOA) in raw form and uniformly processed. DR2 extends DR1 by adding 130 new quasars to the sample, including additional observations of QSOs in DR1. All new data in DR2 were obtained with the single-chip Tektronix TK2048 CCD configuration of HIRES in operation between 1995 and 2004. DR2 is publicly available to the community, housed as a higher level science product at the KOA and in the *igmspec* database (v03).

Key words: intergalactic medium – quasars: absorption lines – surveys

Supporting material: machine-readable tables

1. Introduction

The High-Resolution Echelle Spectrograph (HIRES; Vogt et al. 1994) on the Keck I telescope has been a spectacular instrument for the study of intergalactic and circumgalactic gas around galaxies since its introduction in 1994. The large aperture of Keck I coupled with the high spectral resolution of HIRES ($R \approx 35{,}000$–$100{,}000$) gave rise to a 20-year golden age of QSO absorption line studies. HIRES has allowed the best determinations of the primordial D/H ratio (Burles & Tytler 1998; O'Meara et al. 2001; Cooke et al. 2014), a test of Big Bang nucleosynthesis and fundamental physics, and it provided high precision constraints on the fine-structure constant (Murphy et al. 2001, 2003). HIRES has been instrumental in studying the physics of the intergalactic medium (IGM) and thereby cosmology (Rauch et al. 1997; Simcoe et al. 2004; Tytler et al. 2009; Bolton et al. 2014), including the highest redshifts accessible at the end of reionization (Bolton et al. 2010; Becker et al. 2011). HIRES has played leading roles in the determination of the metallicity distribution in and physics of damped Lyα systems (Wolfe et al. 2005; Rafelski et al. 2012). HIRES has also been critical in the study of circumgalactic gas at all redshifts (Rudie et al. 2012; Lehner et al. 2013; Werk et al. 2013; Wotta et al. 2016), including in setting the scale for the Milky Way's high velocity cloud population (Wakker et al. 2007, 2008; Thom et al. 2008; Smoker et al. 2011).

As part of their contribution to the Keck Observatory, NASA has provided an archival service to give access to the data from HIRES (and now the other Keck instruments as well). The Keck Observatory Archive (KOA[7] holds the vast majority of all HIRES observations of distant QSOs). These data are publicly accessible to all interested researchers. However, they are nominally only available in their raw form. Given the idiosyncrasies of data collection by a large number of observers operating under disparate conditions, there is a significant hurdle to using these data directly for science. This is especially true given the difficulties in co-addition, a process requiring an order-by-order continuum placement for optimal results.

In order to address these issues, we have undertaken a NASA-funded processing of all of the extant HIRES data for the study of QSO absorption lines to provide fully reduced QSO spectra to the community. This first data release (DR1) of our Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) was presented in O'Meara et al. (2015), which included spectra for 170 QSOs. The original science goal of the KODIAQ survey has been to study of the O VI absorption in strong H I absorbers at high $z > 2.2$ (Lehner et al. 2014), revealing unique properties of these absorbers. Besides our original program, the data making up DR1 have motivated our new KODIAQ Z survey aimed to determine the metallicity of the strong H I absorbers ($15 < \log N_{\rm H\,I} < 19$) at $z > 2$, and have been used in a number of new surveys. These include the characterization and interpretation of the small-scale structure in the Lyα forest (Rorai et al. 2017), the metallicity distribution of strong H I systems at intermediate- and high-redshifts (Lehner et al. 2013, 2016; Fumagalli et al. 2016; Wotta et al. 2016), the statistical characterization of Mg II and LLS absorbers (Prochaska et al. 2015; Mathes et al. 2017), characterization of extremely low-metallicity gas, including for the determination of primordial D/H (Cooke et al. 2016; Crighton et al. 2016), and even the study of the extended reaches of extremely low-redshift galaxies (Dutta et al. 2016).

The DR1 data sets were confined to those taken after the HIRES upgrade to a three-CCD mosaic in 2004 due to the shift

[7] http://www2.keck.hawaii.edu/koa/public/koa.php

Table 1
HIRES Deckers Used in KODIAQ DR1

Decker	Width (arcsec)	Spectra Resolution (FWHM)
B2	0.574	72,000
B5	0.861	48,000
C1	0.861	48,000
C2	0.861	48,000
C5	1.148	36,000
D1	1.148	36,000
E3	0.400	103,000

in data reduction approach associated with that upgrade. This excluded the first decade of HIRES data, including many exceptional data sets. Here, we present KODIAQ DR2, which now includes HIRES spectra of QSOs from the years 1995–2004. Our paper lays out our data reduction approach, which varies somewhat from that discussed in O'Meara et al. (2015) due to the difference in detectors, and presents the statistics of the combined DR1+DR2 data sets. Together these releases contain 300 QSOs covering the redshift range $0.07 \lesssim z_{em} \lesssim 5.29$. All of the data presented here are now available through the KOA and will be available with the distribution of v03 of the `specdb` data set[8] (Prochaska 2017).

2. The Data

A full discussion of HIRES and its data, including the process of downloading and ingesting the raw data into the KODIAQ database is presented in O'Meara et al. (2015). The new data presented here in DR2 all stem from HIRES observations by multiple PIs between 1995 and 2004. Table 1 presents the HIRES deckers used across DR2 and their corresponding spectral resolution. As with DR1, the majority of observations were made with the C1 or C5 decker providing ≈6 and ≈8 km s^{-1} FWHM resolution, respectively.

Data obtained prior to the HIRES upgrade in 2004 had slightly reduced total spectral coverage due to the size of the detector. As a result, observations were frequently made in matched sets where the echelle angle was kept fixed, but the cross-disperser angle varied so as to provide fewer spectral gaps, particularly at redder wavelengths. The overall throughput of the pre-2004 detector was lower, resulting in a longer average per-exposure integration time as compared to post-2004 upgrade observations.

2.1. Data Reduction

The data reduction largely followed the same procedures as outlined in O'Meara et al. (2015). However, for the new data in DR2, only data in the single-chip configuration was reduced, whereas in O'Meara et al. (2015), only three-chip mosaic data was involved. Minor differences in the data reduction with HIRedux data reduction package[9] arise from the change in detectors. First, the single-chip detector has a blemish known as the "ink spot" in the center of the detector. The pixels associated with the ink spot were masked in the extracted spectrum and ignored in any further processing so as not to confuse it with real QSO absorption lines. Second, the flat-fielding procedure from O'Meara et al. (2015) was not applied as the instrument was changed in 2004, thus negating the ability to construct the nominal HIRedux pixel flats. Instead, we use the internal flats both for order edge identification and pixel flat-fielding.

Finally, multiple spectra showed significant echelle order overlaps in their bluest orders, typically at wavelengths $\lambda < 3800$ Å, significantly complicating sky subtraction, source identification, and source extraction. In most cases, these orders were trimmed off. Therefore, the spectral coverage of these objects is smaller than is quoted in the KOA. It is also worth noting that the pre-2004 hires detector was far less blue-sensitive than its successor. As a result, observations at very blue ($\lambda < \sim 3400$ Å) wavelengths often had very long exposure times, some in excess of two hours per integration. For these exposures, unless a number of other exposures of the same object were taken with the same setup, cosmic ray rejection will likely not be optimal. As with the DR1, continua were fit on an order-by-order basis in the manner described in O'Meara et al. (2015). The full DR2 sample contains over 15000 echelle orders that have been continuum fit.

2.2. KODIAQ at the KOA and igmspec

As with DR1, the DR2 data products are available for community download from the KOA. DR2 supersedes DR1 at the KOA and contains the full 300 quasar sample. As with DR1, users may search for and download individual DR2 quasar sight line flux and error spectra, or the full DR2 sample all at once. Machine-readable tables and files exist for each spectrum to link back to the raw KOA data as needed. DR2 also makes available all intermediate data reduction steps grouped on an observing run by observing run basis. Spectra from DR2 are also available in the v03 release of the *igmspec* database (see Prochaska (2017) for details).

3. Properties of KODIAQ DR2

The KODIAQ DR2 comprises HIRES observations of 300 quasar lines of sight in total. Of these, 130 quasar sight lines are new since DR1, along with many new additional observations of some of the DR1 quasars. Table 2 presents the new data since DR1. As in O'Meara et al. (2015), quasars are named according to their J2000 R.A./decl. coordinates as resolved by either SIMBAD or SDSS. Many quasars were observed by multiple PIs in multiple instrument configurations. For a given PI+configuration, the data were combined. Table 2 lists the exposure time of the combined spectra, including the wavelength coverage of the spectra after trimming poorly extracted orders when necessary (see above). We note that these exposure times are only a crude measure of data quality, as they do not include the observing conditions.

Table 3 presents the full DR2 sample of 300 quasars. The rest wavelengths listed in Table 3 are given at the quasar emission redshift, i.e., $\lambda_r = \lambda_{obs}/(1.0 + z_{em})$. The full DR2 comprises 1577 individual exposures, grouped into 831 spectral co-adds. The aggregate exposure time of the full DR2 sample is ∼4.9 megaseconds. DR2 represents a significant increase over DR1, which was comprised of 170 quasar sight lines, 240 co-additions, 567 individual exposures, and 1.6 megaseconds of total exposure time.

8 http://github.com/specdb
9 http://www.ucolick.org/~xavier/HIRedux/

Table 2
The KODIAQ DR1 Sample

Object	R.A.[a] (J2000)	Decl.[a] (J2000)	z_{em}[a]	Observation Date	PI	Total Exp, Time (s)	Decker	Wavelength Coverage (Å)
J000150−015940	00:01:50.0	−01:59:40	2.810	1997 Sep	Wolfe	25000	C5	4111−6520
J000520+052410	00:05:20.21	+05:24:10.79	1.887	1999 Oct	Sargent	8000	C1	3307−4849
…	…	…	…	1999 Oct	Sargent	8000	C1	3340−4888
…	…	…	…	1998 Sep	Steidel	3600	C1	3217−4723
J001602−001224	00:16:02.40	−00:12:24.9	2.087	1996 Nov	Roth	9000	C1	3766−6191
…	…	…	…	1996 Nov	Roth	5400	C1	3903−6875
J001708+813508	00:17:08.47	+81:35:08.13	3.366	1995 Aug	Cowie	3600	C1	4414−6873
…	…	…	…	1995 Aug	Cowie	5400	C1	4437−6911
…	…	…	…	1996 Nov	Roth	10800	C1	5353−7791
J002208−150539	00:22:08.0	−15:05:39	4.528	1996 Sep	Wolfe	13500	C5	5293−7792

Note.
[a] J2000 coordinates and quasar redshifts are determined by passing the quasar coordinates from the raw data header through SIMBAD or SDSS and choosing the appropriate match.

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

Table 3
The KODIAQ DR2 Sample

Object	z_{em}	Total Exp, Time (s)	Wavelength Coverage (Å)	Rest Wavelength Coverage (Å)[a]
J000150-015940	2.81	25000	4111–6520	1079–1711
J000520 +052410	1.887	19600	3217–4888	1114–1693
J000931 +021707	2.35	1200	3334–6198	995–1850
J001602-001224	2.087	14400	3766–6875	1219–2227
J001708 +813508	3.366	19800	4414–7791	1010–1784
J002127-020333	2.596	700	3366–6198	936–1723
J002208-150539	4.528	13500	5293–7792	957–1409
J002830-281704	2.4	600	3366–6198	989–1822
J002952 +020606	2.333	1000	3367–6198	1010–1859
J003501-091817	2.418	16200	3123–5980	913–1749

Note.
[a] Rest wavelength calculated at the quasar redshift.

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

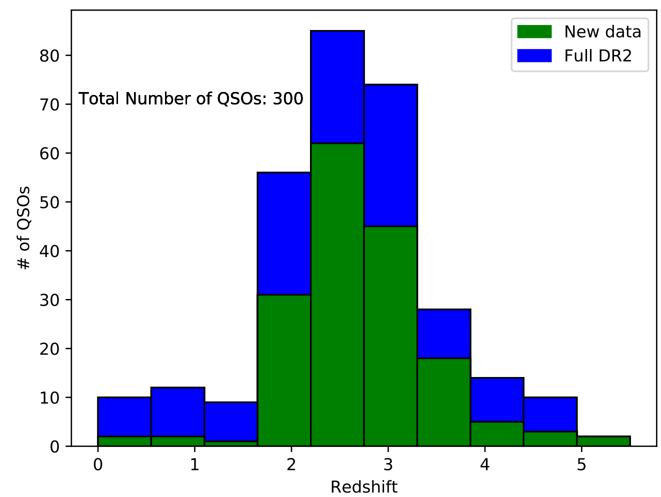

Figure 1. Redshift distribution of KODIAQ DR2 quasars.

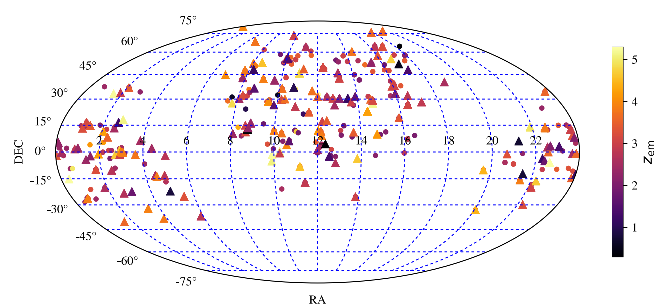

Figure 2. Distribution on the sky of the KODIAQ DR2 quasars. Triangles correspond to new quasars as presented in Table 2, and circles represent quasars presented in DR1.

3.1. General Properties

Figure 1 shows the quasar redshift distribution for the DR2 sample, including the distribution for the 130 new quasars alone. DR2 spans a range in redshift from $0.07 < z_{em} < 5.29$, with a median redshift of $\bar{z}_{em} = 2.586$ and a standard deviation of $\sigma_{z_{em}} = 0.918$. The distribution of DR2 quasars on the sky is shown in Figure 2. To illustrate the properties of DR2, we have performed a further co-addition of the data to produce a single spectrum per quasar. For each quasar, the co-addition resamples the data onto a single binning and resolution. The figures in O'Meara et al. (2015) largely summarize the DR1 in a different manner, applying to the spectra co-added within a particular instrument setup.

Figure 3 displays the median S/N, as calculated by $\frac{\mathrm{S}}{\mathrm{N}} = \frac{1.0}{\sigma}$ per pixel of the DR2 within ±5 Å of three rest wavelengths. The pixels are a constant 2.1 km s^{-1} for pre-2004 observations, 2.6 km s^{-1} for post-2004 observations, or 2.6 km s^{-1} for co-additions of observations mixing pre- and post-2004 data. The rest wavelengths were chosen to illustrate the viability for DR2 to be used for surveys of metal line, Lyα, and Lyman limit absorption. As can be seen in Figure 3, a significant number of quasars have high S/N, with many exceeding S/N 50 per pixel.

3.2. Cosmological Properties

The DR2 represents a significant increase over DR1 in its ability to provide a rich data sample for cosmological studies.

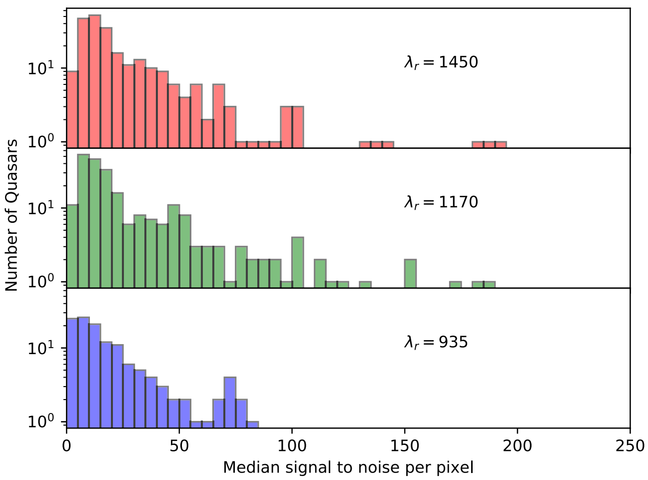

Figure 3. Signal-to-noise (per pixel) distribution of the co-added spectra in the DR2 sample.

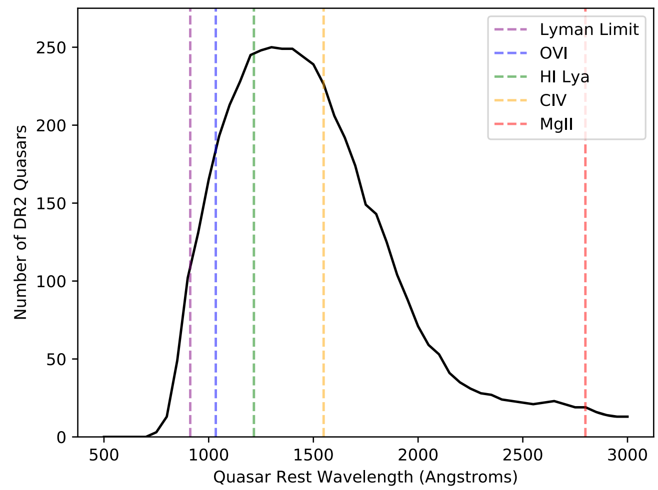

Figure 4. QSO redshift rest wavelength coverage of the DR2 sample. The vertical lines correspond to the rest wavelengths of various commonly studied ions.

We highlight a number of the DR2 properties below with respect to its ability to be used for surveys of metal line and H I absorption.

Figure 4 shows the rest wavelength (quasar redshift frame) spectral coverage of DR2. Well over 200 quasars in DR2 cover rest wavelengths corresponding to Lyα and C IV, and approximately 100 quasars cover the quasar Lyman limit. Figure 5 displays the redshift range over which certain ions may appear in the DR2 spectra. It bears noting that this figure is optimistic in that although certain ions (such as C IV) could appear in DR2 spectra, they may be blended with Lyα forest absorption from higher redshifts. Nevertheless, Figure 5 demonstrates that searches for ionic absorption in DR2 data, particularly at redshifts near $z = 2.5$, are likely to have strong statistical power.

A further illustration of this point is found in Figure 6, which displays the redshift coverage of certain ions and their numbers at or above certain S/N cuts in the data. Again, Lyα and C IV stand out in terms of their numbers and data quality, but we also note that a significant number of high quality spectra exist covering Mg II absorption, albeit subject to the same caveat that some of the Mg II spectral coverage overlaps with higher redshift Lyα forest. Many tens of spectra have S/N in excess of 100, making DR2 the largest collection of high resolution, high signal-to-noise quasar spectra openly available to the community.

Finally, Figure 7 gives the redshift sensitivity function $g(z)$ for DR2. Unlike the previous figures, we consider specific wavelength ranges in the spectra to calculate $g(z)$. As in O'Meara et al. (2015), we select only the regions within a given quasar spectrum between 3000 km s^{-1} to the red of the quasar Lyβ line and 3000 km s^{-1} to the blue of the quasar Lyα line for the $g(z)$ calculation of the redshift sensitivity to Lyα absorption. For C IV, we use the same 3000 km s^{-1} windows, but for the Lyα and C IV emission lines, respectively. The high frequency variations in $g(z)$ stem from the blaze function of the echelle orders, and various sharp features usually arise from detector defects or gaps in spectral coverage.

4. Summary and Future

We have presented here and made public at the KOA the DR2 of the KODIAQ survey. DR2 contains the spectra of 300

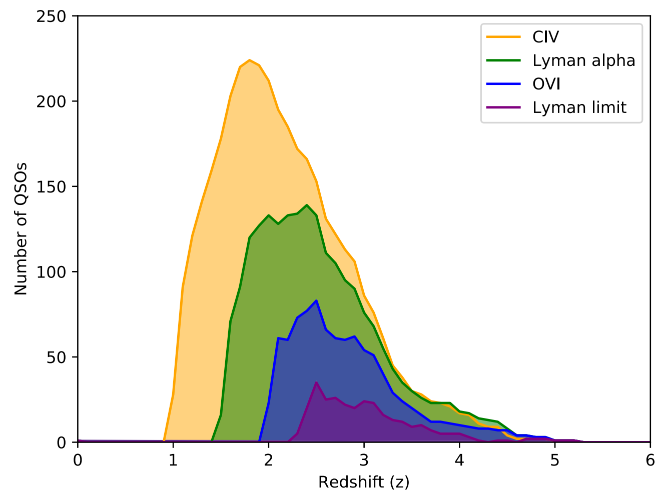

Figure 5. Redshift coverage of intervening absorption from various ions in the DR2 sample.

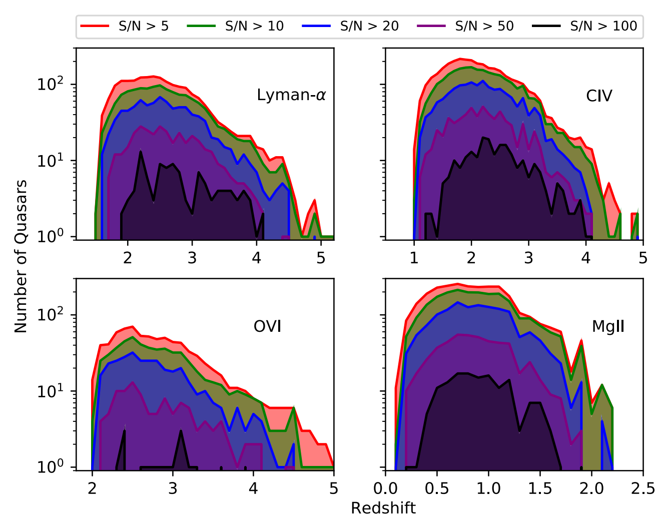

Figure 6. Signal-to-noise distribution vs. redshift for key ions in the DR2 sample.

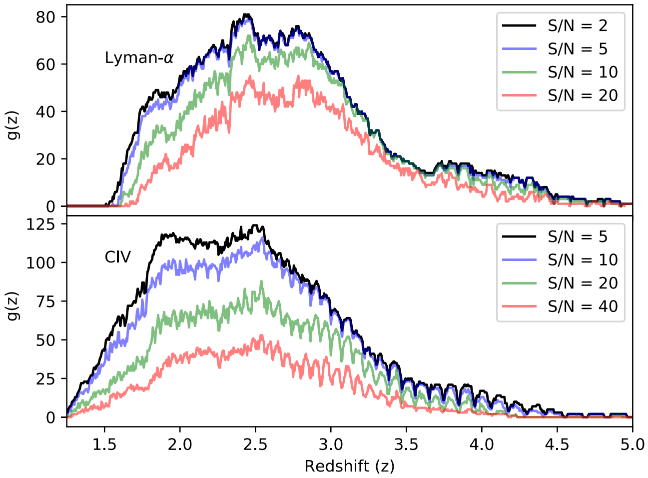

Figure 7. Redshift sensitivity function $g(z)$ for Lyα and C IV in the DR2 as a function of S/N ratio.

individual quasars obtained with HIRES since 1995. The data vary in signal to noise and resolution, but a significant subset of the DR2 is of high enough quality to facilitate a number of precision studies of the intergalactic and circumgalactic medium between $0.15 < z < 5$.

With DR2, a significant fraction of the total number of quasar observations made with HIRES since 1995 has been released to the community. We intend one final data release of HIRES data, which will mix the remaining pre- and post-2004 data available from the KOA that can be nominally reduced. This final data release is anticipated in 2019–2020. As part of KODIAQ Z, we will also reduce and distribute all quasars observed with the ESI instrument on Keck II in ∼2020.

When using data products from DR2, in addition to the standard KOA acknowledgement (including acknowledgement to the original PIs of each program), we request that the community please include the following acknowledgement: "Some/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 NSF award number 1516777" and to cite this paper and O'Meara et al. (2015).

Support for this work 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 Mauna Kea 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. Finally, we continue to dedicate this work to the late astronomers Arthur Wolfe and Wal Sargent, and extend the dedication to the late Jerry Nelson, without whom the Keck telescopes and their great impact on science would not exist.

ORCID iDs

J. M. O'Meara https://orcid.org/0000-0002-7893-1054
N. Lehner https://orcid.org/0000-0001-9158-0829
J. X. Prochaska https://orcid.org/0000-0002-7738-6875
A. J. Fox https://orcid.org/0000-0003-0724-4115
M. S. Peeples https://orcid.org/0000-0003-1455-8788
J. Tumlinson https://orcid.org/0000-0002-7982-412X
B. W. O'Shea https://orcid.org/0000-0002-2786-0348

References

Becker, G. D., Sargent, W. L. W., Rauch, M., & Calverley, A. P. 2011, ApJ, 735, 93
Bolton, J. S., Becker, G. D., Haehnelt, M. G., & Viel, M. 2014, MNRAS, 438, 2499
Bolton, J. S., Becker, G. D., Wyithe, J. S. B., Haehnelt, M. G., & Sargent, W. L. W. 2010, MNRAS, 406, 612
Burles, S., & Tytler, D. 1998, ApJ, 499, 699
Cooke, R. J., Pettini, M., Jorgenson, R. A., Murphy, M. T., & Steidel, C. C. 2014, ApJ, 781, 31
Cooke, R. J., Pettini, M., Nollett, K. M., & Jorgenson, R. 2016, ApJ, 830, 148
Crighton, N. H. M., O'Meara, J. M., & Murphy, M. T. 2016, MNRAS, 457, L44
Dutta, R., Gupta, N., Srianand, R., & O'Meara, J. M. 2016, MNRAS, 456, 4209
Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2016, MNRAS, 455, 4100
Lehner, N., Howk, J. C., Tripp, T. M., et al. 2013, ApJ, 770, 138
Lehner, N., O'Meara, J. M., Fox, A. J., et al. 2014, ApJ, 788, 119
Lehner, N., O'Meara, J. M., Howk, J. C., Prochaska, J. X., & Fumagalli, M. 2016, ApJ, 833, 283
Mathes, N. L., Churchill, C. W., & Murphy, M. T. 2017, arXiv:1701.05624
Murphy, M. T., Webb, J. K., & Flambaum, V. V. 2003, MNRAS, 345, 609
Murphy, M. T., Webb, J. K., Flambaum, V. V., et al. 2001, MNRAS, 327, 1208
O'Meara, J. M., Lehner, N., Howk, J. C., et al. 2015, AJ, 150, 111
O'Meara, J. M., Tytler, D., Kirkman, D., et al. 2001, ApJ, 552, 718
Prochaska, J. X. 2017, A&C, 19, 27
Prochaska, J. X., O'Meara, J. M., Fumagalli, M., Bernstein, R. A., & Burles, S. M. 2015, ApJS, 221, 2
Rafelski, M., Wolfe, A. M., Prochaska, J. X., Neeleman, M., & Mendez, A. J. 2012, ApJ, 755, 89
Rauch, M., Miralda-Escudé, J., Sargent, W. L. W., et al. 1997, ApJ, 489, 7
Rorai, A., Hennawi, J. F., Oñorbe, J., et al. 2017, Sci, 356, 418
Rudie, G. C., Steidel, C. C., Trainor, R. F., et al. 2012, ApJ, 750, 67
Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2004, ApJ, 606, 92
Smoker, J. V., Fox, A. J., & Keenan, F. P. 2011, MNRAS, 415, 1105
Thom, C., Peek, J. E. G., Putman, M. E., et al. 2008, ApJ, 684, 364
Tytler, D., Paschos, P., Kirkman, D., Norman, M. L., & Jena, T. 2009, MNRAS, 393, 723
Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, Proc. SPIE, 2198, 362
Wakker, B. P., York, D. G., Howk, J. C., et al. 2007, ApJL, 670, L113
Wakker, B. P., York, D. G., Wilhelm, R., et al. 2008, ApJ, 672, 298
Werk, J. K., Prochaska, J. X., Thom, C., et al. 2013, ApJS, 204, 17
Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, ARA&A, 43, 861
Wotta, C. B., Lehner, N., Howk, J. C., O'Meara, J. M., & Prochaska, J. X. 2016, ApJ, 831, 95