

## THE FIRST DATA RELEASE OF THE KODIAQ SURVEY

J. M. O'MEARA<sup>1</sup>, N. LEHNER<sup>2</sup>, J. C. HOWK<sup>2</sup>, J. X. PROCHASKA<sup>3</sup>, A. J. FOX<sup>4</sup>, M. A. SWAIN<sup>5</sup>,  
C. R. GELINO<sup>6</sup>, G. B. BERRIMAN<sup>6</sup>, AND H. TRAN<sup>7</sup>

<sup>1</sup>Department of Chemistry and Physics, Saint Michael's College, One Winooski Park, Colchester, VT 05439, USA

<sup>2</sup>Department of Physics, University of Notre Dame, 225 Nieuwland Science Hall, Notre Dame, IN 46556, USA

<sup>3</sup>University of California/Lick Observatory, Santa Cruz, 1156 High Street, CA 95064, USA

<sup>4</sup>Space Telescope Science Institute, Baltimore, MD 21218, USA

<sup>5</sup>Raytheon, 299 N Euclid Avenue, Suite 500, Pasadena, CA 91101, USA

<sup>6</sup>NASA Exoplanet Science Institute, Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

<sup>7</sup>W. M. Keck Observatory, 65-1120 Mamalahoa Hwy, Kamuela, HI 96743, USA

Received 2015 May 13; accepted 2015 August 6; published 2015 September 14

### ABSTRACT

We present and make publicly available the first data release (DR1) of the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) survey. The KODIAQ survey is aimed at studying galactic and circumgalactic gas in absorption at high redshift, with a focus on highly ionized gas traced by O VI, using the HIRES spectrograph on the Keck I telescope. KODIAQ DR1 consists of a fully reduced sample of 170 quasars at  $0.29 < z_{\text{em}} < 5.29$  observed with HIRES at high resolution ( $36,000 \leq R \leq 103,000$ ) between 2004 and 2012. DR1 contains 247 spectra available in continuum normalized form, representing a sum total exposure time of  $\sim 1.6$  megaseconds. These coadded spectra arise from a total of 567 individual exposures of quasars taken from the Keck Observatory Archive (KOA) in raw form and uniformly processed using a publicly available HIRES data reduction package. DR1 is publicly available to the community, housed as a higher level science product at the KOA. We will provide future data releases that make available additional quasars, including those with pre-2004 observations taken with the previous-generation HIRES detectors.

*Key words:* catalogs – intergalactic medium – quasars: absorption lines – quasars: general – surveys

### 1. INTRODUCTION

HIRES (Vogt et al. 1994) at the Keck I telescope on Maunakea has contributed significantly to our knowledge of the intermediate- to high-redshift universe. In particular, observations of absorption lines toward background quasars with HIRES have provided a number of fundamental discoveries and studies. HIRES observations have provided experimental tests of Big Bang nucleosynthesis and have helped to determine the cosmological baryon density (e.g., Burles & Tytler 1998; O'Meara et al. 2001; Cooke et al. 2014). They have discovered pristine gas clouds (Fumagalli et al. 2011) and determined the metal budgets over 10 decades in H I column density from the Ly $\alpha$  forest (Simcoe et al. 2004) to the damped Ly $\alpha$  systems (Wolfe et al. 2005). They have explored the galaxy-IGM connection (Steidel et al. 2010) and have constrained the thermal history of the universe (Bolton et al. 2014). They have helped constrain dark matter models (Viel et al. 2013), have provided critical tests to cosmological simulations (Tytler et al. 2009), and been used to study possible changes in the fine-structure constant with redshift (Murphy et al. 2001, 2003; Webb et al. 2001). Indeed, at the time of submission, HIRES appears in the title or abstract of over 1000 papers totaling over 20,000 citations.

Remarkably, these successful studies have primarily been achieved with single objects, or with samples numbering in the tens of sightlines. This largely reflects the limited allocations of Principal Investigators (PIs) with direct access to the Keck telescopes. The compilation of a significant portion of the quasar sightlines observed with HIRES in the Keck Observatory Archive (KOA<sup>8</sup>), however, now facilitates the study of

hundreds of quasar sightlines by making the data available to the community in raw form. In 2009, we were awarded a NASA Astrophysics Data Analysis Program (ADAP) grant (PI: Lehner) to use this rich quasar database ( $>600$  quasars as of this writing) to study the hot ionized O VI gas in the circumgalactic medium of  $2 < z < 4$  galaxies in a sample of H I-selected strong absorbers ( $\log N_{\text{H I}} \gtrsim 17$ , i.e., the Lyman limit and damped Ly $\alpha$  systems). The first results from our Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) survey are presented in Lehner et al. (2014).

As part of the KODIAQ survey, we have uniformly processed the HIRES spectra of hundreds of quasars. In this work, we present the first data release (DR1) from the survey. We make available to the community continuum normalized 1D spectra for 170 quasars and make them publicly available at the KOA. This work joins and extends the public data releases of Keck data from Prochaska et al. (2001, 2003, 2007), and Rafelski et al. (2014), and is complementary to the European Southern Observatory Advanced Data Products Quasar Sample (e.g., Zafar et al. 2013). Here we present the details of our work to produce this archive. An outline of the paper is as follows: In Section 2, we describe the basic properties of the data, its reduction, and its availability at the KOA. In Section 3, we describe the specific properties of the DR1 sample. Finally, in Section 4 we describe future planned data releases.

### 2. THE DATA

Data taken over the last two decades from the HIRES spectrograph fall into three general categories. The first is the earliest data from 1995 to 1997, where the light is dispersed onto a single Tektronix 2048  $\times$  2048 chip (see Vogt et al. 1994 for details). The second category spans the years 1997–2004, when HIRES could be configured using either a red or blue

<sup>8</sup> <http://www2.keck.hawaii.edu/koa/public/koa.php>

**Table 1**  
HIRES Deckers used in KODIAQ DR1

Decker	Width (arcseconds)	Spectral Resolution (FWHM)
B2	0.574	72,000
B5	0.861	48,000
C1	0.861	48,000
C5	1.148	36,000
E3	0.400	103,000

sensitive cross-disperser to enhance throughput over specific wavelength ranges, but still feeding a single CCD. Finally, the third category, covering data from 2004 to the present, has the original single CCD replaced by a 3-chip mosaic, with dramatically increased blue sensitivity. The data in DR1 are all drawn from this last category.

HIRES has a number of user-selectable instrument element settings. These include detector binning, entrance slit width and length (set simultaneously by choosing a decker), and echelle and cross-disperser angle. The latter two elements determine the wavelength coverage for a given setup. For data taken with the blue cross-disperser, the data frequently have continuous wavelength coverage (with some overlaps between echelle orders), albeit with two small gaps in coverage due to the spacing between the chips in the mosaic. Data obtained with the red cross disperser often has additional gaps in wavelength coverage, due to the increasing spatial extent of the dispersed light becoming larger than the extent of the chip. The slit width sets the resolution of the data. Table 1 gives the spectral resolution associated with the HIRES deckers that span the DR1 sample and the last column of Table 2 gives the wavelengths covered for each spectrum. The vast majority of HIRES data in DR1 were obtained with the C1 or C5 decker that provide FWHM  $R = 48,000$  and  $R = 36,000$ , corresponding to velocity resolutions of 6.2 and 8.3 km s<sup>-1</sup> (FWHM), respectively.

### 2.1. Raw Data

The data for the KODIAQ survey are taken entirely from the KOA, i.e., they constitute a publicly available sample. While a quantity of HIRES observations of quasar absorption lines still remain unavailable,<sup>9</sup> the KOA sample constitutes a high percentage of all quasars observed to date with HIRES, as measured by comparing the data in the sample with the Keck I telescope schedule over the last two decades. Through searching by a combination of the Principal Investigator (PI) and project titles associated with HIRES data at the KOA, we downloaded all the quasar data in raw form for reduction and analysis. Including calibration files, the sample contains many tens of thousands of individual HIRES exposures totaling ~250 gigabytes. The data were grouped by observing run, where a single PI would observe with HIRES for a single night, or over multiple nights close in time. The choice of observing run grouping was made to best facilitate data reduction of objects with nearly identical HIRES configurations.

<sup>9</sup> Some PIs have requested and received proprietary times exceeding the nominal 1.5 years.

### 2.2. Data Reduction

Once grouped into observing runs, the data were uniformly reduced using the HIRedux code<sup>10</sup> developed as part of the XIDL<sup>11</sup> suite of astronomical routines in IDL. HIRedux was developed as a variant of the reduction pipeline<sup>12</sup> for the *Magellan* Inamori Kyocera Echelle (MIKE) spectrograph on the *Magellan* Clay telescope at Las Campanas Observatory.

Briefly, the HIRedux workflow is as follows: First, the raw data are read into the pipeline and sorted into “setups,” where each setup is a unique combination of decker, binning, and dispersing element choice and configuration. Next, for each setup, a bias level and gain is determined from the data, and a flat field image is constructed by a median combine of individual exposures. The flat field is used to determine the echelle order edges via polynomial fitting, and to determine the illumination of light along the slit (i.e., the tilt), which is not constant. Traditional pixel-to-pixel variation flat-fielding is performed using archival pixel flat fields obtained every few semesters by JXP. These were acquired with a clever recipe kindly suggested by R. Kibrick: a series of spectral images obtained with the cover to the cross-disperser in place. This cover is partially reflective and yields a relatively uniform image on one detector of the CCD mosaic when the cross-disperser grating is tilted to specific angles. The resultant, normalized image corrects for small-scale variations in the detector (e.g., defects, “pocks,” etc.) that challenge a traditional approach to flat fielding.

With the orders spatially defined, the next step is to determine a two-dimensional wavelength image for the chip by first extracting down the center of each order an exposure of a ThAr arc lamp. This extracted arc spectrum is then compared to a suite of archival ThAr databases and the algorithm solves for a two-dimensional wavelength solution rejecting outliers or spurious arc line features. The residuals to the arc line solutions are generally at the sub-pixel level, where a single, unbinned HIRES pixel spans 1.3 km s<sup>-1</sup>. If multiple exposures of a science target exist, cosmic rays are identified by analyzing the ratio of each exposure to the median combined frame in standard ways.

Next, the data are sky-subtracted and the object extracted on an order-by-order basis. The specifics of the sky subtraction and object extraction are given in Bochanski et al. (2009) who describes the MagE Spectral Extractor (MaSE) Spectral Extractor extraction pipeline used to process data from the *Magellan* Echelle Spectrograph (MagE), which shares a common codebase with HIRedux. The output of the extraction is an order-by-order spectrum of counts versus vacuum wavelength. For multiple exposures of a science target with the same setup, the individual orders are weighted-mean combined, with the weights coming from the signal to noise of the individual exposures.

### 2.3. Continuum Placement

For most types of analysis, quasar absorption line data are required to be in continuum normalized form. To put the data in this form, the shape of both the spectrograph blaze function and the underlying quasar SED must be removed via the process of continuum-fitting. Many spectrographs are able, via

<sup>10</sup> <http://www.ucolick.org/~xavier/HIRedux/>

<sup>11</sup> <http://www.ucolick.org/~xavier/IDL/index.html>

<sup>12</sup> <http://web.mit.edu/~burles/www/MIKE/>

**Table 2**  
The KODIAQ DR1 Sample

Object	R.A. <sup>a</sup> (J2000)	decl. <sup>a</sup> (J2000)	$z_{em}$ <sup>a</sup>	Observation Date	PI	Total Exp. Time (s)	Decker	Wavelength Coverage (Å)
J000931+021707	00:09:31.43	+02:17:07.5	2.350	2009 Dec	Tytler	1200	C1	3334–6198
J002127–020333	00:21:27.35	–02:03:33.6	2.596	2009 Dec	Tytler	700	C1	3366–6198
J002830–281704	00:28:30.44	–28:17:04.8	2.400	2009 Dec	Tytler	600	C1	3366–6198
J002952+020606	00:29:52.12	+02:06:06.2	2.333	2009 Dec	Tytler	1000	C1	3366–6198
J003501–091817	00:35:01.88	–09:18:17.6	2.419	2009 Dec	Pettini	16200	C5	3123–5980
J004054–091526	00:40:54.65	–09:15:26.8	4.976	2011 Jan	Wolfe	7200	C1	5453–8570
J004351–265128	00:43:51.81	–26:51:28.0	2.786	2009 Dec	Tytler	1500	C1	3370–6198
J004358–255115	00:43:58.80	–25:51:15.5	2.501	2009 Dec	Tytler	900	C1	3365–6198
J004448+372114	00:44:48.32	+37:21:14.8	2.410	2009 Dec	Tytler	900	C1	3368–6198
J004530–261709	00:45:30.50	–26:17:09.0	3.440	2007 Aug	Crichton	14400	C1	3973–8543
J005202+010129	00:52:02.40	+01:01:29.3	2.271	2009 Dec	Tytler	1000	C1	3366–6198
J005700+143737	00:57:00.19	+14:37:37.8	2.643	2009 Dec	Tytler	1100	C1	3366–6198
J005814+011530	00:58:14.32	+01:15:30.2	2.494	2004 Oct	Wolfe	3600	C1	3386–6250
...	...	...	...	2009 Dec	Tytler	1200	C1	3368–6198
J010311+131617	01:03:11.26	+13:16:17.7	2.721	2005 Oct	Steidel	7200	C5	3123–5978
J010741–263328	01:07:41.92	–26:33:28.4	2.460	2008 Jan	Tytler	900	C5	3133–6031
J010806+163550	01:08:06.41	+16:35:50.0	2.652	2005 Dec	Tytler	22204	C5	3048–5895
J010925–210257	01:09:25.19	–21:02:57.0	3.226	2006 Jan	Prochaska	5400	C5	3304–6179
J011150+140141	01:11:50.07	+14:01:41.5	2.470	2008 Jan	Tytler	900	C5	3133–6031
J012156+144823	01:21:56.03	+14:48:23.8	2.870	2004 Sept	Prochaska	7200	C1	3244–6138
J013340+040059	01:33:40.40	+04:00:59.0	4.150	2006 Dec	Wolfe	7200	C1	4159–8742
J014516–094517A	01:45:16.60	–09:45:17.0	2.719	2005 Oct	Steidel	7200	C5	3102–5978
...	...	...	...	2008 Sep	Steidel	4000	C1	3022–5897
J014516–094517B	01:45:16.70	–09:45:18	2.719	2005 Oct	Steidel	17400	C1	3102–5979
...	...	...	...	2008 Sep	Steidel	7200	C1	3022–5897
J014944+150106	01:49:44.43	+15:01:06.7	2.060	2009 Dec	Tytler	1200	C1	3176–6038
J015227–200107	01:52:27.34	–20:01:07.1	2.060	2008 Jan	Tytler	840	C1	3177–6032
J020455+364917	02:04:55.59	+36:49:17.9	2.912	2004 Oct	Wolfe	3600	C1	3157–6031
...	...	...	...	2006 Aug	Ellison	5400	C1	3473–6384
...	...	...	...	2006 Aug	Prochaska	2300	C1	3501–6296
J020950–000506	02:09:50.70	–00:05:06.4	2.828	2004 Sept	Prochaska	13500	C1	3151–5994
...	...	...	...	2004 Oct	Wolfe	3600	C1	4650–9219
...	...	...	...	2008 Sep	Steidel	9830	C1	3022–5897
J020944+051714	02:09:44.62	+05:17:14.1	4.180	2006 Dec	Wolfe	7200	C1	5733–7314
...	...	...	...	2007 Sep	Wolfe	11100	C1	5211–8319
J021129+124110	02:11:29.16	+12:41:10.8	2.953	2011 Jan	Wolfe	10200	C1	3681–6588
J022554+005451	02:25:54.85	+00:54:51.9	2.975	2006 Dec	Wolfe	10800	C1	3571–6485
J022853–033737	02:28:53.21	–03:37:37.1	2.066	2009 Dec	Tytler	800	C1	3177–6039
J023145+132254	02:31:45.89	+13:22:54.7	2.059	2009 Dec	Tytler	1400	C1	3176–6039
J023359+004938	02:33:59.72	+00:49:38.4	2.522	2008 Jan	Tytler	1080	C1	3174–6033
J023924–090138	02:39:24.48	–09:01:38.6	2.471	2008 Jan	Tytler	1080	C1	3133–6032
J025127+341442	02:51:27.74	+34:14:42.0	2.230	2008 Jan	Tytler	780	C1	3133–6031
J025515+014828	02:55:15.21	+01:48:28.7	2.470	2008 Jan	Tytler	840	C1	3157–6030
J025644+001246	02:56:44.69	+00:12:46.0	2.251	2008 Jan	Tytler	900	C1	3157–6030
J030046+022245	03:00:46.02	+02:22:45.2	2.520	2008 Jan	Tytler	900	C1	3133–6033
J030341–002321	03:03:41.04	–00:23:21.8	3.176	2004 Sept	Prochaska	7200	C1	3536–6423
J031003–004645	03:10:03.01	–00:46:45.7	2.114	2008 Jan	Tytler	1200	C1	3133–6032
J033900–013318	03:39:00.90	–01:33:18.0	3.197	2005 Oct	Prochaska	9000	C5	3686–6599
...	...	...	...	2006 Jan	Prochaska	7200	C5	4185–7198
J040241–064124	04:02:41.42	–06:41:37.9	2.432	2011 Jan	Wolfe	3600	C5	2995–5882
J045213–164012	04:52:13.60	–16:40:12.0	2.684	2005 Oct	Steidel	13800	C5	3124–5979
...	...	...	...	2009 Dec	Pettini	3600	C5	3100–5979
J073149+285448	07:31:49.50	+28:54:48.6	3.676	2006 Jan	Prochaska	7200	C5	4160–8768
J074110+311200	07:41:10.70	+31:12:00.2	0.631	2006 Dec	Wolfe	1800	E3	3076–5934
J074521+473436	07:45:21.78	+47:34:36.2	3.220	2007 Oct	Milutinovic	4500	C1	3627–5535
J074749+443417	07:47:49.74	+44:34:17.0	4.435	2011 Jan	Wolfe	10800	C1	4908–7983
...	...	...	...	2011 Jul	Wolfe	3600	C1	4850–9459
J074927+415242	07:49:27.90	+41:52:42.3	3.111	2007 Mar	Hamann	12600	C1	3571–8127
J080518+614423	08:05:18.18	+61:44:23.70	3.033	2006 Jan	Prochaska	5400	C5	3987–8534
J081240+320808	08:12:40.68	+32:08:08.58	2.712	2007 Apr	Prochaska	14400	C1	3331–6190
J081435+502946	08:14:35.19	+50:29:46.30	3.897	2006 Dec	Wolfe	10800	C1	4024–8574
...	...	...	...	2008 Mar	Wolfe	10800	C1	5219–9725
J081740+135134	08:17:40.52	+13:51:34.5	4.389	2011 Jan	Wolfe	7200	C1	4935–8006

**Table 2**  
(Continued)

Object	R.A. <sup>a</sup> (J2000)	decl. <sup>a</sup> (J2000)	$z_{em}$ <sup>a</sup>	Observation Date	PI	Total Exp, Time (s)	Decker	Wavelength Coverage (Å)
J082107+310751	08:21:07.61	+31:07:51.2	2.626	2007 Apr	Sargent	10800	C1	3101–5981
...	...	...	...	2008 Sep	Steidel	5620	C5	3058–5897
J082540+354414	08:25:40.12	+35:44:14.2	3.846	2008 Mar	Wolfe	10800	B2	5180–9716
J082619+314848	08:26:19.70	+31:48:48.0	3.093	2006 Dec	Wolfe	7900	C1	3842–8362
J082849+085854	08:28:49.16	+08:58:54.8	2.271	2012 Apr	Wolfe	1293	C1	3153–5992
J083052+241059	08:30:52.08	+24:10:59.8	0.942	2006 Dec	Wolfe	1800	E3	3051–5936
J083102+335803	08:31:02.55	+33:58:03.1	2.427	2009 Dec	Pettini	18600	C5	3123–5979
J090033+421546	09:00:33.49	+42:15:46.8	3.290	2005 Mar	Prochaska	6754	C5	3545–8085
J092759+154321	09:27:59.78	+15:43:21.8	1.805	2011 Jan	Wolfe	12800	C1	3074–5967
J092708+582319	09:27:08.88	+58:23:19.4	1.910	2011 Jan	Wolfe	21600	C1	2995–5880
J092914+282529	09:29:14.49	+28:25:29.1	3.399	2007 Apr	Wolfe	10800	C1	4036–7424
J093643+292713	09:36:43.51	+29:27:13.6	2.924	2010 Feb	Prochaska	7200	C1	2995–5869
J094202+042244	09:42:02.04	+04:22:44.6	3.275	2005 Mar	Wolfe	7200	C1	3364–6200
J095309+523029	09:53:09.05	+52:30:29.7	1.8756	2005 Mar	Wolfe	7200	C1	3156–6039
J095822+014524	09:58:22.19	+01:45:24.2	1.960	2012 Jan	Wolfe	3600	C1	3169–6030
J095820+322402	09:58:20.95	+32:24:02.2	0.530	2006 Dec	Wolfe	600	E3	3078–5934
J095852+120245	09:58:52.19	+12:02:45.0	3.298	2006 Jan	Prochaska	7200	C5	3608–6495
...	...	...	...	2006 Apr	Prochaska	1800	C1	5444–8564
J100841+362319	10:08:41.22	+36:23:19.3	3.126	2007 Mar	Hamann	10800	C1	3571–7993
J101155+294141	10:11:55.60	+29:41:41.6	2.652	2005 Apr	Sargent	12000	C1	3101–5908
...	...	...	...	2005 May	Steidel	7200	C5	3101–5983
...	...	...	...	2005 Dec	Tytler	12615	C5	3048–5896
J101336+561536	10:13:36.37	+56:15:36.4	3.633	2006 Jan	Prochaska	3600	C1	3899–8368
J101447+430030	10:14:47.19	+43:00:30.0	3.125	2005 Apr	Prochaska	5100	C5	3500–6401
...	...	...	...	2007 Apr	Wolfe	7200	B2	3975–7057
J101723–204658	10:17:23.98	–20:46:58.6	2.545	2007 Apr	Sargent	14400	C1	3128–5981
J101939+524627	10:19:39.15	+52:46:27.8	2.170	2007 Apr	Prochaska	7200	C1	3118–5974
J102009+104002	10:20:09.99	+10:40:02.7	3.168	2006 Apr	Prochaska	7200	C5	3616–6524
J102325+514251	10:23:25.32	+51:42:51.1	3.447	2007 Mar	Hamann	6200	C1	3584–7975
J102410+060013	10:24:10.44	+06:00:13.7	2.131	2012 Apr	Wolfe	3300	C1	3128–5992
J103514+544040	10:35:14.22	+54:40:40.1	2.989	2008 Mar	Wolfe	10800	C1	3299–7801
...	...	...	...	2008 Jan	Prochaska	3600	C1	3368–6262
J104018+572448	10:40:18.52	+57:24:48.1	3.409	2006 Jan	Prochaska	8100	C5	3794–6657
J104213+062853	10:42:13.52	+06:28:53.0	2.035	2012 Apr	Wolfe	2400	C1	3128–5992
J105123+354534	10:51:23.03	+35:45:34.3	4.912	2010 Apr	Wolfe	7800	C5	5127–9713
J105648+120826	10:56:48.69	+12:08:26.8	1.923	2006 Apr	Prochaska	3600	C1	3101–5975
...	...	...	...	2006 Jun	Prochaska	7200	C1	3074–5966
...	...	...	...	2011 Jan	Wolfe	10500	C1	3074–5966
J110045+112239	11:00:45.23	+11:22:39.1	4.707	2011 Jan	Wolfe	17400	C1	5362–8391
J110621+104432	11:06:21.43	+10:44:32.6	1.859	2012 Apr	Wolfe	2700	C1	3325–6184
J110610+640009	11:06:10.74	+64:00:09.6	2.203	2007 Apr	Sargent	4800	C1	3128–5981
J111113–080402	11:11:13.6	–08:04:02.0	3.922	2006 Apr	Prochaska	7200	C1	4158–8751
J111909+211917	11:19:06.71	+21:18:50.3	1.183	2006 Jun	Prochaska	1800	E3	3051–5932
J112442–170517	11:24:42.87	–17:05:17.5	2.400	2005 May	Sargent	12000	C1	3127–5980
...	...	...	...	2007 Apr	Sargent	3600	C1	3126–5981
J113130+604420	11:31:30.40	+60:44:20.6	2.921	2006 Dec	Wolfe	7200	C1	3597–6423
J113418+574204	11:34:18.96	+57:42:04.7	3.522	2006 Jan	Prochaska	6300	C5	3973–8534
J113508+222715	11:35:08.09	+22:27:15.6	2.886	2006 Dec	Wolfe	6900	C1	3433–6311
J115538+053050	11:55:38.60	+05:30:50.5	3.475	2005 Apr	Prochaska	7200	C1	5057–8140
J115940–003203	11:59:40.79	–00:32:03.5	2.035	2012 Apr	Wolfe	1200	C1	3154–5992
J120207+323538	12:02:07.78	+32:35:38.8	5.292	2010 Apr	Wolfe	14400	C1	5123–9709
J120416+022111	12:04:16.68	+02:21:11.0	2.529	2005 Apr	Prochaska	5400	C5	3455–6309
...	...	...	...	2005 May	Prochaska	5400	C5	3568–6422
J120917+113830	12:09:17.93	+11:38:30.3	3.105	2006 Apr	Prochaska	7200	C5	3500–6400
J121117+042222	12:11:17.59	+04:22:22.2	2.526	2005 May	Prochaska	9000	C1	3403–6302
J121930+494052	12:19:30.77	+49:40:52.3	2.633	2007 May	Pettini	4000	C1	3127–5980
...	...	...	...	2009 May	Steidel	8000	C5	3027–5880
J122518+483116	12:25:18.64	+48:31:16.0	3.090	2007 Mar	Hamann	9000	C1	3571–8128
J122824+312837	12:28:24.96	+31:28:37.6	2.200	2007 Apr	Sargent	7200	C1	3101–5980
J123643–020420	12:36:43.12	–02:04:20.9	1.865	2012 Apr	Wolfe	3000	C1	3303–6185
J123748+012606	12:37:48.99	+01:26:06.9	3.145	2012 Apr	Prochaska	4000	C1	3370–6274
J124610+303131	12:46:10.80	+30:31:17.0	2.560	2006 Jun	Prochaska	3600	C1	3189–6086
J124924–023339	12:49:24.86	–02:33:39.7	2.117	2006 Jan	Prochaska	4800	C1	3075–5966

**Table 2**  
(Continued)

Object	R.A. <sup>a</sup> (J2000)	decl. <sup>a</sup> (J2000)	$z_{em}^a$	Observation Date	PI	Total Exp. Time (s)	Decker	Wavelength Coverage (Å)
J130426+120245	13:04:26.15	+12:02:45.5	2.980	2008 Mar	Wolfe	9000	B2	3606–8171
J130411+295348	13:04:11.98	+29:53:48.8	2.850	2010 Feb	Prochaska	7200	C1	2996–5871
J130542+092427	13:05:42.77	+09:24:27.7	2.063	2012 Apr	Wolfe	5400	C1	3128–5992
J131040+542449	13:10:40.24	+54:24:49.6	1.929	2005 Mar	Wolfe	10800	C1	3244–6136
J131215+423900	13:12:15.23	+42:39:00.8	2.567	2007 Apr	Sargent	9687	C1	3102–5983
J131341+144140	13:13:41.19	+14:41:40.5	1.884	2006 Jun	Prochaska	3600	C1	3244–6130
...	...	...	...	2011 Jan	Wolfe	7200	C1	3243–6132
J131855+531207	13:18:55.75	+53:12:07.2	2.321	2006 Jan	Prochaska	3600	C5	3838–6899
J133532+082404	13:35:32.65	+08:24:04.2	1.909	2012 Apr	Wolfe	3000	C1	3326–6185
J134328+572147	13:43:28.73	+57:21:47.2	3.034	2007 Apr	Prochaska	3600	C1	3075–5966
...	...	...	...	2007 Apr	Prochaska	3600	C1	3333–6190
J134544+262506	13:45:44.50	+26:25:06.0	2.031	2005 Apr	Prochaska	4800	C5	3501–6386
J135038–251216	13:50:38.88	–25:12:16.8	2.534	2005 Apr	Sargent	6000	C1	3128–5981
J135317+532825	13:53:17.10	+53:28:25.5	2.920	2008 Mar	Wolfe	8400	C1	3497–7985
J140501+444800	14:05:01.93	+44:47:59.8	2.218	2005 May	Martin	14400	C5	3252–6095
...	...	...	...	2005 May	Ellison	3200	C5	3215–6095
J141719+413237	14:17:19.23	+41:32:37.0	2.024	2005 Mar	Wolfe	14400	C5	3304–6188
...	...	...	...	2006 Jun	Prochaska	10800	C5	3281–6183
J141906+592312	14:19:06.31	+59:23:12.2	2.321	2007 Apr	Prochaska	7200	C1	3108–5966
J143316+313126	14:33:16.10	+31:31:26.3	2.940	2006 Apr	Prochaska	6000	C5	3245–6095
J143500+535953	14:35:00.54	+53:59:53.7	2.635	2005 May	Prochaska	7200	C1	3472–6379
J143916–015627	14:39:16.27	–01:56:27.6	2.162	2007 Apr	Prochaska	5400	C1	3075–5967
J144453+291905	14:44:53.54	+29:19:05.5	2.660	2007 Apr	Sargent	14400	C1	3102–5981
...	...	...	...	2009 May	Steidel	10000	C5	2995–5880
J145435+094100	14:54:35.18	+09:41:00.0	1.946	2012 Apr	Wolfe	2400	C1	3152–5992
J145408+511443	14:54:08.95	+51:14:43.7	3.644	2005 Jul	Wolfe	1800	C5	4075–8565
J150932+111313	15:09:32.11	+11:13:13.6	2.110	2012 Apr	Wolfe	5194	C1	3128–5992
J151224+465233	15:12:24.18	+46:52:33.5	3.360	2007 May	Pettini	11702	C5	3588–6423
J152156+520238	15:21:56.48	+52:02:38.4	2.208	2007 Apr	Sargent	7200	C1	3102–5982
J154153+315329	15:41:53.46	+31:53:29.4	2.558	2008 Mar	Wolfe	4800	C1	4417–7480
J155152+191104	15:51:52.48	+19:11:04.2	2.843	2007 Apr	Sargent	900	C1	3102–5982
...	...	...	...	2009 May	Steidel	6000	C5	2995–5880
J155556+480015	15:55:56.89	+48:00:15.0	3.299	2005 Jul	Wolfe	16200	C5	4078–8565
...	...	...	...	2006 Jun	Prochaska	10800	C5	4018–8564
J155810–003120	15:58:10.16	–00:31:20.0	2.827	2006 Apr	Prochaska	4100	C5	3361–6201
...	...	...	...	2006 Jun	Prochaska	10800	C5	3335–6201
J155814+405337	15:58:14.51	+40:53:37.0	2.635	2006 Jun	Chaffee	25200	C5	3101–5979
J160455+381214	16:04:55.40	+38:12:01.0	2.551	2005 Apr	Sargent	12000	C1	3125–5981
...	...	...	...	2009 May	Steidel	4200	C5	3030–5879
J160413+395121	16:04:13.97	+39:51:21.9	3.130	2007 Apr	Wolfe	9300	C1	4291–8767
J160547+511330	16:05:47.58	+51:13:30.2	1.785	2007 Apr	Prochaska	7200	C1	3118–5975
J160843+071508	16:08:43.90	+07:15:08.6	2.877	2007 Apr	Prochaska	5400	C1	3078–5975
...	...	...	...	2007 Apr	Prochaska	5400	E3	3078–5975
...	...	...	...	2008 Jul	Murphy	1500	C1	4475–7586
J161009+472444	16:10:09.42	+47:24:44.5	3.217	2006 Apr	Prochaska	3300	C1	3928–8507
...	...	...	...	2006 Jun	Prochaska	10800	C5	3834–8356
J162453+375806	16:24:53.47	+37:58:06.6	3.380	2006 Jul	Benn	7200	C5	3561–6591
J162548+264433	16:25:48.00	+26:44:32.6	2.535	2005 May	Steidel	12000	C5	3102–5984
...	...	...	...	2006 Jul	Ellison	1800	C5	5015–8103
...	...	...	...	2008 Sep	Steidel	5400	C1	3022–5896
J162548+264658	16:25:48.79	+26:46:58.7	2.518	2005 May	Steidel	10800	C5	3101–5983
...	...	...	...	2008 Sep	Steidel	6000	C1	3022–5896
J162557+264448	16:25:57.38	+26:44:48.2	2.601	2005 May	Steidel	15889	C5	3102–5983
...	...	...	...	2005 Oct	Steidel	5400	C5	3124–5979
...	...	...	...	2006 Jul	Ellison	7200	C5	5015–8123
J162645+642655	16:26:45.69	+64:26:55.2	2.320	2005 Apr	Sargent	11300	C1	3127–5981
...	...	...	...	2007 Apr	Sargent	3840	C5	3127–5981
J162902+091322	16:29:02.98	+09:13:22.5	1.986	2012 Apr	Wolfe	2400	C1	3153–5992
J170100+641209	17:01:00.61	+64:12:09.0	2.735	2005 Apr	Sargent	5800	C1	3129–5981
...	...	...	...	2005 May	Steidel	7800	C5	3102–5985
...	...	...	...	2005 Jul	Cowie	4800	B2	3571–8125
...	...	...	...	2005 Jul	Cowie	7200	B2	3741–6786
...	...	...	...	2008 Sep	Steidel	2000	C5	3060–5896

**Table 2**  
(Continued)

Object	R.A. <sup>a</sup> (J2000)	decl. <sup>a</sup> (J2000)	$z_{em}$ <sup>a</sup>	Observation Date	PI	Total Exp, Time (s)	Decker	Wavelength Coverage (Å)
J171227+575507	17:12:27.74	+57:55:06.8	3.008	2004 Sept	Prochaska	3600	C1	3471–6309
...	...	...	...	2005 May	Prochaska	3900	C1	3470–6310
...	...	...	...	2006 Aug	Prochaska	7200	C1	3451–6305
J173352+540030	17:33:52.23	+54:00:30.5	3.425	2005 May	Prochaska	5400	C5	3735–6645
...	...	...	...	2007 Aug	Crighton	5400	C1	3990–8543
J175603+574848	17:56:03.62	+57:48:48.0	2.110	2004 Sept	Prochaska	5400	C1	3185–6047
...	...	...	...	2004 Oct	Wolfe	10800	C1	3970–8544
...	...	...	...	2006 Jun	Prochaska	8800	C1	3157–6048
...	...	...	...	2006 Aug	Prochaska	18000	C1	3157–6047
...	...	...	...	2008 Jul	Murphy	18000	C1	4111–7170
J182157+642036	18:21:57.20	+64:20:36.0	0.297	2006 Jun	Prochaska	2400	E3	3105–5932
J193957–100241	19:39:57.25	–10:02:41.5	3.787	2005 Jul	Crighton	21600	B5	3569–6618
...	...	...	...	2005 Aug	Tytler	18800	C5	3843–6759
J200324–325145	20:03:24.11	–32:51:45.0	3.783	2005 Jul	Cowie	4800	B2	3565–8300
...	...	...	...	2005 Jul	Cowie	7200	B2	5217–8700
J203642–055300	20:36:42.29	–05:53:00.2	2.582	2004 Oct	Wolfe	10800	C1	3244–6145
...	...	...	...	2005 Jul	Wolfe	10800	C5	4078–8565
J210025–064146	21:00:25.03	–06:41:46.0	3.137	2005 Oct	Prochaska	1860	C5	3722–6634
...	...	...	...	2004 Oct	Wolfe	18000	C1	4439–8981
...	...	...	...	2007 Sep	Wolfe	10800	C1	3889–6924
J212329–005052	21:23:29.46	–00:50:52.9	2.262	2006 Aug	Prochaska	21600	E3	3027–5896
J212904–160249	21:29:04.90	–16:02:49.0	2.900	2007 Aug	Crighton	14400	C1	3992–8545
J212912–153841	21:29:12.19	–15:38:42.8	3.268	2004 Sept	Prochaska	7200	C1	3571–6422
J220639–181846	22:06:39.70	–18:18:46.0	2.728	2004 Sept	Prochaska	5400	C1	3304–6183
J220852–194400	22:08:52.07	–19:44:00.0	2.573	2008 Sep	Steidel	12000	C5	3022–5896
J222256–094636	22:22:56.11	–09:46:36.2	2.926	2006 Aug	Ellison	10800	C1	3214–6096
J224145+122557	22:41:45.11	+12:25:57.1	2.631	2005 Jul	Wolfe	6300	C5	4095–8551
J225409+244523	22:54:09.34	+24:45:23.4	2.328	2009 Dec	Tytler	1100	C1	3366–6198
J230301–093930	23:03:01.45	–09:39:30.7	3.455	2005 Nov	Wolfe	7200	C5	3837–8352
J231324+003444	23:13:24.45	+00:34:44.5	2.083	2009 Dec	Tytler	600	C1	3176–6038
J231543+145606	23:15:43.56	+14:56:06.3	3.390	2005 Nov	Wolfe	4508	C5	4797–7819
...	...	...	...	2004 Jun	Prochaska	4800	C5	3926–7785
J233446–090812	23:34:46.40	–09:08:12.3	3.317	2006 Dec	Wolfe	3600	C1	3761–8338
...	...	...	...	2007 Sep	Wolfe	10800	C1	4065–7120
J233823+150445	23:38:23.16	+15:04:45.2	2.420	2009 Dec	Tytler	900	C1	3367–6198
J234023–005327	23:40:23.66	–00:53:27.0	2.085	2006 Aug	Prochaska	7800	C1	3048–5896
J234352+141014	23:43:52.62	+14:10:14.6	2.896	2005 Jul	Chaffee	7200	C5	3296–6151
...	...	...	...	2005 Oct	Steidel	1800	C5	3124–5977
...	...	...	...	2006 Dec	Wolfe	7200	C1	3813–6660
J234451+343348	23:44:51.25	+34:33:48.6	3.010	2004 Oct	Wolfe	5400	C1	4306–8912
J234628+124859	23:46:28.20	+12:49:00.0	2.573	2005 Oct	Steidel	7200	C1	3120–5978
...	...	...	...	2008 Jul	Murphy	9000	C1	4476–7587
...	...	...	...	2008 Sep	Steidel	4000	C5	3022–5896
J234856–104131	23:48:56.48	–10:41:31.2	3.173	2006 Jun	Chaffee	5600	C5	3102–5980
J235050+045507	23:50:50.25	+04:55:07.8	2.593	2009 Dec	Tytler	850	C1	3368–6198

**Note.**

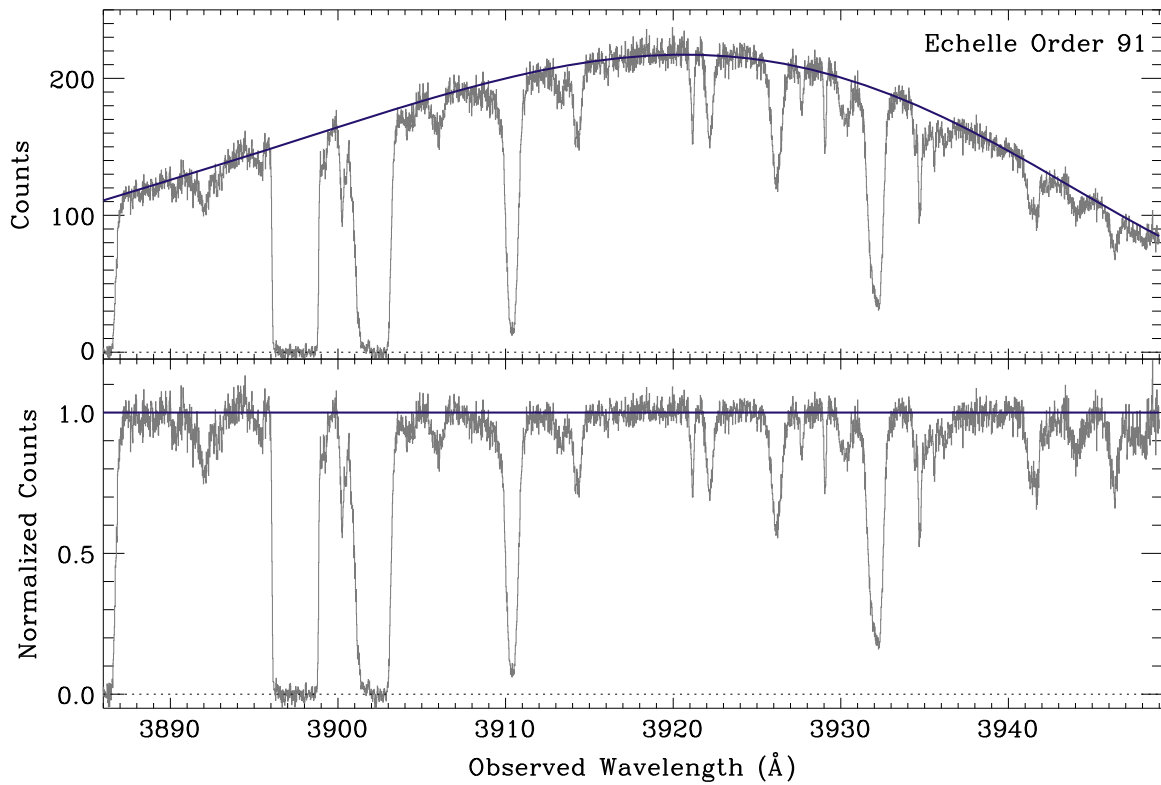
<sup>a</sup> J2000 coordinates and quasar redshifts are determined by passing the quasar coordinates from the raw data header through SIMBAD or SDSS and choosing the appropriate match.

the use of a spectrophotometric standard star, to remove the instrumental response and flux calibrate the data to aid in this process. Unfortunately, HIRES can not be reliably flux calibrated (e.g., Suzuki et al. 2005). As such, the data are fitted by hand, one echelle order at a time.

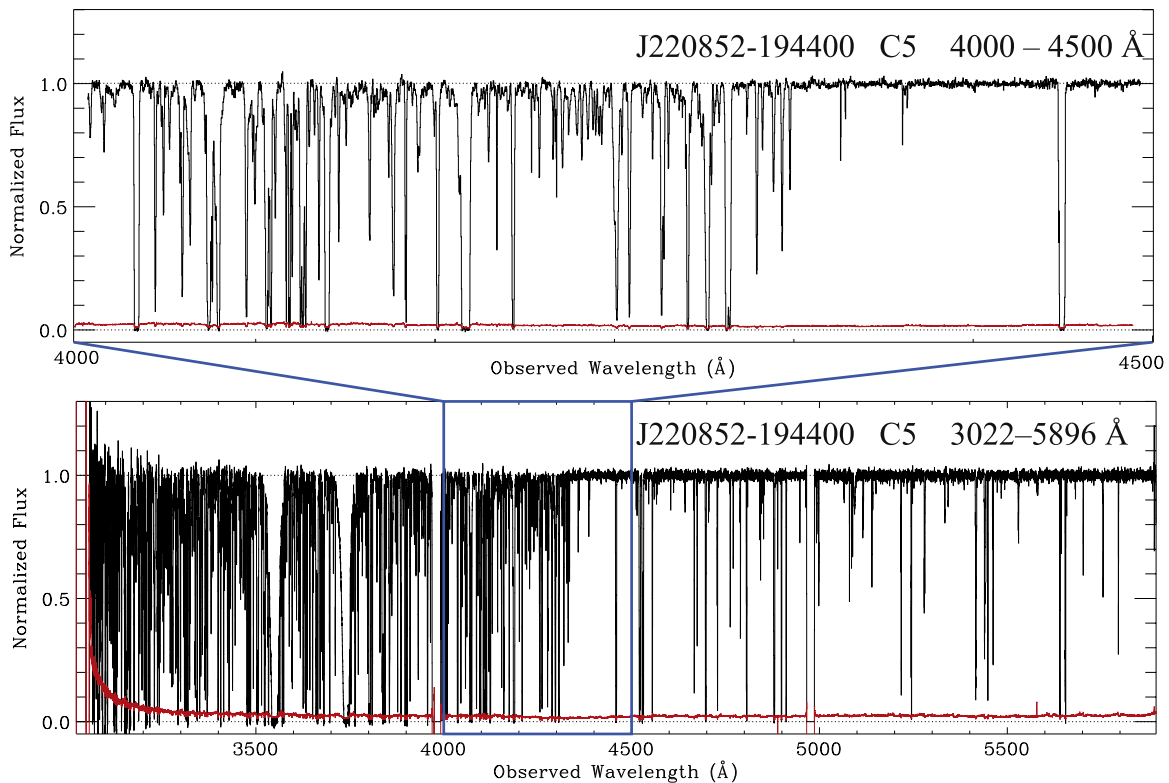
To minimize variation in continuum-fitting methodologies, a single member of the team (J.M.O) performed the continuum-fitting for the DR1 sample. The exact method of fitting varies by quasar to quasar depending on a combination of the signal to noise of the data, whether the regions to be fit are blueward or redward of the quasar's Ly $\alpha$  line (i.e., if the data covers the Ly $\alpha$  forest), the redshift of the quasar (larger

redshift quasars have significantly more Ly $\alpha$  forest absorption), the shape of the blaze function in a given order, and the occurrence of strong features such as quasar emission lines or damped Ly $\alpha$  absorbers. Nevertheless, a basic feature of all the continuum fits is that they were performed using Legendre polynomial fits to each spectral order, with the order of the Legendre polynomial ranging nominally anywhere from 4 to 12 (or more in rare cases), depending on the factors given above. In the Ly $\alpha$  forest, the continuum is anchored at points deemed by eye to be absorption free. In regions redward of the Ly $\alpha$  forest, significantly more regions of the spectrum are absorption-





**Figure 1.** Example of a continuum fit to a single HIRES echelle order (Order 91) of the quasar J220852-194400 ( $z_{\text{em}} = 2.573$ ). The upper panel shows the output of the HIRedux extraction, and the adopted continuum level. The lower panel shows the normalized echelle order.



**Figure 2.** Lower panel: example of a continuum normalized, coadded 1D spectrum from DR1 (the same quasar is shown in Figure 1). The  $1\sigma$  error array is shown in red. The quasar name, HIRES decker used, and the wavelength coverage of the data are listed in the figure. Upper panel: a zoom in of the observed wavelengths surrounding Ly $\alpha$  at  $z = z_{\text{em}}$ .

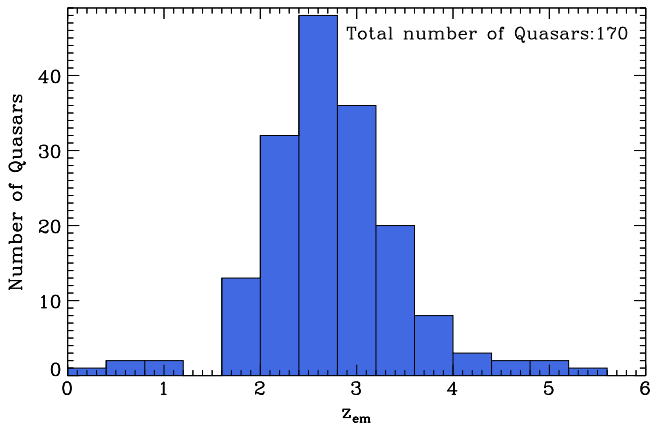


Figure 3. Redshift distribution of KODIAQ DR1 quasars.

free, and the fit is in general more reliable and objective. Given the number of spectra present in the DR1, the number of echelle orders fit is well over 8000. An example of a continuum fit to a single HIRES echelle order is shown in Figure 1.

Since the true quasar SED is unknown, especially at the resolution of HIRES, an exact quantification of errors in the continuum fit cannot be obtained. Previous attempts at estimating continuum level errors (e.g., Faucher-Giguère et al. 2008) typically find the error to be at the few percent level or less in moderate to high signal-to-noise data (e.g., signal-to-noise ratio;  $S/N \gtrsim 10$ ), with the errors increasing as either the signal to noise decreases, or the redshift of the quasar increases. As we will see in Section 3, the DR1 sample covers a wide range of both quantities. We also note that the continuum placement will incur additional uncertainties in regions within  $\sim 1000 \text{ km s}^{-1}$  of strong emission/absorption features such as damped  $\text{Ly}\alpha$  absorbers, or BAL features in the quasar spectrum itself. We recommend that users of DR1 data either refine our continuum estimate, or place their own continuum levels on the extracted data if a more precise continuum estimate is demanded in their analysis. As we show below, the DR1 data are housed at KOA in a manner such as to facilitate a re-assignment of the continuum level as users see fit. After the continuum level is assigned to each order, the data are combined into separate one-dimensional flux and error spectra, with regions of echelle order overlap in wavelength space added together weighted by signal to noise. An example of this final product is shown in Figure 2. The full HIRES header is preserved from the raw data, along with KOA- and KODIAQ-specific header cards appended to the end of the original data header.

#### 2.4. KODIAQ at the KOA

Beginning in 2015 May, data from DR1 and all subsequent data releases from KODIAQ will be available via the “Contributed Datasets” section of the KOA.<sup>13</sup> Users can search for individual quasar sightlines, or can download the entire set of one-dimensional continuum normalized final science products. Each quasar sightline contains a machine readable summary file detailing all observations of the quasar, one-dimensional flux and error files for each (coadded) observation

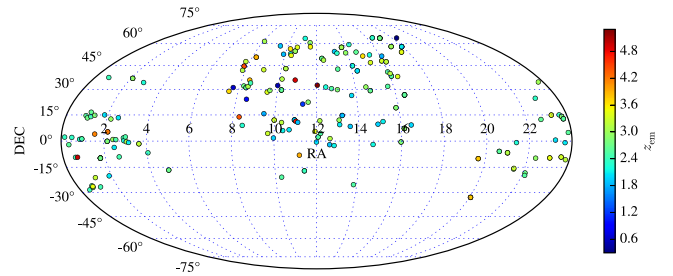


Figure 4. Aitoff sky distribution of the KODIAQ quasars in R.A. (hours) and decl. (degrees). Each quasar is represented by a circle with its color encoding its emission redshift. The large regions on the sky without data stem from the Keck I south pointing limits and the blocking on the sky of the Milky Way.

in FITS format, and a preview image. To facilitate analysis according to the users’ needs, all intermediate data reduction steps will be also available at the KOA, and can be downloaded on an observing run by observing run basis. The intermediate data are grouped by observing run since many quasars can be observed in a single HIRES setup on a given run. Each observing run has a file structure dictated by HIREDUX. In particular, users may wish to download the intermediate data products to assign their own continuum level to the individual HIRES orders.

### 3. PROPERTIES OF KODIAQ DR1

The KODIAQ DR1 comprises 170 quasars observed in the mosaic configuration of HIRES between 2004 and 2012.<sup>14</sup> The DR1 sample is summarized in Table 2. The convention adopted here is to name each quasar after their J2000 R.A./decl. as resolved by SIMBAD or SDSS. For example, the quasar with J2000 coordinates 01:08:06.41+16:35:50.0 receives the name J010806+163550. In the case of very close quasar pairs or lensed quasars, the name is appended with A,B, and so on. Many of the quasars have been observed by multiple PIs, and/or with multiple HIRES setups, such that the total number of spectra available in DR1 is 247 representing coadditions from 567 individual HIRES observations. Exposure times range from the hundreds to tens of thousands of seconds, and the combined total exposure time for DR1 is approximately 1.6 megaseconds.

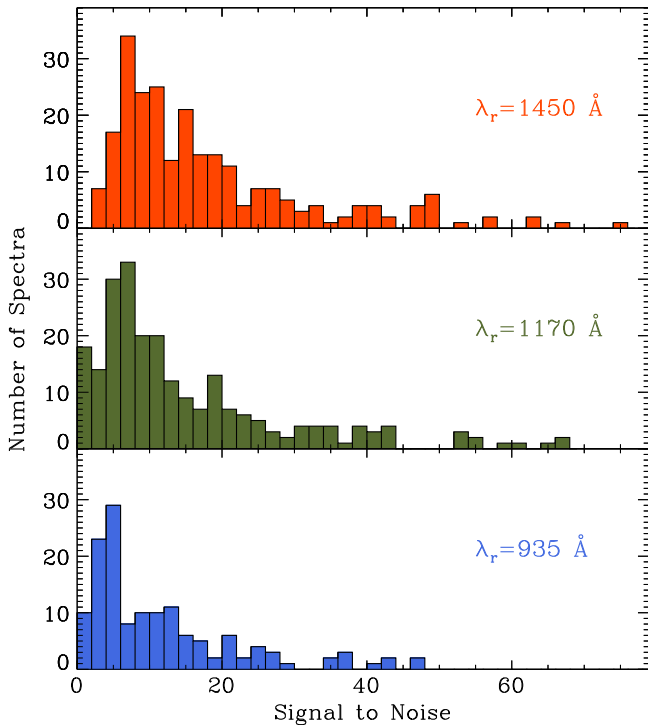
#### 3.1. General Properties

Figure 3 shows the quasar emission redshift distribution for DR1, and shows that the DR1 data span a redshift range of  $0.29 < z_{\text{em}} < 5.29$  in a roughly Gaussian distribution, with a mean value of  $z_{\text{em}} = 2.725$  and a standard deviation of 0.738. The distribution of DR1 quasars on the sky is shown in Figure 4; these primarily cover the North Galactic Pole. The data vary significantly in signal to noise. In Figure 5, we display the S/N of the KODIAQ spectra in the following way. For each of the 247 spectra, we calculate the median S/N per  $2.6 \text{ km s}^{-1}$  interval (i.e., two unbinned HIRES pixels) over a window of  $\lambda = \pm 5 \text{ \AA}$  on either side of 3 wavelength bins in the quasar rest frame (see Figure 5). These wavelengths are chosen to give a sense of the data near the quasar Lyman limit, the regions encompassing the  $\text{Ly}\alpha$  forest and a region sampling  $\text{Ly}\alpha$  forest-free wavelengths for metal line searches. Figure 5

<sup>13</sup> <https://koa.ipac.caltech.edu/Datasets>

<sup>14</sup> One quasar is a lensed system, so the number of individual sightlines is 171.



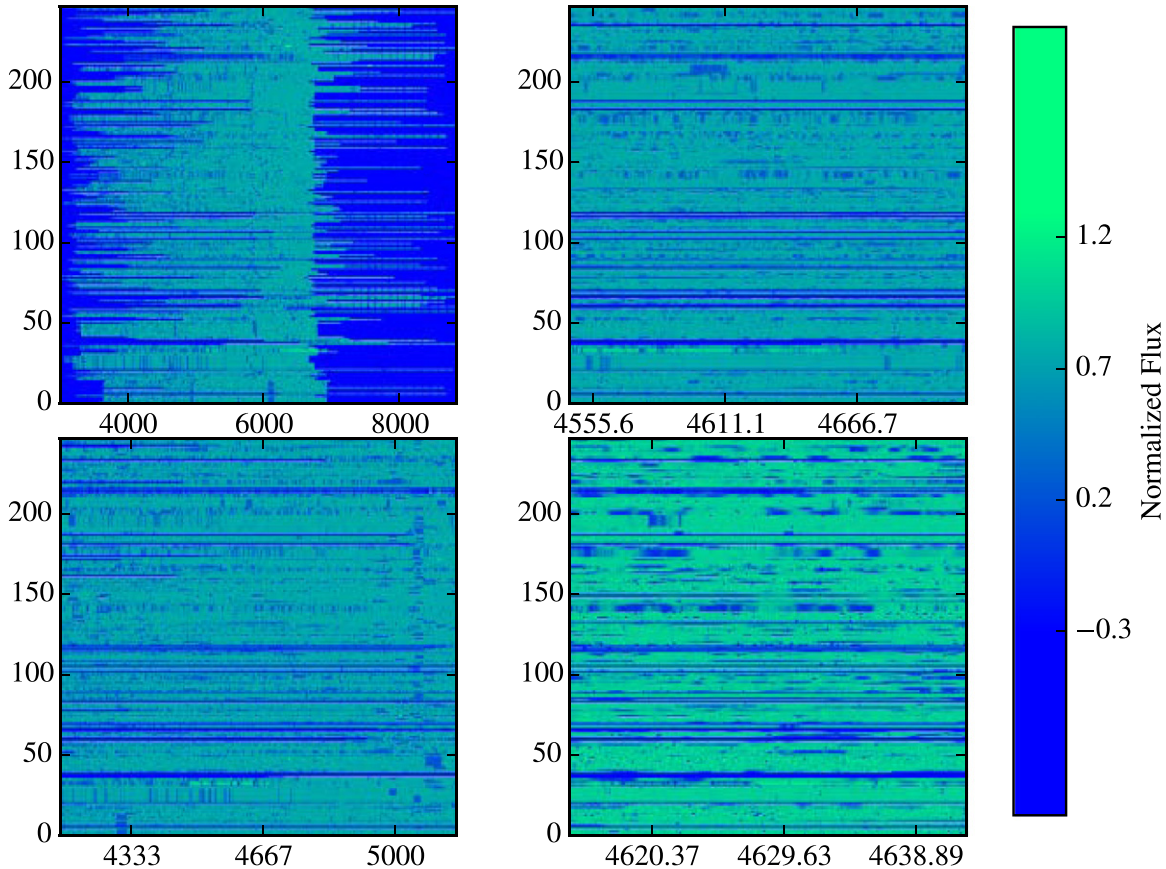


**Figure 5.** Median signal to noise per  $2.6 \text{ km s}^{-1}$  interval in a window  $\pm 5 \text{ \AA}$  about rest wavelengths  $\lambda = 935, 1170, 1450 \text{ \AA}$  for the 247 spectra in the KODIAQ sample.

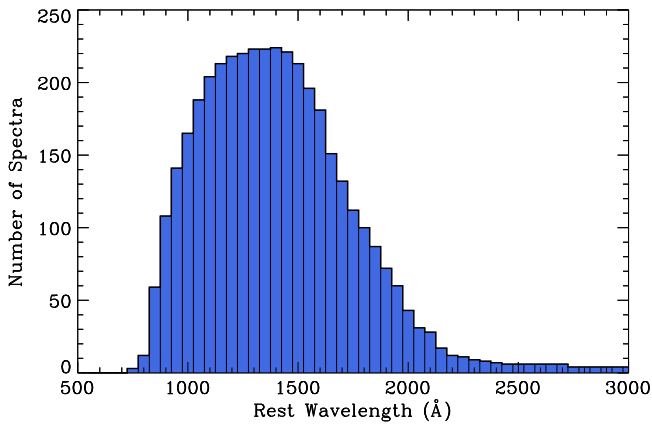
illustrates that a significant fraction of the data has  $S/N > 10$ , facilitating a wide variety of future studies. For further illustration of the data, Figure 6 shows the full spectral dataset rebinned to 500 spectral pixels ordered from top-to-bottom by quasar emission redshift. The echelle order gaps and HIRES setups are evident on large scales, as well as the large-scale influence of the  $\text{Ly}\alpha$  forest on the smaller scales.

### 3.2. Cosmological Properties

We anticipate that the data made available in the DR1 and future releases will have a multitude of uses beyond those presented in Lehner et al. (2014). To give a sense of the size of the DR1 in terms of possible cosmological applications, we present several quantities summarizing this rich dataset. In Figure 7, we display a histogram of the wavelengths covered in the 247 spectra of DR1, placed in the quasar rest frame to best display what ionic species could be searched for in the data. Figure 8 shows how many spectra in DR1 provide coverage of a number of ions common in quasar absorption line studies, namely  $\text{Ly}\alpha$ ,  $\text{O VI}$ ,  $\text{C IV}$ , and  $\text{Mg II}$ . For  $\text{Ly}\alpha$ , we restrict ourselves to sampling only the wavelengths between the quasar  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  emission lines when counting sightlines. For  $\text{C IV}$  and  $\text{Mg II}$ , we sample only available wavelengths larger than that of the quasar  $\text{Ly}\alpha$  line and less than the quasar  $\text{C IV}$  and  $\text{Mg II}$  emission lines, respectively. Finally, for  $\text{O VI}$ , we consider all available wavelengths lower than that of the quasar  $\text{O VI}$  emission line.



**Figure 6.** Spectral image of the full DR1 dataset. The data are rebinned to 500 spectral pixels. The panels show the dataset at various wavelength scales (on the  $x$ -axis) to illustrate the effects of instrument setup at the largest scales (upper-left panel) down to the level of the influence of the  $\text{Ly}\alpha$  forest (lower-right panel).



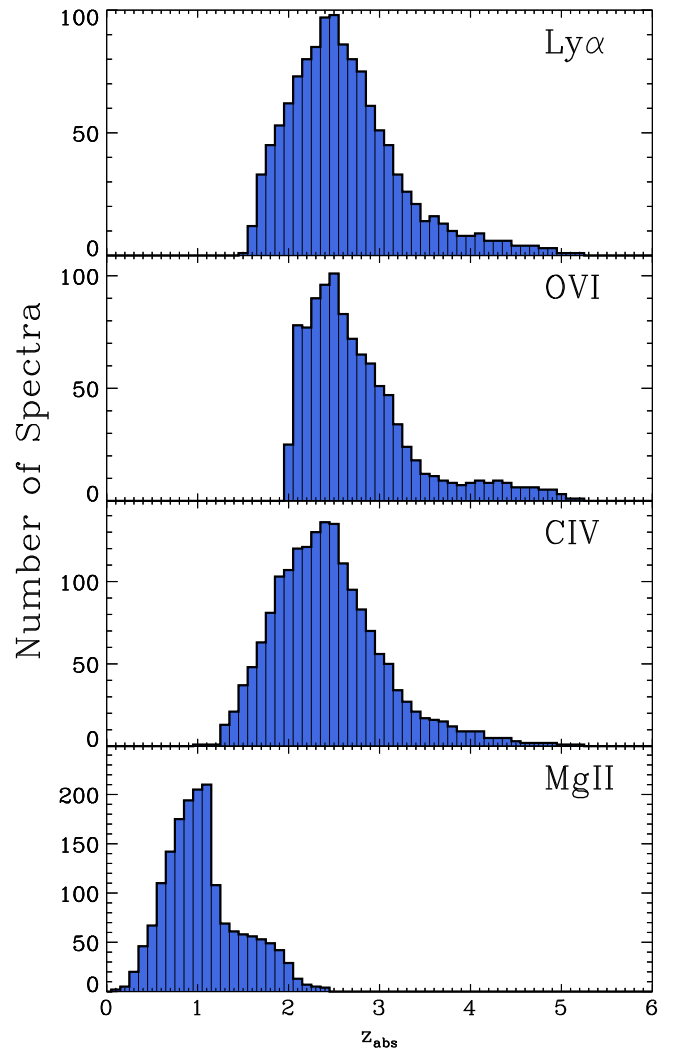
**Figure 7.** Rest frame wavelength spectral coverage of the 247 spectra in KODIAQ DR1.

We can also cast the sample in terms of the redshift sensitivity function  $g(z)$  used in studies of the incidence frequency  $\ell(z) = dN/dz$  of an absorption feature or for the column density distribution function  $f(N_x, z)$  of ion  $x$ . In Figure 9, we present two estimates of  $g(z)$ . First, for each of the 170 quasars in the sample we estimate the sensitivity to searching for Ly $\alpha$  forest absorption by determining which quasars cover the Ly $\alpha$  forest between  $3000 \text{ km s}^{-1}$  to the red of the quasar Ly $\beta$  line and  $3000 \text{ km s}^{-1}$  to the blue of the quasar Ly $\alpha$  line, so as to avoid sampling the proximity regions of the quasars themselves. We then calculate  $g(z)$  for this sample subject to four different S/N criteria, namely  $S/N > 2, 5, 10,$  and  $20$ , where the S/N is estimated by taking the median S/N (per pixel) over a window of 150 pixels. Second, we perform the same process, but for C IV absorption. For C IV, we maintain our proximity zone exclusions, (this time between the Ly $\alpha$  and C IV emission lines), but change our S/N criteria to be  $S/N > 5, 10, 20,$  and  $40$ , so as to best demonstrate the high quality of the DR1 data. Both  $g(z)$  estimates show sharp downward spikes (e.g., at  $z \sim 2.3$  for Ly $\alpha$ ). These arise due to the wavelength gaps between the CCDs in the HIRES mosaic. In the case of C IV, additional ripples are present in  $g(z)$ . These arise from the blaze sensitivity of the echelle. Figure 9 demonstrates that DR1 offers a significant dataset for any number of studies of the intergalactic and circumgalactic medium over a significant swath of cosmic time.

#### 4. SUMMARY AND FUTURE

We have presented here and made available to the public at the KOA the DR1 of the KODIAQ survey. DR1 is comprised of 247 spectra of 170 quasars obtained with HIRES between 2004 and 2012. The data vary in signal to noise and resolution, but a significant portion of the spectra are of sufficient quality as to facilitate a number of precision studies of the inter- and circum-galactic medium at  $z \sim 2.5$ .

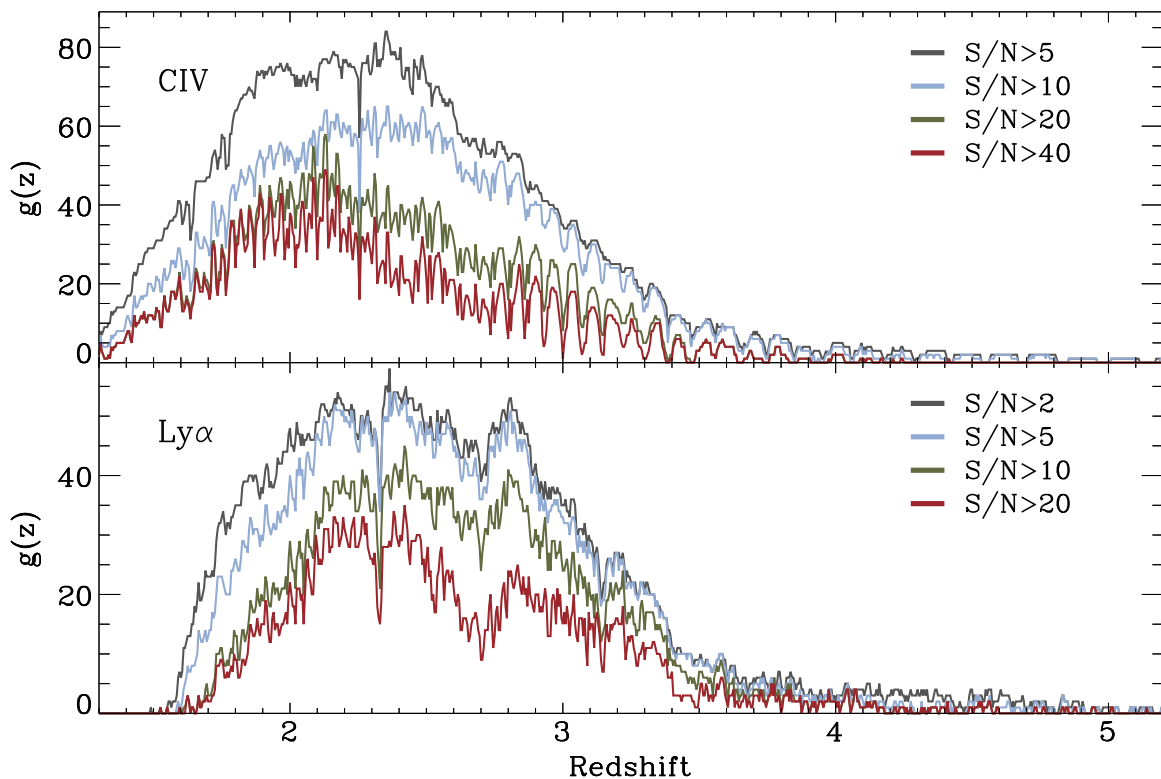
Additional data releases for KODIAQ are planned for the future. DR2 will incorporate pre-2004 spectra into the sample for another  $\sim 150$  quasars in addition to augmenting some spectra in DR1. Finally, DR3 will make available the full KODIAQ sample of all HIRES spectra in the KOA that could be processed. We anticipate that DR2 will be released in 2017, and DR3 in early 2018. All data releases will be housed at the KOA, in addition to minor updates between data releases.



**Figure 8.** Redshift coverage of specific ions in the 247 spectra of KODIAQ DR1.

When using data products from DR1, 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 the data presented in this work were obtained from the Keck Observatory Database of Ionized Absorbers toward Quasars (KODIAQ), which was funded through NASA ADAP grant NNX10AE84G,” and cite this paper and Lehner et al. (2014).

Support for this research was made by NASA through the Astrophysics Data Analysis Program (ADAP) grant NNX10AE84G. This research has made use of the 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



**Figure 9.** Redshift sensitivity function  $g(z)$  for C IV and Ly $\alpha$  absorption of the 170 quasars in DR1. The sharp downward spikes result from the effect of the mosaic detector gaps. The high frequency variations result from the effects of the echelle blaze function.

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. We thank the anonymous referee for their suggestions which improved this work. Finally, the authors recognize the major contributions by the late astronomers Wallace L. W. Sargent and Arthur M. Wolfe to this data set, and indeed to the entire discipline of quasar absorption line studies. We dedicate this work to the memory of these great scientists.

#### REFERENCES

- Bochanski, J. J., Hennawi, J. F., Simcoe, R. A., et al. 2009, *PASP*, **121**, 1409
- Bolton, J. S., Becker, G. D., Haehnelt, M. G., & Viel, M. 2014, *MNRAS*, **438**, 2499
- 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
- Faucher-Giguère, C.-A., Prochaska, J. X., Lidz, A., Hernquist, L., & Zaldarriaga, M. 2008, *ApJ*, **681**, 831
- Fumagalli, M., O’Meara, J. M., & Prochaska, J. X. 2011, *Sci*, **334**, 1245
- Lehner, N., O’Meara, J. M., Fox, A. J., et al. 2014, *ApJ*, **788**, 119
- 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., Tytler, D., Kirkman, D., et al. 2001, *ApJ*, **552**, 718
- Prochaska, J. X., Gawiser, E., Wolfe, A. M., Cooke, J., & Gelino, D. 2003, *ApJS*, **147**, 227
- Prochaska, J. X., Chen, H.-W., Bloom, J. S., et al. 2007, *ApJS*, **168**, 231
- Prochaska, J. X., Wolfe, A. M., Tytler, D., et al. 2001, *ApJS*, **137**, 21
- Rafelski, M., Neeleman, M., Fumagalli, M., Wolfe, A. M., & Prochaska, J. X. 2014, *ApJL*, **782**, L29
- Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2004, *ApJ*, **606**, 92
- Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, *ApJ*, **717**, 289
- Suzuki, N., Tytler, D., Kirkman, D., O’Meara, J. M., & Lubin, D. 2005, *ApJ*, **618**, 592
- Tytler, D., Paschos, P., Kirkman, D., Norman, M. L., & Jena, T. 2009, *MNRAS*, **393**, 723
- Viel, M., Becker, G. D., Bolton, J. S., & Haehnelt, M. G. 2013, *PhRvD*, **88**, 043502
- Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, *Proc. SPIES*, **2198**, 362
- Webb, J. K., Murphy, M. T., Flambaum, V. V., et al. 2001, *PhRvL*, **87**, 091301
- Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, *ARA&A*, **43**, 861
- Zafar, T., Popping, A., & Péroux, C. 2013, *A&A*, **556**, A140