Journal of
Astronomical Telescopes, Instruments, and Systems

AstronomicalTelescopes.SPIEDigitalLibrary.org
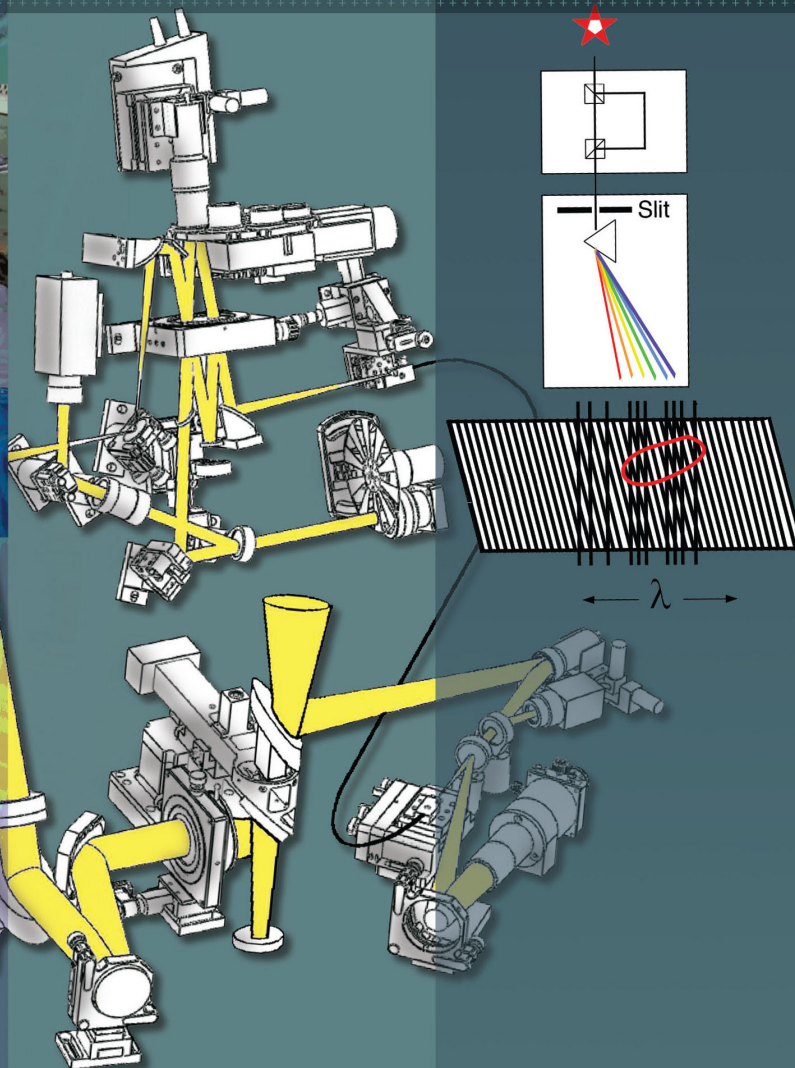
SPIE.

High-resolution broadband spectroscopy using externally dispersed interferometry at the Hale telescope: Part 1, data analysis and results

David J. Erskine,[a,*] Jerry Edelstein,[b] Edward H. Wishnow,[b] Martin Sirk,[b] Philip S. Muirhead,[c] Matthew W. Muterspaugh,[d] James P. Lloyd,[e] Yuzo Ishikawa,[b] Eliza A. McDonald,[b] William V. Shourt,[b] and Andrew M. Vanderburg[f]

[a]Lawrence Livermore National Laboratory, Mailstop L-487, 7000 East Avenue, Livermore, California 94550, United States
[b]University of California, Space Sciences Laboratory, 7 Gauss Way, Berkeley, California 94720, United States
[c]Boston University, Department of Astronomy, 725 Commonwealth Avenue, Boston, Massachusetts 02215, United States
[d]Tennessee State University, Boswell Science Hall, Nashville, Tennessee 37209, United States
[e]Cornell University, Carl Sagan Institute, Astronomy Department, 230 Space Sciences, Ithaca, New York 14853, United States
[f]Harvard Smithsonian Center for Astrophysics, 60 Garden Street, MS-10, Cambridge, Massachusetts 02138, United States

Abstract. High-resolution broadband spectroscopy at near-infrared wavelengths (950 to 2450 nm) has been performed using externally dispersed interferometry (EDI) at the Hale telescope at Mt. Palomar. Observations of stars were performed with the "TEDI" interferometer mounted within the central hole of the 200-in. primary mirror in series with the comounted TripleSpec near-infrared echelle spectrograph. These are the first multidelay EDI demonstrations on starlight, as earlier measurements used a single delay or laboratory sources. We demonstrate very high (10×) resolution boost, from original 2700 to 27,000 with current set of delays (up to 3 cm), well beyond the classical limits enforced by the slit width and detector pixel Nyquist limit. Significantly, the EDI used with multiple delays rather than a single delay as used previously yields an order of magnitude or more improvement in the stability against native spectrograph point spread function (PSF) drifts along the dispersion direction. We observe a dramatic (20×) reduction in sensitivity to PSF shift using our standard processing. A recently realized method of further reducing the PSF shift sensitivity to zero is described theoretically and demonstrated in a simple simulation which produces a 350× times reduction. We demonstrate superb rejection of fixed pattern noise due to bad detector pixels—EDI only responds to changes in pixel intensity synchronous to applied dithering. This part 1 describes data analysis, results, and instrument noise. A section on theoretical photon limited sensitivity is in a companion paper, part 2. © *2016 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JATIS.2.2.025004]

Keywords: dispersed interferometry; Doppler radial velocimetry; high-resolution spectroscopy.

Paper 15076P received Nov. 8, 2015; accepted for publication Apr. 27, 2016; published online May 27, 2016.

1 Introduction

1.1 Wide Bandwidth Desired

Wide bandwidth high-resolution spectroscopy is one of the most valuable diagnostic tools for astronomy, from Doppler radial velocimetry in the search for exoplanets to studying the characteristics of stars through the size, breadth, and position of spectral features. A wide bandwidth is needed in radial velocimetry to collect as many photons as possible to reduce Doppler velocity noise. Wide bandwidth is desired for analysis of species population because the characteristic wavelengths are often distributed over a wide band.

However, it is difficult to achieve both wide bandwidth and high spectral resolution for a dispersive spectrograph, since this requires not only a large and expensive instrument but also a detector with many pixels. This is especially so for infrared wavelengths since the spectrograph size scales with resolving power times wavelength.

1.2 Stability Against Point Spread Function Drifts Desired

Second, a serious challenge for high-resolution spectrographs is the unwanted shifting of the focal spot (point spread function, PSF) along the dispersion direction, such as due to air convention or thermomechanical drifts. While use of a calibrant spectrum measured in parallel or comingled with the stellar spectrum can remove some lower order components of the drift, the higher order components corresponding to gaps between strong features in the calibrant spectrum can remain, producing a wavenumber error. These environmental insults are frequently the limitation, rather than photon noise, to achieving the highest desired Doppler precision, or sufficiently accurate spectra to use as a template for subsequent Doppler calculations.

1.3 Externally Dispersed Interferometry

A technique called externally dispersed interferometry[1] (EDI) can dramatically reduce the necessary size of spectrographs

for Doppler radial velocimetry[1–9] and high-resolution spectroscopy.[10–15] Other workers have adopted the EDI method from our laboratories and demonstrated a Doppler planet detection.[5] We have demonstrated EDI in both applications in the near-infrared at the Hale 5-m telescope at Mt. Palomar Observatory (Fig. 1) with a version called TEDI, (the T denotes the TripleSpec[16] spectrograph to which it was coupled). An earlier report[2] describes the Doppler velocimetry, and this report describes the spectroscopy.

This is the first time multiple delays have been used on starlight to recreate a high-resolution spectrum. Previous EDI use on stellar spectroscopy used a single delay.[11] Multiple delays had been used previously[12] only on laboratory sources.

The TEDI project was designed primarily to test Doppler measurements, and the selection of interferometer glass etalons and observing schedule were prioritized for this purpose. These were not optimal for testing high-resolution spectroscopy (having a gap in coverage of delay space when contiguous coverage is needed to avoid ringing in the lineshape). Nevertheless, the preliminary results for high-resolution spectroscopy using multiple delays are a resounding success in two important areas.

1.4 Resolution, Robustness Demonstrated

We have demonstrated on starlight that the EDI technique can produce both high resolution, a factor of 4× to 10× beyond the native spectrograph resolution of ∼2700, and have extremely wide simultaneous bandwidth, limited only by the bandwidth of the native spectrograph, in this case (0.95 to 2.45 μm).

Second, the native spectrograph in our study suffered from extremely large and irregular PSF drifts that would normally

have precluded the high resolution we obtained, even if by some other means the slit width and focal spot had been reduced by several times and the detector pixel density increased. We realize now that a great advantage of the EDI technique, especially when used with multiple delays rather than a single delay as used previously, is an order of magnitude or more improvement in the stability against PSF drifts.

We observe an approximate 20-fold decrease in the translational reaction of a ThAr lamp line to a translational insult along the dispersion direction. This robustness to PSF errors is explained theoretically and demonstrated with measurements. A recently realized method of further reducing the PSF shift sensitivity to zero is described theoretically and demonstrated in a simple simulation which produces a 350× times reduction.

Third, we demonstrate EDI's dramatic robustness to fixed pattern (FP) noise such as from bad detector pixels. These pollute the native spectrum but not the EDI-derived spectrum measured simultaneously.

The TEDI data set is instrument limited rather than photon limited. A theoretical description of EDI photon-limited sensitivity compared to other spectroscopy techniques is presented in a companion paper.

2 How Externally Dispersed Interferometry Works

The EDI technique uses a field-widened Michelson interferometer (Fig. 1) crossed with (in series with) a dispersive spectrograph to heterodyne unresolvable narrow features to a detectable low frequency moiré pattern. Detailed instrument description is in Ref. 2 and instrument theory for a single delay spectroscopy is
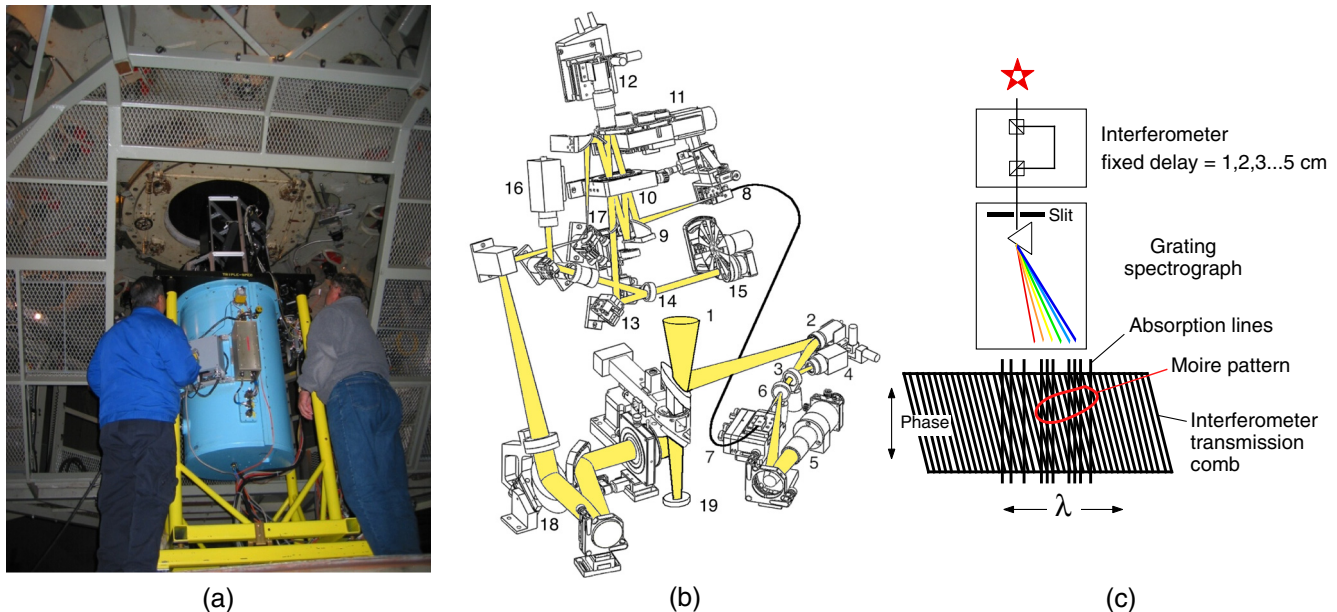

(a) (b) (c)

Fig. 1 (a) The T-EDI interferometer (black and silver) sits atop TripleSpec (native) spectrograph (blue cryogenic cylinder) which bolts to Cassegrain output (39 in. diameter cavity) of the Mt. Palomar 200 in. mirror. (b) TEDI 2.0[1] intercepts starlight beam (1) heading for TripleSpec and uses a dichroic split input tracking system (3) to mixing in ThAr lamp (5) to feed a symmetric, collimated interferometry cavity (beamsplitter 10). One interferometer output (13) is dichroic split (14) to a sensitive chopped IR diode (15) for flux maximization and a fringe tracking camera (16) for cavity nulling. The complementary interferometer output (17) is open-path relayed (18, 19) to the TripleSpec slit. (c) Externally dispersed interferometer (EDI) scheme—a fixed delay interferometer is crossed with a disperser. The sinusoidal transmission of the interferometer creates moiré patterns whose phase encodes the Doppler velocity and whose shape encodes the spectrum's shape. (Schematic graphic from Ref. 1 reproduced with permission © The Astronomical Society of the Pacific. All rights reserved.)

in Ref. [11]. The heterodyning effectively shifts the native spectrograph sensitivity peak psf_{conv} to higher frequencies (Fig. [2])

$$psf_{conv} = psf_0(\rho), \tag{1}$$

$$psf_{edi} = \frac{1}{2}\gamma psf_0(\rho - \tau), \tag{2}$$

by the amount of the interferometer delay τ, to become the fringing sensitivity peak (psf_{edi}). The frequencies ρ are the Fourier space variable in units of cm delay, complementary to wavenumber ν in cm^{-1}. The interferometer visibility γ is assumed unity.

Early EDIs used linear spectrographs in which the crossing dimension was orthogonal to the dispersion, where interferometer phase varied along the slit length. In more recent EDIs using an echelle spectrograph (having the important advantage of extremely wide bandwidth), the spacing between orders is limited and the crossing dimension is time. A series of phase-stepped exposures are taken with a sequence of one to a few fixed delay values, which map out delay-space contiguously up to a maximum value τ_{max}, currently 3 cm, but which could easily be higher.

After Fourier processing to shift the frequencies to their original value, a spectrum with a resolution 4 to 10 times higher than the existing grating spectrograph has been recovered. Higher boost factors are possible with purchase of more glass etalons to allow a larger contiguous delay range. The EDI resolving power is not determined by the focal spot size, but the maximum delay τ_{max} relative to the wavelength, $Res \sim \tau_{max}/\lambda$, and thus can greatly exceed the ceiling of the conventional spectrograph. In contrast with a dispersive spectrograph, it is now independent of the slit width, focal blur, pixel Nyquist limitations, and grating dispersion.[10,12,13]

2.1 Externally Dispersed Interferometry Has Corrugated Point Spread Function

The EDI PSF[10] (instrument response) is a corrugated peak or wavelet (bold curve of Fig. [3]), having the envelope set by the native spectrograph (dashed curve), multiplied by a sinusoid set by the interferometer delay chosen by the user. The key point is that now the PSF slope can be much higher than the low slope of envelope (native spectrograph), being proportional to the interferometer delay. It is easy to implement the 1- to
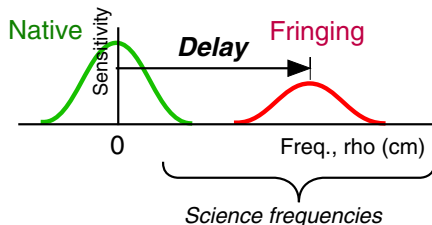

Fig. 2 Heterodyning shifts the native spectrograph sensitivity peak from zero to a higher frequency where science frequencies typically reside, by amount equal to the interferometer delay τ, and an amplitude 50%. Frequency in units of features per wavenumber (cm^{-1}), conveniently has units of delay, units of cm.
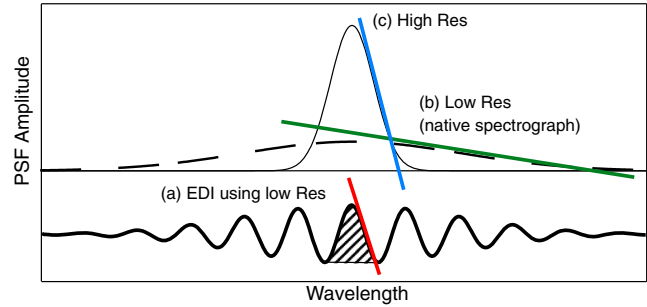

Fig. 3 The EDI PSF (instrument response) is a corrugated peak (bold curve), which has a much higher slope (and thus higher Doppler sensitivity) than the low resolution native spectrograph alone (dashes). The slope of the sinusoid is proportional to the interferometer delay. For use in high-resolution spectroscopy, the subject of this report, the central area (cross-hatched) is made unambiguous by combining several wavelets of different periodicities (delays), as shown in Figs. [4] and [5].

5-cm delay needed to match the slope of a high-resolution spectrograph.

The PSF slope is now much more mathematically controlled, largely independent of the envelope shape. The interferometer has only three degrees of freedom (phase, period, and intensity offset), rather than the many degrees of freedom typical of a grating PSF (contributed by the aggregate of tens or hundreds of grating grooves). Second, since the interferometer component is small ($\sim$50 cm), it is much easier to put in a stable environment than a stand-alone disperser, which for a 50,000 resolution can be several meters in size.

2.2 Combining Several Delays Makes a Peaked Point Spread Function

For use in high-resolution spectroscopy, the subject of this report, the many redundant peaks in the corrugated PSF for a single delay would be a problem. The solution is to use several delays. The central area (cross-hatched) is made unambiguous by combining several wavelets of different periodicities (delays), as demonstrated in simulation in Fig. [4], and in practice in Fig. [5]. The latter shows measurements of two ThAr lamp lines resolved by the EDI technique, but which otherwise are unresolved by the native spectrograph.

3 Data Analysis

3.1 Data Analysis Overview

The data analysis pipeline is described in several sections:

- (Stage 1) Order extraction [converts the two-dimensional (2-D) echelle data into a manifold of one-dimensional (1-D) spectra],
- (Stage 2) Phase-stepping arithmetic (separates fringing and nonfringing components),
- (Stage 3) Rebinning and regularization (linearizes horizontal axis to wavenumber and compensates for etalon glass dispersion),
- (Stage 4) Spectral reconstruction (shifts data in Fourier space, sums over etalons, and equalizes to final resolution).
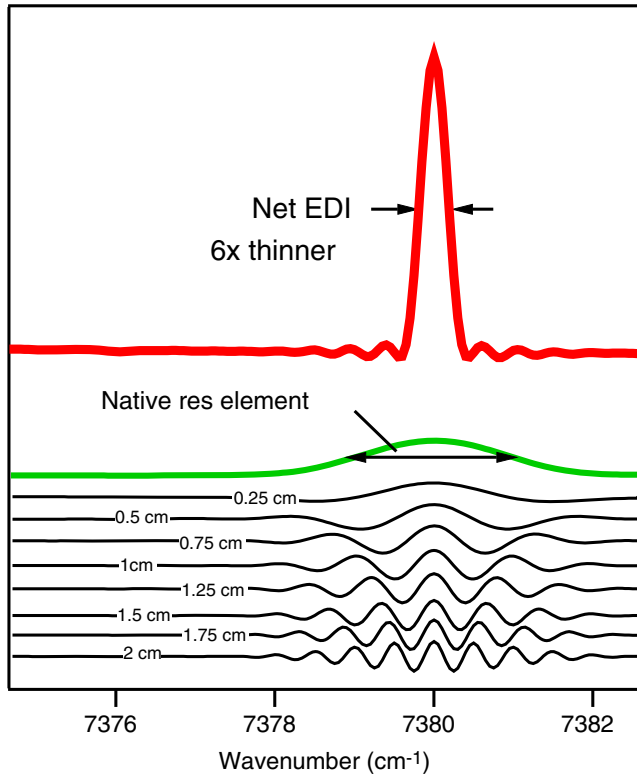
Fig. 4 Simulated summation of several simulated wavelets, weighted in Gaussian equalization scheme, produces a single Gaussian PSF peak with minimal ringing (bold red) and thinner than the native spectrograph (green curve) by factors of several.

4 Stage 1: Order Extraction

4.1 Two Stripes (A, B) for Each Echelle Order

After optional flat fielding to remove pixel-to-pixel gain variations, the 2-D data from echelle spectrograph detector, which are an FITS file, are converted to 16 1-D spectra (4 fibers × 4 orders) by collapsing vertically across each of four stripes for each order. There are four useful orders, enumerated A to D starting from the top (high Y pixel), and the fifth (E) is weak and mostly redundant in range with the D. These have wavenumber (wn or ν) ranges of A: 4100 to 5300 cm^{-1}, B: 5400 to 7100, C: 6700 to 8800, D: 8100 to 10,600, and E: 9400 to 10,700. These are wavelength ranges of A: 2.44 to 1.89 μm, B: 1.85 to 1.41, C: 1.49 to 1.14, D: 1.23 to 0.94, and E: 1.06 to 0.93.

In the first version of the TEDI apparatus,[6] both complementary interferometer outputs were sent through the spectrograph and detected. However, that interferometer design suffered from noncommon path errors. We rebuilt it in 2010 (version 2) using collimated beams, which eliminated the path errors. We also decided to use one of the two interferometer outputs to monitor the flux for maximizing starlight alignment at the input side (beam path described in Ref. 2). Although this prevents summing the two complementary outputs to create a nonfringing ordinary spectrum for each exposure, the latter is created by summing the 10 phase stepped exposures in the phase-stepping math process (Sec. 5) to create one nonfringing spectrum per set of exposures. The increase in flux due to better alignment of the starlight into the input fiber made up for the loss of half the outputs.

Hence we recorded four stripes per order, which we denoted sky A, fiber A, fiber B, and sky B. But since the companion sky stripes (which sensed blackbody background) were soon optionally subtracted from the data-carrying fiber stripes, we effectively had two stripes per order. (For the bluer orders, we could omit the subtraction.) The fibers A and B could contain either ThAr spectral lamp or starlight combined with ThAr light. We explored several data taking configurations. Typically, one fiber contained ThAr-only and the other ThAr + star. Often these roles were alternated during the night.

4.2 Fringe Comb Visible for Small Delay

Figure 6 shows a portion of an FITS echelle spectrograph output around the C-order, and Fig. 7 shows three closeups showing the effect of interferometer delay stepping. In Figs. 6 and 7, note the periodic fringes imprinted on the stellar spectrum by interferometer acting on the continuum portion of the input spectrum. A small interferometer delay (∼0.1 cm, "E1") was used here, which produces fringes wide enough to be resolvable by the native spectrograph (res ∼2700). The sinusoidal fringes are not required to be resolvable, and for most of the higher delays (>0.3 cm), are not. The heterodyning occurring here creates moiré patterns that manifest high frequency information beaten down to lower, resolvable frequencies.

Figure 7 is a closeup in C-order of three exposures of the same set as Fig. 6 but having different phase. The top fiber (A) has only ThAr spectral lamp, and the bottom fiber (B) is a sum of stellar and ThAr. The change in fringe phase between the two exposures can be seen in the different intensities of the ThAr lines, comparing top to bottom panel, and in the apparent sideways shift of the periodic comb in fiber B stripe.

4.3 Summing Across Stripes

For each of the four stripes of an order, the counts are summed across 15 vertical pixels to form a set of 1-D spectra outputted from stage 1. The pairs of white and red curves in the figures demark the 15-pixel vertical width of the stripes.

The stripe shape does not change significantly from night to night, and it merely needs to be offset vertically. The vertical summing does not correct for the effect of the slight slant in the spectrograph slit across the order, but the blurring this creates is insignificant since most of the counts are concentrated in 3 or 4 vertical pixels. However, the effect of the slant between the A and B stripes is more significant. Our routines compensate for this slight wavenumber shift between the two fibers in the pixel to wavenumber mapping done in stage 3, by having individual calibrations for each fiber. (However, our sky subtraction is currently unshifted.)

4.4 Detector Artifact Cleaning and Flatfielding

An optional first step, before order extraction, is to run a routine that removes electronic ghosting artifacts specific to this detector. This routine was supplied by TripleSpec spectrograph group at Cornell.[16] Then an optional flatfielding exposure divides out some FP noise due to pixel gain variation. Flatfielding is not critical for the fringing component since a built-in flatfielding-like process occurs already during phase-stepping mathematics that rejects stationary patterns. However, flatfielding benefits the ordinary spectrum.
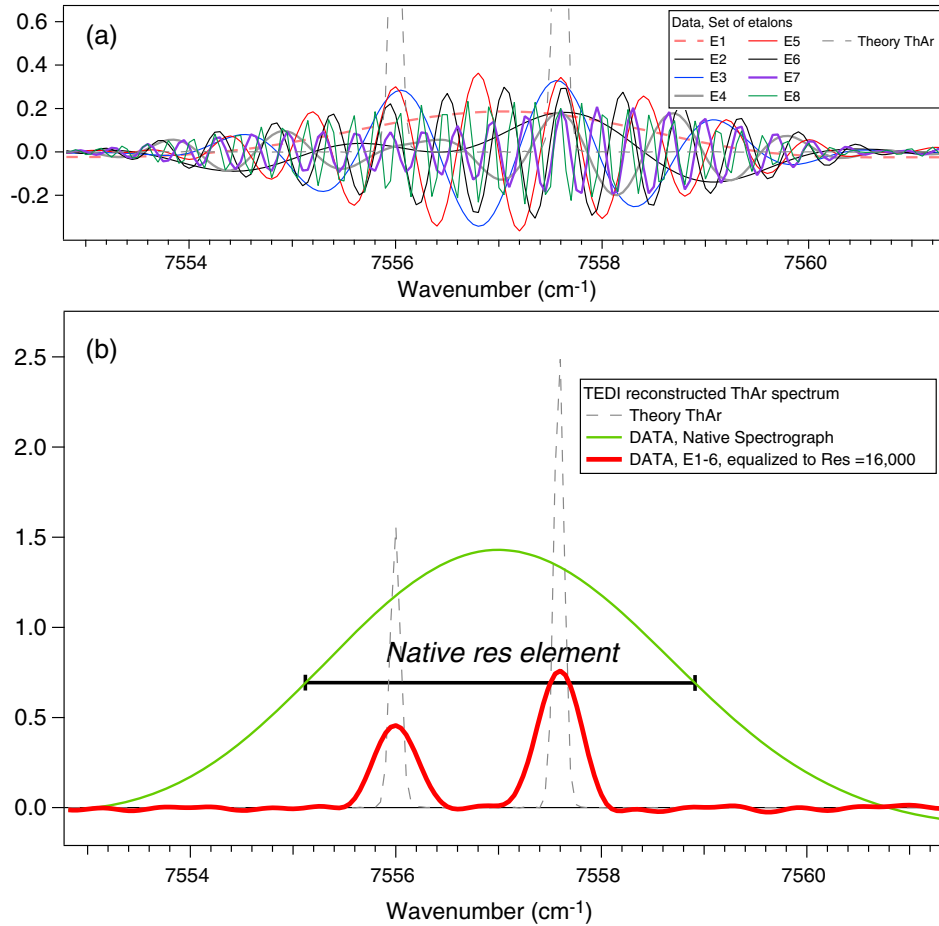
Fig. 5 (a) Reconstruction of an otherwise unresolved ThAr doublet (7556, 7557.6 cm^{-1}) from many wavelets measured at multiple delays. (b) The native spectrograph (green curve) cannot resolve the doublet. The EDI reconstructed spectrum (red curve), which is sum of wavelets, then equalized to RES = 16,000, easily resolves the doublet.
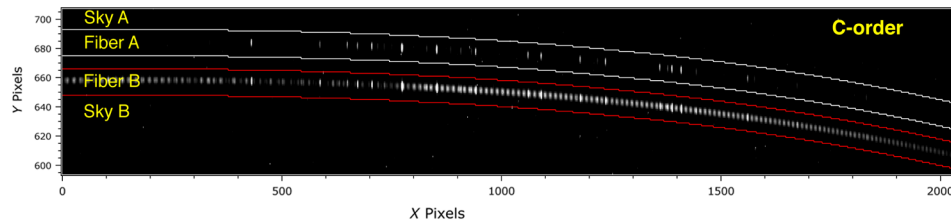

Fig. 6 Example C-order of the 2-D echelle fringing spectrum covering 6700 to 8800 cm^{-1} (left to right) over 2048 pixels. The three other orders (A, B, D) spanning 4100 to 11,000 cm^{-1} (0.9 to 2.45 μm) are shown in Fig. 2 of Ref. 2. Fiber A (white curves) is ThAr lamp alone, fiber B (red curves) is Star + ThAr. Note the sinusoidal comb from the continuum. Data from 2010Sep19, #0756, star HD962, etalon E1 ($\sim$0.1 cm).

4.5 Stage 1: Output Is Two-Dimensional Array Stack Versus Delay and ν

For each of the stripes, the 10 sequentially recorded phase-stepped 1-D spectra $B_n(\nu)$ are stacked into an array versus exposure number along Y-axis. Since exposure number is essentially delay, which is related to phase, this forms a pseudomultiphase format (Fig. 8). These arrays of 1-D uniphase data are thus in the same format as 2-D multiphase data of earlier EDIs (such as Fig. 5 of Ref. 1), where the phase spatially varies along slit length in each exposure. This is a visually intuitive and algorithmically convenient way of presenting the phase set data, and routines written earlier for multiphase data can be applied.

Echelle spectrographs require uniphase mode because they usually do not have enough vertical pixels across an order to support varying phase spatially. Instead, the phase is varied temporally via interferometer mirror movement in 10 exposures. But there are some advantages to the uniphase mode that makes it attractive in any case. In uniphase mode, the same pixel is used repetitively, so FP noises are easily suppressed (because each pixel has a constant gain). Second, it allows
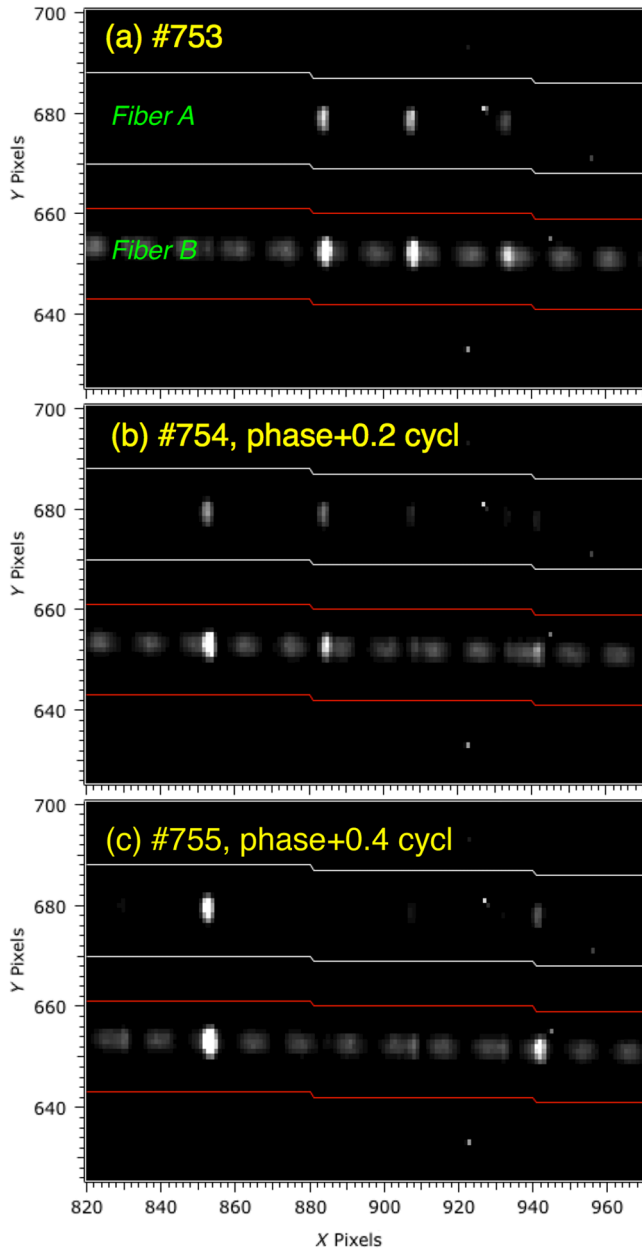
Fig. 7 Three closeups (a), (b), and (c) of the echelle C-order for ThAr lamp + star HD962, as interferometer delay changes for the E1 etalon (~0.1 cm). Wavenumber is ~7525 cm^{-1} or 1.33 μm, so the 0.25 μm commanded delay change creates a 0.19 cycle phase change. Horizontal scaling ~1 cm^{-1} per pixel. Each feature's intensity varies sinusoidally with interferometer phase, but having various relative phases and magnitudes to neighboring features.

using the spatial dimension to splay across the target, for imaging spectroscopy.

Comparing the two different delays shown in Figs. 8(a) and 8(b) note that the pattern of phase shifts for the ThAr lines is different. These constitute a "fingerprint" that reveals during analysis the precise delay that was used for each exposure, so we do not need to rely on knowledge of the commanded interferometer mirror position.

4.6 Cosmic Ray Removal

For EDI, it is more reliable to do the cosmic ray (CR) removal when the data are in pseudomultiphase format, rather than

earlier when individual exposures are in 2-D format. (In the latter, a bright ThAr line can be mistaken as a CR by a median filter.) We fit a sinusoid along each wavenumber column versus the phase. Unusually large residuals are deemed CRs and their values replaced by the sinusoidal fit value.

4.7 Wavenumber Map

A pixel to wavenumber map is made from the ThAr line positions. This is used in stage 2 to convert from delay step to phase step through $\Delta\phi = \Delta\tau/\lambda = \nu\Delta\tau/10^4$, where phase ϕ is in units of cycles or fringes, λ is wavelength in microns, τ the delay in microns, and ν is wavenumber in cm^{-1}.

5 Stage 2: Phase-Stepping Math to Separate Fringing from Nonfringing

5.1 Brief Theory

Reference 11 describes single delay EDI theory regarding spectroscopy. Fringing $\mathbf{W}$ and nonfringing B_{ord} components can be computed from a series of phase stepped exposures B_n of regular phase ϕ_n (evenly distributed around the phase circle) by

$$\mathbf{W}(\nu) = \frac{1}{N}\sum B_n e^{-i2\pi\phi_n}, \tag{3}$$

$$B_{\text{ord}}(\nu) = \frac{1}{N}\sum B_n. \tag{4}$$

For illustrative purposes, for four exposures every 90 deg

$$\mathbf{W}(\nu) = \frac{1}{4}\left[(B_0 - B_{180}) + i(B_{90} - B_{270})\right], \tag{5}$$

$$B_{\text{ord}}(\nu) = \frac{1}{4}(B_0 + B_{180} + B_{90} + B_{270}). \tag{6}$$

The goal of this section below is to solve this for the general case of irregular and unknown ϕ_n.

5.2 Overview

We now describe the phase-stepping algorithm we use to separate fringing ($\mathbf{W}$) and nonfringing (B_{ord}) spectra from the spectral array data (B_n) of Fig. 8. This algorithm can handle unknown and irregular steps and intensities, and therefore it is useful to other interferometry applications generally (the first author uses it fruitfully for analyzing Doppler velocimetry interferograms having irregular visibilities and intensities in shock physics[17]). This distinguishes our algorithm from most others in the literature,[18] which have trouble handling unknown and irregular steps.

5.2.1 Alternate solving X and Y dependencies

The problem we face is to solve both X and Y behaviors. The X axis is wavenumber ν, and the Y axis is delay step position $\delta\tau_n$. We alternate between analyzing data along a column using provisional knowledge of delay steps to produce "template" spectra ($\mathbf{W}$ and B_{ord}), and then analyzing along rows using provisional knowledge of the template spectra to produce detailed delay step values (flowchart Fig. 9).
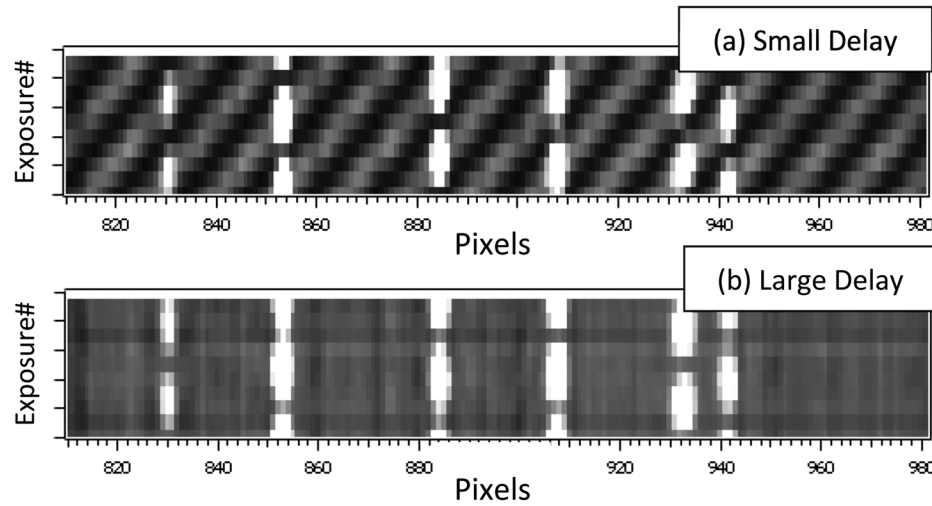
Fig. 8 The 10 sequentially recorded phase stepped 1-D spectra are stacked versus exposure # (~phase) along the vertical. (a) Small delay (E1, 0.1 cm), the comb period is wide enough to be resolved. (b) Larger delay (E3, 0.7 cm), the comb is unresolved. ThAr phase pattern constitutes a "fingerprint" that reveals the precise interferometer delay. Un-normalized intensity fluctuations due to atmospheric seeing cause horizontal artifacts in (b). Horizontal scaling ~1 cm^{-1} per pixel.

The observed TEDI delay steps [Fig. 10(b)] are reasonably uniform so the initial calculation of the provisional templates by assuming constant 0.25 μm steps produces such an accurate result for W and B that the whole process converges rapidly in only 1 to 3 iterations, using only a single order. (A much tighter constraint on the delay step would result with the increased bandwidth of multiorder processing in parallel, but this has yet to be implemented.)

Figures 10 and 11 show example results for $B_{ord}(\nu)$ and $W(\nu)$ for star *11UMi plus weak ThAr lamp, for delay E4 (~1 cm). The magnitude of the whirl is typically just a few percent of the ordinary nonfringing spectrum for the higher delays

where the sinusoidal comb is unresolved. For the lowest delay (E1, ~0.1 cm) where the sinusoidal comb is fully resolved the magnitude can approach 100%.

Pathologies in the iterative process can be detected by monitoring the plot of $\delta\tau_n$ in Fig. 10(b). Since the typical TEDI delay step is fairly uniform, if the observed $\delta\tau_n$ is erratic this indicates a problem in the data. Often this is due to an uncorrected CR.

5.2.2 *Phase step irregularity from wide bandwidth*

The phase step $\delta\phi$ can be irregular because even with a regular delay step $\delta\tau$, the phase varies across the wide band as
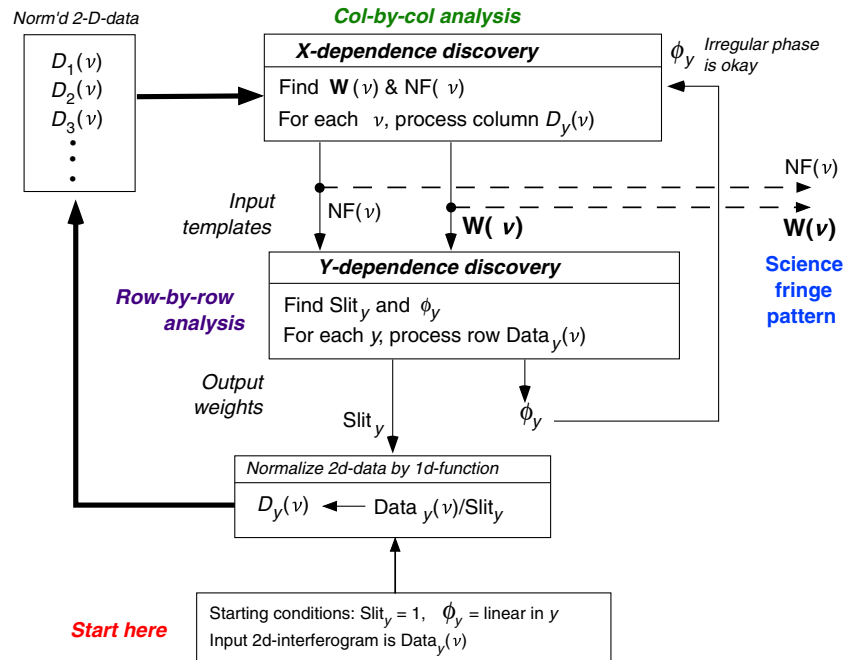

Fig. 9 Flowchart of iterative strategy for finding phases and amplitudes of each row (phase step, Y) of the interferogram, while also finding the ordinary spectrum (nonfringing intensity NF) and the fringing spectrum value $W(\nu)$ for each column (each X or ν). Slit$_y$ is the phase (or channel) dependent intensity variations, due to atmospheric seeing, which we normalize out.
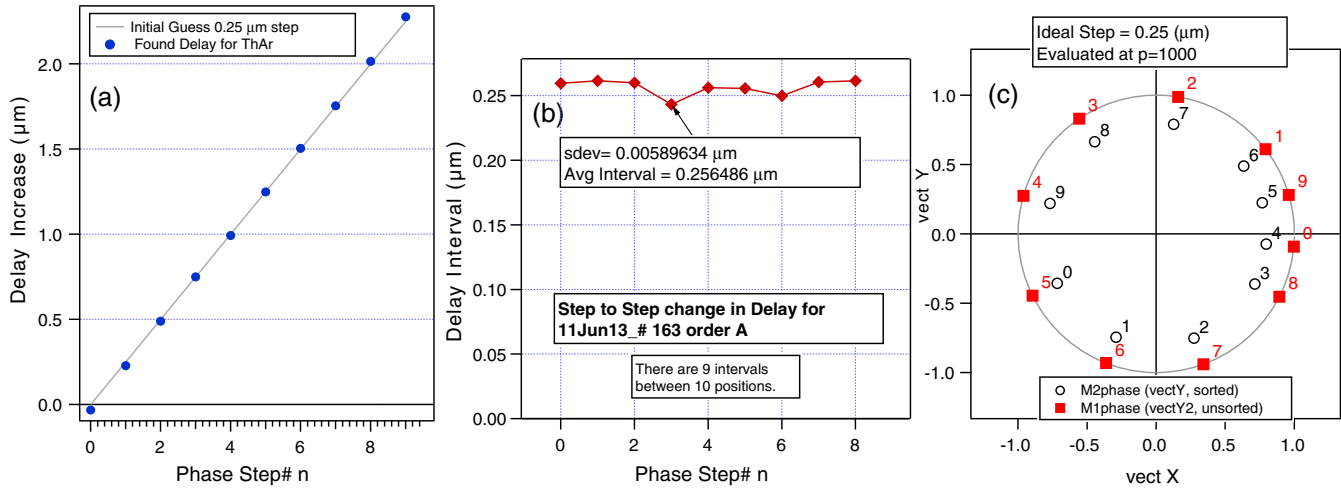
Fig. 10 Example diagnostics during phase-stepping math, etalon E4 (∼1 cm) A-order data. (a) Delay values (blue dots) $\delta\tau$ found from ThAr data. Gray line is initial guess having 0.25 μm steps. (b) Interval in delay $\Delta\tau_n = \delta\tau_{n+1} - \delta\tau_n$. Deviations typically ∼0.01 μm. Our algorithm does not require uniform steps, but larger deviations flag presence of CRs or other data failure. (c) Phase at $x = 1000$, showing that after sorting (solid squares to open circles), even a constant step in delay causes an irregularly spaced phase step, due to wrapping around circle.
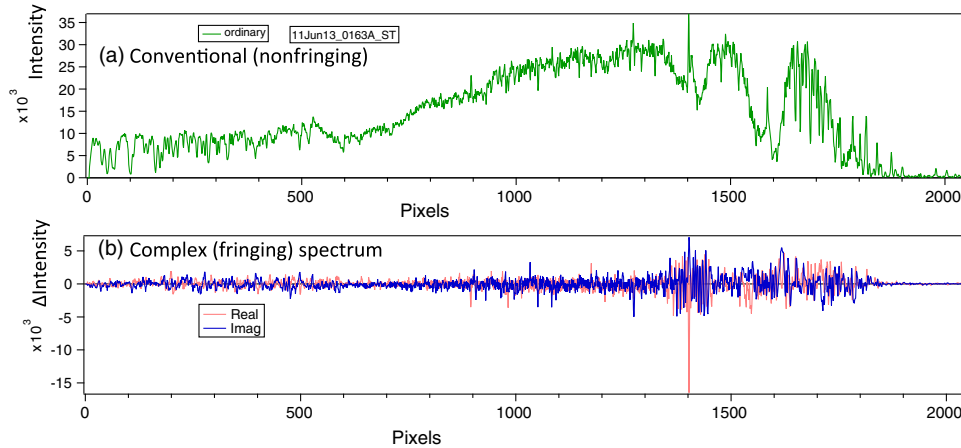

Fig. 11 (a) An ordinary (nonfringing) spectrum and (b) complex spectrum (fringing) called a "whirl" are both produced from phase-stepping math for each multiphase input array. Red and blue designate real and imaginary parts. Whirl magnitude is typically a few percent of ordinary spectrum. The target is star *11UMi plus weak ThAr lamp, and etalon E4 (delay ∼1 cm) is used. Vertical axis is intensity, in thousands of counts, and for bottom panel is change in intensity. Note a ThAr line at pixel $X = 1402$ (4849 cm^{-1}).

$$\Delta\phi = \Delta\tau/\lambda = \nu\Delta\tau. \qquad (7)$$

The phase step wraps around the circle producing an unbalanced distribution [Fig. 10(c)], that also changes across the band. For example for the A-order, ν varies 5300/4100 = 30%. Our nominal 0.25 μm delay step produces a phase change on the red side of 0.102 cycle and blue side of 0.132 cycle. After 10 steps they will be mismatched by 0.3 cycle, which is significant.

Lastly, mechanical vibrations, air convection, and atmospheric seeing can potentially alter step delay and exposure flux, and smear time averaged fringe visibility, so for most generality we assume that the phases and visibilities are unknown and irregular.

5.3 Fitting Along a Column Assuming Y-Knowledge

5.3.1 Sine fitting versus vector fitting

For each (wavenumber) column of Fig. 8, the intensity versus phase is essentially fitted to a sinusoid. The sine and cosine

amplitudes represent the imaginary (blue) and real (red) parts of the complex spectrum **W** [Fig. 11(b)], called "whirl," which represents the moiré pattern. The vertical offset to the sine fit represents the contribution of the nonfringing ordinary spectrum B_{ord}, also called the native spectrum [Fig. 11(a)].

In the case of uniform visibility fringes and knowledge of phase step positions, a standard sine fitting routine can be used. But if the visibility is nonuniform [as in Fig. 12(a)], such as under atmospheric seeing induced intensity variations, then a standard sine fitting routine can give an erroneous result. (Magnifying a weak signal also magnifies the photon noise, so fixing severe drops in fringe visibility only through normalization is not recommended.)

Rather than sine fitting, we use a vector combination method instead [Fig. 12(b)], which has greater flexibility in handling irregular fringe visibility and irregular phases, and has the ability for massively parallel computations over many columns
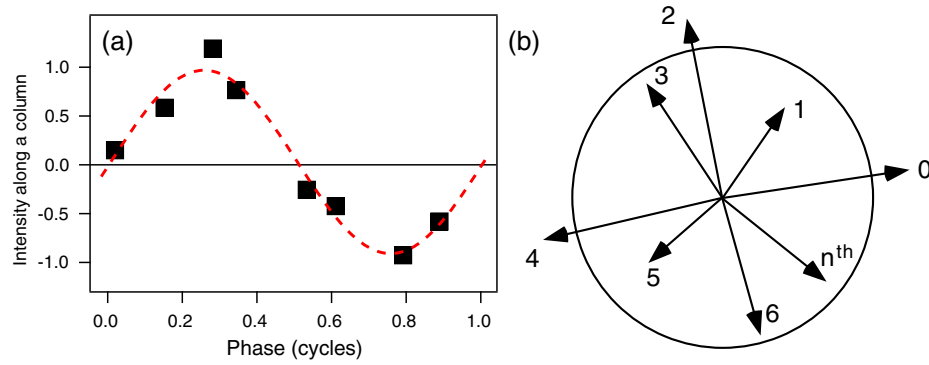
Fig. 12 (a) Sine fitting is equivalent to a (b) vector construction where step phase and visibility are designated by vector length and angle, which can both be irregular. To isolate the nonfringing from fringing components the vector sum must be zero. The values of weights to achieve this are calculated from phase of ThAr features. Whatever arithmetic is performed on vectors is also performed on data.

simultaneously—useful for rapidly processing the wide bandwidths of our data. To separate the fringing and nonfringing component, the vector sum must be zero, and weightings are applied to achieve this.

The method of choosing the optimal weightings is discussed in Sec. 5.3.3 on "healing." Notably, a "clumping" technique is used so that if some of the vectors have near zero magnitude, they do not spoil the value of the clump. So dropouts in signal due to a passing cloud, e.g., are tolerated well in our scheme.

5.3.2 Designating vectors

It is useful to introduce a set of designating vectors (Fig. 13), whose phase and magnitude represent the interferometer comb phase and visibility for a given pixel or group of pixels. Whatever operation is performed on the designating vectors is also performed on the underlying data it represent.

We visualize Eq. (4) for the ordinary spectrum B_{ord} as a sum of vectors [Fig. 13(d)] such that the fringing component is cancelled because the vector sum is zero. We are allowed to weight the vectors. We find a combination of weights (there are many) that produces vector cancellation, and then apply the same weights and arithmetic to the data to produce the ordinary nonfringing spectrum.

Analogously, we visualize Eq. (3) for the fringing component $\mathbf{W}$ as a series of rotations $e^{-i2\pi\phi_n}$, tailored to each exposure, to align all the vectors to point in the same direction. These software rotations are ideally exactly counter-rotating to the
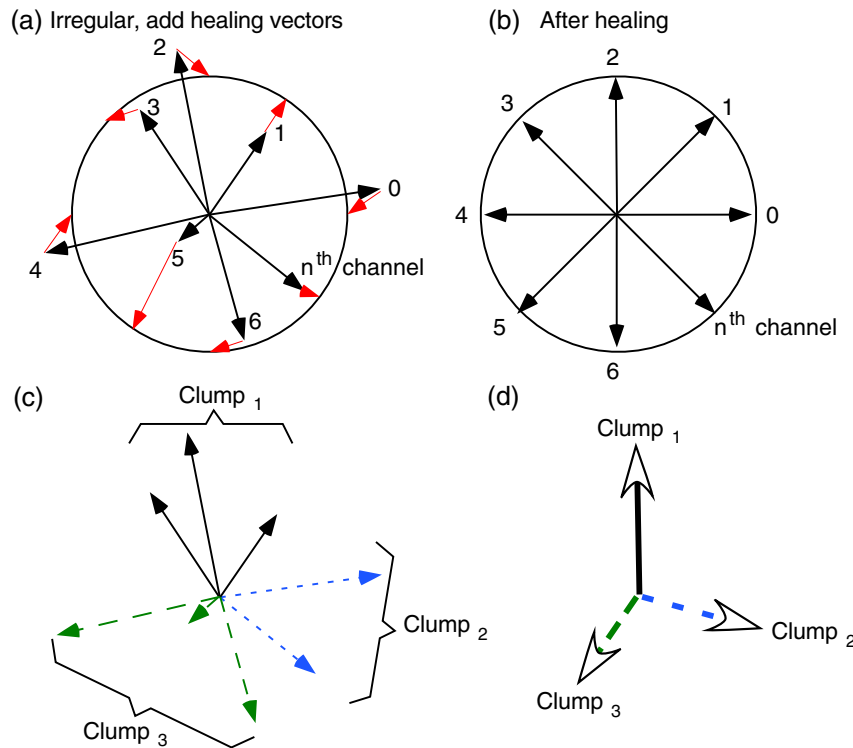

Fig. 13 Regularizing an irregular phase set of fringing data (a) into a regular set (b) involves, for each input vector, adding small amounts of other data from the same set. In the presence of noise, it is beneficial to add averages (d), "clumps," of those data (c) facing similar directions.

delay-induced rotations. Now, it is the nonfringing terms which appear to rotate and vector cancel, leaving the purely fringing component.

5.3.3 Conversion of irregular to regular phase steps

Since there are more phase steps than degrees of freedom, there are many allowable weighting schemes that solve the phase-stepping vector problem. The scheme described below works well and is programmatically straightforward, but it is not a unique solution, nor necessarily universally optimal.

We handle irregular phase steps by first regularizing the data by an arithmetic recombination we call "healing," which is a scheme that converts irregular to regular data using only adjustment of weightings and no rotations. Then to this ersatz regular data, we can apply any of a wide variety of methods designed for regular phase steps, such as the direct equations Eqs. (3), (9), and (10), or by applying a standard sine fitting algorithm column by column through the array.

We use a rule of thumb that we never rotate real data (e.g., a single exposure), but only adjust its weight. This is because real data B_n contains at least two phasors of opposite polarity, one clockwise (negative frequencies) and the other counterclockwise (positive frequencies), as in Eq. (5) of Ref. 11, which is

$$B_n(\nu) = B_{\text{ord}}(\nu) + \frac{1}{2}[S_0(\nu)e^{i2\pi\phi_n}e^{i2\pi\tau\nu}$$
$$+ S_0(\nu)e^{-i2\pi\phi_n}e^{-i2\pi\tau\nu}] \otimes \text{PSF}(\nu). \quad (8)$$

For irregular phase positions around the circle, rotations that would satisfy the clockwise component are unlikely to also exactly simultaneously satisfy the counterclockwise component. However, weightings (being scalars) can satisfy both. The healing process below produces the desired vector cancellations by only adjusting weightings.

Figure 13 shows the healing process that adds or subtracts other uniphase data from the same multiphase array, to each uniphase row. Fundamentally, this is allowed mathematically only if the nonfringing component has already been removed from each uniphase data, so that the so-modified data have a sinusoidal behavior exactly proportional to its designating vector, i.e., no intensity offset. So this healing process takes place after an ordinary spectrum B_{ord} is calculated (described in next section), then subtracted from each of the uniphase rows B_n to produce a purely fringing array of

$$B(\nu)'_n = B_n(\nu) - B_{\text{ord}}(\nu), \quad (9)$$

which gives us an improved whirl

$$\mathbf{W}(\nu) = \frac{1}{N}\sum B'_n(\nu)e^{-i2\pi\phi_n}. \quad (10)$$

For each column (X-pixel), the phase (and underlying data) is sorted so it is monotonically increasing. Then we create a regular set of goal vectors $\phi_n = n/N$, evenly distributed about the circle having constant magnitude [Fig. 13(d)]. We compute the difference between the goal and nearest actual vectors and call these healing vectors [red vectors in Fig. 13(c)]. These represent how much fraction of other channel data we need to add to a given channel data to bring it to regularity. Note, we are not allowed to rotate a given channels data directly, and we can

only modify it by adding or subtracting various amounts of other data channels.

It is best to select data vectors that lie in similar direction to the needed healing vector. We create three vector primitives, labeled Clump₁, Clump₂, Clump₃ [Fig. 13(b)], by averaging together those vectors pointing in similar directions, grouped in thirds of a circle [Fig. 13(a)]. The advantage of using clumps is that it avoids using a single data that by chance is weak and noisy.

Then we form each healing vector from a linear combination of two of the three clump vector primitives. The weightings are then calculated easily since all the angles and magnitudes are known.

Once all the arithmetic operations have been identified that convert the irregular to regular vectors, the same operations are applied to the data. This forms a corrected multiphase array having regular phase steps $\phi_n = n/N$. Then one applies a sine fitting or Eqs. (3), (9), and (10) to form the whirl $\mathbf{W}(\nu)$ and ordinary spectrum $B_{\text{ord}}(\nu)$.

5.3.4 Calculation of B_{ord} from three clumps

The B_{ord} used in Eq. (9) was calculated by weighting and then summing the data associated with the three clump primitive vectors [generalization of Eq. (4)]. With one vector held constant, the two weightings of the other vectors needed to produce vector cancellation can be found directly through algebra, since there are two equations (x and y) and two unknowns.

5.3.5 Delays relative to first exposure

The first exposure (bottom row of the multiphase array) is the designated representative exposure for the set, when finding the absolute delay τ, which would be τ_0, or outputting the SR spectrum. We may use $\delta\tau$ and $\delta\phi$ to denote the change offset from the first exposure, e.g., $\delta\tau_n \equiv \tau_n - \tau_0$ and $\delta\phi_n \equiv \phi_n - \phi_0$.

(Out of habit, we may use term "phase step" generically where "delay step" is more proper. Fundamentally, it is the delay which is commanded to change and which is constant across the band. Also, "steps" may refer to delay position, but in other places refers to change in delay position.)

5.3.6 ThAr provides step and calibration info

We can determine the individual exposure's relative $\delta\tau_n$ from either stellar or ThAr spectra, but it is more accurate to get it from ThAr since it has narrow isolated lines that have very high visibility, near unity, rather than the few percent typical of the stellar whirl. So we process the ThAr-only fiber data first, and use that ThAr step information to process the other fiber containing StarThar. The presence of some starlight with the ThAr does not hinder finding accurate phase step.

5.4 Fitting Along a Row Assuming X-Knowledge

5.4.1 Fitting data to twisted whirl

The purpose of this section is to fit each row of 1-D uniphase data $B_n(\nu)$ to a three component equation, where the unknowns are the relative delay change τ_n and magnitude γ_n of the whirl, plus the amplitude c_n of the ordinary nonfringing spectrum.

$$B_n(\nu) = \Re[\gamma_n W_0(\nu)e^{i2\pi\nu\delta\tau_n}] + c_n B_{\text{ord}}(\nu), \quad (11)$$

where W_0 is the whirl template and B_{ord} is the ordinary template, which we get iteratively from the previous section. A useful approximation to this (provided the twist is less than one revolution across the band) is to replace the twisting of the whirl caused by $\nu\delta\tau$ with a fixed rotation which we will call an average phase over the band $\langle\phi\rangle \sim \langle\nu\rangle\delta\tau$

$$B_n(\nu) = \Re[\gamma_n W_0(\nu)e^{i2\pi\langle\phi_n\rangle}] + c_n B_{\mathrm{ord}}(\nu). \tag{12}$$

The advantage of the above equation is that it can be solved directly with Cramer's rule to yield $\langle\phi\rangle$ (and γ and c). We can use it to estimate a better guess for $\delta\tau$ as we fit for $\delta\tau$ iteratively.

Anticipating that the whirls for higher delays than the 0th are twisted, we form a twisted template whirl

$$W_n(\nu) = W_0(\nu)e^{i2\pi\nu\delta\tau_n}. \tag{13}$$

Then we use that and seek the average phase shift that remains. We search for the $\delta\tau_n$ that yields $\langle\phi_n\rangle = 0$

$$B_n(\nu) = \Re[\gamma_n W_n(\nu)e^{i2\pi\langle\phi_n\rangle}] + c_n B_{\mathrm{ord}}(\nu). \tag{14}$$

An initial value for $\delta\tau_n = n(0.25\ \mu m)$ is used in a direct evaluation of Eq. (12) by the projection method described next (Sec. 5.4.2) to yield $\langle\phi\rangle$. We decrease our guess for τ by about $\langle\phi\rangle/\langle\nu\rangle$ and repeat. This converges rapidly to zero phase remainder and our output value for $\delta\tau_n$. (This iteration loop is inside the box labeled "Y-dependence discovery" of Fig. 9 and separate from the overall iteration loop.) This produces the two template spectra $\mathbf{W}$ and B for the next section used to determine $\delta\phi_n$ for every channel n.

5.4.2 Projection method for determining individual step phase

Now that we have found templates $B_{\mathrm{ord}}(\nu)$ and $\mathbf{W}(\nu)$ that characterize the set of channels, the next step is to find the phase $\delta\phi_n$, fringe visibility γ_n, and amplitude of ordinary spectrum c_n of each individual channel n (each exposure). This is done using a method of projections.

This method only works when the templates are approximately orthogonal to each other, meaning their overlap integral is small relative to the self overlap integral. This occurs when the bandwidth is sufficiently large so that $B_{\mathrm{ord}}(\nu)$ and $\mathbf{W}(\nu)$ have dissimilar shapes.

The guiding equation Eq. (14) is re-expressed into a linear model having three components: two from the fringing and one from the nonfringing

$$B_n(\nu) = [\gamma_n \cos 2\pi\phi_n]\Re W_n(\nu)$$
$$+ [\gamma_n \sin 2\pi\phi_n]\Im W_n(\nu) + c_n B_{\mathrm{ord}}(\nu). \tag{15}$$

To make the equations more systematic we re-express $\mathbf{W}(\nu)$ in its real (cosine-like) and imaginary (sine-like) parts

$$\mathbf{W}(\nu) \equiv F_c(\nu) + iF_s(\nu) \tag{16}$$

and rename the nonfringing component B_{ord} as F_{NF} to be consistent. Then we have

$$B_n(\nu) = a_n F_c(\nu) + b_n F_s(\nu) + c_n F_{\mathrm{NF}}(\nu) \tag{17}$$

and we will solve for the coefficients a_n, b_n, c_n given data B_n. From these we get the channel phase (and magnitude) through

$$\tan 2\pi\phi_n = a_n/b_n, \tag{18}$$

$$\gamma_n = \sqrt{a_n^2 + b_n^2}. \tag{19}$$

The magnitudes γ_n, $|c_n|$, of the fringing and nonfringing components are useful for normalizing the exposure data $B_n(\nu)$, which obviously is also done iteratively after an initial guess. For the ThAr, the initial guess is a constant since it is a steady source. For the Star or Star + ThAr data, the initial guess is obtained by a simple average of the signal over all the pixels of the order, relying on the wide bandwidth to wash out the fringes.

We use overlap integrals to generate from Eq. (17) three equations for the three unknowns. Let the integral between two functions $F_1(\nu)$ and $F_2(\nu)$ be represented by the more compact expression

$$\langle F_1 F_2 \rangle \equiv \int F_1(\nu)F_2(\nu)\mathrm{d}\nu. \tag{20}$$

To both sides of Eq. (17), we apply the overlap integral, three times using each of the template functions $F_c(\nu)$, $F_s(\nu)$, and NF(ν). This creates three linear equations

$$\langle F_c B_n \rangle = a_n\langle F_c F_c \rangle + b_n\langle F_c F_s \rangle + c_n\langle F_c F_{\mathrm{NF}} \rangle, \tag{21}$$

$$\langle F_s B_n \rangle = a_n\langle F_s F_c \rangle + b_n\langle F_s F_s \rangle + c_n\langle F_s F_{\mathrm{NF}} \rangle, \tag{22}$$

$$\langle F_{\mathrm{NF}} B_n \rangle = a_n\langle F_{\mathrm{NF}} F_c \rangle + b_n\langle F_{\mathrm{NF}} F_s \rangle + c_n\langle F_{\mathrm{NF}} F_{\mathrm{NF}} \rangle. \tag{23}$$

The integrals are quickly computed into values, and these three equations Eq. (23) are easily solved by Cramer's rule to yield the a_n, b_n, and c_n coefficients. Then we apply Eq. (17) to yield channel phase ϕ_n and γ_n is the angle and length of the designating vector.

Note that because the equations include a third component for the nonfringing B_{ord}, these can be used either before or after the subtraction of B_{ord} from each B_n, in Eq. (9). Also, it is okay if the templates $\mathbf{W}(\nu)$ and $B_{\mathrm{ord}}(\nu)$ contain multiple components, such as (star + ThAr), since all the components are locked in phase with respect to each other, regarding how they respond to an interferometer delay change.

5.4.3 Normalization

As seen in the flowchart Fig. 9, a channel-dependent normalization (Slit$_y$) divides the data to become more uniform against fluctuations such as passing clouds. The challenge is to learn the source intensity variation that is apart from the sinusoidal variation seen naturally from fringes. There are two methods of finding Slit$_y$.

The first method uses the amplitude c_n from the three component decomposition (Sec. 5.4.2). This method is the best choice when the spectral features are sparse, such as for the ThAr lamp. For ThAr, the second method of simple summation is much worse because there are too few features to have the statistical benefit of averaging. Then summation does not sufficiently diminish the net fringe visibility, so the sinusoidal nature

"prints through" to the Slit$_y$ and produces a wobble-like effect on the whirl, which prevents perfect separation of ordinary and whirl components in the two outputs.

Initially, we also applied this first method to the Star + ThAr data, but discovered that it produced a slight residual of the ordinary spectrum in the whirl, which produced a ringing artifact. The remedy was to use a second normalization method, which is simple summation over all X-pixels. This method works well for the Star component because the number of features is high so the net fringe is washed away statistically. In the case of StarThAr data, we temporarily mask away the bright ThAr lines prior to the summing.

5.4.4 Normalization problems lead to pollution of whirl with ordinary

We found that if the normalization was not working properly, this harms the phase-stepping math, and pollutes the whirl with a small fraction of ordinary component, causing the whirl to have a pedestal appearance [Fig. 14(c)], seen between the real and imaginary parts and having the same shape as the ordinary spectrum. This is not critical for the ThAr because it is large. But for the Star whirl which is only a few percent, a small fractional pollution of the ~50× larger ordinary component produces a significant absolute amount of pedestal.

This pedestal is a DC-like (zero frequency) component that gets upshifted during heterodyning reversal to contribute a ringing artifact [Fig. 14(a)]. This was most easily detected in places of the telluric spectrum which are heavily absorbing and thus should be relatively flat (zero intensity). When we eliminated the pollution of the whirl by changing normalization schemes,

we got a beautiful ring-free result [Fig. 14(b)]. (We could also solve this by manually subtracting an adjustable amount of ordinary spectrum from the whirl, but that did not address the root problem.)

The pollution and ringing problem disappeared when we changed the method of computing normalization for the StarThAr to the simple sum method (after masking away ThAr lines).

5.4.5 Normalizing challenge: Star twinkles but ThAr does not

The ThAr lamp intensity is steady with time, but the stellar intensity is not, due to atmospheric seeing fluctuations. For our present algorithm and StarThAr data, we do not attempt the mathematically challenging task of normalizing both simultaneously. Instead, we first use the ThAr-only signal to provide the calibrations for phase step values and for equalization used for all signals. Then we process the Star-ThAr signal, normalizing the stellar component, since it is the weaker and more valuable signal. This is adequate.

6 Stage 3: Rebinning and Regularization

Stage 3, called rebinning and regularization, linearizes the horizontal axis, allowing comparison of the data to theory when plotting in mathematically more convenient wavenumbers (ν). The horizontal variable is rebinned ~10× more finely, to a uniform grid of 0.05 cm^{-1} so that it can hold the high frequency information of the eventual high-resolution output spectrum. We also remove the phase twist versus ν due to the changing refractive index of glass etalons comprising each delay. This
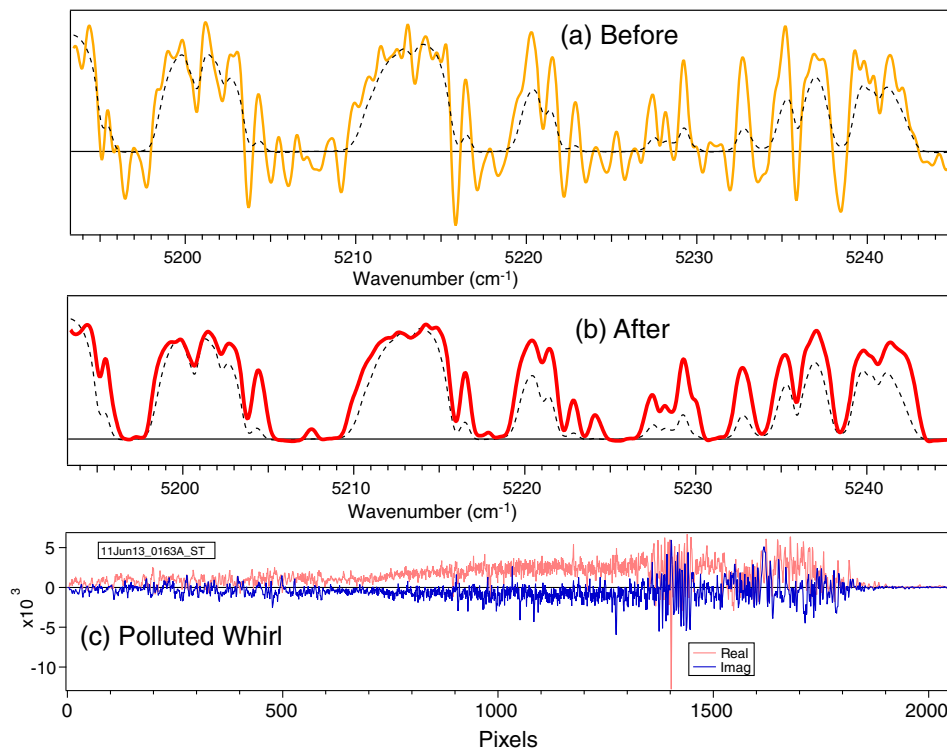

Fig. 14 (a) Earlier spectral reconstruction results had excessive ringing (yellow curve), compared to telluric theory[19] (dashed). Bad normalization polluted the fringing signal (whirl) with nonfringing component, which created a pedestal-like offset in the whirl (c), which became ringing during heterodyning reversal frequency shift. (b) Ringing problem was solved (red curve) using a different normalization method.

regularization is applied to each of the eight etalon channels of the data, to form two output arrays sent to the next stage 4 (spectral reconstruction), which are a stack of whirls $\mathbf{W}$ and a stack of ordinaries B_{ord}, eight high versus etalon #. Two arrays are thus produced for each of source: ThAr, StarThar, calculated ThAr, and calculated Telluric + ThAr.

The latter two sources are simulations of data which can be injected into the data analysis pipeline here at the beginning of stage 3, on simulated pixels. This allows comparison of data to theory both in detector pixel space and in uniform wavenumber space. The comparisons are used on ThAr data to compute the glass dispersion detwisting, and tweaks to the pixel to wavenumber calibrations.

6.1 Rebinning

The horizontal pixels are rebinned more finely $\sim 10\times$, to a uniform grid of 0.05 cm^{-1}. The criteria are that half the reciprocal of the final grid spacing (called the Nyquist frequency) is 10 cm, and this needs to be larger than the largest delay we intend to use (4.6 cm of E8).

We found that a standard cubic spline does a better job than our custom-written FFT method, regarding avoiding any overshoot artifacts produced by the narrowest features.

6.2 Pixel to Wavenumber Map Calibration

During rebinning, we map the data from pixels to wave numbers using an improved calibration curve (map) which we fine-tune here. We use individual calibrations for fibers A and B that accounts for the slit slant between those fibers. We add to the original polynomial fit a line segmented "tweak map" which forces the map to exactly pass through each important ThAr line or other feature. We primarily use the magnitude of the whirl for the comparison because it is much cleaner than the ordinary spectrum (since the former is insensitive to FP noise). In sparse ThAr regions we use telluric features.

6.3 Bulk Shift of Map versus (Time or Etalon)

We find that the map has an etalon-dependent offset, which we call shear. This is due to a time-dependent drift, probably thermomechanical from the changed gravity vector after repointing the telescope. The magnitude of the shift is ~ 0.3 cm^{-1} over 1 or 2 h.

After removing the monolithic shift in each ordinary spectrum, we notice that there remains a significant misalignment which changes its amount rapidly along the band, and even changes polarity. Examples are shown in the instrument noise section.

6.4 Lowpass Filtering to Prevent Nyquist Artifacts

Figures 15(a) and 15(b) show aliasing artifacts in data from etalon E5 (1.27 cm), that occurs when the detector pixel density is too sparse. Energy from the continuum fringe comb spike leaks beyond the Nyquist frequency (0.5 pixel^{-1}, or $\rho \sim 1$ cm), wrapping to the other frequency branch on the negative frequency side. This causes an erroneous signal.

Our remedy [Fig. 15(c)] is to restrict frequencies to <0.5 cm, then apply a standard cubic spline interpolation routine to rebin. The latter has good behavior similar to a Hanning-shaped lowpass filter.
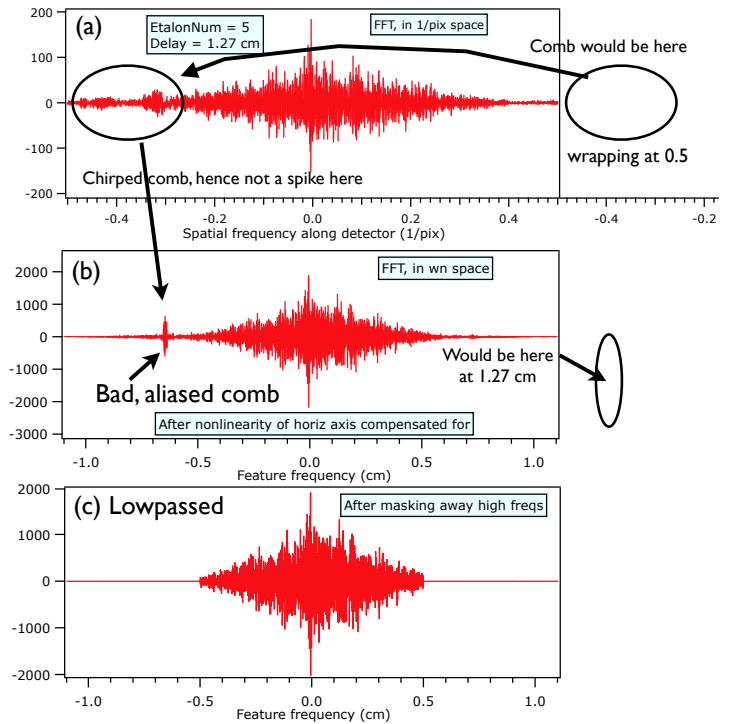
(I fixed a missing Fig. 15 in published version --DJE)

Fig. 15 Aliasing problem because detector pixels too sparse. (a) For delay E5 the FFT has spatial frequencies above 0.5 pix^{-1}, so it wraps to the negative frequency axis. (b) FFT of data after horizontal linearization—aliased spike easier to see at -0.65 cm (which belongs at 1.27 cm). (c) Lowpass filtering removes most aliased spikes.

6.5 Detwisting the Whirl, Removing Glass Dispersion

The interferometer delay is a function of wavenumber due to the refractive index dispersion of the glass etalon. This creates a phase twist versus wavenumber, $\theta_{glass}(\nu)$, associated with each etalon. Data whirls are untwisted via $\mathbf{W}(\nu)e^{-i2\pi\theta_{glass}(\nu)}$. Otherwise, the FFT will not produce a narrow spike [Fig. 16(b)] for the fringe comb.

For each etalon, we find $\theta_{glass}(\nu)$ [Fig. 17(a)] by comparing measured to calculated phases of ThAr lines. Fitting a polynomial to $\theta_{glass}(\nu)$ produces a baseline fit good to ~ 0.1 cycle in chunks of 20 to 200 pixels. The precision is further improved later to better than 0.01 cycle using the ThAr or telluric features of each target's data (Fig. 18).

For the spectral reconstruction, we only need this phase curve. However, it is useful to produce a delay versus ν curve [Fig. 17(b)] by taking the ν-derivative. This is useful for detecting skipped integer fringes of phase, lost spanning band regions that lack features.

6.6 Fine Delay Determination

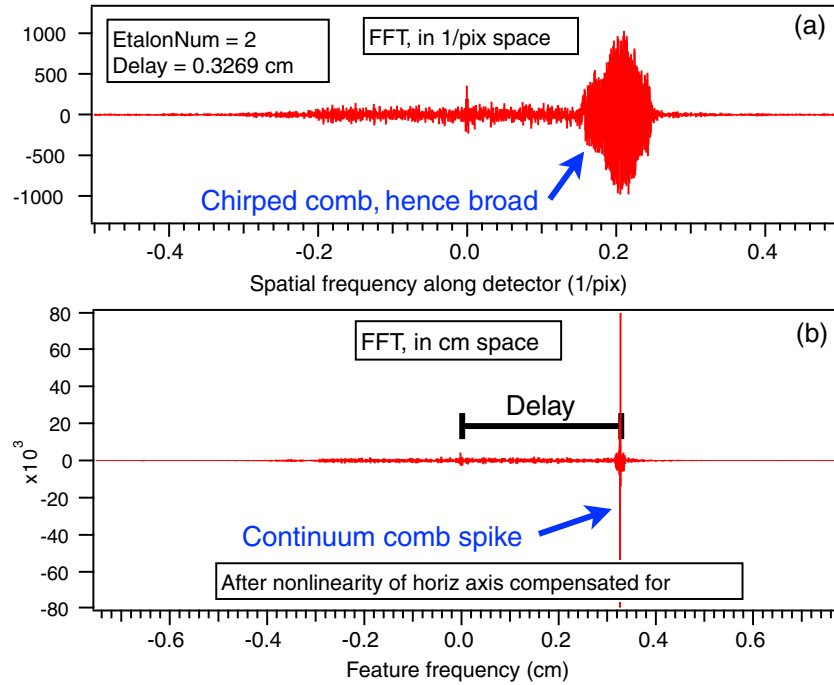Due to our wide bandwidth, by observing the twist in phase along the band, we could determine the delay unique to the

Fig. 16 The effect of linearizing the horizontal axis from pixels to wavenumbers, and compensating for glass dispersion, is to narrow the comb spike. (a) The Fourier transform of whirl for delay E2 (0.33 cm) in 1/pixel frequency space, showing a broad continuum comb spike. (b) Linearized data, the spike is extremely narrow.

nearest fringe (nearest 2 μm), i.e., with no integer ambiguity. Then within that fringe we further determined the delay to a fraction of a fringe and removed the last fractional phase shift.

Due to mirror repositioning errors, the delays would change a few microns from night to night; we could also detect a delay difference of a fraction of a micron between A and B fibers. Since these were so small, we kept the original delay value and instead slightly adjusted the whirl phase.

The values of the eight delays in our quiver for 2010 September were for E1-E8 0.083, 0.34, 0.66, 0.96, 1.27, 1.75, 2.92, 4.63 cm, and 2.4 cm for E6.5, added in June 2011. (The latter was used to fill the gap between 1.7 and 3 cm, and swapped for the E1 delay on the eight-position filter wheel.)

6.7 Calculating Phase Difference Between Theory and Data

We have two methods for calculating the phase difference between theory and data. (1) The dot product method is automated and thus used initially. However, it has the drawback that it blindly uses the entire theory spectra without regard to any incorrect components to it. (2) Then we fine-tune the result using a manual Lissajous method which has the ability for the user to ignore components suspected to be incorrect.

6.7.1 Dot product method of phase evaluation

The method of dot products uses an overlap integral

$$\mathbf{W}(x) \cdot \mathbf{U}(x) \equiv \sum_{x'=x}^{x+199} W_{\mathrm{Re}}(x') U_{\mathrm{Re}}(x') + W_{\mathrm{Im}}(x') U_{\mathrm{Im}}(x') \quad (24)$$

over a small region (chunk) of X-pixels typically 20 to 200 pixels. The order of the summation is immaterial, and the pixels can be nonlinear, such as being in wavenumber or wavelength. We presume that $\mathbf{W}$ rotates between two orthogonal versions of the reference whirl $\mathbf{U}$ and $\mathbf{U}_\perp$, ($\mathbf{U}_\perp = i\mathbf{U}$), as

$$\mathbf{W} = a\mathbf{U} + a_\perp \mathbf{U}_\perp. \quad (25)$$

The vector angle ϕ is found by

$$\tan 2\pi\phi = a_\perp/a = (\mathbf{W} \cdot \mathbf{U}_\perp)/(\mathbf{W} \cdot \mathbf{U}). \quad (26)$$

To reduce noise, one can mask away regions having whirl magnitude (theory or data) below a threshold.

We calculated simulated whirl data from theoretically supplied high-resolution spectra of our sources, such as ThAr from the Nat. Inst. Standards and Techn. (NIST) group[20] or a telluric model from Roe.[19]

6.7.2 Fine tuning phase and ν, alignment via Lissajous

A fine tuning alignment of data to theory uses a Lissajous presentation to compare two wavelets for a small region of the bandwidth, data versus theory. This style of presentation has the advantage of showing small changes in both phase and ν somewhat independently. The user observes the Lissajous shape (Fig. 18) while adjusting phase and ν translations. We use a rainbow color change of the curve to indicate the approximate ν, so that features can be identified. Potentially this method could be automated but is currently manual.

An important advantage of this style of presentation is the ability for the user to disregard features in theory or data which are suspected to be unreliable. Examples include blended lines, containing multiple species such as both Ar and Th within
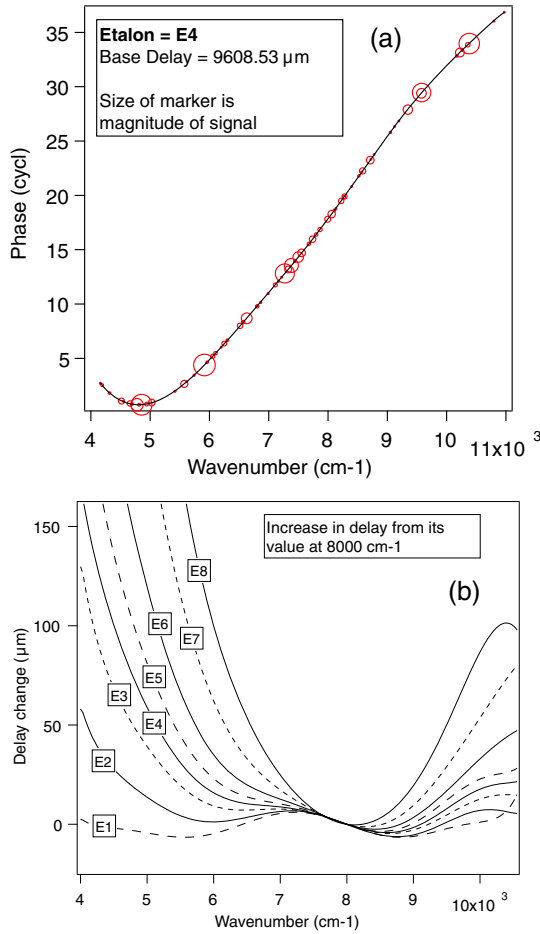
Fig. 17 The dispersion characteristics of set of glass etalons were calibrated over a 4000 to 11,000 cm^{-1} bandwidth by phase shift comparison with ThAr emission line model.[20] (a) The resulting phase versus wavenumber curve is then used to "untwist" stellar fringing spectra so that a constant base delay τ_0 can be used during reversal of heterodyning effect. (b) The derivative of the phase data provides delay versus wavenumber curves. The dispersion characteristic of the beamsplitter substrate (most obvious in the smallest delay E1) is a common component of all. Delay values for E1-8 are ~0.1, 0.3, 0.7, 1, 1.3, 1.7, 2.9, 4.6 cm, and 2.4 cm for E6.5 (not shown).

the resolution element. We have discovered (Sec. 8.4.1) that our ThAr lamp differs from the NIST ThAr lamp—our Th species appears several times lower in magnitude, whereas our Ar species has the expected magnitude.

Figure 18 shows that a small misalignment in phase manifests a different appearance than a misalignment in ν. Thus both can be corrected independently. (a) In perfect-alignment each loop becomes a line. The phase misalignment (c) makes the individual loops more open and elliptical. The ν misalignment (b) tends to spread the loops out like a fan.

7 Stage 4: Spectral Reconstruction

7.1 Summary of Spectral Reconstruction

In stage 4, "spectral reconstruction," the whirls are numerically shifted in frequency space (heterodyning reversal, Fig. 19) to undo the analog frequency shift imparted by the interferometer. This produces a wavelet signal $\Omega_n(\nu)$ for each etalon, covering

a different spatial frequency range centered on its delay value. These wavelets are summed to produce a provisional output spectrum $SR_{prov}(\nu)$. This is equalized to remove the bumps in the net frequency response, so that the latter is Gaussian with a user-chosen final resolution, to form a final spectral output $SR_{final}(\nu)$ having minimal ringing. Comparison of ThAr data to theory for an isolated peak provides the equalization curve used for both ThAr and Star. The choice of final resolution is limited by the maximum delay which produces a contiguous range of delays, as $\sim \tau_{max}/\lambda$.

7.2 Bell-Like Weighting to Reduce Net Photon Noise

Prior to heterodyning reversal, frequencies outside a set value (0.52 cm) were deleted, since they contain little signal. In addition, a bell-shaped weighting $h_0(\rho)$ was sometimes applied $\mathbf{w} \rightarrow h_0(\rho)\mathbf{w}(\rho)$.

For photon-limited noise, the ideal shape is that of the native sensitivity peak, but for instrument limited noise (the TEDI situation) there is little benefit. If we used a weighting, it was a Hanning function since it goes to zero absolutely at its edges (set to 0.52 cm).

7.3 Reversal of Heterodyning

Figure 19(b) shows the effect of reversal of the heterodyning, where the signal has been shifted in the direction that moves the continuum comb to zero. The effect is illustrated in frequency space, but is more accurately computed in the software in ν space as

$$W'(\nu) = e^{-i2\pi\nu\tau} W(\nu), \tag{27}$$

because this presents no issues with continuous values of τ. (Shifting in ρ space led to quarter cycle-sized errors due to discrete nature of the FFT.)

Our convention has the continuum fringe comb starting at a positive frequency ρ. Hence during numerical heterodyning reversal, Figs. 19(a) and 19(b), the signals shift to left toward negative ρ. This convention is arbitrary since after the real part is taken (d) both frequency branches (c) have energy.

7.4 Wavelets Produced from Real Part of Shifted Whirl

The output spectrum we are reconstructing must be purely real. We take the real part of the heterodyne reversed signal (in ν space) to form what wavelet signal

$$\Omega(\nu) = \mathbf{W}'^* + \mathbf{W}' = 2\Re\mathbf{W}'(\nu), \tag{28}$$

so called because of its appearance for narrow lines (such as 18). Figure 19 shows that $\Omega(\nu)$ for a small delay of E2 has similar frequency content as the native spectrum, for the whole A-order.

In practice, the above equation might actually require using the imaginary part of $\mathbf{W}$, depending on how the phase is defined. The test is that the continuum comb portion of $\mathbf{W}$ (produced by a small delay) should be, after heterodyning reversal, mostly real and positive.
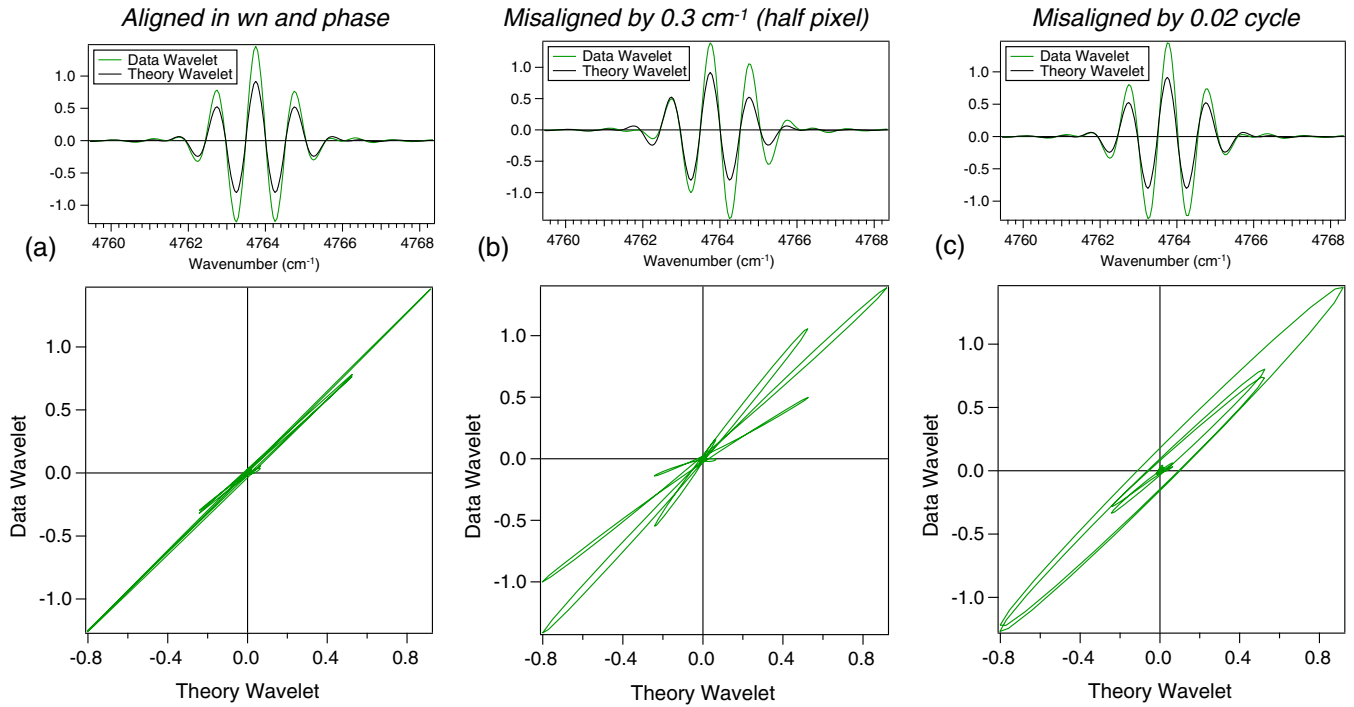
Fig. 18 Aligning a data wavelet (light green) to theory wavelet (black) in wavenumber and phase is easy to do precisely by using Lissajous presentation (bottom row), which plots data versus theory, rather than data and theory versus wavenumber (top row). One restricts bandwidth to a small (~10 cm^{-1}) region otherwise the loops obscure each other. (b) Effect of a 0.3 cm^{-1} change in wavenumber. (c) Effect of a 0.02 cycle phase change. (a) Perfect alignment. Data of ThAr lamp from September 19, 2010 etalon E4, A-order.
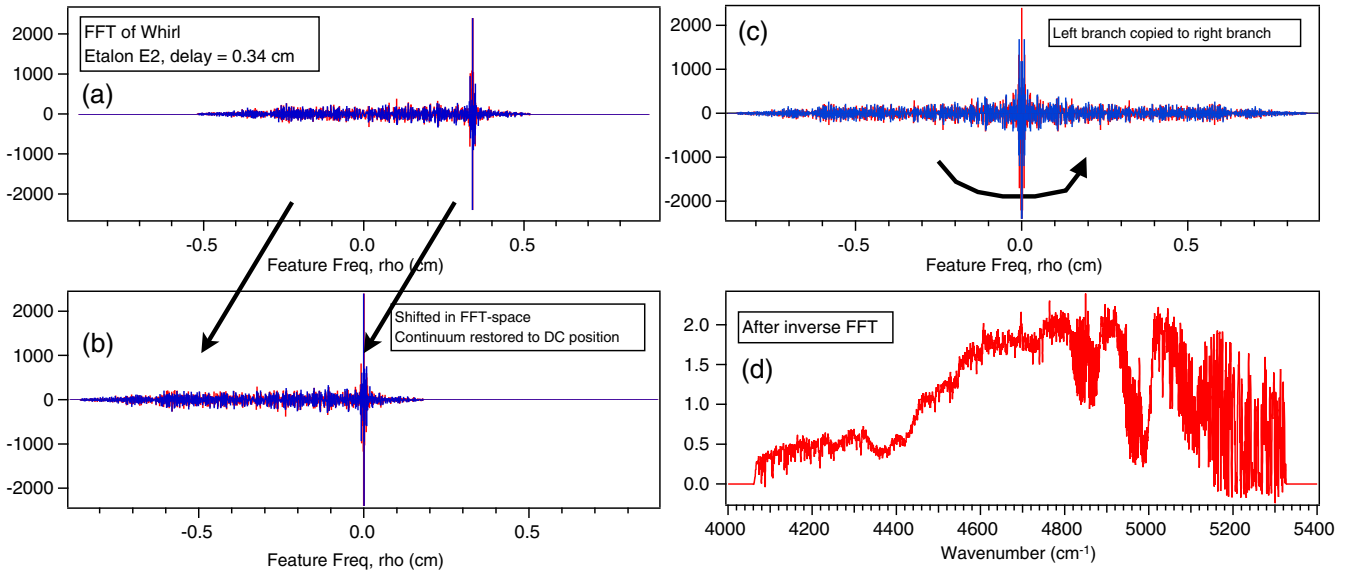
Fig. 19 Stages of heterodyning reversal processing. (a) FFT of whirl, here for a small delay E2 (0.34 cm). Continuum comb spike at 0.34 cm is within range of native spectrograph ($-0.5 < \rho < 0.5$ cm^{-1}). Horizontal axis is frequency in features per cm^{-1}, having units of cm. Vertical axis limited to ± 2400 to more easily see details, and the spike is about $\pm 10,000$ high. (b) The FFT is shifted left by the same amount as the delay, to move the spike to zero frequency and other signal energy to left branch. (For larger delay etalons there may be no comb spike.) The real part of the inverse FFT of (b) produces the spectrum (d), which may look like wavelets for higher delays, but for lower delays will contain the continuum shape. Equivalently, one could copy the complex conjugate to the opposite branches to form (c) then apply the inverse FFT to form (d).

7.5 Summation of Wavelets to Form Provisional Spectral Reconstruction

The wavelets are summed to form a provisional SR spectral output called SR_{raw}

$$SR_{raw}(\nu) = \sum_n \Omega_n(\nu), \tag{29}$$

$$SR_{raw}(\nu) = \sum_n B_{ord,n}(\nu) + \sum_n \Omega_n(\nu). \tag{30}$$

This has not yet been equalized to create a Gaussian tapering. Thus it would produce a large ringing for each narrow spectral feature. The B_{ord} term is optional, and we omit it for TEDI processing (Sec. 7.5.2).

7.5.1 Ordinary spectrum useful for photon limited performance

Each of the $M = 8$ etalons produces an ordinary spectrum $B_{ord\,n}(\nu)$. We average them together to form a cleaner version $\langle B_{ord}(\nu) \rangle = (1/M) \sum_n^M$. This is equivalent to an ordinary spectrum exposed over the total exposure time of all etalons. Hence we expect this to have $\sim 0.5\sqrt{M}$ times less fractional photon noise than the signal from the E1 delay which covers a similar range of low frequencies. Hence for a photon-limited situation, it is optimal to weight it stronger in the sum, by $1/(0.5\,M^{-1/2})^2 = 4M$ so we have

$$SR_{raw}(\nu) = 4M\langle B_{ord}(\nu) \rangle + \sum_n \Omega_n(\nu), \tag{31}$$

so that the absolute amount of noise produced by the nonfringing and fringing components is about the same.

7.5.2 Ordinaries not included in TEDI output spectra

Initially, we included the ordinary native spectrum in the output spectrum, to provide the lowest frequencies, but we discovered for TEDI data we get cleaner results NOT including the ordinary spectrum. Instead, we rely on fringing from the low delay etalons E1 and E2 for the lowest frequencies information. (However, we still include it in plots of the SR output, for comparison.)

Figure 20 demonstrates in stellar data that the fringing-derived spectrum (red curve) agrees extremely well with the telluric model (black curve). In some cases, it is more accurate than the native spectrum (green curve), such as when FP noise is present. The EDI is extremely robust to FP noise because only changes between exposures contribute to the whirl.

7.5.3 Each native resolution element is independent

The heterodyning and SR process occurs independently for each native resolution element, massively parallel across the whole bandwidth of the instrument. Thus noise or a data defect such as a CR in a given resolution element only effects that resolution element.

7.5.4 Errata

In Fig. 6 of Ref. 15, similar to Fig. 20 here, we noted that the bumps of the conventional signal looked different than the model. Now we realize that this was due to our use of a lowpass filter on the conventional signal having a sharp cutoff frequency (i.e., rectangular) which introduced ringing. The lowpass filtering prevented aliasing artifacts when shifting the native signal to compensate for spectrograph drifts versus time. Since we now usually omit the ordinary signal in the SR sum we also omit lowpassing it.

7.6 Modulation Transfer Function Displays Net Frequency Response

The net frequency response of the EDI system is probed by a modulation transfer function (MTF), which is the ratio out versus in of complex amplitude, for each frequency ρ

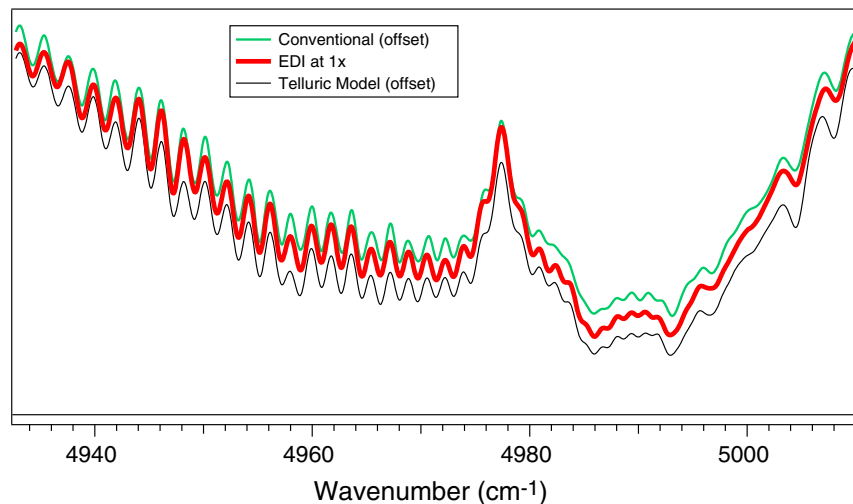$$mtf(\rho) = sr_{raw}(\rho)/s_{thry}(\rho). \tag{32}$$

Fig. 20 The EDI result (red, 1×) can provide an accurate measurement of the low-resolution information usually provided by the native spectrograph (green). Here we deliberately blur the fringing EDI result to the same resolution as native spectrograph. It accurately follows the details of telluric model (black). Green and black curves offset vertically. Target is star HD219134 + ThAr lamp, September 19, 2010.

The MTF is evaluated for a well-known source, such as a ThAr spectral lamp. Our mtf(ρ) can be a complex value, whereas the traditional definition for MTF is the absolute value $|\text{mtf}(\rho)|$.

7.6.1 Using a narrow bandwidth for modulation transfer function

Evaluating the MTF over the full bandwidth of the order has the disadvantage of producing an MTF that has too many points (thousands) in its curve. We found it advantageous to evaluate the MTF for an extremely small bandwidth, centered over a single prominent feature such as an isolated ThAr line. We use a subset spectrum 100 final-pixels spanning 5 cm^{-1}. (By final-pixel we mean the 0.05 cm^{-1} rebinned points). We then scale this 100 point MTF over any size bandwidth up to the bandwidth of the entire order, which has 240 times as many points, by linear interpolation in frequency space.

7.6.2 Measured MTF of externally dispersed interferometry data

Red curve in Fig. 21 shows an MTF measured for a star. Note that the positions of the peaks (associated with each delay) are not ideal. We desire shoulder-to-shoulder coverage to produce a ring-free net lineshape. Unfortunately, the set of TEDI delays had a gap between E7 and E8 (3 and 4.6 cm), since TEDI delays were selected for Doppler radial velocimetry purposes, not

spectroscopy, and there were only eight positions in the rotary etalon holder.

The presence of the 3 to 4.6 cm gap limits the effective resolution, because attempting to use a wide Gaussian MTF that extends over this region (such as the Res = 36,000 curve in Fig. 21) would produce ringing in the final lineshape in ν-space. If additional etalons filled the gap between 3 and 4.6 cm, a resolution of 45,000 at 7500 cm^{-1} is possible.

For data taken prior to June 11, 2011, there was a gap at 2.4 cm between E6 and E7. This was remedied by a purchase of a new etalon, called E6.5, of 2.4 cm delay which occupied the position in the eight-position "filter wheel" previously held by the smallest etalon E1. (E2 can cover the range of E1.) Only a few exposures have been taken with the E6.5, during what turned out to be the last run of TEDI, June 11 to 19, 2011.

7.7 Equalization Applies Taper to Modulation Transfer Function

The purple curve in Fig. 21(b) is the equalization curve used to reshape the MTF into a smoother Gaussian. It is the ratio between the desired MTF (black curve) and the actual MTF (red curve)

$$\text{eq}(\rho) = \text{Gaussian}(\rho)/\text{mtf}(\rho). \tag{33}$$

The eq(ρ) is optionally limited to a ceiling value so that in the wings where the response tends toward zero the EQ does not

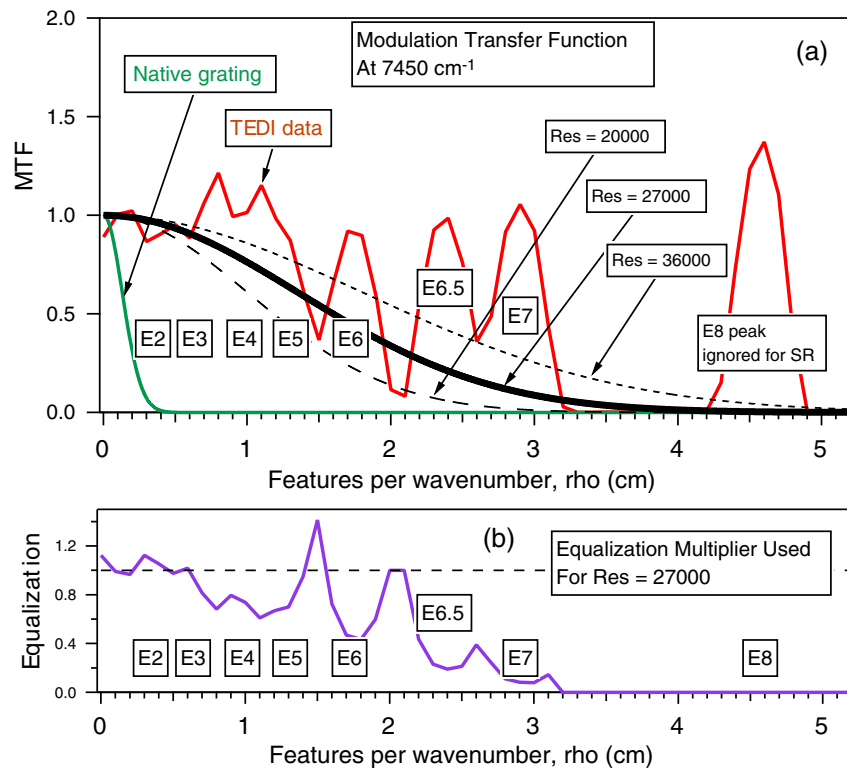**Fig. 21** (a) Net MTF versus frequency ρ for the TEDI system near 7540 cm^{-1} (1.34 μm) shows peaks for each delay (red), having same shape as native peak (green) but shifted by τ_n. Some of peaks have blended together (the ideal situation). (b) Equalization curve (purple) is Gaussian shape of res = 27,000 (thick black curve) divided by red curve MTF. If additional etalons filled the gap between 3 and 4.6 cm, resolution of 45,000 at 7500 cm^{-1} is possible. Data of for star κ_{CrB} on June 19, 2011. The E6.5 etalon (2.4 cm) swapped with E1 (0.1 cm) in the rotary holder. Most data were taken without E6.5, thereby producing a gap at 2.4 cm which limited ring-free resolution.

blow excessively magnify noise. We zero out $eq(\rho)$ in regions where the sensitivity is expected to be zero.

The EQ is applied to the FFT of the data, $sr_{raw}(\rho)$ via

$$sr_{out}(\rho) = sr_{raw}(\rho)eq(\rho). \tag{34}$$

The choice of goal resolution for the Gaussian in Eq. (33) is a tradeoff between resolution and ringing. In the presence of gaps in the delay coverage, increasing resolution increases the ringing.

7.7.1 Selecting etalon values in a future instrument

As the average ν increases, the width of each etalon peak in the MTF narrows, so a designer needs to select etalon values to produce contiguous coverage at the blue end of the native spectrograph bandpass. It is also beneficial to have enough overlap of etalons to produce two independent sets of SR results, one for the odd numbered etalons and the other for even numbered (using different equalizations). These can be used to confirm fidelity of the SR process, since it is unlikely an artifact would have the same appearance for different delays.

8 Example Reconstructed Spectra

Observations of many stars were performed with the TEDI interferometer producing reconstructed resolution as high as ~27,000 (10× boost), using 3 cm maximum delay with TripleSpec having a native resolution of ~2700. Figure 22 is an example of 10× resolution boosting. Figure 23 shows the extremely wide bandwidth we can achieve, covering all four orders of the native spectrograph from 4100 to 10,500 cm^{-1}. Figures 24 and 25 show closeups for the reddest order.

Figure 22 shows 10-fold boost in resolution (2700 becoming 27,000) observing mostly telluric features in the spectrum of star κ CrB along with ThAr calibration lamp emission lines, observed June 2011. Seven contiguous delays were used [as seen in MTF Fig. 21(a)], including a new etalon 2.4 cm E6.5 that filled a gap between 1.7 and 3 cm delays allowing a higher goal resolution without excessive ringing. The E1 (0.1 cm) position on the eight-position rotary holder was used for E6.5. [The E1 is not critical since low frequencies can be supplied by E2 (0.3 cm) or ordinary spectrum.] If additional delays were

available to contiguously fill the gap between E7 and E8 (3 to 4.6 cm), then the equalized output resolution could be as high as 36,000 [dotted curve of Fig. 21(a)] or more without significant ringing.

Figure 23 shows the wide bandwidth capabilities: star HD219134 measured simultaneously over the four orders, 4100 to 10,500 cm^{-1}. The EDI resolution scales with wavenumber (or as τ_{max}/λ), so the resolution is not constant across this wide bandwidth.

Figure 24(a) shows the whole A-order at 4× resolution boost and Fig. 25 shows closeups. The black curve is the theoretical telluric + ThAr model. The data (red and green curves) are attenuated significantly on the red end of the spectrum, probably due to transmission loss of the native spectrograph.

Note how the broad features of the red and green curves follow each other accurately, even though the red curve is obtained entirely from the fringing signal component and the green curve from the nonfringing portion. These are independent measurements extracted from the same data set. The nonfringing data were not included in the sum which forms the red curve.

8.1 Comparing Two Fibers on Same Star

Figures 26 and 27 show EDI measurements comparing the two fibers A (blue curve) and B (red curve) measuring the same source, their difference curve (gold), and compared to the telluric model (gray). The two fibers used different detector pixels in the Y direction and were exposed at times ~10 min different because the role of the fibers was being alternated between Star and Star + lamp. Hence the fibers experienced different PSF distortions to the native spectrograph (see Sec. 9.4). Yet the resulting absorption lines yield reasonably accurate lines because it sums over several wavelets.

8.2 Measurement Comparing Star to Literature Spectrum

Figure 28 compares our EDI measurement of HD219134 with a spectrum obtained by another group with a conventional spectrograph, the NIRSPEC at Keck Observatory.[21] All the stellar NIRSPEC features are seen in the EDI output, but many are unresolved in the TripleSpec native spectrum (green curve).
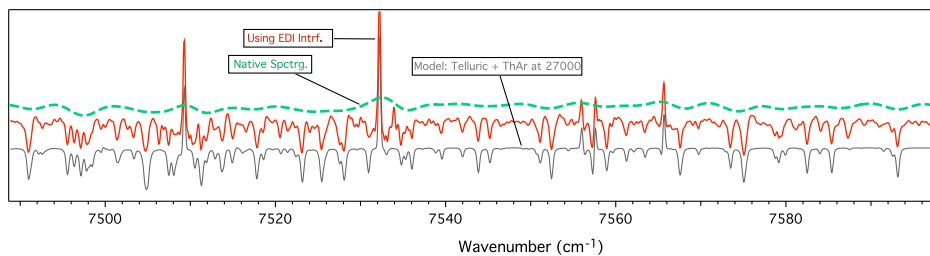

Fig. 22 Demonstration of a 10-fold resolution boost observing telluric features mixed into spectrum of star κ CrB along with ThAr calibration lamp emission lines. The green dashed (top) curve is the "ordinary" spectrum measured without the interference, having native resolution 2700. It cannot resolve any of the telluric features. The red (middle) curve is the EDI (TEDI) reconstructed spectrum measured with seven contiguous delays, up to 3 cm, and equalized to a Gaussian resolution of 27,000. The gray (bottom) curve is a model of telluric[19] and ThAr[20] features blurred to resolution of 27,000, showing excellent agreement with EDI reconstructed data. Resolution boosting occurs simultaneously across the full bandwidth (0.9 to 2.45 μm) of the native spectrograph (final resolution varies linearly with wavenumber times largest delay). Y-axis is intensity, vertically offset for clarity. Data from June 19, 2011.
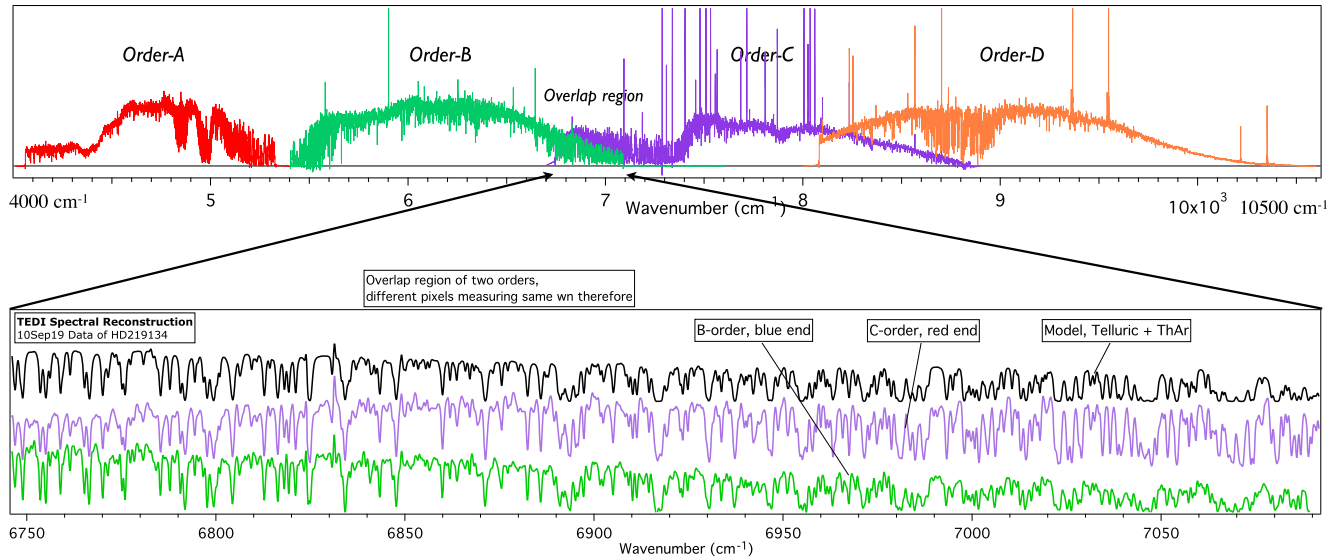
Fig. 23 The TEDI reconstructed resolution boosted spectrum spans the four orders of TripleSpec spectrograph in NIR (4100 to 10,500 cm^{-1}), here observing HD219134 on September 19, 2010. The lower graph is zoomed in region of overlapping orders B and C, showing agreement between two very unrelated groups of pixels (on opposite sides of detector), and with telluric model[19] (black curve).
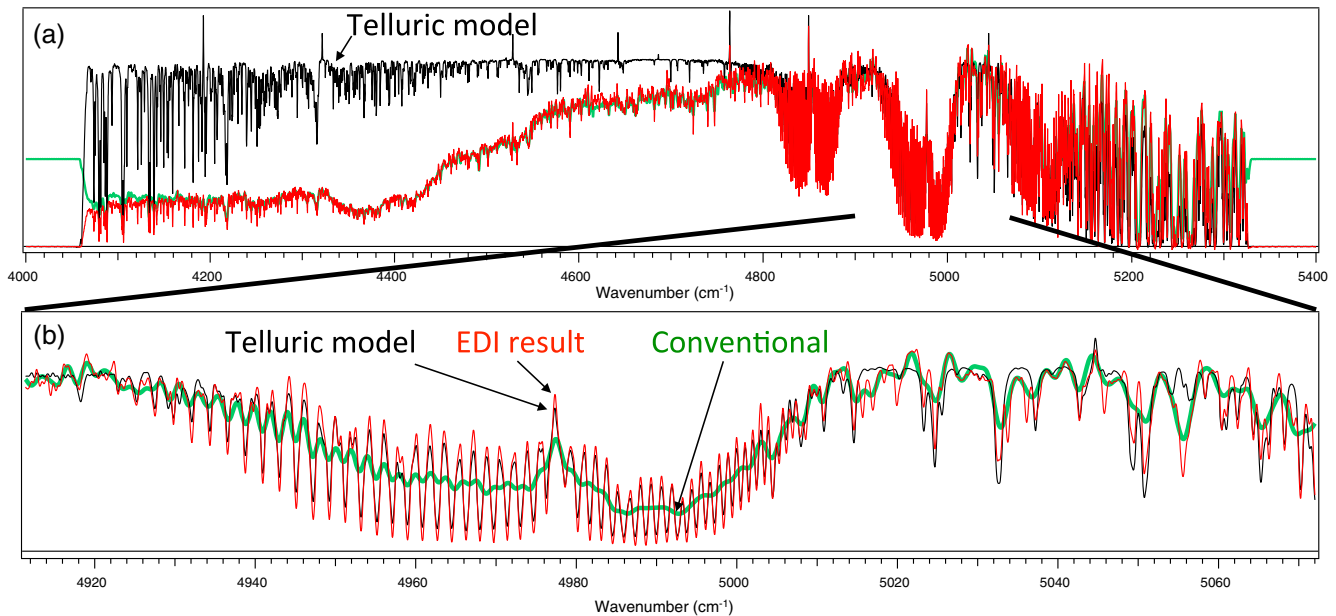

Fig. 24 A-order spectral reconstruction (SR) results having 4× boost (11,000 resolution) for star HD219134 + ThAr lamp, September 19, 2010. Black is telluric model, red is SR results, and green is conventional (native, ordinary, and nonfringing) spectrum. (a) Full order 4100 to 5300 cm^{-1}. The shape of the continuum is accurately reproduced (comparing red to green curves) in spite of the signal coming from two very different kinds of components, fringing versus nonfringing. The red and green curves are lower than the black curve at the low wavenumber side because of coloration of TripleSpec disperser. (b) Zoom in on a telluric feature near ∼4980 cm^{-1} due to CO_2 molecule.

8.3 Measurement Comparing Two Stars: Unexpectedly Deep Lines and Doppler Shift Value

Figures 29 and 30 compare two stars measured on same night, star HD219134 (blue curve) and HD962 (red curve). We observed the latter star because the SIMBAD database described it as a featureless A-type, and thus could indicate the telluric component of the spectrum. However, our observations disagree with SIMBAD in two ways for HD962, while agreeing for HD219134.

First, the depth of narrow stellar absorption features of HD962 is similar to that of HD219134, which is a K-type. So the designation of A-type for HD962 is suspect. This is
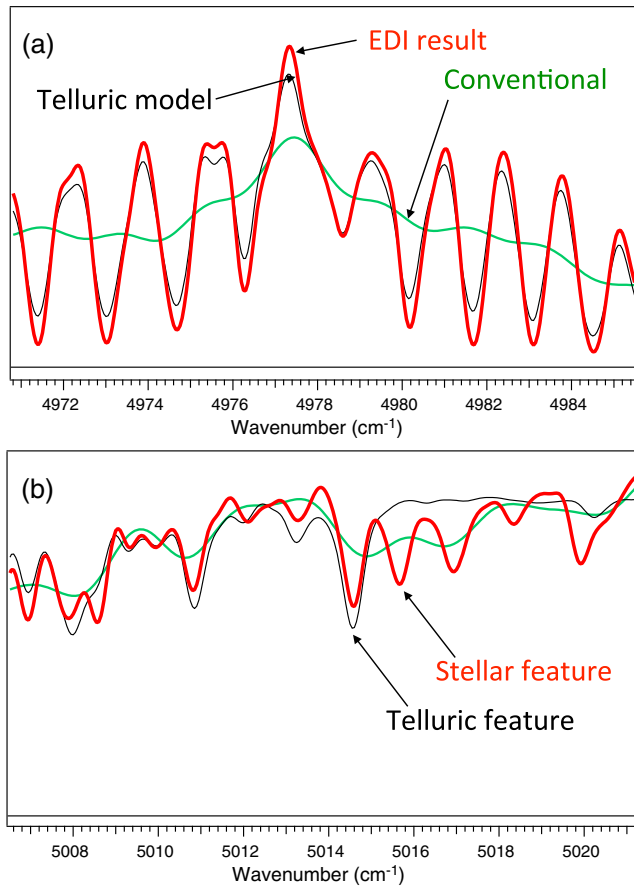
Fig. 25 Zoom-ins of A-order SR result (bold red) of Fig. 24. (a) Subtle features of telluric model (thin black), such as the tiny dips and detailed shapes of the peaks of CO_2 feature at ~4980 cm^{-1}, are beautifully reproduced. The ordinary spectrum (green) of native spectrograph cannot resolve these. (b) The current model does not yet contain stellar features, such as at 5016 to 5018 cm^{-1}, because of individuality of each star's Doppler velocity, elemental constituents, photosphere temperature, and pressure broadening. Target is star HD219134 + ThAr lamp, September 19, 2010.

seen clearly in Fig. 31 which compares the native spectra of HD962 to several A- and B-type stars (HR1178, Gam Oph, HD7891) and the K-type star HD219134.

All were measured on the same night September 19, 2010. Strikingly, HD962 has more similarity to the K-type star than the A- or B-type stars, in that it shares the narrow stellar absorption lines at 9450 cm^{-1}, which are absent for the three telluric calibrator stars. Interestingly, the three telluric calibrator stars all show (Fig. 31) a broader absorption feature at 9140 cm^{-1} not seen in the K-type, and the dispute star HD962 also shows this, although weakly.

Second, we observe a relative Doppler shift between the stellar lines of the two stars, but it is of the wrong polarity. According to SIMBAD, we should be seeing HD962 blue shifted relative to HD219134. Instead we see HD962 redder. The Doppler velocity of HD219134 that we measure agrees perfectly with the SIMBAD database when the barycentric motion is accounted for. The two stars were measured within an hour of each other, using the same data analysis procedure, and the telluric and ThAr features match up perfectly (see dips at 9188 and 9206 cm^{-1}). We checked the RA and Dec values of the FITS file header to confirm the telescope was pointed

at the right star. Thus we believe the SIMBAD database for HD962 is in error.

Some details of the Doppler calculation–Solar spectrum from FTS measurements at NIST are shown shifted by velocities 5.7 and 26.4 km/s to match the position of absorption lines. We observe geocentric blueshifts of 5.7 and 26.4 km/s for HD962 and HD219134, respectively. Given Earth's motion of 11.679 and 7.566 km/s toward HD962 and HD219134 at UT times 10:49 and 9:50, respectively, this implies heliocentric blueshifts of 5.7 to 11.679 = −5.98 km/s for HD962 and 26.4 to 7.566 = 18.83 km/s for HD219134. SIMBAD blueshift values are 23.0 and 18.83 km/s for HD962 and HD219134. Thus they agree nicely for HD219134, but are off by 23 + 5.98 = 29 km/s for HD962.

8.4 High Dynamic Range and No Fixed Pattern Noise

Figure 32 shows the extremely high dynamic range and robustness to FP noise of which EDI is capable. These are of our ThAr lamp to a 6× boosted resolution of 19,000 in the D-order. We resolve weak peaks of heights ~0.04% of the brightest line in that order. (Note the weak but distinct peaks at 9171 and 9240 cm^{-1}, for example.) The data are normalized so that the brightest line of the order, at 9548 cm^{-1}, defines unity height. The ~1e^{-4} noise at the base of the 9188 line (used to calibrate the equalization) is a dynamic range of ~10,000, which is excellent. There is very little ringing (~0.01%).

This data also illustrates EDI's robustness to FP noise. The inset is a small section of the 2-D raw detector image at the D-order associated with this data, showing a couple of bad pixels. These produce erroneous bumps in the native spectrum (green curve) at 9174 and 9214 cm^{-1}. The red EDI output spectrum does not react to these bad pixels, because they are constant and thus do not respond to the sinusoidal interferometer phase dithering.

8.4.1 Individuality of each ThAr hollow lamp

Interestingly, we see a discrepancy between the measurement of our lamp, and the NIST measurement[20] of their lamp, in the very weak lines (heights < ~ 1%). Using the line identifiers (purple letters) provided by NIST, this is explained by the Th of our lamp being weaker to the Ar lines by ~6×, compared to NIST. This is consistent with the reports[20] that each ThAr hollow cathode lamp can be very individual in character.

This stresses the importance of measuring each ThAr lamp and demonstrates that EDI is scientifically useful for this.

We now understand earlier noticed discrepancies in fringe phase between data and "theory" (NIST measurement) for some weak features, such as during calibration of our glass etalon. For blended lines, changes in relative amplitude of components (Th relative to Ar) can create phase errors.

9 Instrument Noise Behaviors

The TEDI data set is dominated by instrument noises at ~3%. Photon noise, at ~0.1%, is about 30 times smaller for our typical exposures. The detector also suffers from occasional bad pixels which give false peaks on the native spectra (as in Figs. 32 and 33). Figures 34–38 describe the extreme PSF drifts that afflict the native spectrograph. The EDI is robust to all these types of insults to the native spectrograph. This is evident by the high-resolution spectra it produces, which would not be possible with
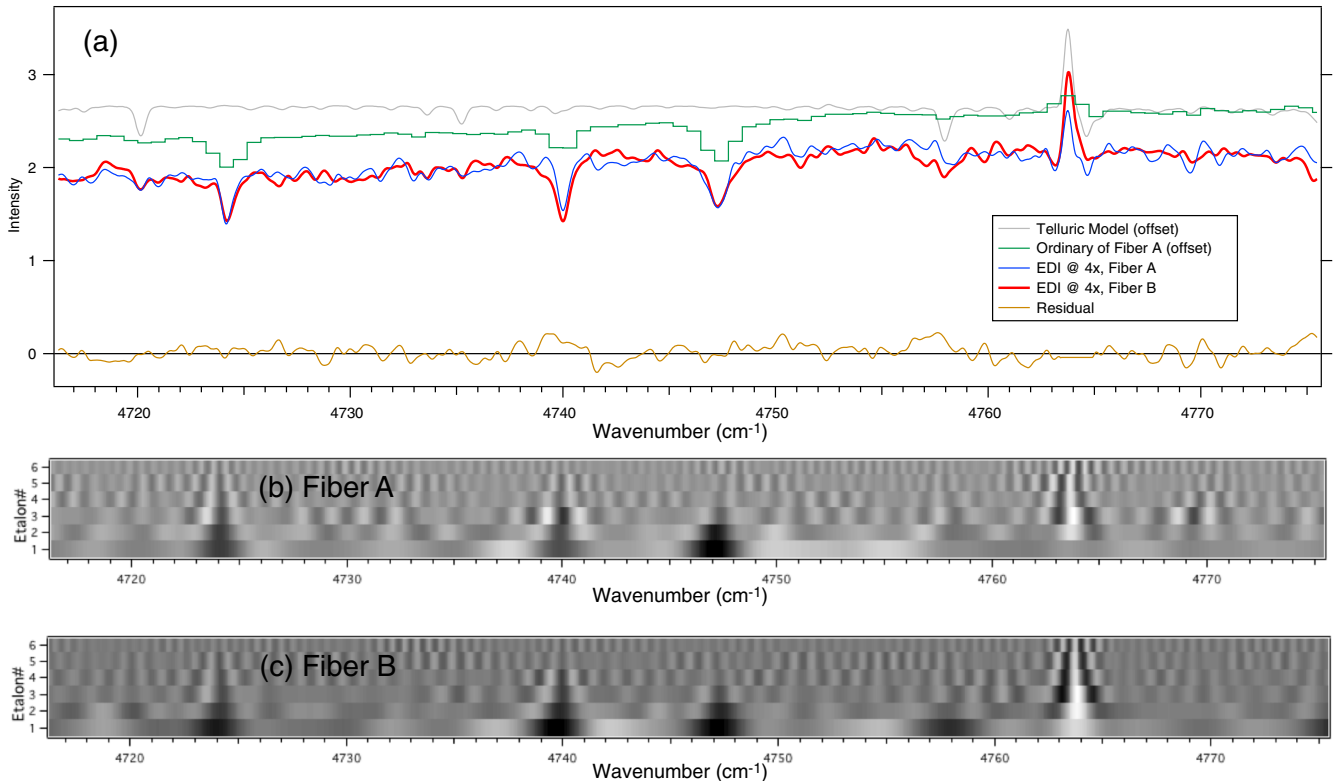
Fig. 26 Comparison of SR result from two fibers A (thin blue) and B (thick red), thus on different pixels, looking at same stellar source HD219134 in A-order at 4x boost, resolution of 11,000. The difference between fibers is the tan curve. The ordinary spectrum in native pixels is green curve. That and the telluric + lamp model (gray) are offset vertically for clarity. SR results detects the telluric and lamp features in the model, as well as many stellar features not in the model. The two fibers were measured a few minutes apart, on different exposures, since the roles (ThAr-only, ThAr + star) were alternated between the fibers. This figure shows 4715 to 4775 cm^{-1}. (b) and (c) show the wavelet arrays for the 6 significant etalons used. Fiber B appears to have better phase alignment than fiber A.

a conventional spectrograph under the same conditions. We will also explain the theoretical basis of the robustness.

9.1 Fixed Pattern Noise Rejection Advantage

9.1.1 Additive fixed pattern noise

Figures 32 and 33 are examples of EDI's extraordinary rejection of FP noises—only intensity changes synchronous to the phase dithering contributes to the whirl. Any constant intensity pixel will cancel in Eqs. (3) or (5) for $\mathbf{W}(\nu)$. Even if the user decides not to increase the resolution, i.e., a boost of 1×, the EDI result can be cleaner than the native spectrum. (If the photon noise dominates over FP noise, the native spectrum information should be included with the EDI fringing information.)

9.1.2 Multiplicative fixed pattern noise

If the data are not properly flatfielded, there can be a pixel-to-pixel gain variation artifact that changes the magnitude of the whirl. [However, the phase is not affected because of the constant error term cancels in Eqs. (3) or (5) for $\mathbf{W}(\nu)$ discussed above]. This is the same magnitude error imposed on the ordinary spectrum. For different delays, this magnitude error pattern will get shifted to different final frequencies. For Doppler velocimetry, a purely magnitude error is less harmful than a purely

phase error, since the location of an isolated spectral feature is mostly determined by the phase of the wavelets.

9.2 Severe Distortions of Native Spectrograph

Figure 34(a) shows that high-resolution spectra can still be produced in spite of native PSF drifts that are extremely large and irregular—locally as much as ~0.6 cm^{-1} as shown in Fig. 35, and averaged over 10 cm^{-1} in Fig. 37. Figure 35 is a stack of the eight native spectra showing drifts along the dispersion direction. There is a bulk drift versus time (Fig. 36) which we call shear, which is removed prior to heterodyning reversal. Note the significant residual irregular component, which even causes the shear polarity to flip over small wavenumber distances. Figure 38 shows how this horizontal noise is equivalent to an intensity noise of ~3%.

9.3 Robustness to Point Spread Function Widening

The EDI is very robust to both changes in the PSF position and width, such as due to atmospheric seeing disturbances or thermomechanical drifts. Let us first discuss the effect of an unexpected change in the width of the native lineshape. For concreteness, let us assume the net focal spot size is a sum in quadrature of the native slit width and some fixed additional blur, fixed in detector space, and that we have a 10× boost so
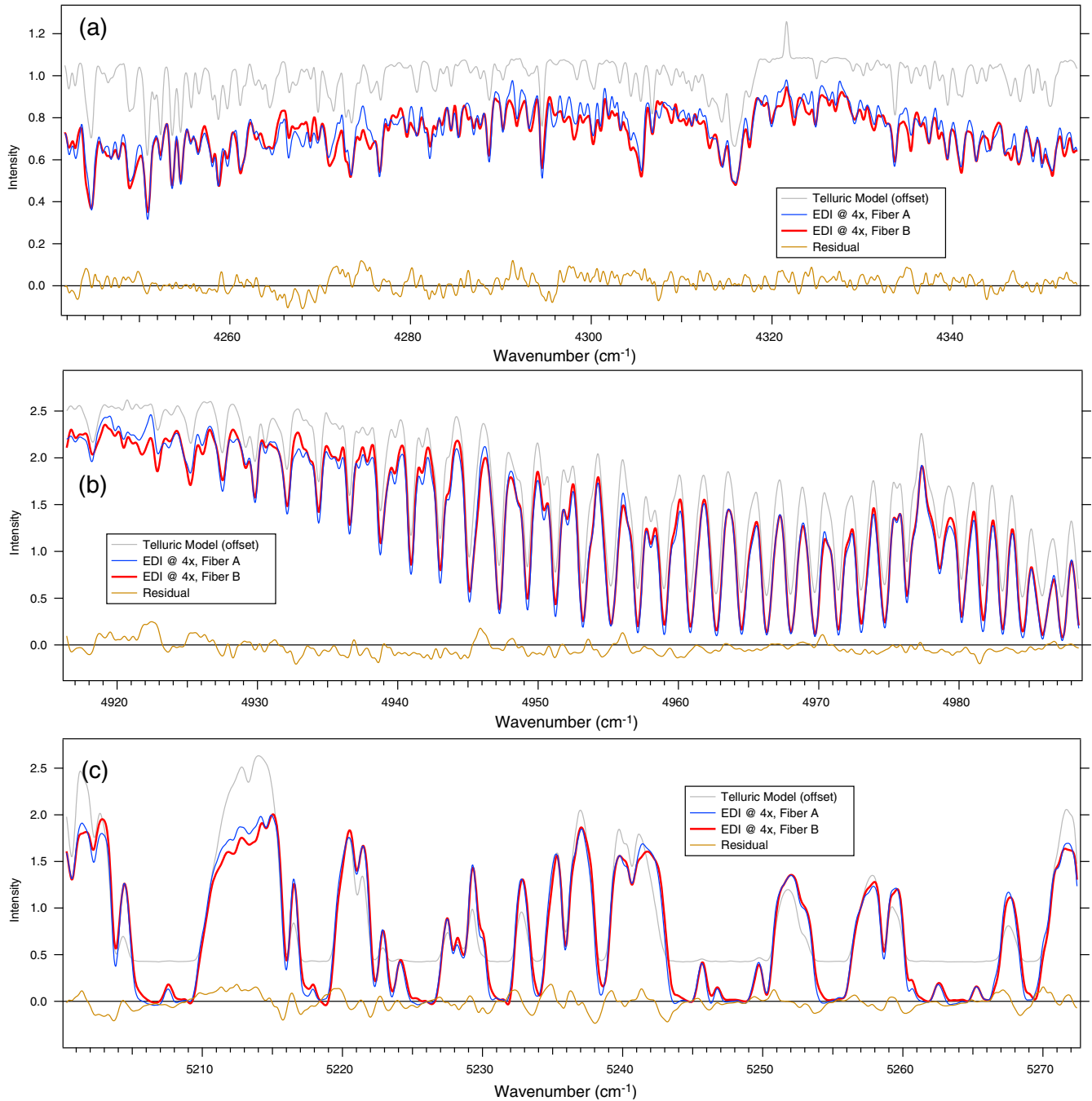
Fig. 27 Additional regions of the A-order when comparing two fibers A (thin blue) and B (thick red) and their difference (tan); same data as Fig. 26. (a) 4240 to 4354 cm^{-1}, (b) 4916 to 4989 cm^{-1}, (c) 5200 to 5272 cm^{-1}. Telluric model (gray) offset vertically.

that the slit width of the EDI native spectrograph is 10× wider than the conventional spectrograph it is trying to match.

This is numerically simulated in Fig. 39. The result is that EDI lineshape (red) is dramatically more robust than the conventional spectrograph (black) because EDI data are sensed at spatial frequencies defined by a physical delay, not controlled by focal spot size or slit width. The conventional spectrograph lineshape (black solid to dash) doubles and its MTF is halved in width. Whereas the effect on the EDI MTF is just to have slightly narrower bumps [solid to dashed red curves of (b)]. Since the original bumps are equalized

away, the EDI final resolution is mostly unchanged by the blurring.

9.4 Robustness to Point Spread Function Drift of Native Spectrograph

9.4.1 Translational reaction coefficient for single delay externally dispersed interferometry

Now we will discuss, at length, the EDI robustness to an unwanted translation or "insult" Δx (in wavenumber units) of
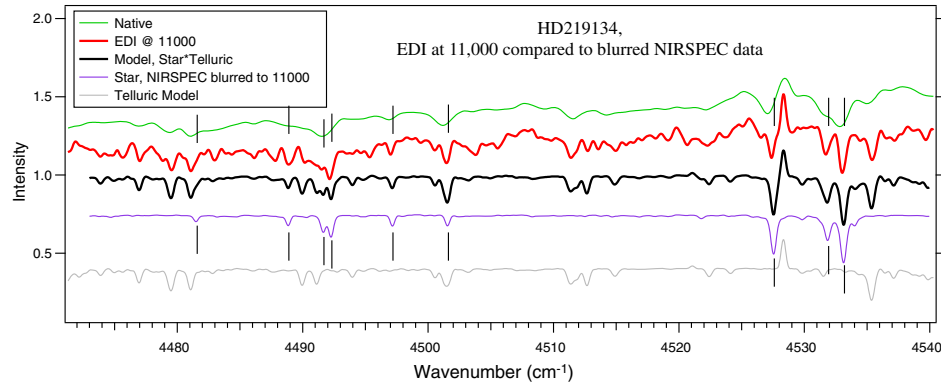
Fig. 28 Comparison of SR result (red) at 11,000 resolution on star HD219134 on September 19, 2010, to data measured independently by the NIRSPEC conventional spectrograph[21] at Keck Observatory on June 21, 2003. Black curve is the (star*telluric + lamp) model, where the star curve is NIRSPEC data blurred from its original 18,000 resolution to 11,000 resolution (purple), and is multiplied by the telluric and ThAr lamp model (gray). Native TripleSpec spectrum (green) cannot resolve stellar features (vertical lines) now resolved by the EDI. Wavelength shifts due to Earth and solar system motion not corrected.
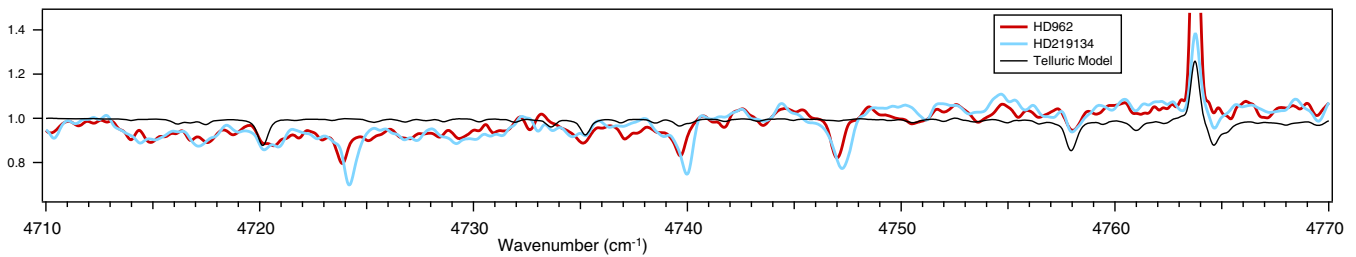
Fig. 29 Science result comparing spectra of two stars, in A-order, measured nearly simultaneously star HD219134 (blue curve) and HD962 (red curve). The panel shows only a very small 53 cm^{-1} region (4715 to 4768 cm^{-1}) of the 7200 cm^{-1} wide spectrum measured simultaneously. Telluric + lamp model[19] is black curve. This region is interesting because it contains a ThAr spectral emission line (4764 cm^{-1}) which appears in all three curves, telluric absorption features (e.g., 4758 cm^{-1}) which appear in all curves, and three stellar absorption features (4724, 4740, and 4747 cm^{-1}) which appear only in the stellar data not the model and which appear Doppler shifted one star relative to the other. Data were taken on both fibers and averaged. Effective resolution here was set to only 4× boost or 11,000 to reduce ringing to an extreme minimum, which aids identifying weak lines. Higher resolution can be used if some ringing is tolerated.
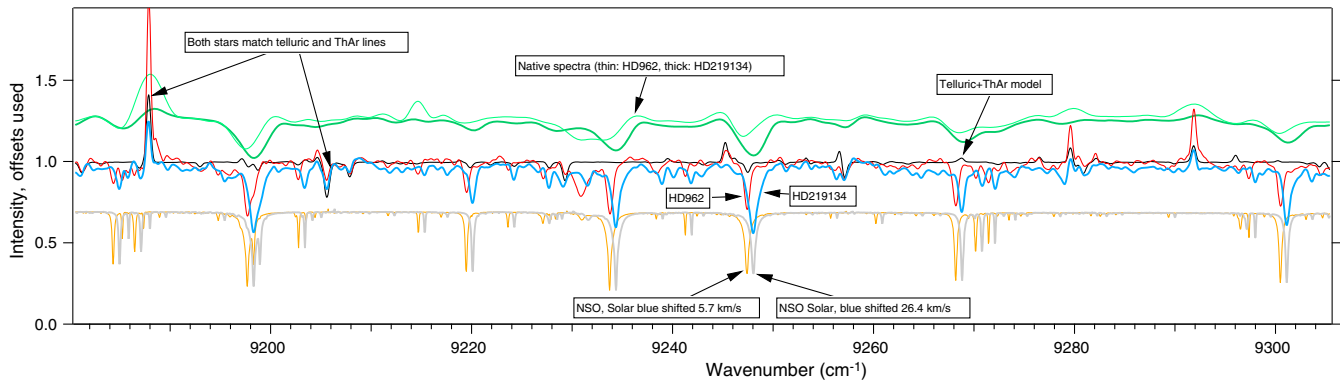

Fig. 30 D-order data version of the two stars of Fig. 29, comparing two stars measured nearly simultaneously, HD962 (red curve), and HD219134 (blue curve) on September 19, 2015, around UT of 10 h. They have resolution of 19,000, a 6× boost over native spectra (thin and thick green curves are of HD962 and HD219134). The SIMBAD database classifies HD962 as a featureless A-type star. However, we find stellar absorption lines of similar depth as to HD219134, which is K-type. Second, we measure for HD962 a Doppler velocity that differs significantly from its SIMBAD value, (29 km/s too red—the HD962 lines should be bluer than HD219134, not redder), whereas for HD219134 it agrees perfectly. The ThAr and telluric features of HD962 agree nicely with theory (see 9188 and 9206 cm^{-1}), and the native spectra are consistent with the EDI result, so we believe the measurement is accurate. The difference in Doppler velocity observed here in the D-order is the same as observed in the A-order in Fig. 29.
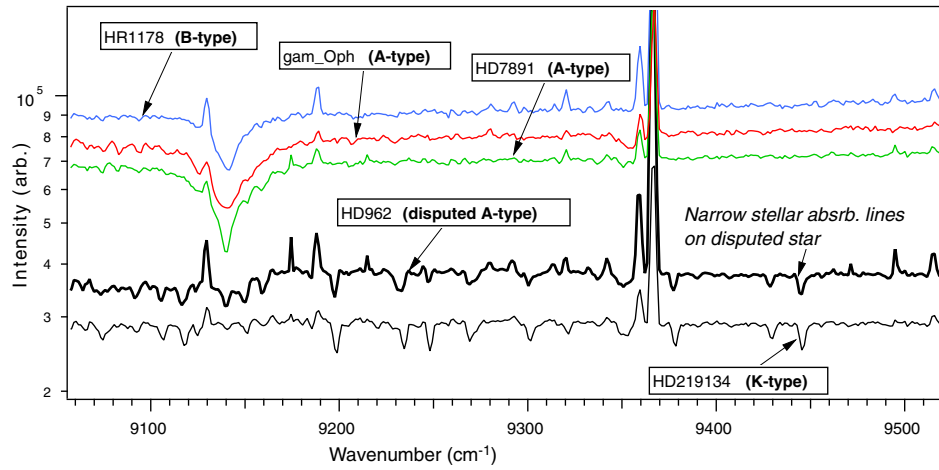
Fig. 31 Comparing the native spectra (star + ThAr lamp) of the disputed A-type star HD962 (bold black) with three other "featureless" (B and A type) stars (thin blue, red, green for HR1178, gam Oph, HD7891), and with K-type star HD219134 (thin black) shows that the disputed star is more similar to the K-type star regarding the narrowest stellar absorption lines (such as at 9450 cm^{-1}). However, there is a wider feature at 9140 cm^{-1} in the undisputed A stars that is weakly evident in the disputed star and not in the K-star. Curves are vertically offset for clarity, and the original continuum intensities for a single phase stepped exposure are 13.5k, 75k, 26.5k, 7.4k, 28.5k counts, for HR1178, Gam Oph, HD7891, HD962, HD219134.
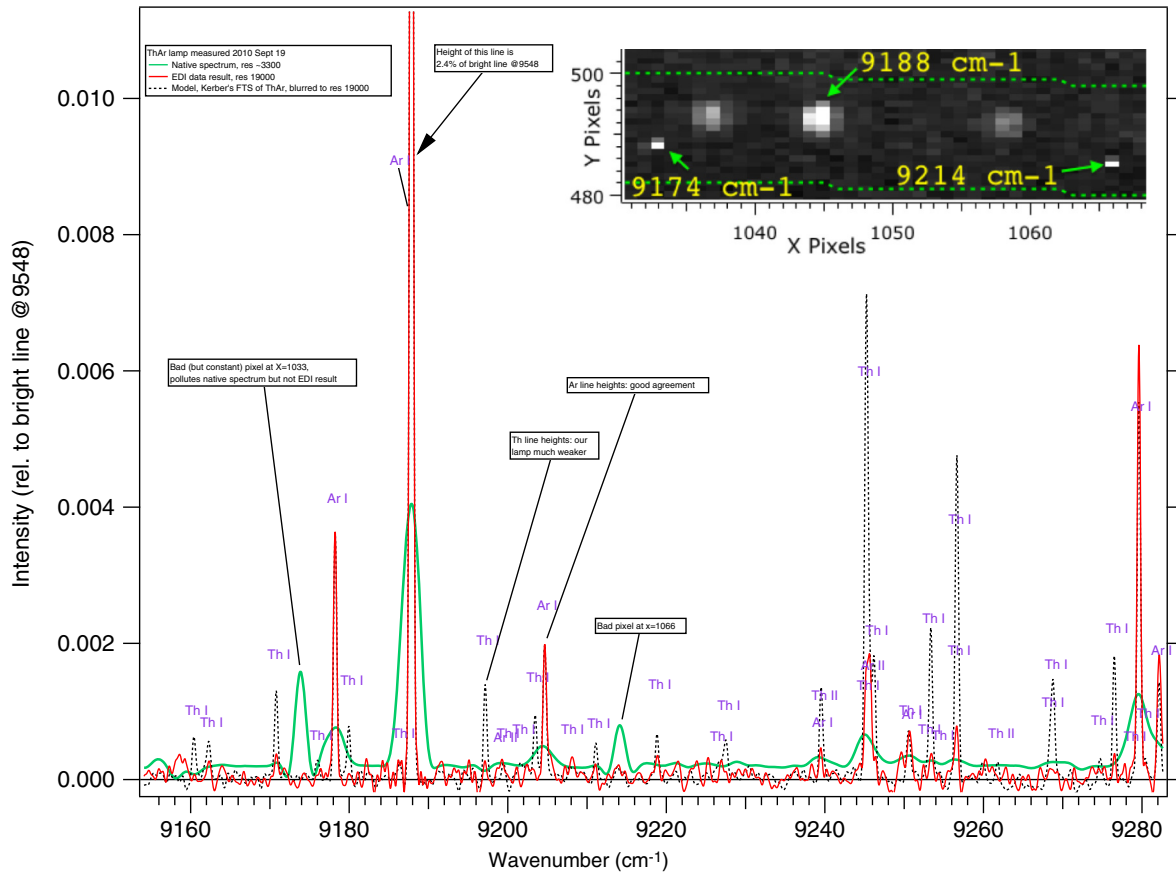
Fig. 32 Measurement (red curve) of our ThAr lamp to a 6× boosted resolution of 19,000, compared to NIST measurement[20] of a different ThAr lamp blurred to same resolution (black dashes), D-order. Purple text is their assignment of species. Native spectrograph (green curve) has resolution ∼3300. Note disagreement in heights of Th lines, whereas Ar lines agree well, due to individuality of each ThAr lamp. Note extremely high dynamic range of measurement—∼0.1% lines are easily observed (heights are fraction of 9548 cm^{-1} line). The EDI curve is robust to FP noise such as false peak in the native spectrum at 9174 and 9214 cm^{-1} due to bad pixels at $X = 1033$ and 1066 (inset shows detector there). Bad pixels are constant and thus do not affect whirls which look at changes between exposures.
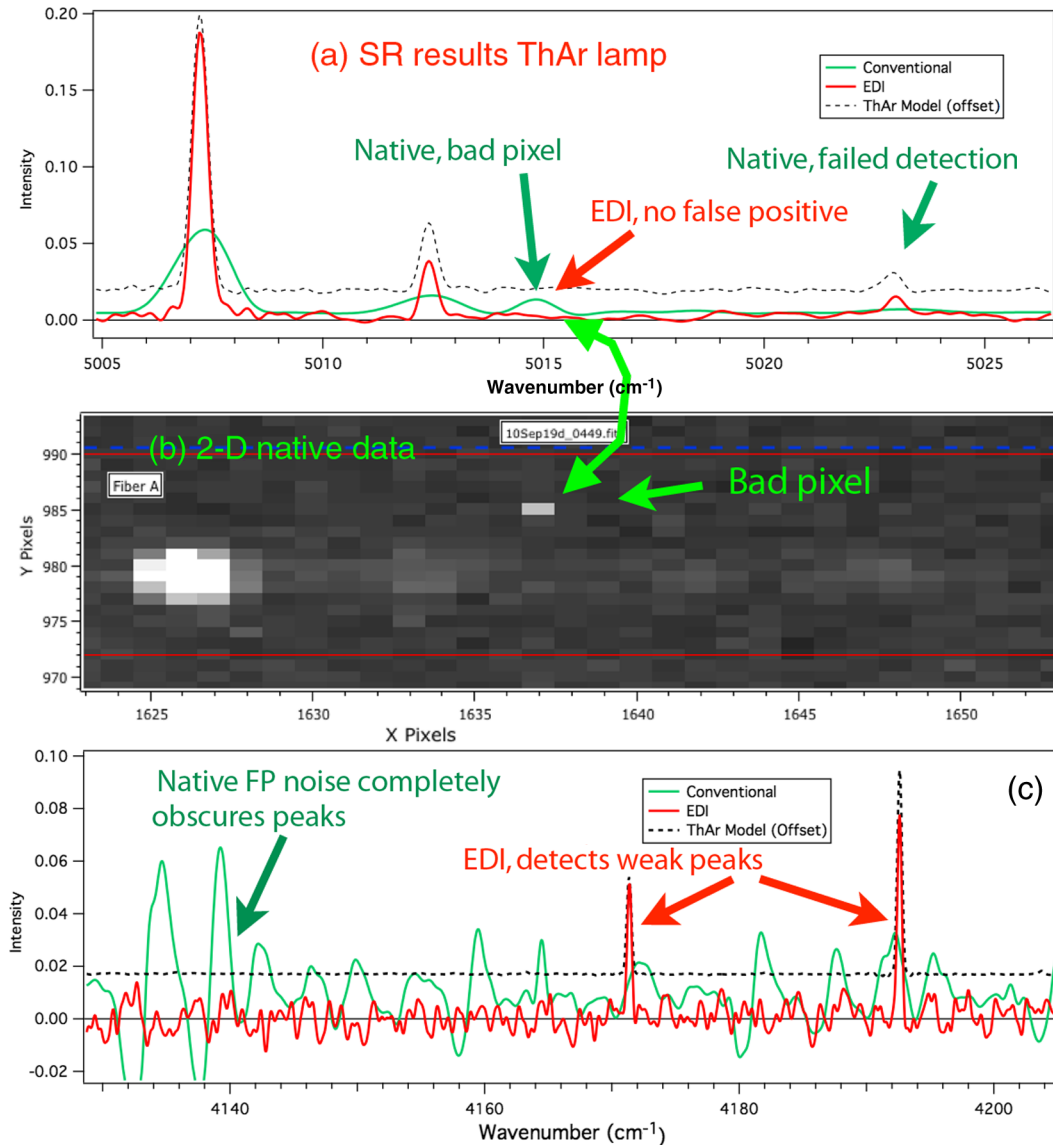
Fig. 33 EDI is robust to FP errors that afflict the native spectrum, such as the false peak at 5015 cm^{-1} (a) caused by a bad pixel (1,637,985) shown in raw 2-D data (b) of A-order. The EDI result (red curve) does not respond to this bad pixel because it does not vary sinusoidal with phase in the 10 exposures. (c) Another example is at extreme red end of the A-order (<4200 cm^{-1}). The EDI is able to detect two very weak ThAr lines at 4172 and 4193 cm^{-1}, whereas the native spectrum is not. Intensity axis is fraction of the 4849 cm^{-1} line.

the native spectrograph PSF along the dispersion direction on the detector, producing a perceived shift $\Delta\nu$ in the final output spectra. We define a translational reaction coefficient (TRC) $\equiv \Delta\nu/\Delta x$. For a conventional dispersive spectrograph, TRC = 1.

For a single delay EDI, the TRC is smaller than unity by a factor of several, and can be zero, depending on the shape of the spectrum where TRC is being evaluated. For example, at an isolated narrow feature such as a ThAr line, it is zero. This is because the moiré pattern has a phase which has no slope along the dispersion direction, so translating the moiré pattern does not change the perceived phase, and hence perceived ν.

However, for a pair or more of lines that are partially blended together by the native blur, there can be a slope to the moiré pattern. This slope, acting against a translation Δx, can produce

a phase change which leads to a nonzero TRC (but still much smaller than unity by a factor of several). How TRC behaves on average for an arbitrary spectrum is best seen in Fourier space, by multiplying the peak shape at a delay τ of psf$(\rho - \tau)$ by a line $(\rho - \tau)$ running through the center of the peak [Fig. 40(b)]. A similar line $(\rho - 0)$ runs through the native spectrograph psf(ρ) in (a) to calculate its reaction to Δx. We have colored green similar regions of ρ in the panels to suggest how the net reaction for the native spectrograph (a) is several times that of the EDI peak in (b), because the location of multiplier line is at much higher ρ for the EDI, because the heterodyning effect moves it, along with the peak, to be located at $\rho = \tau$.

Figure 40 is notional and does not yet multiply the peak shape against the line. That is done in Fig. 41(a). This latter graph actually depicts the net reaction (bold black curve) for
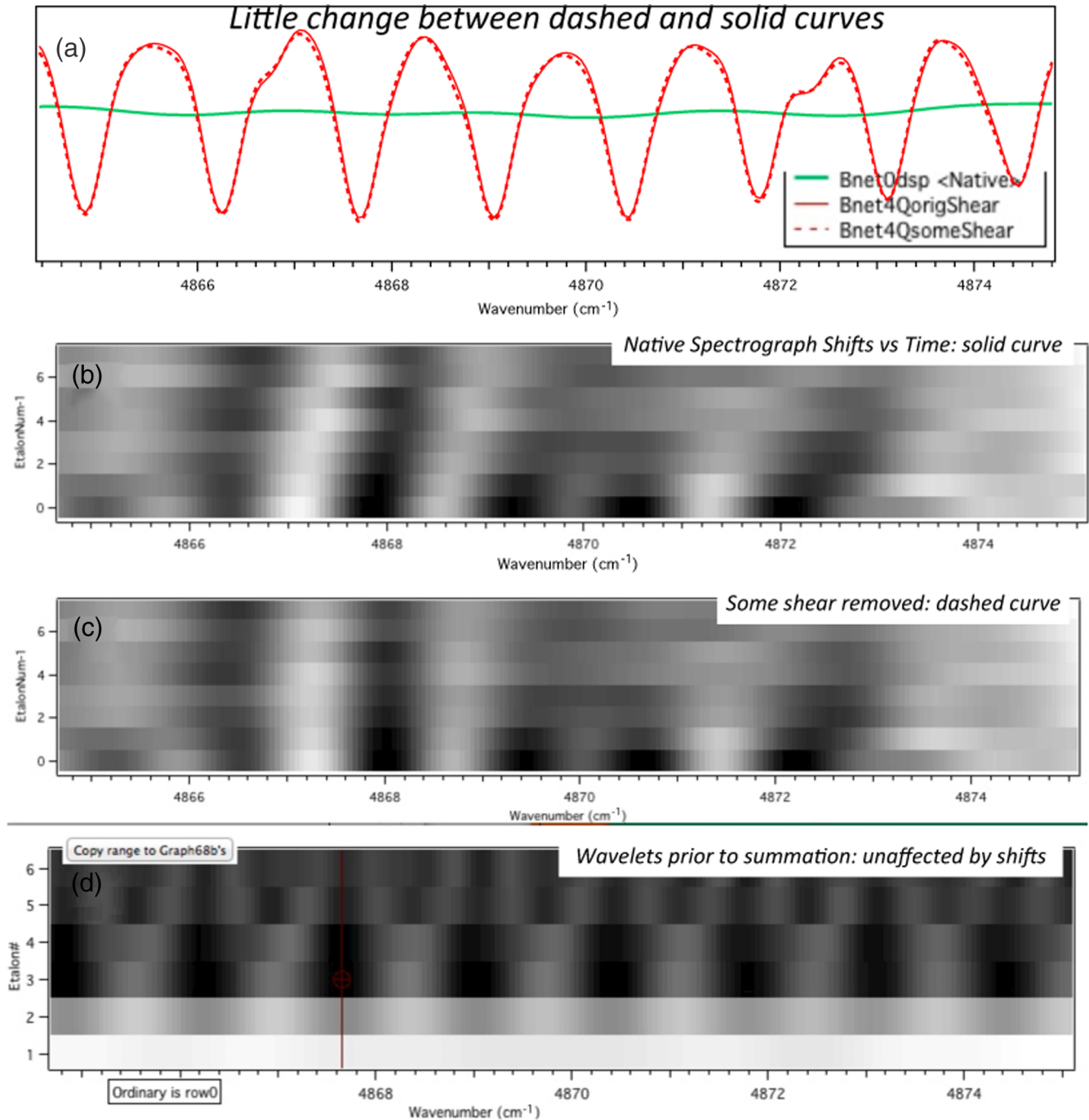
Fig. 34 Demonstration in TEDI data of extreme robustness against irregular PSF shifts to the native spectrograph, due to use of multiple delays. (a) The sixfold boosted spectral resolution (16,200 from 2700) is achieved for the reconstructed spectrum (dashed and solid red curves) despite significant irregular wavenumber shift versus etalon number (Y), hence time, and versus wavenumber (ν) of the data. (b) and (c) Native spectra versus etalon, (c) has bulk shear removed compared to (b). (b) Yielded solid red curve in (a), (c) yielded dashed red curve—little difference. Wavelets (d) summed to produce outputs (a). Green curve is native spectrum. Data of star HD219134, September 19, 2010 A-order.

multiple delays, but if you look at the reaction curve under one of the two highest delays, E7 or E8, they are isolated, and thus depict what happens for a single delay. This net reaction curve (bold black) in Fig. 41 is a phase change in ρ space, which manifests translations in ν space. The black and gray dashed slopes depict how that phase would change for TRC = 1 and TRC = 0.05, respectively. Note that the reaction curve for E7 at 3 cm is roughly below the TRC = 0.05 slope, and would

be under TRC = 0.02 for the highest delay E8 at 4.5 cm. The TRC scales as $1/\tau$ for a single delay, and would be of order 0.15 for the smaller delay E4 at 1 cm.

9.4.2 *Translational reaction coefficient for multiple delays*

We find that with the use of multiple delays, and without any special lineshape reshaping, we can reduce the $\Delta\nu$ per Δx by
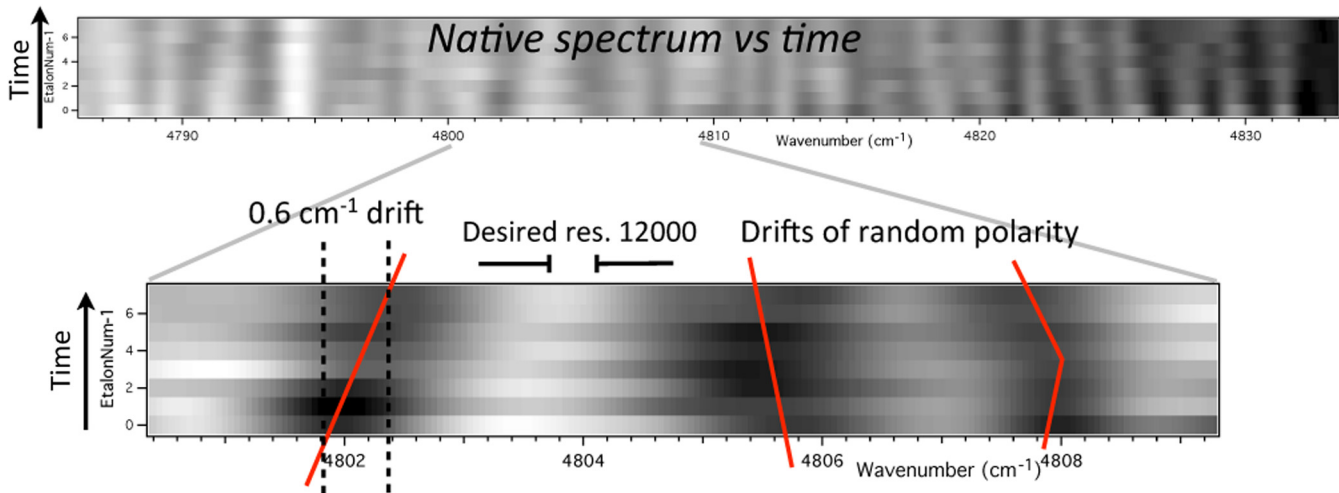
Fig. 35 Large PSF irregularity in the native spectrograph to which EDI is robust. Row is native spectra of A-order versus etalon # (~10 min apart). The PSF shift ~0.6 cm^{-1} versus etalon # is called "shear." Bulk shear has already been removed here, and an irregularity remains. Note irregular change in shear polarity. Fortunately, for the EDI the native lineshape only affects the envelope of the wavelets, but not the internal shape.
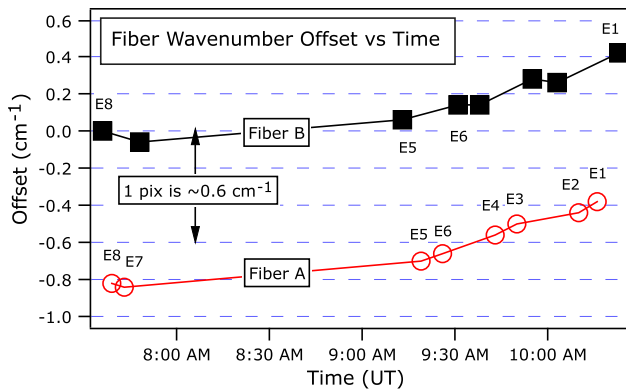

Fig. 36 Bulk translation of the native spectrum along the dispersion axis due to thermomechanical or changing gravity vector, and due to slit slant acting on vertical separation of fibers A and B, for star HD219134 on September 19, 2010.
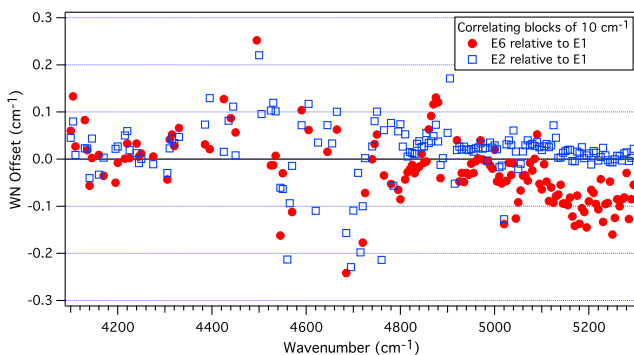

Fig. 37 Irregular PSF drift of native spectrograph, versus wavenumber, correlated in 10 cm^{-1} chunks between E6 and E1 (red dots), and E2 and E1 (blue open squares), for the A-order and September 19, 2010, data of HD219134. The time-dependent bulk offset (Fig. 36) has already been removed.

roughly a factor of 20. The idea is that having multiple overlapping peaks will cause the negative reaction on the left side of one peak to cancel with the positive reaction on the right side of a neighboring peak. This is notionally shown in Fig. 40(c), and calculated for actual TEDI peak positions and original shapes in Fig. 41(a). The bold black curve is consistent with TRC = 0.05 in the region of the overlapped peaks E1 to E6, 0 to 2 cm. This agrees with the simulation of Fig. 42, even though this was done in a different manner—it calculates movement of a specific feature rather than estimate for an arbitrary spectrum, and does it in wavenumber space rather than Fourier space. It follows a ThAr line in TEDI data and computes $\Delta\nu$ when all etalon rows of the moiré data are artificially shifted by $\Delta x = 0.5$ cm^{-1} to the left. The result is a $\Delta\nu = 0.025$ cm^{-1} shift, for a TRC = 1/20. This means the Doppler noise would be 20× smaller than if using the native spectrograph alone (if Δx is the only noise insult type).

10 New Crossfading Method to Cancel Point Spread Function Drift

The factor of 20× improvement in the robustness to a PSF shift Δx is remarkable, since it can potentially reduce radial velocity Doppler noise by the same factor (if Δx is the only noise insult mode). What is even more remarkable is that with a simple reshaping of the lineshapes, we can theoretically achieve TRC = 0. We have realized this only recently. We present the following theoretical explanation and simple simulation below.

10.1 General Lineshape for Two Crossfading Peaks

To engineer perfect cancellation of the net reaction of overlapped peaks, the shape and width of the peaks are modified into a so-called "crossfading" shape. The original psf is reshaped to a new psf by multiplying the data in Fourier space with a multiplier $[\mathrm{psf}_{new}(\rho)/\mathrm{psf}_{orig}(\rho)]$, as in Fig. 43(a) from a typical peak with gradual wings into a triangle crossfading shape having a well-defined width. For the dual peak crossfade discussed here (mixing more peaks is possible), the most
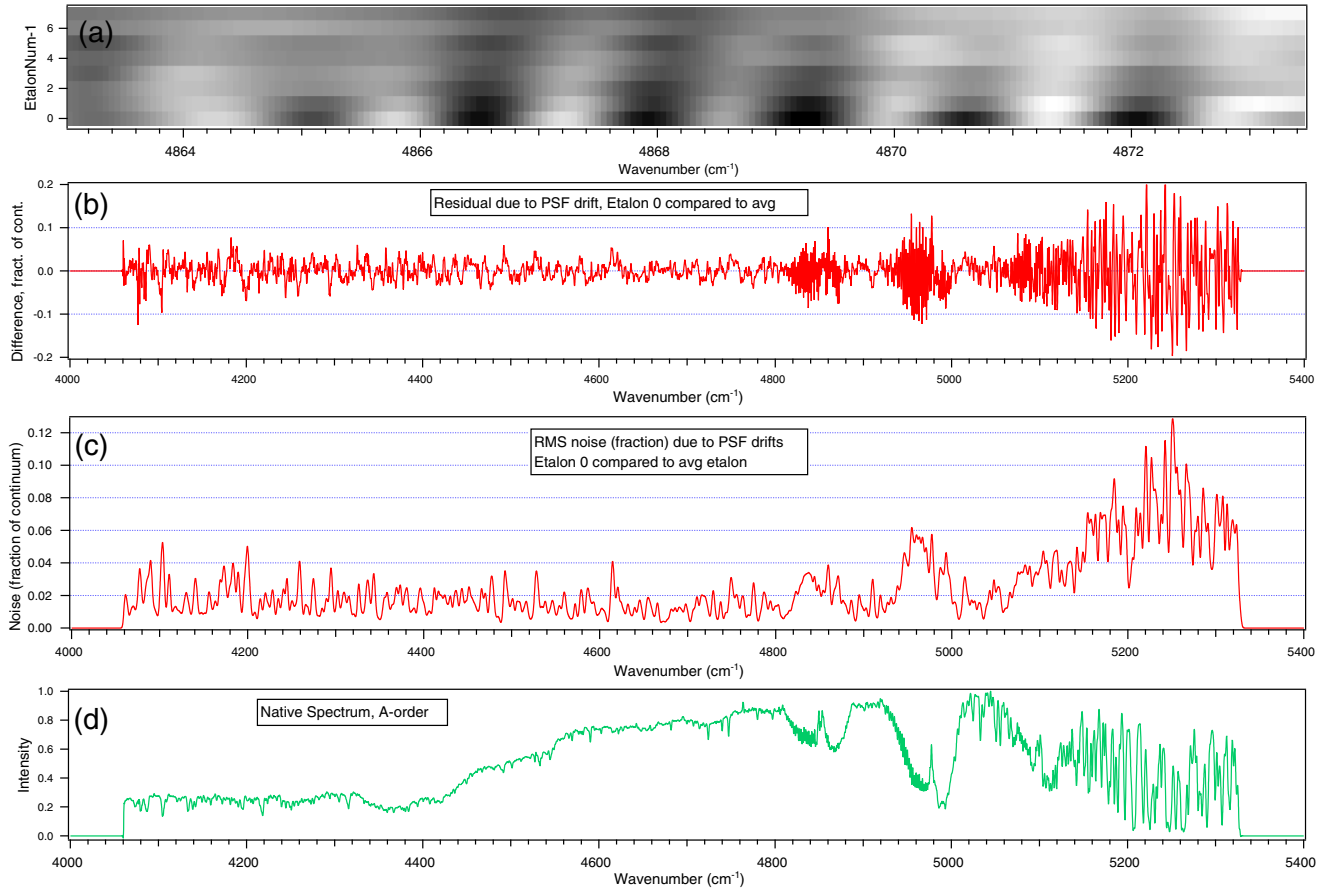
Fig. 38 (c) Measured amount of (vertical) instrumental noise for A-order, as a fraction of continuum, due to (horizontal) PSF drift or width change, by subtracting the native spectrum averaged over all eight etalons (d) from each spectrum (a) measured individually for each etalon to produce about a 3% residual (b).

important consideration is that the edge of the peak$_1$ (called a "foot") goes to zero magnitude at the center position τ_2 of the neighboring peak$_2$, and stays zero beyond that.

From the isolated peaks E7 and E8 of Fig. 41(a), one can see that each peak produces an odd symmetry sine-like contribution to the reaction curve, with the right side being positive and the left is negative. The new idea is to reshape the peaks to overlap so that the positive contribution of one peak perfectly cancels the negative contribution of the neighboring peak. This is shown in Figs. 43(b) and 43(c) for the simple case of a triangle lineshape. Between the two peaks, i.e., between τ_1 and τ_2, a crossfading is said to occur.

There are a family of curves that satisfy this condition of producing TRC = 0 in this crossfade region between the peaks. Let us derive the equation that describes them. The net reaction to a horizontal PSF drift is the product of the line $(\rho - \tau_n)$ through the center of the peak, times each peak. For n peaks this is

$$\text{Reaction}(\rho) \propto \sum (\rho - \tau_n)\text{psf}(\rho - \tau_n). \tag{35}$$

To produce perfect cancellation between two overlapping peaks of same shape $\text{psf}(\rho - \tau_1) \equiv f(r)$, and $\text{psf}(\rho - \tau_2) = f(r - 1)$, and using a dimensionless frequency $r = (\rho - \tau_1)/|\tau_2 - \tau_1|$ centered at the left peak, #1, we must satisfy

$$rf(r) + (r - 1)f(r - 1) = 0 \tag{36}$$

which is equivalent to satisfying, when $F(r) = rf(r)$,

$$F(r) + F(r - 1) = 0. \tag{37}$$

10.2 *Case of Triangle Lineshape*

The simplest case is a triangle lineshape, separated by distance 1 and height 1, which thus has linear crossfading of unity slope between the peaks. Then the right side of the left peak is $f(r) = 1 - r$, so the first term of Eq. (36) is $r(1 - r)$, and the left side of the right peak is r so the second term is $(r - 1)r$. Hence Eq. (36) is satisfied since $r(1 - r) + (r - 1)r = 0$, and in the region between the two peaks TRC = 0.

We apply this crossfade reshaping to every pair of overlapping peaks to fill a contiguously overlapped ρ region entirely with TRC = 0. For the case of irregular τ_n, the peaks will be asymmetrical, treated as half-peaks, in order to always place the foot on the center of the neighboring peak and vice versa. That is, the left half of peak at τ_{n+1} is the reflection of right half of peak at τ_1, and these can be a different shape than neighboring pairs of half-peaks. So one should view a manifold of peaks as instead a manifold of crossfading pairs of half-peaks.
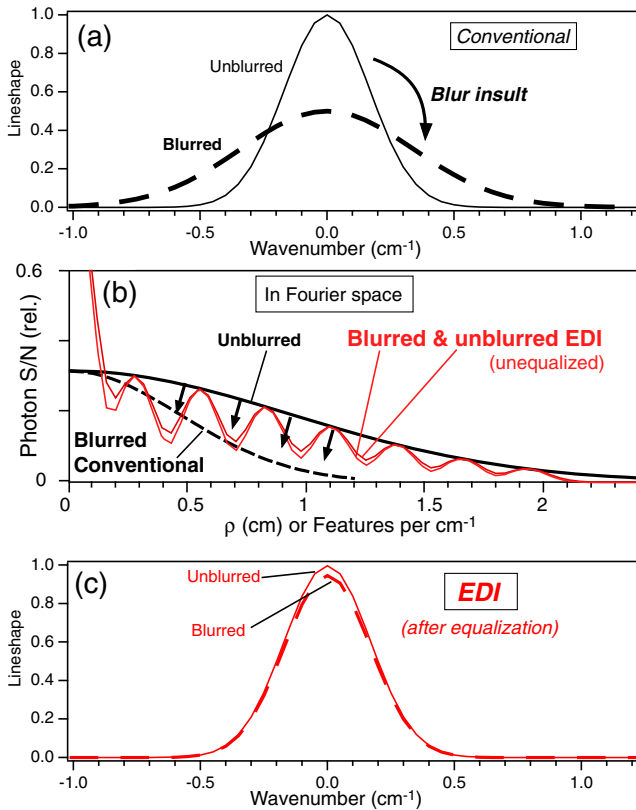
Fig. 39 Simulated effect of adding focal blur to the native instrument lineshape for conventional spectrograph (black), and an EDI (red) having a 10× boost—the EDI is much more robust. The conventional spectrograph lineshape (black solid to dash) doubles (a) and its MTF (b) is halved in width. For EDI native spectrograph, already 10× wider slit width, relative broadening is very small ($\sqrt{1^2 + 0.173^2} = 1.015$), so change in bump shapes (b) also small. Note, the EDI still retains significant high frequency sensitivity. The 1.5% change in bump width creates a 5% error in equalization in removing the bumps, creating two ~1% height satellite peaks at 3.5 cm^{-1} from central lineshape (off graph).

10.3 Other Crossfading Lineshapes

Colleague Eric Linder pointed out[22] that a sinc function, $f(r) = \sin(\pi r)/r$, for just the region between its first zeros, also satisfies Eq. (36). This produces $F(r) = \sin(\pi r)$, which has the required property that shifting by 1 flips its polarity. This has the advantage of a rounded peak that avoids some possible ringing produced by the sharp tip of a triangle. Interestingly, the sinc function (but with its many satellite peaks) is the Fourier transform of a rectangle and hence is naturally the psf of a slit opened much wider than its spectrograph focal blur. We are studying the utility of other crossfading shapes that satisfy Eq. (36).

10.4 Crossfaded Multiple Delays Produce TRC = 0

Figure 41(b) is the same calculation of net reaction as (a), but using sinc function peaks instead of original peak shapes, and having asymmetrical shapes to satisfying the foot to center condition for the irregular delay spacing. Remarkably, this produces TRC = 0 over the contiguously overlapped region from E1 to E6 (up to 1.8 cm). This means that for resolutions that only need 0 to 2 cm of frequency (6× boost, or ~16,000 in the A-order), the interferometer is completely decoupled from the native
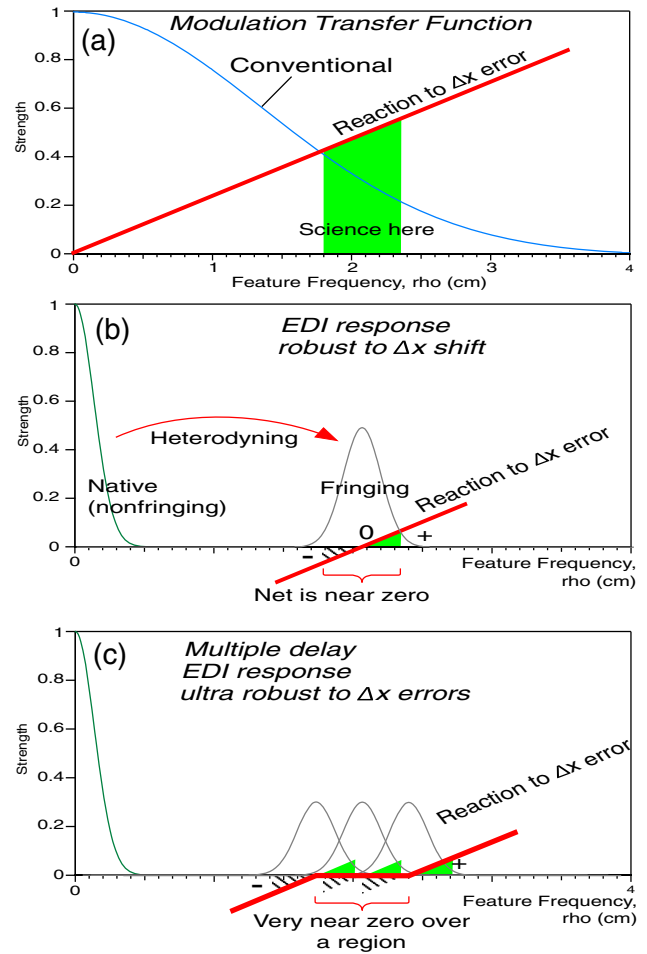

Fig. 40 (a) The reaction of a conventional spectrograph to an unwanted displacement Δx along the dispersion direction at the detector, in Fourier space versus frequency ρ, is to multiply the psf(ρ) or MTF(ρ) by ρ (red line). Science frequencies typically are high (green band), producing a large reaction. (b) In contrast, for the EDI both the sensitivity peak and red line are shifted to high frequency by the interferometer delay, and the reaction flips polarity as ρ crosses the center of the peak. This can reduce the integrated reaction by an order of magnitude. (c) With a multiple delay EDI the net reaction is further reduced when the positive and negative portions of the peaks overlap.

spectrograph and not disturbed by changes in the native PSF location. (In principle, three new etalons at 2.3, 3.4, and 4.1 cm to fill gaps between E6, E7, and E8 could be added to extend the contiguous delay region to 4.6 cm, greatly increasing resolution for TRC = 0 operation to ~45,000).

Crossfading can include the native peak, but that is not done here since the smallest delay E1 provides a cleaner signal of the lowest ρ than native spectrograph peak (robust to the FP noise seen in the native).

Crossfading has not yet been applied to bulk TEDI data. How well one can model the actual peak shape, and how much it varies across the band is under study.

10.5 Simple Crossfading Simulation Using Two Delays

Figure 44 shows a simple demonstration using a synthetic source spectrum consisting a group of evenly spaced lines (a),
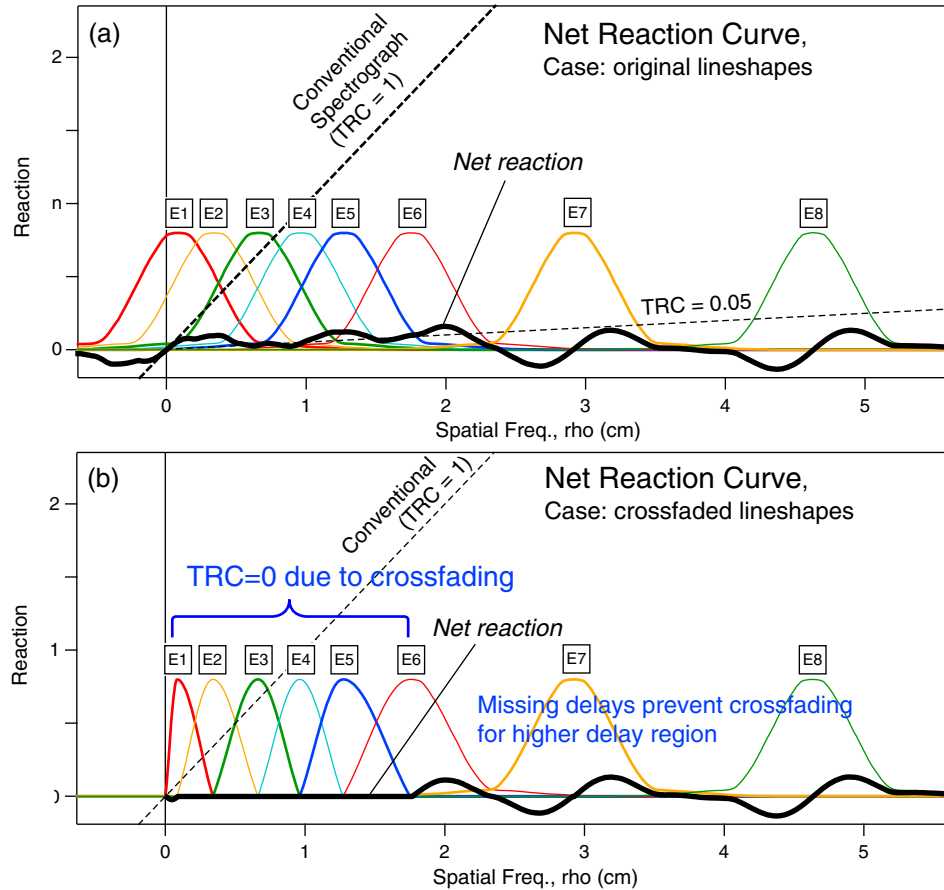
Fig. 41 Net reaction (bold black curves) to a PSF Δx shift for set of TEDI peaks, for (a) original lineshape and (b) peaks reshaped into a sinc function crossfading lineshape (triangle shape also works). Reaction is peak n shape (colored) multiplied by line $(\rho - \tau_n)$ passing through peak center, then summed over n. (a) For original peak shapes, net reaction is consistent with same slope (gray dashes) of TRC = 0.05 $(\Delta\nu/\Delta x)$ seen in Fig. 42. (b) Reshaped peaks cancel each other's reaction so TRC = 0. Remarkably, PSF can now drift in wavenumber without shifting the output spectrum—not possible in conventional spectroscopy where TRC = 1 (black dashes). Missing delays at 2.3 and 3.8 cm prevent crossfading above 2 cm (but TRC still less than 0.05). The TRC = 0 region 0 to 1.8 cm allows 6× boost (16,000 resolution) in the A-order at 4700 cm^{-1}. Crossfaded lineshapes are asymmetrical to achieve foot to center positioning for irregular delays τ_n.

so that a reasonably narrow frequency range around 1.4 cm is produced, so that only two peaks that straddle this frequency, at 1.2 and 1.6 cm, are needed. The source spectrum and moiré pattern are calculated twice, one in the original position, and again after artificially shifted by $\Delta x = 0.75$ cm^{-1} to the left.

A key point is that the moiré pattern changes its slope (phase versus wavenumber) when the delay passes through 1.4 cm, from 1.2 cm (b) compared to 1.6 cm (c). Since the moiré flips its slope polarity, the idea is to mix the two results to produce a net cancellation to the insult of a Δx translation. The mixing happens not to the data in moiré form, but after it has been processed into an output spectrum (d). Both a triangle and sinc function crossfading shape were tested (sinc results are shown), and the triangle produces slightly more ringing.

The average of the 1.2 and 1.6 outputs form the final outputs shown in (e), for original (blue dash) and forcibly shifted (black) cases. These two cases overlap in position extremely well, producing a perceived shift of only $\Delta\nu = 0.0022$ cm^{-1} for a TRC = 1/350, which is remarkably small.

10.6 Correlated Point Spread Function Drifts Needed

The above crossfading method presumes that the same Δx PSF drift applies to the two overlapped delay peaks. Otherwise the cancellation will be incomplete. This is naturally satisfied by long-term drifts in a TEDI style apparatus which measures delays sequentially on the same pixels using different glass etalons that are rotated into the interferometer beam, where "long" means longer than the time to measure each delay (typically several minutes).

We propose that shorter term drifts would be more correlated between peaks using a stepped mirror (Fig. 45) which can be designed to manifest several delays simultaneously from the same source, although on different detector pixel rows. Then comparison of original (having shift) to crossfaded data (having no shift) yields information on the PSF shift Δx, which thus could be at short times scales as well as long. However, this would be susceptible to any row dependence to the Δx drifts (such as thermally induced on the detector, and thus likely to be long time scale).
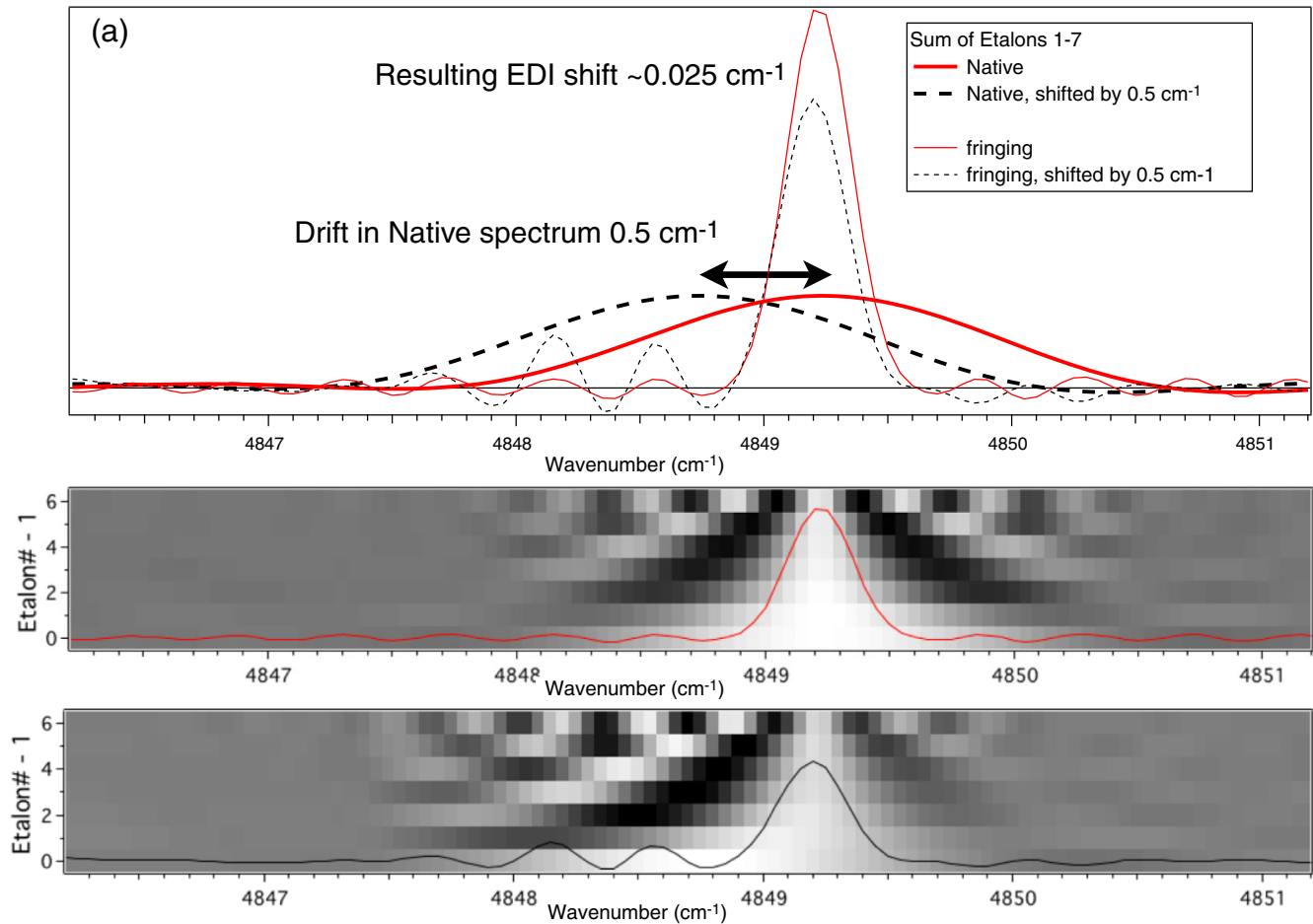
Fig. 42 Simulated EDI reaction to a native spectrograph PSF translation using ThAr data and TEDI data processing. Moiré input data become wavelets—a stack of seven is shown in (b), which sum to produce a narrow peak in (a) (thin red curve) at ~16,000 resolution (6× boost). (Unequalized gap between E6 and E7 delays produces slight ringing.) The effect of a PSF shift is simulated by translating every moiré input by $\Delta x = 0.5$ cm^{-1} to the left (c). Note shift in the envelope of the wavelet stack yet internal peaks are stationary. The summed peak (thin black) moves only $\Delta\nu = 0.025$ cm^{-1} relative to the unshifted version (red peak), hence the TRC, $\Delta\nu/\Delta x = \text{TRC} = 0.05$, a 20× reduction.

By combining in the same apparatus a stepped mirror with a sequentially changed etalon similar to TEDI, information on short timescale variations from the stepped mirrors but on different pixel rows could be joined with that sequentially measured on the same rows. We speculate that this could produce a near zero TRC apparatus for Doppler velocimetry or high-resolution spectroscopy that was robust both on long and short timescales. The details of how to combine the data to best tease out the desired information are still under investigation.

10.7 *Implications for Precision Radial Velocimetry*

In light of the great advantage of having at least two overlapping delays to cancel the effect of PSF drift, we encourage users of single delay EDI for velocimetry to consider modifying their apparatus to measure two or more delays quasi-simultaneously, such as in Fig. 45. Or if their existing single delay overlaps their native peak, they can apply crossfading to mix native and single delay signals, for frequencies up to but not beyond the delay. The decrease in TRC should directly reduce a horizontal (wavelength) noise, translating into reduced Doppler velocity noise if

PSF translation is the dominant type of insult, and if not dominated by photon or read noise. (Example noise sources not addressed by the TRC could be a telluric contribution that is time varying and blended into the spectrum, and spatial drifts in the reference ThAr beam relative to the stellar beam, causing slight differential phase shifts in the interferometer.)

11 Discussion and Conclusions

The TEDI is dominated by instrumental noise of about 3% due to severe drifts of the native PSF, which would otherwise prevent high-resolution spectroscopy without the interferometer. The photon shot noise is about 30 times smaller at about 0.1%. A companion paper[23] describes an EDI theory of photon noise for various configurations of multiple delays and final resolutions.

An important ability of EDI is its great robustness to distortions to the native spectrograph instrument lineshape PSF. This robustness is valuable because in many high-resolution spectroscopy applications, particularly radial Doppler velocimetry, the limiting noise is not photon noise but fluctuations produced
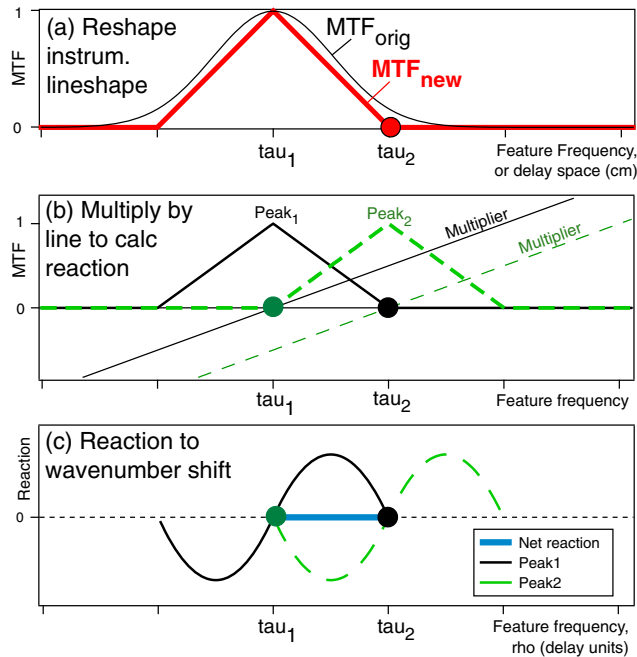
Fig. 43 Achieving perfect PSF drift robustness between two delays τ_1 and τ_2 by crossfading using a triangular lineshape. This achieves zero net translational reaction (TRC = 0) for frequency ρ between the peaks (frequency and delay have same units). (a) The $\mathrm{psf}(\rho)$ or MTF for each peak is reshaped to a crossfading shape, here a triangle, by multiplication of data in ρ-space by [$\mathrm{MTF}_{new}(\rho)/\mathrm{MTF}_{orig}(\rho)$]. (b) For dual-peak crossfading, foot of peak$_1$ (black dot) must be 0% at τ_2, which is maximum of neighboring peak$_2$ (green dashed triangle), and vice versa for peak$_2$ foot (green dotted) at τ_1. Reaction to a wavenumber shift Δx is estimated by multiplying MTF by line through center of each peak. (c) For triangular peak this creates a sine-like odd-symmetry shape consisting of two opposite parabolas. (For sinc function, it makes an exact sinusoid.) When overlaid with the identical but shifted sine-like shape of the neighboring peak (dashed), the net reaction (bold blue segment) is zero between τ_1 and τ_2. (The Δx presumed same for both peaks.) Scheme is extended to all pairs of peaks τ_n to τ_{n+1} to cover all delay space with TRC = 0. For unevenly spaced peaks, reshaped peaks are asymmetrical.

by instrumental irregularities. This is why many radial velocity spectrographs are put in vacuum tanks, built with thermomechanically stable materials, and fiber optically scrambled (in spite of the flux loss at fiber injection). The deleterious effects of instrument noises on the conventional spectrograph are so significant that it is worth the costs of these remedies.

The instrument error that results is a product of two factors: the magnitude of applied insult (e.g., Δx) times the sensitivity (e.g., TRC) of the instrumental technique to such insult. Optimally, both of these factors should be reduced for a precision radial velocimeter or high-resolution spectrograph, but we have focused our attention in this project on the latter (sensitivity) factor. Previously, it was not possible to reduce the TRC of a conventional dispersive spectrograph—it is fixed at unity, so naturally the effort by the community was put into reducing the former factor, such as by environmental stabilization and fiber scrambling. Remarkably, we have shown using EDI with multiple delays we can reduce TRC by at least a factor of 20 using our standard processing, with the promise of an order of magnitude further reduction of TRC possible with strategic reshaping of the data in a so-called crossfading method.

11.1 Disadvantages of Externally Dispersed Interferometry Implementation

Disadvantages of implementing EDI include increase number of optics that the beam must pass through or reflect, and thus increases the parasitic loss, which decreases the flux. This is only relevant for photon-limited applications. We note that researchers have passed starlight through fiber scramblers[24] to reduce focal spot mode variation, while accepting some flux loss such scrambling entails. The increase in PSF stability is worth some flux loss, because it is critical to attaining precision high-resolution spectroscopy or Doppler velocimetry.

Another EDI disadvantage is that there are two interferometer outputs, which (unless polarization is used) necessarily travel on two separate beam directions. So additional optical engineering is required to detect both outputs, otherwise a 50% flux hit is suffered. This could be with optics that redirects the two outputs to different portions of the same spectrograph slit, or to two different spectrographs. For TEDI version 1 both outputs were detected,[6] but for version 2 the complementary output was found[2] more useful as a monitoring port to maximize the amount of flux entering the input fiber.

Third, additional read noise is generated because multiple exposures (at least three) are needed. However, this disadvantage may often be immediately outweighed by the advantage of FP noise rejection that such phase stepping creates (where two exposures recently taken are compared for differences), since FP noises can often be more significant than photon noise. (Read noise behaves like detector noise, which is described in the companion paper.[23]) Also, a recent paper[9] proposes using a readout noise-free detector (EMCCD) with fast scanning of interferometer phase to improve Doppler precision.

11.2 Advantages of Externally Dispersed Interferometry

Summarizing, the EDI method offers three significant advantages: (1) resolution increases beyond classical limits (Fig. 46), and (2) dramatic robustness to PSF irregularities (Figs. 41 and 42), and (3) superb rejection of FP noises such as from bad pixels (Fig. 33). While the resolution increase tends to be discussed most often, what may prove to be the greatest advantage is the robustness to PSF irregularities, since this is the most serious practical limitation to many precision Doppler velocimeters or high-resolution spectrographs in astronomy today.

11.3 Summary

In summary, the EDI allows one to make high-resolution measurements not possible or practical with your current apparatus due to physical laws of optics, such as a nonzero focal spot size, sparse pixel spacing, bad pixels, and severe drifts and irregularities of the focal position and diameter. The bandwidth is set by the native spectrograph, which tends to be wide because it has lower resolution. Hence the bandwidth resolution product can be several times larger than in a conventional spectrograph, which is limited by the number of pixels. Spacecraft and aircraft have severe weight restrictions that make large spectrographs impractical. Since the cost and weight of a dispersive spectrograph grows in a power law with resolution, any technology that reduces the native spectrograph
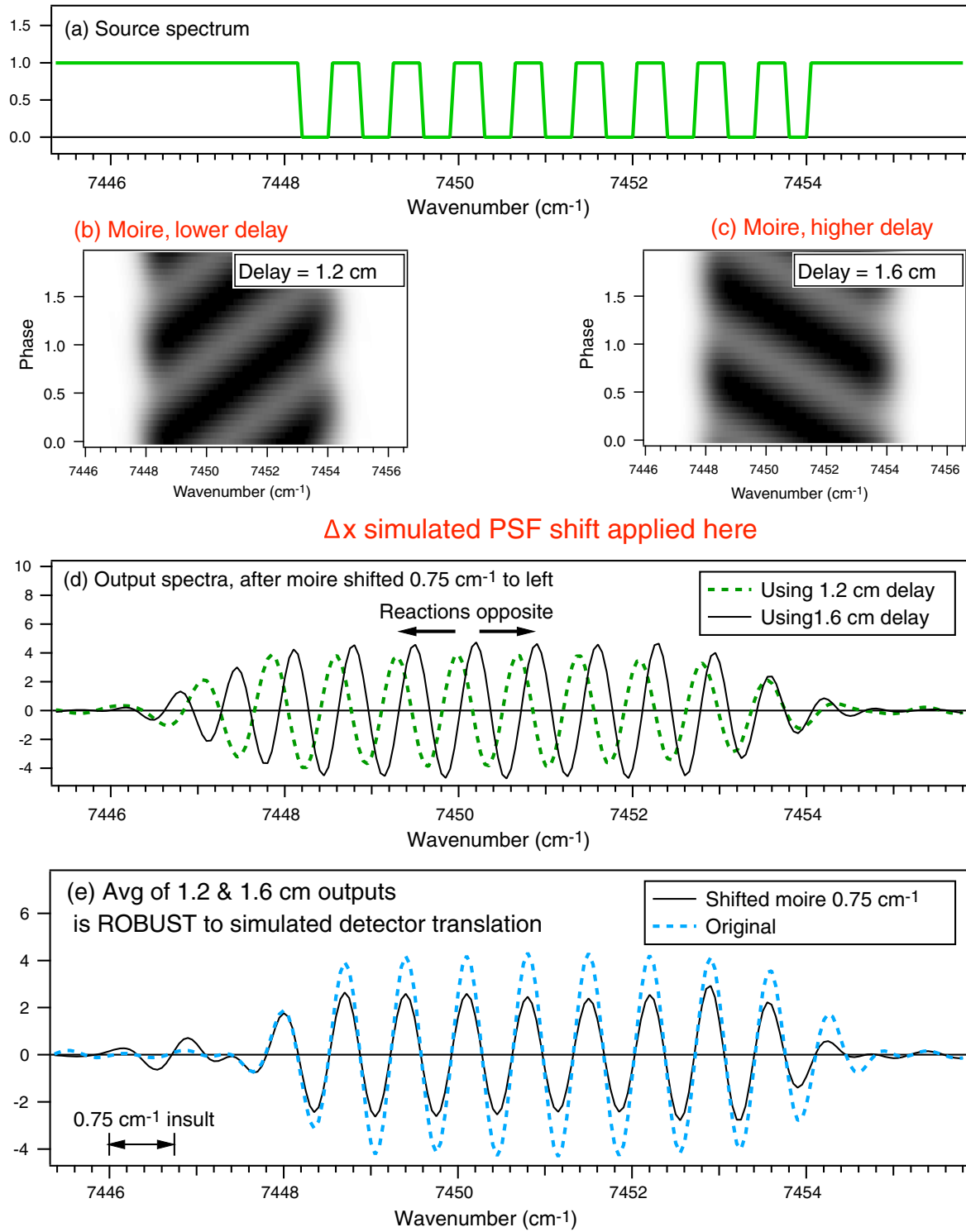
Fig. 44 Demonstration of the crossfading weighting technique on a simulated spectrum (a) consisting of several rectangular lines spaced 0.7 cm^{-1}, so that Fourier energy (at ~1.4 cm) is between overlapping delay peaks at 1.2 and 1.6 cm. Moirés for the (a) smaller and (c) larger delays have opposite polarity slopes, and thus react oppositely in phase to a simulated PSF drift of $\Delta x = 0.75$ cm^{-1} to left. The net reaction can be cancelled by strategic shaping of lineshape prior to combining (sinc function shape was used, triangle produces slightly more ringing). (d) Peaks in output spectra for 1.2 cm (green dash) and 1.6 cm (black) delays react oppositely (and differently than the envelope which shifts to left). Average of the 1.2 and 1.6 curves forms the final outputs (e), for original (blue dash) and artificially shifted (black) moiré positions. Peaks appear in their original positions, robust to the moiré shift. Change in peak position between blue dashed and black curves is only 0.0022 cm^{-1}, 350 times smaller than the simulated PSF drift of 0.75 cm^{-1}. That is, TRC ~0.003.
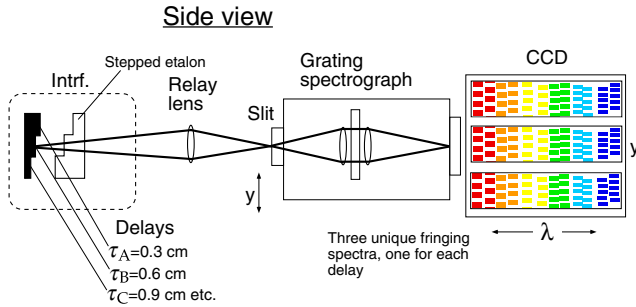
Side view

Fig. 45 Notional diagram of stepped mirror scheme[4] for measuring multiple delays simultaneously with the same disperser, so that short-term PSF spectrograph drifts can be shared among at least two delays, to cancel drift using crossfaded data mixing. Staircase stepped shape of mirror and glass etalon of interferometer (dashed square) images different delays to different detector pixels, but passes through same disperser. This is undesirable if drifts are pixel row dependent. However, the stepped mirror could be combined with TEDI's swapped delay scheme that uses same pixel rows, to discover both long- and short-term drifts on the same pixels for at least two delays.
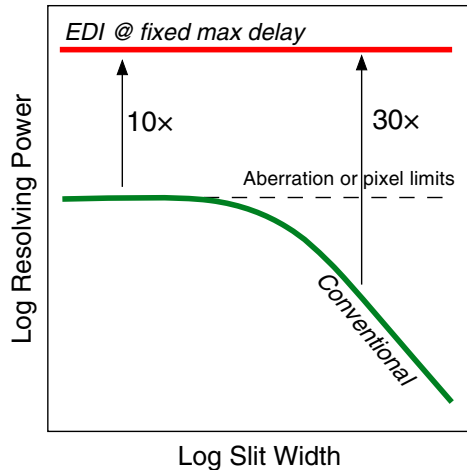
Fig. 46 Notional diagram: the EDI surmounts classical limits to resolution from native spectrograph lens aberrations or detector pixel Nyquist sampling requirements. EDI resolution Res $\sim \tau_{max}/\lambda$ is independent of slit width.

size by a factor of several to achieve a given resolution goal, offers a significant opportunity.

Acknowledgments

References

1. D. J. Erskine, "An externally dispersed interferometer prototype for sensitive radial velocimetry: theory and demonstration on sunlight," *Publ. Astron. Soc. Pac.* **115**, 255–269 (2003).
2. P. S. Muirhead et al., "Precise stellar radial velocities of an M dwarf with a Michelson interferometer and a medium-resolution near-infrared spectrograph," *Publ. Astron. Soc. Pac.* **123**(904), 709–724 (2011).
3. D. J. Erskine and J. Ge, "Novel interferometer spectrometer for sensitive stellar radial velocimetry," in *Imaging the Universe in Three Dimensions: Astrphysics Advanced Multi-Wavelength Imaging Devices*, ASP Conf. Series, W. van Breugel and J. Bland-Hawthorn, Eds., Vol. **195**, pp. 501–507 (2000).
4. D. J. Erskine, "Combined dispersive/interference spectroscopy for producing a vector spectrum," U.S. Patent No. 6, 351, 307 (2002).
5. J. Ge et al., "The first extrasolar planet discovered with a new-generation high-throughput Doppler instrument," *Astrophys. J.* **648**, 683–695 (2006).
6. J. Edelstein et al., "TEDI: the TripleSpec exoplanet discovery instrument," *Proc. SPIE* **6693**, 66930W, (2007).
7. S. Mahadevan et al., "An inexpensive field-widened monolithic Michelson interferometer for precision radial velocity measurements," *Publ. Astron. Soc. Pac.* **120**, 1001–1015 (2008).
8. J. C. van Eyken, J. Ge, and S. Mahadevan, "Theory of dispersed fixed-delay interferometry for radial velocity exoplanet searches," *Astrophys. J. Suppl. Ser.* **189**, 156–180 (2010).
9. R. Jensen-Clem et al., "Attaining Doppler precision of 10 cm/s with a lock-in amplified spectrometer," *Publ. Astron. Soc. Pac.* **127**(957), 1105–1112 (2015).
10. D. J. Erskine and J. Edelstein, "High-resolution broadband spectral interferometry," *Proc. SPIE* **4854**, 158–169 (2003).
11. D. J. Erskine et al., "High resolution broadband spectroscopy using an externally dispersed interferometer," *Astrophys. J.* **592** L103–L106 (2003).
12. D. J. Erskine and J. Edelstein, "Interferometric resolution boosting for spectrographs," *Proc. SPIE* **5492**, 190–199 (2004).
13. D. J. Erskine et al., "Ten-fold spectral resolution boosting using TEDI at the Mt. Palomar NIR Triplespec spectrograph," *Proc. SPIE* **8146**, 81460M (2011).
14. J. Edelstein et al., "Enhanced spectral resolution via externally dispersed interferometry," *Proc. SPIE* **8446**, 84464J (2012).
15. D. J. Erskine et al., "High resolution broad-band spectroscopy in the NIR using the triplespec externally dispersed interferometer at the Hale telescope," *Proc. SPIE* **9147**, 914717 (2014).
16. J. C. Wilson et al., "Mass producing an efficient NIR spectrograph," *Proc. SPIE* **5492**, 1295–1305 (2004).
17. D. J. Erskine, "Speckle-adaptive VISAR fringe analysis technique," in *19th APS Topical Conf. Shock Compression of Condensed Matter (SCCM15)*, AIP Conf. Series, R. Chau, T. Germann, and T. Sewell, Eds., H3.004 (2015).
18. J. Greivenkamp and J. Bruning, "Chapter 14 phase shifting interferometry," in *Optical Shop Testing*, D. Malacara, Ed., 2nd ed., pp. 501–598, John Wiley & Sons, New York (1992).
19. H. G. Roe, "Titan's atmosphere at high-resolution," *Publ. Astron. Soc. Pac.* **15**(812), 1262 (2003).
20. F. Kerber, G. Nave, and C. J. Sansonetti, "The spectrum of Th-Ar hollow cathode lamps in the 691-5804 nm region: establishing wavelength standards for the calibration of infrared spectrographs," *Astrophys. J. Suppl. Ser.* **178**, 374–381 (2008).
21. G. W. Doppmann et al., "The physical natures of class i and flat-spectrum protostellar photospheres: a near-infrared spectroscopic study," *Astron. J.* **130**, 1145–1170 (2005).

22. E. V. Linder, "Lineshape discussion," personal communication (2014).
23. D. J. Erskine et al., "High resolution broad-band spectroscopy using externally dispersed interferometry at the Hale telescope: Part 2, photon noise theory," *J. Astron. Tele. Instr. Syst.* (2016) (submitted).
24. S. Halverson et al., "An efficient, compact, and versatile fiber double scrambler for high precision radial velocity instruments," *Astrophys. J.* **806**, 61 (2015).

David J. Erskine received his PhD from Cornell in 1984. He is an experimental physicist at Lawrence Livermore National Laboratory with experience in femtosecond lasers, semiconductor physics, super-conductivity, diamond anvil cell high-pressure physics, shock physics, high-speed recording techniques, Doppler interferometry, white-light interferometry, digital holography, Fourier signal processing, image reconstruction, and phase-stepping algorithms for interferogram analysis. He collaborates with astronomers on interferometric techniques for Doppler planet search and high-resolution spectroscopy. He is a member of SPIE.

Edward H. Wishnow received a PhD in physics from the University of British Columbia. He is now a research physicist at the Space Sciences Lab at UC Berkeley. He is working on stellar interferometry and spectroscopy in the mid-infrared and visible.

Martin Sirk has over 30 years experience in the design, construction, calibration and science analysis of astronomical instrumentation. This has included working with Digicon detectors on the Hubble Space Telescope, CCD detectors on ground based telescopes, micro-channel plate detectors and optics on 6 NASA missions (EUVE, FUSE, ORFEUS, CHIPS, SPEAR, and ICON), and photographic plates at Lick Observatory.

Biographies for the other authors are not available.