High-resolution broadband spectroscopy using externally dispersed interferometry at the Hale telescope: part 2, photon noise theory

David J. Erskine
Jerry Edelstein
Edward Wishnow
Martin Sirk
Philip S. Muirhead
Matthew W. Muterspaugh
James P. Lloyd

SPIE.

High-resolution broadband spectroscopy using externally dispersed interferometry at the Hale telescope: part 2, photon noise theory

David J. Erskine,[a,*] **Jerry Edelstein,**[b] **Edward Wishnow,**[b] **Martin Sirk,**[b] **Philip S. Muirhead,**[c]
Matthew W. Muterspaugh,[d] **and James P. Lloyd**[e]
[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, 60 Garden Street, MS-10, Ithaca, New York 14853, United States

Abstract. High-resolution broadband spectroscopy at near-infrared (NIR) wavelengths (950 to 2450 nm) has been performed using externally dispersed interferometry (EDI) at the Hale telescope at Mt. Palomar, with the TEDI interferometer mounted within the central hole of the 200-in. primary mirror in series with the comounted TripleSpec NIR echelle spectrograph. These are the first multidelay EDI demonstrations on starlight. We demonstrated very high (10×) resolution boost and dramatic (20× or more) robustness to point spread function wavelength drifts in the native spectrograph. Data analysis, results, and instrument noise are described in a companion paper (part 1). This part 2 describes theoretical photon limited and readout noise limited behaviors, using simulated spectra and instrument model with noise added at the detector. We show that a single interferometer delay can be used to reduce the high frequency noise at the original resolution (1× boost case), and that except for delays much smaller than the native response peak half width, the fringing and nonfringing noises act uncorrelated and add in quadrature. This is due to the frequency shifting of the noise due to the heterodyning effect. We find a sum rule for the noise variance for multiple delays. The multiple delay EDI using a Gaussian distribution of exposure times has noise-to-signal ratio for photon-limited noise similar to a classical spectrograph with reduced slitwidth and reduced flux, proportional to the square root of resolution boost achieved, but without the focal spot limitation and pixel spacing Nyquist limitations. At low boost (∼1×) EDI has ∼1.4× smaller noise than conventional, and at >10× boost, EDI has ∼1.4× larger noise than conventional. Readout noise is minimized by the use of three or four steps instead of 10 of TEDI. Net noise grows as step phases change from symmetrical arrangement with wavenumber across the band. For three (or four) steps, we calculate a multiplicative bandwidth of 1.8:1 (2.3:1), sufficient to handle the visible band (400 to 700 nm, 1.8:1) and most of TripleSpec (2.6:1). © *2016 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JATIS.2.4.045001]

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

Paper 16010P received Feb. 13, 2016; accepted for publication Nov. 2, 2016; published online Dec. 2, 2016.

1 Introduction

High-resolution broadband spectroscopy at near-infrared (NIR) wavelengths (950 to 2450 nm) has been performed using externally dispersed interferometry (EDI) at the Hale telescope at Mt. Palomar, with the TEDI interferometer mounted within the central hole of the 200–in. primary mirror in series with the comounted TripleSpec[1] NIR echelle spectrograph. These are the first multidelay EDI demonstrations on starlight. We demonstrated very high (10×) resolution boost and dramatic (20× or more) robustness to point spread function (PSF) wavelength drifts in the native spectrograph.

A companion paper[2] (part 1) describes how to extend single delay spectroscopy[3,4] into multiple delay spectroscopy, emphasizing data analysis, results, and instrument noise. EDI theory for radial velocimetry (RV) has been described,[5,6] but not for multiple delay EDI used for general spectroscopy. (Single delay EDI has been called dispersed fixed-delay interferometry by other researchers[6] using this technique for RV and used to discover a new exoplanet HD 102195b.[7]) This part 2 describes

theoretical photon-limited and readout noise-limited behaviors, for both single- and multiple-delay spectroscopy, using simulated absorption and emission spectra and instrument model with noise added at the detector.

The EDI forms Moire patterns by multiplying sinusoidal comb against the input spectrum $S_0(\nu)$ in a heterodyning effect:

$$B_n(\nu) = S_0(\nu)[1 + \gamma \cos 2\pi(\tau\nu + \phi_n)] \otimes \text{PSF}(\nu) + \text{noise}_n,$$
(1)

where wavenumber ν has unit cm^{-1}. Then, subsequent blurring by the native spectrograph PSF_0, which removes high frequencies, does not significantly affect the Moire patterns, which are primarily at low frequencies. Equation (1) describes a single exposure (B_n) of a set of N phase stepped exposures for a given delay, where phase ϕ_n (units of cycles) increments around the circle. Figure 1(b) shows simulated Moire patterns, where phase is splayed vertically continuously, whereas Fig. 2 shows how sinusoidal fitting along columns extracts fringing (W) and nonfringing (B_{ord}) components. (In TEDI, the phase was varied
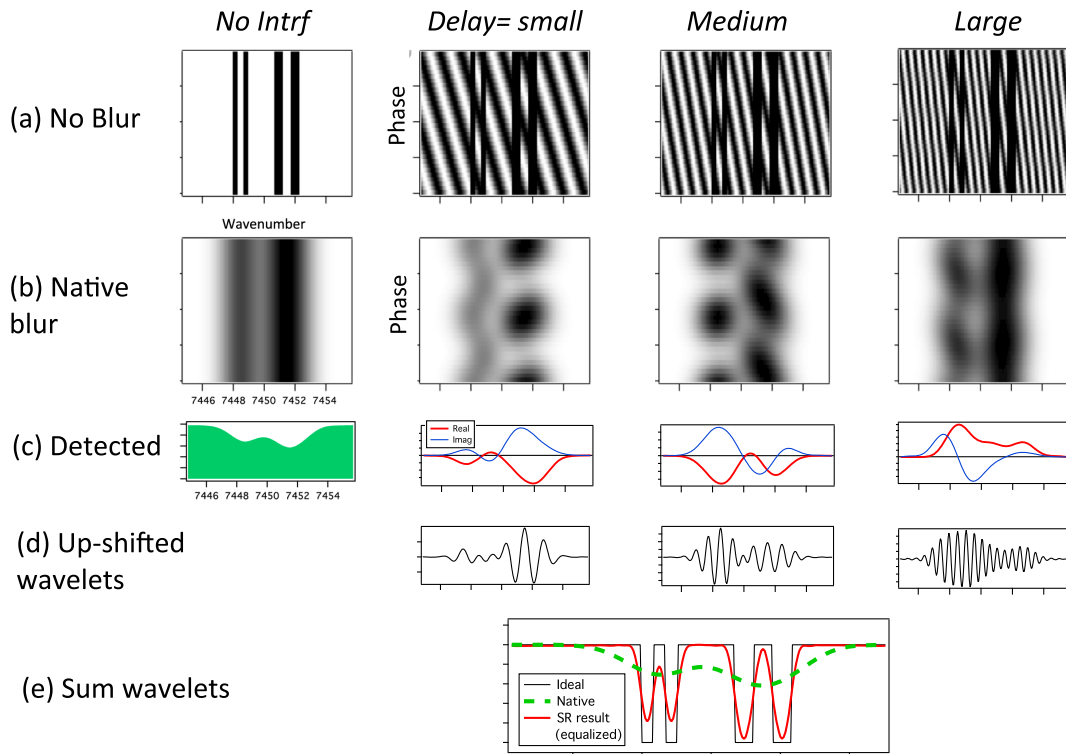
2329-4124/2016/$25.00 © 2016 SPIE

Fig. 1 Numerically simulated Moire patterns using a test absorption spectrum of two pairs of black lines [black curve in (e)]. (a) Without blur, sinusoidal comb multiplies input spectrum, comb pitch proportional to delay τ. Only three delays (0.75, 1.25, 2 cm) of eight shown. (b) With blur, the comb unresolved but Moire pattern remains. (c) Complex expression of Moire (whirl or **W**) from Fig. 2, red (real), blue (imaginary). (d) Whirls upshifted in frequency; real part taken to form wavelets. (e) Sum of wavelets forms reconstructed output (red curve). An EQ step weights the wavelets to eliminate ringing. Native spectrum (dashed green) has insufficient resolution (2 cm^{-1}) to resolve test pair. Graphs are ~10 cm^{-1} across average wavenumber of 7450 cm^{-1}.

temporally, since transverse detector space was needed to record several light beams. But it is useful to plot phase vertically analogous to the earliest EDIs, which splayed phase spatially across the detector.)

The interferometer visibility γ ranges from 0.85 to 0.95 for TEDI. For simplicity, in the models we assume $\gamma = 1$. Instrumental factors that could reduce γ include intensity imbalance between the arms (e.g., beamsplitter interface reflectivity changing) and imperfect optical surfaces. (Optics having $\lambda/20$ flatness make this latter factor insignificant.)

Heterodyning shifts (beats) high feature frequency (ρ) information down to lower frequencies, by the interferometer delay τ (unit: cm), where it is detected. It is later restored to its original high frequency mathematically. Figure 3 shows heterodyning for a single delay creating a new EDI sensitivity peak (red, psf$_{edi}$) that is a copy of the native spectrograph sensitivity peak (green, psf$_0$ or psf$_{conv}$):

$$\mathrm{psf_{conv}} = \mathrm{psf_0}(\rho), \qquad (2)$$

$$\mathrm{psf_{edi}} = \frac{1}{2}\gamma\,\mathrm{psf_0}(\rho - \tau), \qquad (3)$$

but shifted to higher frequency ρ by delay τ and having half the amplitude. (Feature frequency space ρ has units of features per cm^{-1}, i.e., cm, i.e., same units as delay space.) Lower case denotes ρ space or Fourier transform versions of upper case functions in pixel or ν space.

Under a Doppler shift, the phase of the entire Moire pattern shifts. Thus, EDI can sensitively measure Doppler radial velocity, even when the spectrograph has a very low resolution otherwise insufficient for this task. This was the motivation for inventing EDI. The primary reason for the TEDI project was to demonstrate this in the NIR. It uses the TripleSpec[1] spectrograph, which bolts to the Cassegrain output of the Hale telescope (and hence, suffers changing gravimetric drifts). Its resolution of ~2700 is otherwise insufficient for precision RV measurements. But with the TEDI interferometer, it was able to make precision RV measurements[8] of M-stars.

An equally useful application of EDI became apparent, which is wide bandwidth (BW) high-resolution spectroscopy. The same Moire pattern produced by the instrument for Doppler velocimetry is studied for its shape rather than its overall phase. The goal is to go backward through the heterodyning process to discover what high-resolution spectra would produce the measured Moire patterns. This is a method of measuring much higher effective resolutions than allowed by the native spectrograph, if the interferometer delay is larger than the width of the native response peak psf$_0(\rho)$.

Since EDI spectroscopy starts with the same Moire data as Doppler measurements, the secondary purpose of the TEDI project was to explore using EDI to make wide BW high-resolution stellar measurements. Single-delay EDI spectroscopy had already been performed,[3] and multiple delays promised even greater resolution boost but had been performed only on laboratory sources.[4] The TEDI project presented the opportunity to test

(a) Data source

(b) Fitting example

Example, fixed wn = 7448 cm-1
● Intensity vs 10 phase steps
— Sine Fit
- - - Nonfringing value

(c) Complex representation
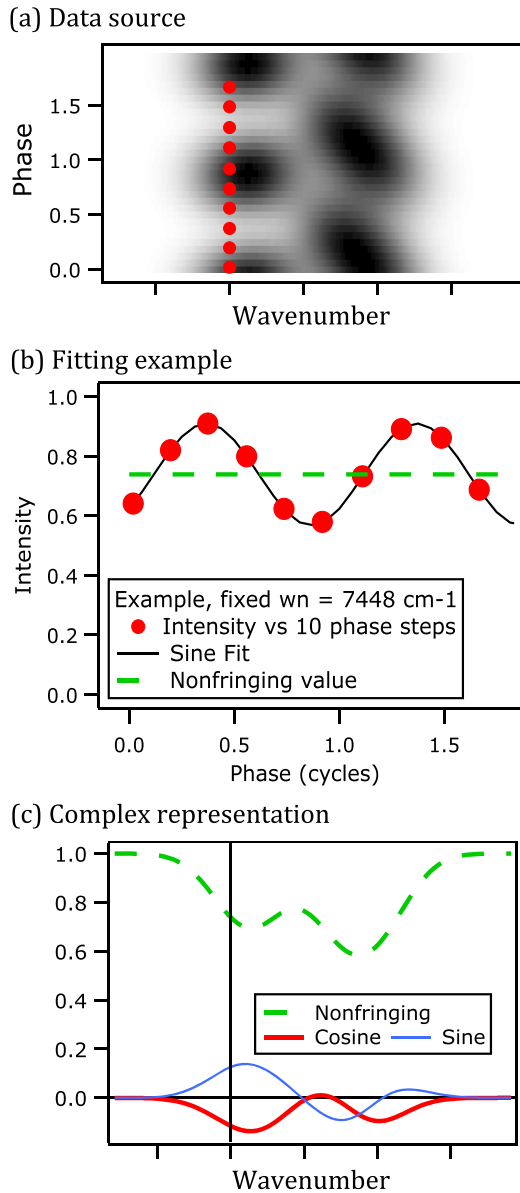

- - - Nonfringing
—— Cosine ——— Sine

Fig. 2 How to convert a row along Moire pattern to complex data. Vertical lineout (a) across multiphase stack for a given wavenumber produces an intensity versus phase plot (b), which is fitted to a sinusoid (black curve). (c) The sine and cosine amplitudes (red and blue curves) are the whirl's (**W**) imaginary and real complex values, for a given wavenumber. The vertical offset of the fit is the ordinary spectrum (green, B_{ord}) at that wavenumber.
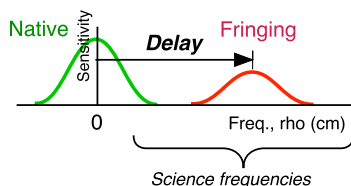

Fig. 3 Heterodyning shifts the native spectrograph sensitivity peak from zero to a higher frequency, where science frequencies typically reside, by the interferometer delay τ, and to 50% amplitude. Frequency in units of features per wavenumber (cm^{-1}) conveniently has units of delay (cm).

this concept on stellar measurements. (However, the set of delay values was not ideal for spectroscopy–some were not evenly spaced over the delay range, having been selected for the RV purpose for targets having various rotational broadening.)

We discovered that, indeed, wide BW high-resolution spectroscopy is quite practical. Part 1 (Ref. 2) shows the data analysis methods and results of many example high-resolution spectra, having resolution boosts up to 10× (using delays up to 3 cm). Proportionately higher boosts could be achieved by purchasing several more glass etalons to allow a delay range up to 5 cm without gaps.

We learned that an important advantage of EDI is not only the resolution boost but that the output spectrum is impressively insensitive to the native spectrograph instrumental distortions (that distort the shape and wavenumber position of the native PSF). Sections 9 and 10 of Ref. 2 describe how we observed a 20× reduction to PSF drift insult using original lineshapes and a 350× reduction with optimized lineshapes.

The TEDI data were dominated by severe instrument noise, which was chiefly an irregular and large wavenumber PSF shift, rather than photon-limited or readout-limited noise, and this was the subject of part 1.

However, naturally, we are also interested in how EDI responds to low flux environments, where shot noise and readout noise dominate. This is the subject of part 2, which is to study how adding simulated noise$_n$ in Eq. (1) at the detector affects the final high-resolution spectrum. A detector noise, such as readout noise, has a constant magnitude (standard deviation), while photon (shot) noise is simulated by scaling noise$_n$ by the square root of the noise-free version of B_n. Both absorption and emission spectra have been studied.

2 Phase Stepping and Bandwidth

2.1 Sine Fitting Along a Column

The process of fitting a sinusoid to a column of data (intensity versus phase at a given ν) is called phase stepping arithmetic, and the particular case of four steps is called pushpull arithmetic. Its purpose is to separately extract fringing (**W**) and nonfringing (B_{ord}) components from the set of phase stepped data $B_n(\nu)$, where the exposure number index n is also along the phase axis.

Consider it to be a sine fit along the column at a particular ν. Then, the sine amplitude is assigned to the imaginary part of **W**, and the cosine amplitude is assigned to the real part of **W**. The average value of the sinusoid is assigned ordinary spectrum's value, or B_{ord}, at that wavenumber.

2.2 Generic Expression for N Ideal Steps

The generic expressions for many (N) regular steps that evenly fit around a circle ($\phi_n = n/N$, in units of cycles) are

$$\mathbf{W}(\nu) = \frac{1}{N} \sum B_n e^{-i2\pi\phi_n}, \tag{4}$$

$$B_{\mathrm{ord}}(\nu) = \frac{1}{N} \sum B_n, \tag{5}$$

where **W** is called a "whirl" and manifests the fringing information we seek. Note the similarity of Eq. (4) to a discrete Fourier transform, which evaluates the sine and cosine amplitudes.

2.3 Healing Method of TEDI Handles Many Irregular Steps

The TEDI project used 10 steps that were irregularly spaced around the circle. They were irregular because their value is proportional to the wavenumber, which changes across the band as $\Delta\phi = \nu\Delta\tau$. The change in delay τ is the physical constant related to mirror displacement (and glass dispersion causes a small ν dependence that is corrected for). Hence, the phases may be in an ideal configuration at one particular ν and then slowly move into an irregular arrangement, including wrapping around the phase circle, as ν changes across the wide band.

To handle this irregularity, a "healing" algorithm was used, which adjusts the weights of each B_n so that effectively one is mixing fractions of other phase steps into a given phase step, to alter its angle and length to bring it into the nearest ideal configuration. Then, Eq. (4) can be applied.

2.4 Low Readout Noise Motivates Use of Three or Four Steps

Since readout noise combines in quadrature, the total readout noise grows as the square root of N. Hence, in this part 2, when discussing minimizing noise for low flux observations, we are motivated to use the minimum number of step 3 to define a sinusoidal fit. Four steps produce more elegant equations and a slightly wider BW as we will demonstrate.

For four exposures, every 90 deg:

$$\mathbf{W}(\nu) = \frac{1}{4}[(B_0 - B_{180}) + i(B_{90} - B_{270})], \qquad (6)$$

$$B_{\text{ord}}(\nu) = \frac{1}{4}(B_0 + B_{180} + B_{90} + B_{270}). \qquad (7)$$

For three phase steps, symmetrically positioned (trigonal symmetry):

$$B_{\text{ord}}(\nu) = \frac{1}{3}(B_0 + B_{120} + B_{240}), \qquad (8)$$

$$\mathbf{W}(\nu) = \frac{1}{3}[e^{i2\pi 0}B_0 + e^{i2\pi(1/3)}B_{120} + e^{i2\pi(2/3)}B_{240}], \qquad (9)$$

$$\mathbf{W}(\nu) = \frac{1}{3}[B_0 - 0.5B_{120} - 0.5B_{240})]$$
$$+ i\frac{1}{3}[0.866B_{120} - 0.866B_{240}]. \qquad (10)$$

2.5 How Wide is the Bandwidth for Three or Four Steps?

As the ν varies across the band, the phase step will change, gradually bringing it out of ideal configuration into irregularity. An irregular configuration will tend to increase the noise since the weights applied to each phase step to cancel out the nonfringing component, leaving the fringing component behind, tend to increase. The increased weights produce increased net noise (sum in quadrature of all the weights). We therefore have studied (Fig. 4) how the weights and hence net noise varies across the band, for three or four steps, in order to find the useful bandwidth BW.

The results are shown in Fig. 4, which supposes that the ideal configuration is at $\nu = 10000$ cm^{-1}, and that for other wavenumbers, the phase steps are proportionately different. The minima of these curves, i.e., the value of the noise (relative to a single readout) at the ideal ν is $\sqrt{3}$ and $\sqrt{4}$ for the case of readout noise, panel (a). For the case of photon-limited noise (b), there is no dependence on the number of steps since only the total number of detected photons matters; hence, the minima of the curves in (a) merge together in (b).

The circular wheel-like diagrams show the configurations at several places across the band. Interestingly, note how the four-step configuration at 13000 cm^{-1} appears approximately as a three-step configuration, because step 3 overtakes step 0. In this region, we average steps 3 and 0 together and use the three-step formula [Eqs. (9) or (10)]. Then, for larger ν, step 3 passes counter-clockwise away enough from step 0 that once again a four-step formula [Eq. (6)] produces a smaller noise. (However, note that for the use of the formula, we would relabel the steps so that they increase in index counter-clockwise. We do not relabel them in the figure, so that the reader can track the movement of the steps.)

2.5.1 Bandwidth of three steps is usefully wide, 1.8:1

We define the practical BW as where the noise does not increase more than 20% over the minimum four-step value (of $\sqrt{4}$). We observe that the multiplicative BW for four steps is 2.3:1, and for three steps using the same absolute noise threshold is 1.8:1. This is comfortably large, sufficient to handle the visible band (400 to 700 nm, 1.8:1) and most of TripleSpec[1] BW (950 to 2450 nm) of 2.6:1.

2.6 Details on How Irregular Weights Were Calculated

Lets us describe how we discovered the weights (that multiply each exposure) for the three or four steps configuration, as it becomes irregular. When the phase angles deviate from ideal, the weights deviate from unity. The net noise is the sum in quadrature of these weights.

A simple method for discovering these weights is to treat the four data as pair differences, and this style of analysis is called "pushpull," and it has been used extensively by the first author in analyzing other interferometric data (e.g., Secs. A and B of Ref. 9). Code already developed for pushpull analysis was used. The four steps are divided into pair differences:

$$\text{Horiz} = 0.5(B_0 - B_{180}); \quad \text{Vert} = 0.5(B_{90} - B_{270}), \qquad (11)$$

which are assigned to represent the real and imaginary parts of $\mathbf{W}$, as shown in Fig. 5. Then, Vert versus Horiz is plotted to form a Lissajous, which is elliptical generally since the weights are inappropriate for the phase step angles. We desire it to be circular of unity radius to indicate the weights are now correct. (For elegance, we temporarily use a factor of 1/2 instead of 1/4 in Eq. (6), so that the radius can be unity, and then apply another factor of 1/2 later in the processing.)

Rather than adjusting individual weights for the four B_n, we adjust other gains that we have created that manipulate the difference horiz and vert, and another gain to apply a linear transformation that corrects for the so-called obliquity of the data. These gains were adjusted until a circular Lissajous was obtained from the data, when test data were used that created a helical $\mathbf{W}$. That is,
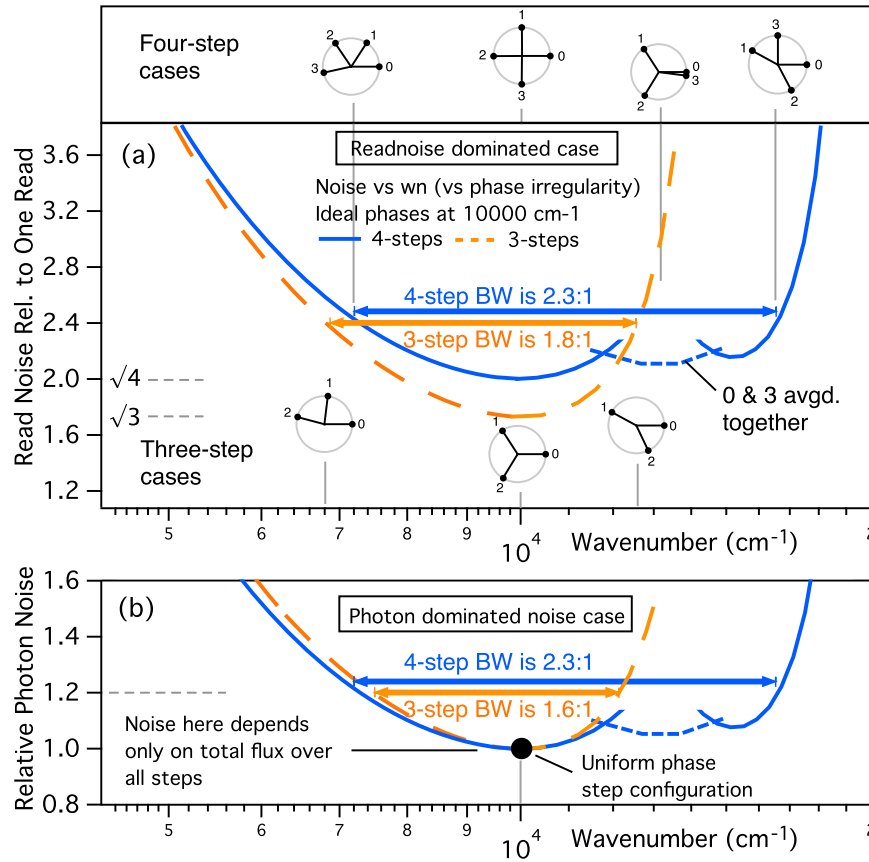
Fig. 4 Change in relative noise of fringing component **W** due to irregular spacing of the phase steps as they grow across the band, for cases of 4 and 3 steps, and (a) for readout noise and (b) photon dominated noise. For a constant physical delay change as the interferometer mirror is stepped, $\Delta\tau$, the corresponding phase change varies with wavenumber as $\Delta\phi = \nu\Delta\tau$. We define the ideal regular phase configurations (four at 90 deg, three at 120 deg) to occur at ν of 10,000 cm^{-1}. When the phases deviate severely from their ideal uniform configuration and cluster in two groups on opposite sides of the phase circle, then the obliquity of a Lissajous plot of **W** increases dramatically, increasing the weights needed on data to achieve circularity, increasing the net noise. (a) Vertical axis is readout noise relative to a single readout. An arbitrary definition of practical BW is the noise increasing 20% over 90 deg configuration. Relative BWs of 2.3:1 and 1.8:1 are produced for 4 and 3 exposure cases, respectively. These are sufficiently wide for most kinds of spectroscopy (visible band 400 to 700 nm is 1.75:1).

$$B_0 = 1 + \cos 2\pi\tau\nu; \quad B_{90} = 1 + \sin 2\pi\tau\nu;$$
$$B_{180} = 1 - \cos 2\pi\tau\nu; \quad B_{270} = 1 - \sin 2\pi\tau\nu \tag{12}$$

over a small test range of $\tau\nu$, where the helical data make a single revolution. When the correct weights are found, the Lissajous will be a circle of radius unity. Initially, when unity weights are used for the irregular condition, an elliptical Lissajous is produced, as shown in Fig. 5 having obliquity.

The idea is to temporarily replace the actual data with the helical data, find the appropriate gains or weights that make a circular Lissajous, and then apply these gains on actual data.

The more irregular the phase configuration, the larger the obliquity. As "Vert" deviates from orthogonality of "Horiz," the obliquity grows. Also, as the magnitudes of Vert and Horiz decrease, as illustrated by panel (c), where the phase steps are dramatically far from ideal, the size of the Lissajous proportional decreases. Then, the overall gain must be increased to restore Lissajous radius to unity. This increases all the weights, which increase the net noise. The obliquity is related to how far from orthogonality is the Vert compared to Horiz.

The gains of the differences and the gain of the obliquity operation are adjusted to produce a circular Lissajous. Then,

the equivalent values of the weights (assigned to each B_n) are calculated and summed in quadrature to yield by what factor the noise increased.

As the phase increases from the ideal configuration, the Lissajous deviates from a circle to an ellipse having obliquity different from unity. We define obliquity as the ratio of the major and minor axis of the ellipse, and they align along the $y = x$ and $y = -x$ axes when the phases are nonideal. We created an obliquity operation that temporarily transforms the data into two new axes along $y = x$ and $y = -x$, applies a gain g that diminishes the length along the $y = x$ axis, and then restores the data to the original axes y and x. This linear operation has the effect of mixing phase step components:

$$\Re W' = \Re W(g+1)/2 + \Im W(g-1)/2,$$
$$\Im W' = \Re W(g-1)/2 + \Im W(g+1)/2. \tag{13}$$

Hence, it is related to the Healing method, which also mixes components. The $g = 1$ is the nominal situation not requiring a change in obliquity. One could redefine obliquity to change the minor instead of major axis, by substituting $g \to 1/g$ in

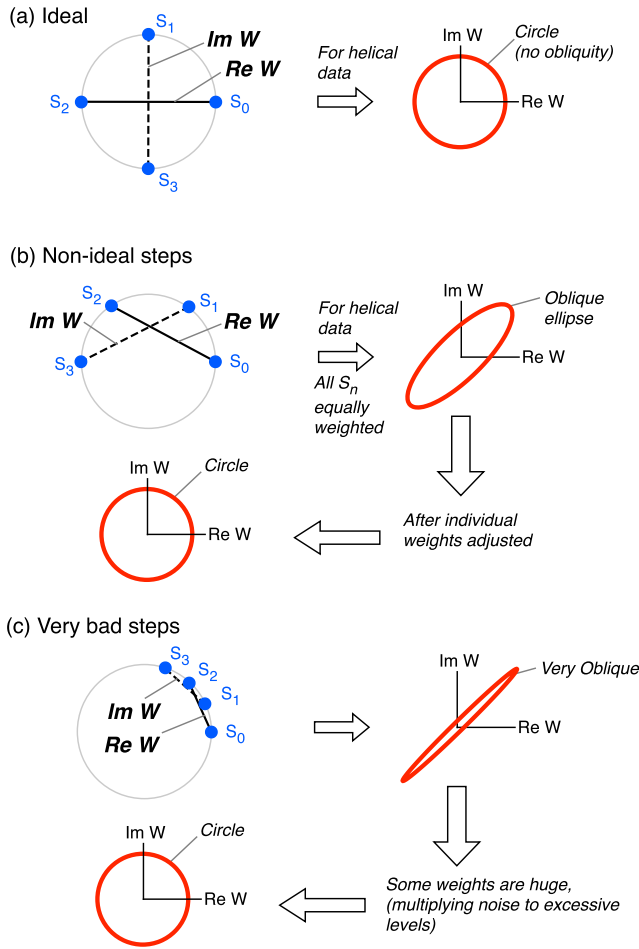**Fig. 5** How nonideal phase steps can still measure the whirl **W**, but with magnified noise. (a) Four ideal 90-deg phases, and intensities B_n associated with the vector positions of steps $\mathbf{S}_n$. Real and imaginary parts are formed from intensity differences $\mathbf{W} = 0.5(B_0 - B_2) + i0.5(B_1 - B_3)$ associated with vector differences $(\mathbf{S}_0 - \mathbf{S}_2)$ and $(\mathbf{S}_1 - \mathbf{S}_3)$. For discovering proper weights, "helical" data are used, $B_n = \cos i2\pi(\nu\tau + \phi_n)$, which normally creates a circular Lissajous, imaginary part versus real part (red circle). (b) Nonideal phase steps, and $(\mathbf{S}_0 - \mathbf{S}_2)$ is no longer orthogonal to $(\mathbf{S}_1 - \mathbf{S}_3)$, and Lissajous has obliquity. After weights are adjusted, circularity is achieved. But sum in quadrature of weights has increased, increasing noise. (c) Extremely poor phase step configuration producing very oblique ellipse. Increase of dot product between $(\mathbf{S}_0 - \mathbf{S}_2)$ and $(\mathbf{S}_1 - \mathbf{S}_3)$, and decrease of their lengths, increase the noise.

the equation. Since the obliquity operation changes the radius of the circular Lissajous that is achieved, the final step is to apply an overall gain adjustment that affects all the individual weights the same to bring the circle radius to the desired value of unity. Then, from these operations, we calculate their equivalent effect on the individual weights, and sum these in quadrature to yield the ratio of increase in noise over the ideal value.

For configurations that had three inputs, we used Eqs. (8) and (10) to create **W** and B. These, of course, produce an elliptical Lissajous for irregular phase steps. To correct this by conveniently using the existing code written for four steps, we map the three steps to four ersatz steps through

$$B_0 = B_{\text{ord}} + \Re W; B_{180} = B_{\text{ord}} - \Re W,$$

$$B_{90} = B_{\text{ord}} + \Im W; B_{270} = B_{\text{ord}} - \Im W, \quad (14)$$

and adjust the gains to make a circular Lissajous of unity radius.

3 Single Delay Noise Behavior

3.1 Overview

We will first discuss a single delay behavior, because multiple delays is a simple extension.

3.1.1 Concrete case of noise reduction for 1× boosting evaluated

For a concrete example, we will examine the 1× boosting case, where the new information (fringing information) is used to reduce high frequency noise rather than increase the resolution. The motivation is to provide a demonstration and detailed analysis for the leftmost data point of a later graph [Fig. 15(a)] that claims the 1× boosted photon-limited noise for an EDI spectrograph is $\sim\sqrt{2}$ times less than a conventional spectrograph, which some readers may find surprising.

Examining this case also illustrates the relative contributions to a $\sim$2× boosting case, since the amount of boost is merely a matter of a different choice in the equalization (EQ) multiplier, which is the final step, and the EQ does not change the signal-to-noise ratio (SNR) curve since both signals and noises are multiplied by the same factor.

3.1.2 Numerical simulator used to study noise propagation

To study the propagation of noise, we use a numerical simulator, which uses similar algorithms for processing the data as the actual TEDI code, after the process of phase stepping or sine fitting. Here, the phase stepping process uses four constant steps at 90 deg because the issues of phase step irregularity and its ν dependence have been separately studied in Sec. 2. The numerical simulator has been used to study various types of noises (shot, detector) under various spectra types (absorption, emission), using various processing choices (bell weighting yes/no, various types of EQ). The simulator equations are in Appendix A.

3.1.3 Externally dispersed interferometry calculator produces smooth theory curves quickly

A second useful tool is the "EDI calculator," which is a set of equations that can be quickly evaluated that lack the stochastic variations of the simulation and do not require a specific input spectrum. We have confirmed that these replicate the numerical simulator. The calculator displays the responses of the various components, signals and noises and their ratio, versus feature frequency ρ. The calculator equations in Appendix B are based on theory, except for the case of the smooth change between uncorrelated and partially correlated behavior found empirically by the numerical simulation. In that case, the EDI calculator uses a best fit [Fig. 8(b)] modeling that observed behavior, since we have not yet developed a complete analytical explanation for it.

3.2 Simulator Results for 1× Boosting and Emission Spectroscopy

Figure 6 shows a portion of the hypothetical emission spectrum made by spikes of random heights (a) prior to blurring and (b) after blurring to resolution 3725 at 7450 cm^{-1} and adding simulated shot noise (nonfringing component is shown). (The
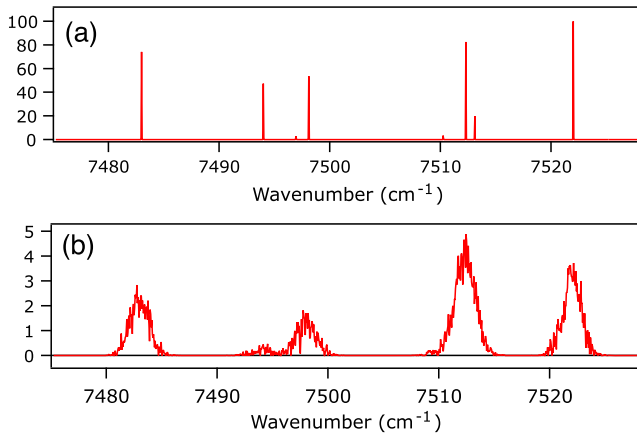
Fig. 6 (a) Small portion (50 out of 450 cm⁻¹) of emission source spectrum S_0 for numerical simulation for studying photon noise, consisting of spikes having no continuum background, using calculational pixels of 0.05 cm⁻¹. (b) Blurred version with added simulated shot noise (native resolution ∼3725 at average 7450 cm⁻¹). For modeling absorption spectroscopy, a spectrum having a continuum was used.

actual TEDI native resolution was closer to 2700, but we are not trying to strictly reproduce TEDI, but use calculationally or graphically convenient parameters.) When simulating absorption spectroscopy, we used a hypothetical spectrum having a continuum. We simulated photon noise by scaling the noise magnitude as square root of local intensity and detector noise by having fixed standard deviation.

The sinusoidal interferometer transmission comb was multiplied against the input spectrum S_0 [Eq. (1)] to form four signals in 90-deg phase relation and blurred. The blur and interferometer delays were adjustable. Figure 1(b) shows the appearance of similar signals, but for an absorption spectrum and displaying continuous phase along Y-axis that produces a smoother appearance. The nonfringing signal was produced by summing the four signals [Eq. (7)] to cancel fringes. The fringing component was obtained using Eq. (6), which subtracted exposure pairs and assigned them to real and imaginary parts of **W**.

The computational grid had a spacing of 0.05 cm⁻¹, a sufficiently small spacing to hold high-resolution signals boosted 10 times the native resolution. The boundaries of the Fourier space (Nyquist frequency) were $0.5(1/0.05) = 10$ cm, which is many times greater than the 0.22-cm half width at half max (HWHM) of the native response peak $\mathrm{psf}_0(\rho)$. This comfortably accommodates delays at least up to 5 cm.

Figure 7(a) shows the simulation results after the fringing component was shifted in frequency by the delay τ [via Eq. (22)], to restore the high frequency information to its original high frequencies. This creates a wavelet appearance to the fringing result in red.

The source spectrum had a small section 7420 to 7436 cm⁻¹ that was a pure sinusoid of frequency 0.22 cm. This made it easy to visually confirm that the relative component magnitudes from the simulation output agreed with the EDI calculator at a specific ρ of 0.22 in (c), indicated by heights of the colored dots.

The difference between the results having added noise at the detector and no added noise was subtracted to produce residuals (b). Note that the EDI 1× result (purple) has less high frequency noise than the conventional result (black dashes), which confirms our claim that adding an interferometer to the spectrograph can reduce high frequency noise, for the same detected flux. We

calculate noises on the components "bass" (nonfringing) and "treble" (fringing), and their combination "net." Fourier transform of the noises shows how residual noises vary with frequency.

3.2.1 *Fringing versus nonfringing noises, correlated or uncorrelated?*

Figure 8 shows how repeated instances of the numerical simulation with photon noise were used to study how the type of noise varied between the two extremes of correlated (sum linearly, open circles), or uncorrelated (sum in quadrature, open squares), depending on the delay value. We discovered that the behavior has a peak, which is well fitted by the native response peak $\mathrm{psf}_0(\rho)$. However, the peak does not reach the perfectly correlated level when the delay approaches zero. Instead, it is 0.7 of this distance. Since this was only recently discovered, we have not yet found an analytical source for the 0.7 factor but suspect that it is $\sqrt{0.5}$.

When the type of noise was switched from photon to detector, the peak disappeared and the noise acted uncorrelated for all delays. This result can also be argued analytically. A correlation between nonfringing $\sim(B_0 + B_{180})$ and fringing-like signals $\sim(B_0 - B_{180})$ involves integrals over products like $(N_0 + N_{180})(N_0 - N_{180}) = (N_0^2 - N_{180}^2) \approx 0$, where N represents the noise apart from the signal. Hence for the detector case the correlation is approximately zero, because the detector noise variance N^2 is nominally the same between the exposures. This argument does not require the heterodyning which involves the delay.

For the photon limited case the noise depends on flux, so N_0^2 could differ from N_{180}^2. This produces some correlation, confirmed by the peak at small delays in Fig. 4(b). The heterodyning contributes the delay dependence by providing another mechanism for noncorrelation at large delays.

3.3 *1× Boost Simulation Result in Frequency Space*

Figure 9 shows stages in the calculation of the SNR curves (versus ρ) for a delay of 0.4 cm, which is on the wing of the native response peak and a typical placement of the peak. The EQ step is omitted to better show high frequency behavior. Figure 10 is the same as Fig. 9 but with EQ applied, to achieve the same final resolution for the net EDI as in the conventional result, i.e., achieving 1× boost. A different EQ could have produced a boosted resolution of ∼2. The EQ step does not change the SNR curve shape since it multiplies both signal and noise.

3.3.1 *Photon-limited noise case*

The simulation output in Figs. 9 and 10 is the photon-limited noise case. The SNR for the net EDI is plotted as a thick purple curve in (d) showing a hump at 0.4 cm, where the fringing peak (red curve) contributes. This hump exceeds the conventional SNR (black dotted Gaussian curve) significantly for frequencies above the midpoint of the conventional response, say 0.3 cm and higher. Figure 11 evaluates the root mean square (RMS) average noise in more detail.

3.3.2 *Readout noise-limited case*

The simulation uses four steps, whereas a conventional measurement only a single readout (unless the conventional is dithering by two exposures to reduce fixed pattern noise, something that EDI fringing signal is immune to since it already dithers). Then, for the readout noise-limited case, we divide the SNR
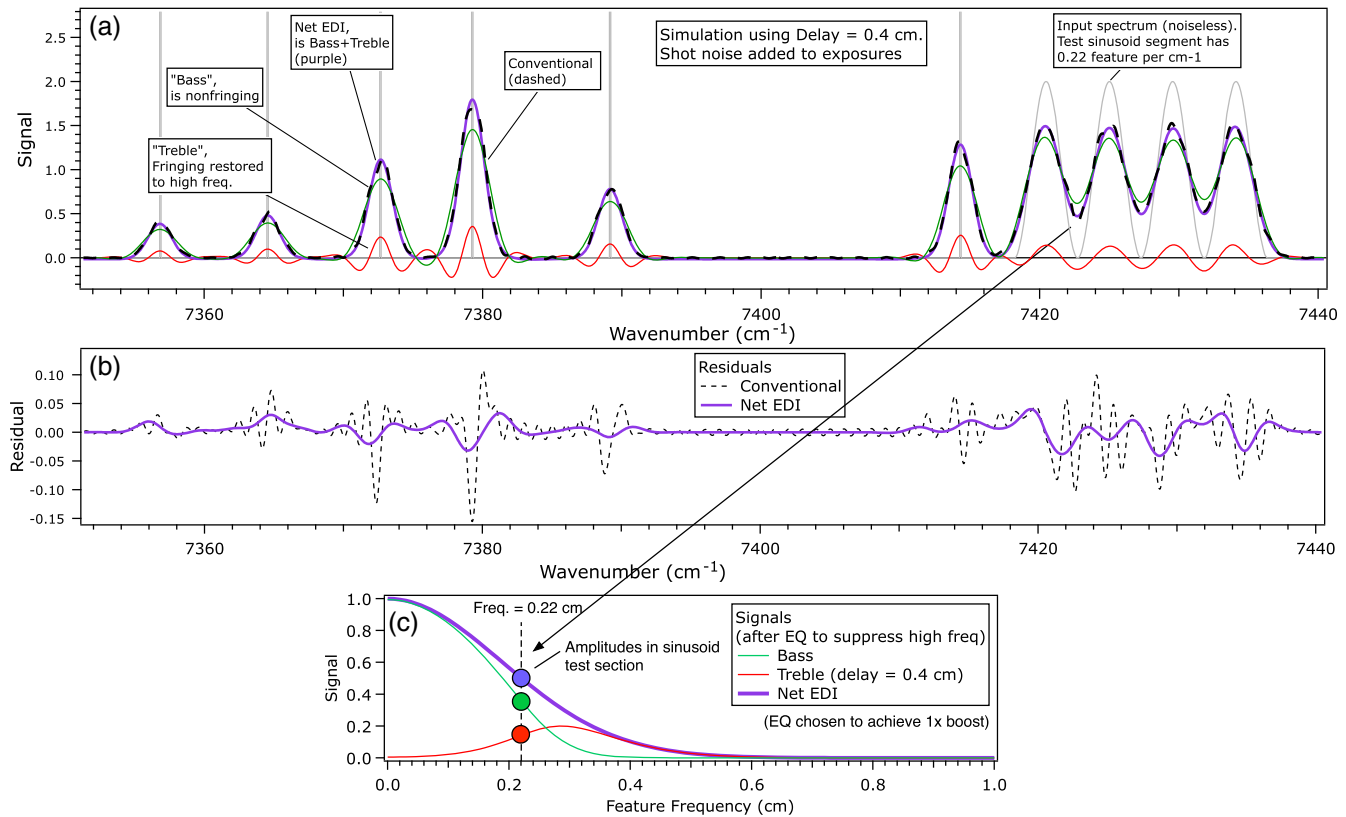
Fig. 7 (a) Components of the numerical simulation plotted versus ν, using simulated shot noise at the detector, delay $\tau = 0.4$ cm, and EQ designed to achieve $1\times$ boost, i.e., same resolution as conventional. Outputs include "Bass" (nonfringing, green), "Treble" (fringing, red), "Net EDI" (purple), and conventional (black dashed). Unblurred input emission spectrum (gray) has a portion of 7420 to 7436 cm^{-1} as sinusoid with a frequency of 0.22 cm, making convenient linkage of amplitudes to height of dots in Fourier response plot (c). (b) Residuals of conventional (black dashes) and net EDI (purple) when in photon-limited regime. EDI has less noise at high frequency. Treble peak in (c) does not appear centered at delay = 0.4 cm due to EQ multiplication, which diminishes high frequencies (both signal and noise). [Fig. 9(c) shows same curves as (c) but with no EQ.]

output by $\sqrt{4}$ or $\sqrt{3}$ to simulate the readout noise-limited case for four (thin purple) or three (long dash thin purple) reads [Fig. 9(d)]. Similarly, in Fig. 10(a), we divide the dotted black photon-limited curve by $\sqrt{4}$ to represent the single readout noise dominated case (black long dashes).

Also, a recent paper[10] proposes using a readout noise-free electron multiplying charge coupled detector (EMCCD), with fast scanning of interferometer phase to improve Doppler precision.

3.3.3 Fixed pattern noise rejection

In the presence of significant fixed pattern (FP) noise, it is fair to assume the conventional technique will dither at least two exposures (such as shifting to an adjacent row of pixels) to cancel this artifact. Since EDI automatically rejects FP noise by its uses of differences between exposures, then for three exposures EDI readout noise would be only $\sqrt{3/2} \sim 1.22$ higher than conventional.

3.3.4 Effective flux increase

We saw in Fig. 10(a) that the EDI (purple) suppresses high frequency noise relative to the conventional (dotted) photon noise. Figure 11 shows RMS averages evaluating noise fraction relative to the conventional photon-limited noise, for various

regions of integration related to the number of pixels per resolution element in ν space. The latter sets the Nyquist frequency, which is the right-hand limit. (These plots are similar to Fig. 10(a) but using the EDI calculator curves.)

Figures 11(a) and 11(b) show that the net EDI has less noise than the conventional, 60% (or 73%) of noise, for 3 (or 2) pixels per resolution elements.

Figures 11(c) and 11(d) show that noise fractions are even smaller (34% and 46% for 3 or 2 pixels) when only high frequencies above 0.22 cm (HWHM the native response) are considered. This is justified since science signals typically require the highest resolution of an instrument (e.g., detecting and locating presence of smaller neighboring peaks).

This amount of noise reduction is conventionally accomplished by increasing flux by a factor $1/0.34^2 = 8.6$ times. Hence, this an effective flux benefit.

In the readout noise-limited case, the conventional noise is either $\sqrt{4}$ or $\sqrt{3}$ times smaller. Assuming four reads for EDI, then the 0.34 fraction becomes 0.68. This means it still has less noise than the conventional, when considering just the higher frequencies that typically manifest the science.

This virtual flux benefit of EDI must be weighed against the decrease in flux due to insertion of the interferometer from parasitic reflections. The topic of how to best construct an ultralow
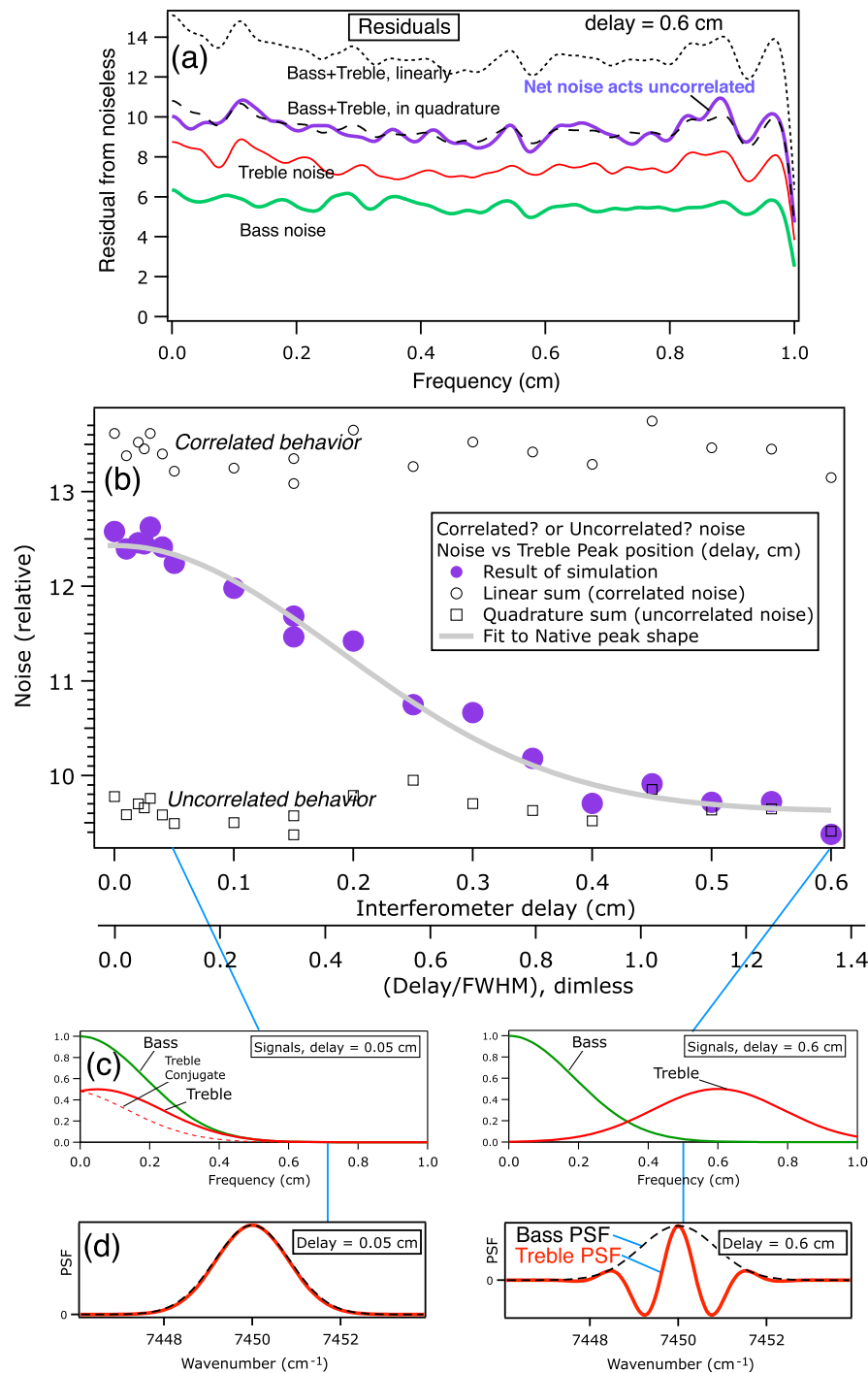
Fig. 8 (a) Example instance of numerical simulation (purple curve) versus frequency with photon noise probes whether the noise type is correlated (sums linearly, dotted black) or uncorrelated (sums in quadrature, dashed black). Fringing (treble, red) and nonfringing (bass, green) components. (b) For photon noise, repeated instances (purple dots) while changing delay fits a peak (gray curve) having same shape as the native $psf_0(\rho)$ or "bass" peak, but with relative height 70% between perfectly correlated (open dots) and uncorrelated (open squares). Detector noise produces uncorrelated behavior independent of delay. (c) Instrument lineshapes in frequency space, small (0.05 cm) delay (left) producing partial correlation, and large (0.6 cm) delay (right) producing noncorrelation. (d) Same instrument lineshapes in wavenumber space. The small delay has similar lineshape components for fringing (red) and nonfringing (black dash). Large delay produces orthogonal lineshapes because of oscillation under peak envelope. Dimensionless delay horizontal axis is normalized by 0.44 cm FWHM of native $PSF_0(\nu)$. Dashed red curve in (c) is wing of conjugate treble peak on the negative frequency branch (only significant for small delays).
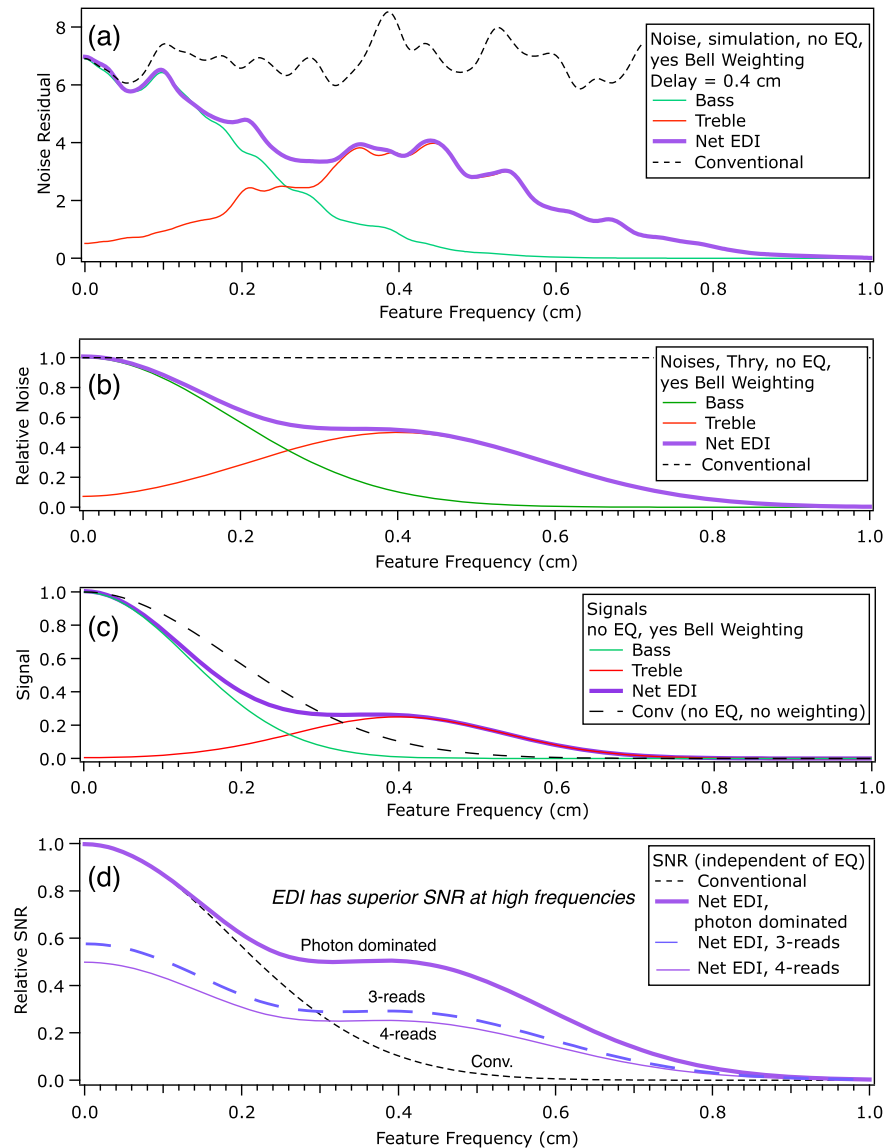
Fig. 9 (a) and (b) Noise, (c) signal, and (d) SNR curves versus feature frequency, ρ, for readout dominated noise and photon dominated noise cases, for a delay of 0.4 cm. (a) The average of 10 simulations (black dotted) with noise confirms the analytical result (b) from "EDI calculator." The conventional photon dominated noise (dashed) in (a) and (b) is nonfringing noise without bell weighting, and hence is uniform with ρ. (c) Signals after bell weighting, for both bass (nonfringing) and treble (fringing). Weighting improves SNR by deleting noise from the high frequency portions of the bass curve from overlapping the 0.4-cm region of the treble. Bell weighting shape is same as original peaks, hence reduces widths of bass and treble signal peaks to 70%, and halves height of the treble. Shaping effects of weighting are normally compensated during EQ process. But EQ step is skipped to better show high frequency behavior. (EQ does not change SNR since it multiplies both noise and signal.) (d) SNR for EDI is higher in the 0.4 cm and higher region than conventional, even when output is divided by $\sqrt{3}$ (long dashes) or $\sqrt{4}$ (thin purple) to account for multiple readouts of EDI relative to conventional, when readout noise dominates.

insertion loss interferometer is interesting and saved for a future paper. We speculate that the monolithic prism interferometer, similar to Refs. 11 and 12, is a fruitful avenue.

3.3.5 *Conclusions on single delay case*

We show that a single interferometer delay can be used to reduce the high frequency noise at the original resolution ("1× boost" case), and that except for delays much smaller than the native response peak half width, the fringing and nonfringing noises act uncorrelated and add in quadrature. This is due to the frequency shifting of the noise due to the heterodyning effect.

4 Multiple Delay Noise Behavior

4.1 *Uniform Exposure Schedule*

Having explored the single delay case, we can extend our understanding on multiple delays. The total input flux will be subdivided into M delays (Fig. 12). We will show below that the peaks add in quadrature to a fixed total sum. Thus, for the case of uniform exposure times for all delays, the height 0.5 of the single fringing peak is subdivided to a height of $0.5/\sqrt{M}$.

Note how the higher delay peaks extend above a hypothetical conventional spectrograph (cross hatching above blue dashed
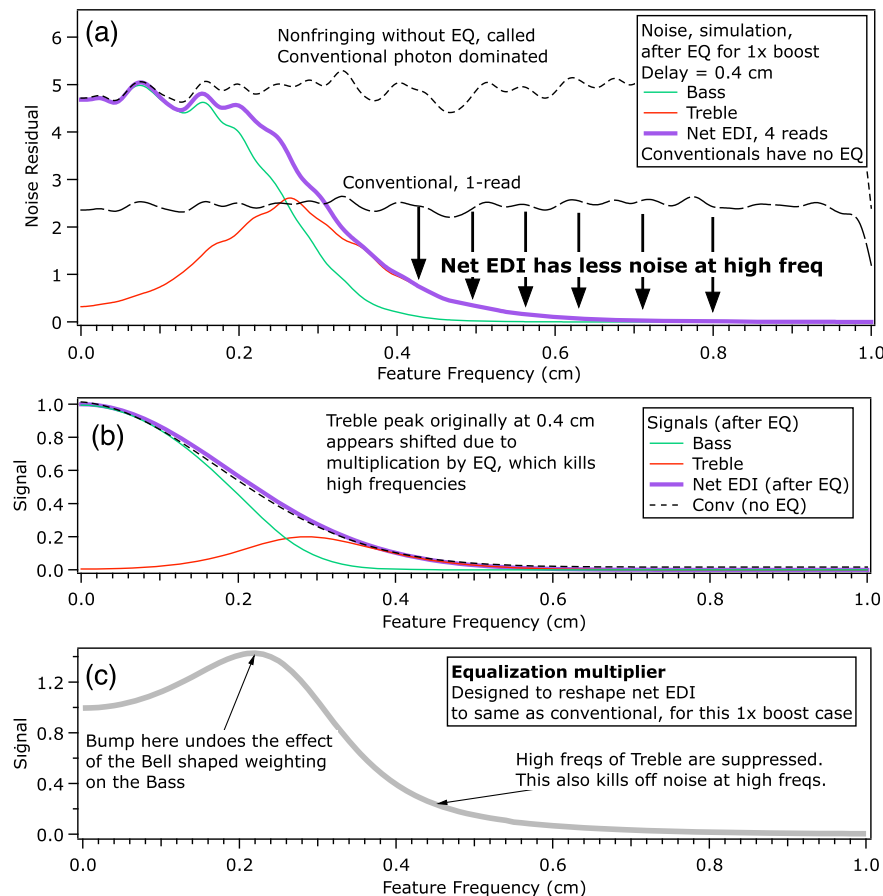
Fig. 10 Simulation output when EQ is applied to achieve "1×" boost, i.e., same final resolution for net EDI as conventional. (a) The photon-dominated conventional noise (short dashes) is divided by $\sqrt{4}$ to estimate single readout noise (long dashes), since simulation uses four exposures. (a) Net EDI noise (purple), bass (green), treble (red). EQ'ing deletes high frequency noise, so net EDI has less noise than conventional photon-limited (or single readout noises), for ρ above 0.2 (or 0.3 cm). (b) Net EDI signal (purple) overlays the conventional signal (black dashes) due to the EQ'ing. (c) EQ(ρ) used.

curve) having the goal 10× boosted native resolution. This shows that the EDI can produce SNR that exceeds the classical at the very highest frequencies, which are the most important frequencies.

This figure and discussion assume that the grating throw length is fixed in size, which is relevant for airborne and space-borne platforms, where volume and mass are critically limited. The blue dashed comparison curve was calculated as if it was an ideal conventional spectrograph with no intrinsic lens blur, achieved by reducing the slitwidth by 10× and decreasing the flux by 10× for an extended source (also requiring 10× more pixels). This is just an artifice to remember EDI behavior, and the native spectrograph could be, for example, an adaptive optics (AO) enhanced spectrograph without a slit.

For an AO spectrograph to increase its resolution conventionally, it would increase its throw length, since its focal spot size cannot be made smaller. If the instrument throw length grows by a factor 10, then native spectrograph (AO or not) has its weight and volume grow by roughly 1000× (ignoring material properties and if all dimensions scaled).

In this case, the comparison response curve would have the same width as the blue dashed curve, but have the same height as the native, i.e., unity on that figure, not $1/\sqrt{10}$. Then, the EDI has

less SNR than the AO result by $\sqrt{10}$, but it needs 10× the pixels and 1000× the volume to accomplish this, a heavy cost to pay.

4.2 Gaussian Exposure Schedule

Figure 13 illustrates a Gaussian schedule of exposure time (flux) per delay. This is an optimal distribution of exposure time to produce a Gaussian final lineshape for the signal, while producing a white (uniform versus frequency) distribution for the noise. This is useful because a Gaussian lineshape lacks ringing and simplifies comparison to conventional spectroscopy, which typically has approximately Gaussian lineshape. The sum rule for peak heights squared was used to redistribute the flux in a Gaussian schedule while preserving the total flux. Due to the square root relationship between noise and flux, this means the exposure time Gaussian is 1.4× narrower than the desired SNR Gaussian (blue dashes).

4.3 Velocimetry Exposure Schedule

For certain kinds of spectroscopy, such as Doppler RV, or for elucidating the most narrow features of a spectrum, neither a Gaussian nor a uniform flux schedule is optimal. Instead, it is best to concentrate the exposure time for a certain range of
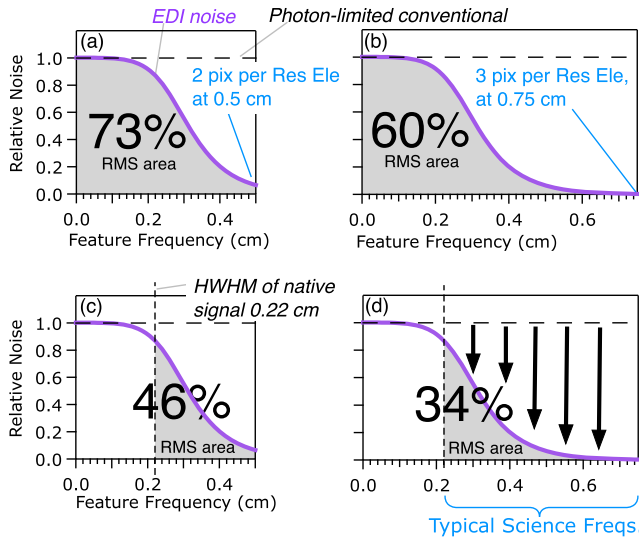
Fig. 11 Fraction of average EDI noise (area under purple) compared to conventional (black dashed line) for case of Fig. 10(a). A root mean square (RMS) integration up to a Nyquist imposed limit set by pixel density. (a) 73% for 2 pixels per resolution element (FWHM in ν-space) or (b) 60% for 3 pixels per resolution elements. The 2 pixel per resolution elements definition is approximately same as 5% signal height definition used in Fig. 15. Thus, the 1.4× improvement of leftmost datum of EDI to conventional in Fig. 15 is justified. (c) and (d) Noise fractions are even smaller (46% and 34%) when only high frequencies are considered, this is where science signals typically reside. The low ρ integration limit is set at 0.22 cm by the HWHM of $psf_0(\rho)$. Note that a 0.34× photon-limited noise reduction is conventionally obtained by increasing flux $1/(0.34)^2 = 8.6$ times.
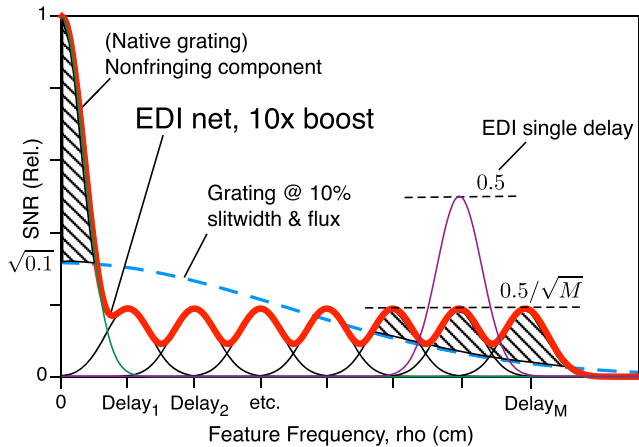

Fig. 12 Multiple EDI sensitivity peaks (black, red is net) of the same height, when each delay has same exposure time, so total flux is evenly divided among M delays. Native spectrograph (green peak at origin) defines unity SNR. An equal-area rule for SNR^2 causes peak height to be $0.5/\sqrt{M}$ for M delays. Blue dashed Gaussian response is for grating with 10× resolution increase and 10× less flux (hypothetically due to slit narrowing 10×). In this figure, the spectrograph throw length is considered fixed, so that avenue for increasing resolution for AO and classical spectrographs is not shown here.

high delays (frequencies), where the most Doppler science lies [see gold dashes of Fig. 13(a)]. The optimal frequencies for Doppler velocimetry are found by taking the derivative of the stellar spectrum and finding the maximum in its Fourier transform. For sunlight, this is a broad peak between 0.5 and 1.5 cm (see Fig. 9 of Ref. 5).
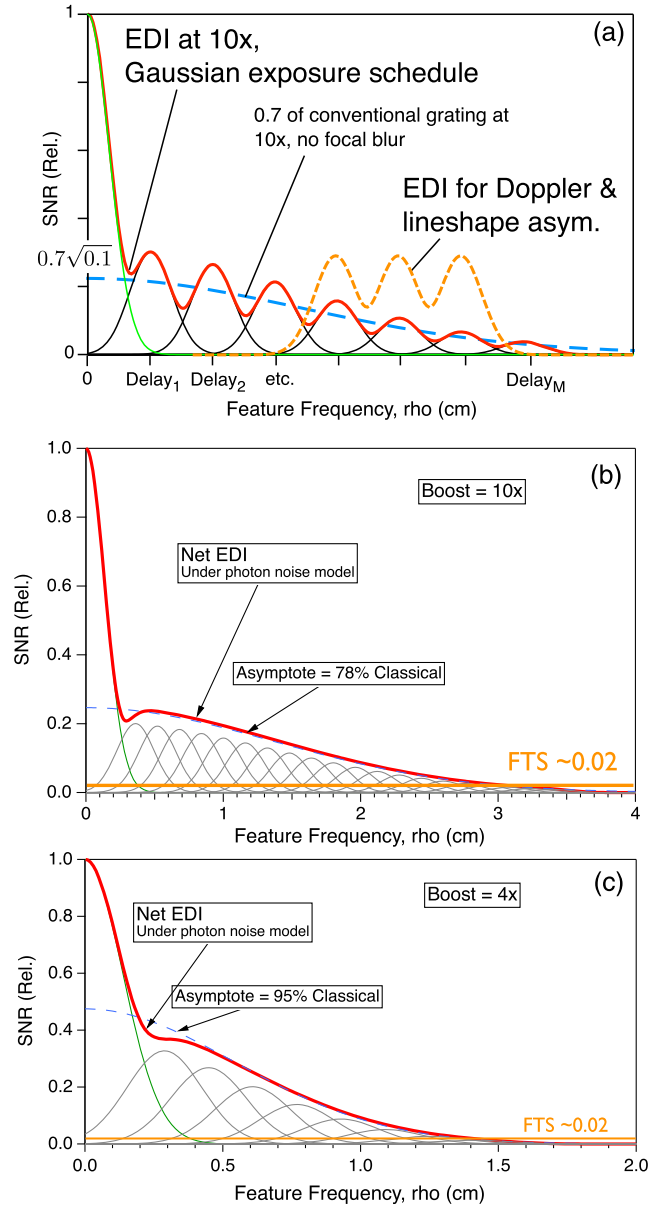

Fig. 13 (a) Net photon SNR behavior (red) when a Gaussian distribution of exposure time versus delay # is used. This is superior to uniform exposure time if white noise behavior is desired, such as in general spectroscopy. For Doppler spectroscopy, concentrating flux in a few high delays (gold dashes) is better. The native spectrograph (green peak at origin) defines unity photon SNR. (b) Asymptote of many overlapped delays, for 10× boost, is ~78% of a "classical" spectrograph having resolution a factor boost larger, and height $\sqrt{boost}$ smaller, as if slitwidth and flux boost times smaller. (c) Asymptote for 4× boost is 95% of classical. Gold line at 0.02 is photon SNR of Fourier transform spectrometer, reduced from native peak by square root of the number (~2500) of native resolution elements, since in Fourier transform spectroscopy (FTS), fringes of different phases sum on a single pixel. In this figure, the spectrograph throw length is considered fixed, so that avenue for increasing resolution for AO and classical spectrographs is not shown here.

4.4 Sum in Quadrature Rule for Peaks

Let us demonstrate that when we use bell-shaped weighting, the net SNR versus ρ curve produced by combining fringing components of different delays and the native nonfringing component is a sum in quadrature. The SNR is a ratio. For uniform

flux, the noise denominator is the same magnitude between different delays. And for other flux schedules, we normalize to force the denominators to be the same magnitude. Hence, we need only to discuss the numerator, which is the EDI sensitivity plots. Therefore, we use plots, such as Fig. 12, that show the fringing response peaks to also represent the photon SNR peaks.

Consider two overlapping peaks 1 and 2, which could include the native peak or different delays. We calculate net SNR by summing the signal S linearly, but combine the noise N in quadrature. We use weightings k_1 and k_2 associated with each peak. Hence,

$$S = k_1 S_1 + k_2 S_2, \qquad (15)$$

$$N = \sqrt{(k_1 N_1)^2 + (k_2 N_2)^2}. \qquad (16)$$

By bell-shaped weighting, we mean that the weight has the same shape as the signal: $k_1(\rho) = S_1(\rho)$ and $k_2(\rho) = S_2(\rho)$. Because we normalize the noise denominators to be the same, $N_1 = N_2 = N_0$. Hence,

$$
\begin{aligned}
\text{SNR} = S/N &= \frac{(S_1^2 + S_2^2)}{N_0 \sqrt{S_1^2 + S_2^2}} \\
&= \sqrt{(S_1/N_0)^2 + (S_2/N_0)^2} \qquad (17)
\end{aligned}
$$

and similarly for multiple peaks. Thus, we have shown the SNR sum in quadrature. (We calculate that the net noise level is not very sensitive to the weighting shape. Rectangular weightings of 1.2 to 1.8 times peak full width at half max (FWHM) produce ∼95% of bell weighting case.)

The total exposure time may be allocated among the delays in various schedules, which affects their heights while representing the SNR(ρ). Since the square of SNR peak height is proportional to the number of photons detected for a delay, and the sum of these is fixed to the total exposure flux, then we have a sum rule for SNR^2 peak heights [Figs. 12 and 13(a)]. The sum rule also works for the area under the SNR^2 curves.

4.5 Distribution of Noise, White or Pink?

Suppose we are not doing velocimetry and thus desire Gaussian final frequency response. Then what is the final frequency distribution of the noise, after any EQ step? Figure 14 answers this question, showing Fourier transforms of the noise (residuals from ideal) from a numerical simulation, for cases of (a) uniform or (b) and (c) Gaussian flux scheduling, and also comparing (c) photon and (b) detector types of noise.

4.5.1 Uniform exposure schedule

For the case of (a) uniform exposure time for each peak, the noise distribution is initially uniform (white). However, a Gaussian EQ is eventually applied to produce the desired Gaussian behavior in the sensitivity. Since the noise is embedded with the signal, the noise also receives this EQ shaping. Hence, the final frequency distribution of the noise in the uniform schedule is Gaussian. It could be called pink noise, having more noise energy at lower frequencies.
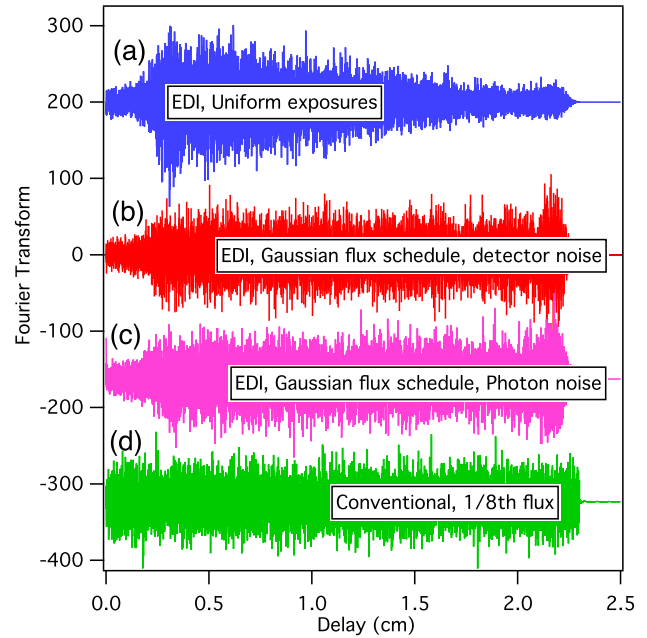

Fig. 14 Fourier transforms of the residuals, after EQ, for different schedules of exposure time (flux) per delay, showing propagation of simulated noise injected at the detector for the case of 10× resolution boost. (a) Uniform exposure schedule (Fig. 12) produces pink noise since a Gaussian EQ is later applied to produce a Gaussian lineshape. (b) and (c) Gaussian schedule of exposures (Fig. 13) using (b) detector and (c) photon noises. After an EQ divides by square root of the Gaussian, the noise becomes a uniform (white) distribution, and signal becomes Gaussian with 1.4× wider width. (d) Native spectrum from one of eight delays (1/8th flux), consistent with rule of thumb that EDI noise is roughly conventional noise at 1/boost flux, i.e., 1/10th flux. (a)–(c) Low noise section (<0.25 cm) due to the cleaner contribution of the native averaged over eight exposure times. Simulated noise was 3% of continuum on four exposures on 0.05 cm^{-1} pixels.

4.5.2 Gaussian exposure schedule

For the case of a Gaussian exposure time schedule versus peak #, the signal distribution is initially Gaussian, and the noise is the square root of that Gaussian. We apply an EQ that divides by this square root Gaussian. This leaves the signal as a Gaussian having a 1.4× wider width and leaves the noise uniform (white noise), as shown by Figs. 14(b) and 14(c). (Hence, the exposure time Gaussian is narrower than the blue dashed curves of Fig. 13.)

This simulation also shows that (c) photon noise, which involves the square root of the flux, and (b) detector noise, which is independent of flux, produce similar overall magnitude of noise, after compensating for the average continuum level of the native spectrum. Hence, there are no surprises for EDI when estimating photon noise by inspecting the average continuum level.

4.6 Selecting Your Delays at Blue End of Band

For a fixed resolving power $R = \nu/\Delta\nu$, the width in ρ-space of the fringing peaks decreases as one moves to the blue (increasing ν), since $\Delta\nu$ increases and thus $\Delta\rho$ decreases. Hence, there is a danger that gaps may open up between the different delay peaks in the blue while being adequately overlapped in the red. Since there is little penalty for having too much overlap (other than excessive number of delays and hence readout

noise), but a severe penalty for a gap (which causes a divide by zero blow up in the EQ and hence increases noise), we recommend selecting the delay positions at the blue end of the band, by subdividing the delay range needed to produce a certain final resolution by the width of $psf_0(\rho)$.

4.7 Readout Noise Case for Multiple

Analogous to the single delay case, the calculation output here for multiple delays is for photon-limited noise, neglecting readout noise. Then to account for readout noise relative to conventional single readout, we divide the EDI SNR peaks by factors of $\sqrt{3}$ or $\sqrt{4}$ for three or four phase steps. We also have to create two different versions for the native peak, one assembled from three or four exposures, and hence having a SNR reduced by $\sqrt{3}$ or $\sqrt{4}$. (This would be the EDI one.) The other would be the original native peak, which would represent the conventional measurement done in a single read but four times longer exposure. The final factor to consider is the increase in readout noise versus ν due to changing phase step size, which is from Fig. 4(a).

4.8 Actual Delay Positions for TEDI

In contrast to the uniform spacing of delays used in this theoretical discussion, the actual delay values used in the TEDI interferometer were irregularly positioned across delay space and had gaps. This was due to the delays being primarily chosen for precision RV, anticipating different sources having different rotational broadening. This required different delay positions over a wide range, and we had only eight positions in our rotary "filter" holder that held the glass delay etalons.

Figure 21 of part 1[2] plots the fringing peak positions in delay space for the TEDI instrument and indicates the TripleSpec native peak as a green peak at the origin. This plot for modulation transfer function is essentially a plot for SNR, since the noise denominator is uniform.

The eight delay values for September 2010, labeled E1 to E8, are 0.083, 0.34, 0.66, 0.96, 1.27, 1.75, 2.92, 4.63 cm. In June 2001, the E1 position was swapped for a new delay called E6.5 having 2.4 cm to fill the gap between E6 and E7 (1.75 to 2.92 cm). This still left a ∼1 cm gap between E7 and E8. In principle, a 10-position rotary holder holding two more delays at 3.5 and 4.0 cm could have made a contiguous coverage. This would have allowed minimal ringing (Gaussian) resolutions up to 36,000 (at 7450 cm^{-1}), rather than the 27,000 we produced. (Processing 36,000 with a delay gap would have produced significant ringing in the lineshape.)

5 Quantifying Performance Relative to Classical

Figures 13(b) and 13(c) red curves show that in the limit of numerous, heavily overlapped, multiple delays having a Gaussian flux schedule, the asymptotic behavior of the EDI is similar to a classical spectrograph of boosted resolution, but having a reduced SNR as $1/\sqrt{\text{boost}}$, consistent with either more read noise from the required more pixels (to maintain fixed number of pixels per resolution element), or more fractional photon noise from less flux (as if the flux was decreasing from a reduced slitwidth acting on an extended source).

Let us quantify the performance of the EDI result by the fraction of this "classical" result. Let the effective Gaussian height (EGH) be the Y-intercept (SNR), where the Gaussian intersects

as the number of delay increases to asymptotically produce a smooth curve, and let FC be the fraction relative to the classical result $1/\sqrt{\text{boost}}$.

For example in (b), the boost is 10×, and the classical value would have EGH at SNR of $1/\sqrt{10}$. The net EDI (red curve) is fitted by a Gaussian shape (blue dashes) of fixed resolution having an SNR EGH of 0.25. This is the classical result diminished to $0.25/0.316 = 78\%$ height, so FC is 0.78. Then, the EGH intercept 0.25 is plotted in Fig. 15(a) as a purple dot at boost = 10.

5.1 Plotting Fraction of Classical versus Boost

How does the EDI compare to the classical result, and how does that vary with boost? Figure 13(b) shows that for a 10× boost,
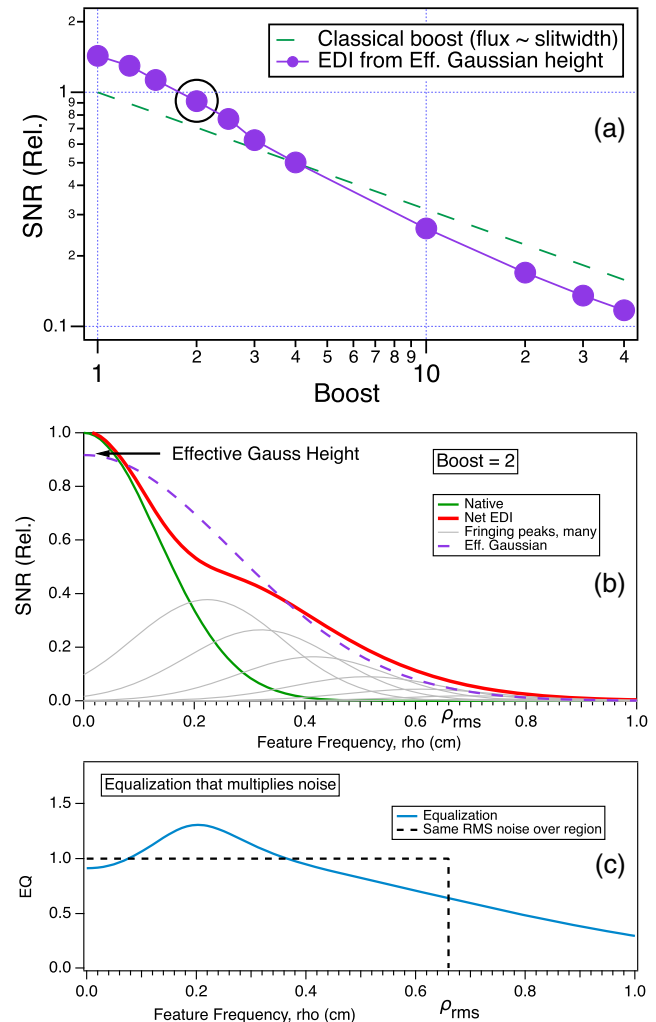

Fig. 15 (a) EGH for EDI versus boost (purple dots). Green dashed line is square root dependence of "classical" standard, having signal-to-noise ratio $1/\sqrt{\text{boost}}$ due to hypothetical reduced flux from reduced slitwidth. (b) Inflected shape of net EDI curve (red) is native response (green) summed in quadrature with EDI peaks (gray). We fit net EDI curve by a single Gaussian (purple dashes) of height EGH and width set by boost. (c) EGH is found by requiring RMS of EQ curve (blue) to be unity–range right limit at 5% height of Gaussian. For boost = 2, EGH is 0.92 and plotted as black circle in (a). Leftmost datum is boost = 1 (the original resolution is retained), showing EGH is 1.4 and SNR is improved relative to a conventional, for the case of photon limited noise (when readout noise insignificant). This is confirmed in more detail by Figs. 10 and 11.

we achieve 78% of the classical SNR, and [c] for 4× boost, we achieve 95% of classical SNR.

Figure 15(a) shows in purple dots how the EGH or Y-intercept for SNR (FC/$\sqrt{\text{boost}}$) varies with a variety of boost values. The green dashed line is the "classic" standard, which has a square root dependence because we are assuming its flux decreases as 1/boost. Note the crossover near boost of 4×. For larger boosts, the asymptotic behavior is that the EDI has a square root dependence like the classical, but a factor $\sim\sqrt{2}$ worse, and for smaller boosts than 4, the asymptotic behavior is a factor $\sim\sqrt{2}$ better than classical.

The reason EDI can be slightly better than classical for lower boosts is that then the native sensitivity peak is much more included in the Gaussian fit of the combination. For very high boosts, the native peak is so much higher than the EDI fringing peaks that it is not effectively included in the Gaussian that must fit both, in the manner described in the caption of Fig. 15(c), which requires the RMS of an EQ curve to be unity.

5.2 Including Focal Blur in Native Model

A more realistic spectrograph will act differently from our classic standard. Namely, it will have a minimum focal blur (FB) that convolves with the slitwidth to put a ceiling to the resolution even while the flux decreases because of the decreasing slitwidth. This behavior is shown in Fig. 16 in the green curve, in comparison to the EDI result in black diamonds. This is a theoretical comparison between EDI and conventional photon noise behaviors versus final resolution, when the native

spectrograph used for both is operated near its resolution limit, and where the slitwidth is summed in quadrature with a constant FB to calculate the native resolution. It illustrates the key point that at some point, every dispersive spectrograph will reach a resolution limit, which is controlled by the FB or the detector pixel density.

6 Comparing TEDI Instrument and Photon Noises

6.1 TEDI Instrument Noise Roughly 3%

For TEDI instrument noise, we observe a large shift (along wavenumber) of the native PSF versus time, as much as 0.6 cm^{-1} in the A-order, as shown by Fig. 35 of part 1.[2] Since the resolution is about 2 cm^{-1} in this order, this is a very large relative instrumental insult. Because it varies in magnitude and even polarity across the band, it cannot be removed by a simple monolithic shift. To express this as a vertical (intensity) noise to compare it to the photon noise, we subtract each of the individual spectra from the average spectra. Figure 38(c) of part 1[2] shows the residual is between 2 and 5%, which we call roughly 3%.

6.2 TEDI Photon Noise Roughly 0.1%

Figure 11 of part 1[2] shows that a typical single exposure of the phase stepping set has of order 10,000 to 30,000 counts per pixel for the continuum portion. (It represents an average single
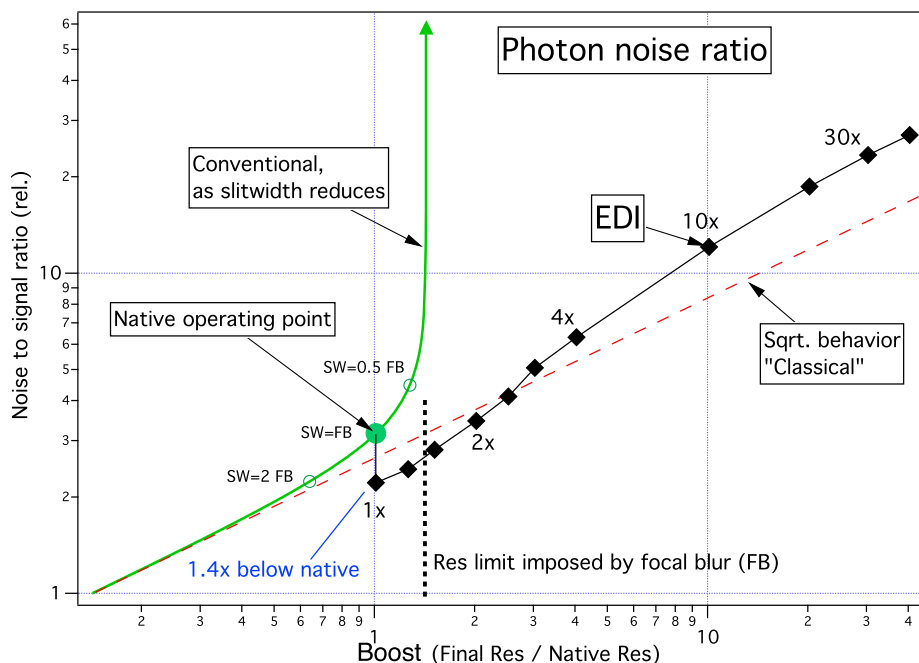

Fig. 16 Theoretical comparison of photon noise-to-signal ratio (arb. units) between the two techniques versus boost (final/native resolution), for EDI (black diamonds), and native-alone (green curve). Hypothetical native spectrograph has a slitwidth (SW) and a FB of 0.70, that sum in quadrature. Red dashed line having square root behavior is idealized "classic" spectrograph standard, whose flux decreases as 1/SW but without any FB. We suppose native spectrograph operating point is at knee of curve (solid green dot), where SW = FB. From operating point, we drop 1.4× to set net EDI noise behavior at boost of 1×, and the rest of the EDI curve follows based on this instance of native spectrograph. The factor of 1.4 is based on reciprocal of Fig. 15(a), and confirmed by simulation Fig. 11. The EDI for low boosts has lower net noise than native-alone because it combines native signal with fringing signal, and they have uncorrelated noises.

exposure, not yet summed over 10 steps. This may not have been clear from that figure caption.)

Summing over the 10 exposures per phase set, the total count is 1 to 3×10^5, and since there are 3.8 photons per count, this is 4 to 11×10^5 photons per pixel. Since there are ~3 pixels per native resolution element, there are about 1 to 3×10^6 photons per native resolution element. Taking the square root, this yields a photon noise of order 0.1%.

6.3 Noise Contribution Plot

From what we have learned about TEDI noise from this and the companion paper (part 1),[2] we assemble a very approximate picture in Fig. 17 of the various types of noise contributions and how they might vary versus resolution boost ratio. For TEDI, the native spectrograph instrumental noise is about 3% due to drifts of the PSF. By contrast, the photon shot noise is about 30 times smaller at about 0.1%. This plot is suggestive and not rigorous — useful for identifying which issues to further investigate.
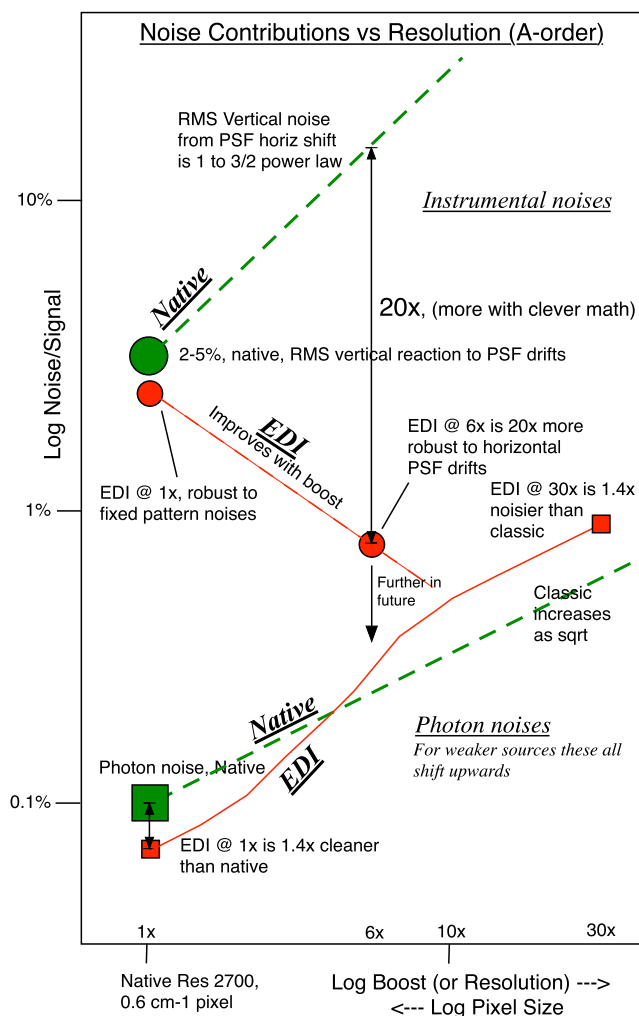

Fig. 17 Estimated relative contributions of various noise contributions for absorption spectroscopy in the A-order of TEDI, compared to native spectrograph (if it could increase its resolution), and how these two vary with boost ratio (EDI resolution over native). Green denotes the conventional "native" spectrograph and red denotes EDI. Circles are instrument noises and squares are photon noises.

The red and green symbols denote EDI and native spectrograph. When boost varies for EDI, it uses the same 1× behavior of the native but with different delay arrangements. But when boost varies for the native, it is used alone and the EDI behavior is not recomputed to use the higher resolution of the new native. Readout noise is neglected here since the photon flux of TEDI was very high.

Native photon: The native classic photon noise would grow as square root of resolution as the flux is assumed to decrease linearly with resolution (we are considering the case in which the spectrograph throwlength is fixed).

EDI photon: The EDI photon noise has the S-shaped behavior, crudely similar to the square root classic behavior, but better than classic by 1.4× at low boosts and worse than classic at high boosts (from Fig. 15).

Native instrumental: Measured vertical error from TEDI's native spectrograph PSF drift in A-order is 2% to 5%. Regarding its dependence on the boost, this approximately follows a power law of 3/2 for blended lines (hence low res) and of power of 1 for isolated lines (at high res). For TEDI data at resolution ~3000, it is in between these values.

EDI instrumental: For low boosts of ~1×, the interferometer comb in the wavelet is as nearly as large as the wavelet envelope. This makes the EDI result more susceptible to native PSF changes. Hence, we place the 1× EDI dot near but slightly below the native dot—below it because the EDI also eliminates fixed pattern noise, so that counts for something.

By contrast, at higher boosts such as 6×, the interferometer comb period is so much finer than the wavelet envelope that it is insensitive to the PSF drifts. Calculations artificially shifting the Moire data of a ThAr line show that TEDI has at least 20× less horizontal reaction to a PSF drift, (and 350× less using a more sophisticated "crossfading" process that modifies the lineshapes. See Sec. 10 of Ref. 2). But at a 6× higher boost, the slope of lineshape that connects between vertical and horizontal errors has increased by 6×, so the net downward movement of the red dot is $20/6 = 3\times$. This reduces the EDI instrumental noise to a 1% level, 20× less than it would be with the conventional alone. Instrument noise is still dominant over estimated EDI photon noise of 0.2% at 6× boost.

7 Zoology of Different Spectroscopy Methods

Figure 18 is a notional plot comparing several methods of spectroscopies in Fourier space (delay space), including FTS. The general goal is essential to map the Fourier information of the source spectrum over a delay range. The higher the maximum delay, the higher the achieved resolution. Apparatus photos and example data of several different kinds of dispersed interferometers are shown in Ref. 13. Different techniques accomplish the mapping in different manners. Either (a) all at once (purely dispersive), (b) subdividing the delay space in a series of discrete chunks (EDI), (c) one chunk that is continuously scanned (dispersed FTS), (d) spatially recording the delay range at once [spatial heterodyning spectroscopy (SHS)], or (e) an extremely narrow spike that is scanned (pure FTS).

7.1 Comparison to Dispersive Spectrographs

The purely dispersive method (a) has a peak at the origin whose width is proportional to spectral resolution. Let the FWHM be τ_{max} (which is also the approximate the rightmost extent of the delay range for a FTS or EDI to map out to the wing of the Gaussian). Then, from the uncertainty principle:
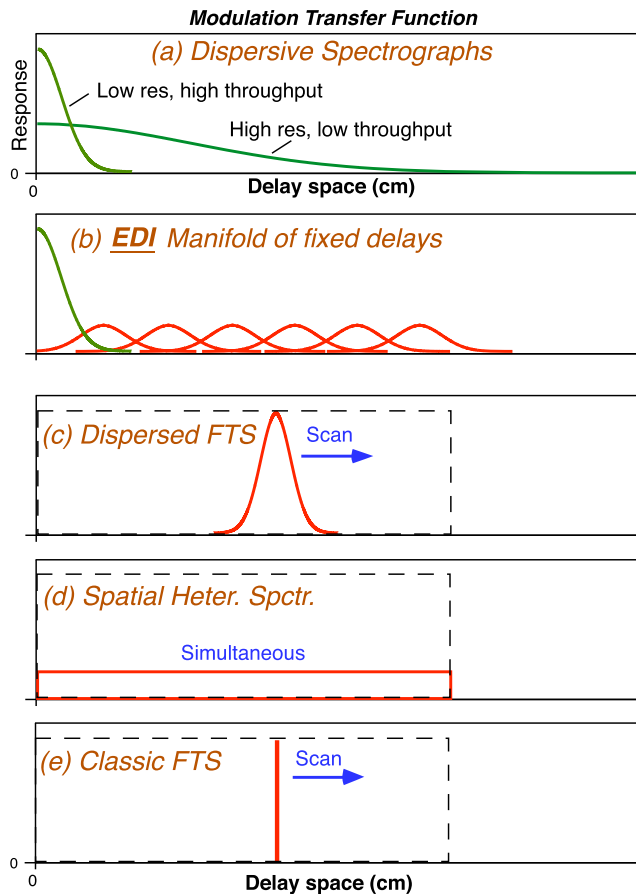
Fig. 18 Notional arrangement of astronomical spectrograph systems, from purely dispersive (top) to purely interferometric (bottom), in suggesting photon-limited SNR in Fourier space (frequency or delay space). (a) Dispersive spectrographs need a wide peak for high resolution– the goal is to cover maximal delay. (e) Purely interferometric FTS maps out delay space directly (with a narrow peak); (d) 1-d spatial heterodyning spectroscopy (SHS) splays a range of delays spatially along a detector at once, recording interferogram on an integrating detector. Since fringes of many phases overlap on pixels, SNR is reduced from dispersive case. (b) and (c) EDI and dispersed FTS are hybrids having a medium wide peak. The EDI measures delay space in chunks; the dispersed-FTS scans the peak.

$$\text{Res} \sim \nu \tau_{\max} \sim \tau_{\max}/\lambda, \tag{18}$$

which can be easily remembered as the number of wavelengths that fit into $\tau_{\max}$. The Res $= \delta\lambda/\lambda = \delta\nu/\nu$, wavenumber $\nu = 1/\lambda$ when wavelength is in cm, and maximum interferometer delay $\tau_{\max}$ in cm. So, a Res 50,000 dispersive spectrograph at 1 μm wavelength has a peak 2.5 cm in half width, or ~5 cm to the wing of the Gaussian.

Panel (b) shows the EDI, which maps delays space in chunks (one peak per delay) set by the native peak (green), which now can be narrower (lower resolution) than in the purely dispersive case (a). The heights of the peaks are $0.5/\sqrt{M}$ relative to the native. Importantly, and the subject of the companion paper,[2] the center of each peak is the most stable region against PSF translations, and the EDI places these at high frequencies, where the science information resides. By contrast, the classic spectrograph has this stable region at zero frequencies, which does less good for the science signals. Hence, the EDI can be an order of magnitude more robust to PSF drifts for the important high frequencies.

7.2 Comparison to Fourier Transform Spectroscopy

The purely interferometric method (e, FTS) scans the interferometer delay continuously over the delay range, recording the Fourier information with an extremely narrow peak. It then Fourier transforms this into a spectrum. The scanning delay requires a time responsive detector capable of recording high frequencies and prevents use of integrating detectors.

By contrast, the EDI does not scan a delay continuously but sits in several discrete positions. It can use the slow (integrating) but sensitive CCD detectors already present in astronomical spectrographs. Thus, EDI can be an add-on unit to enhance existing spectrographs.

For measuring single shot or rapidly changing phenomena, we have designed an EDI using multiple delays in parallel on different detector regions to make snapshot measurements (see Figs. 12A and 12B of Ref. 14).

The EDI signal has lower photon noise[15] than the FTS by the square root of the number of native spectral resolution elements, because the disperser isolates adjacent wavelengths on the detector having independent phases. For an echelle spectrograph that is of order $10^3 - 10^4$, the EDI can have a better photon SNR over the FTS by a factor of 30× to 100×.

7.3 Comparison to Dispersed-Fourier Transform Spectroscopy

In a dispersed-FTS[16,17] shown in Fig. 18(c), a FTS is in series with a disperser. The latter increases fringe visibility, so its photon limited SNR is intermediate between purely dispersive and purely interferometric cases. It is sufficiently high to allow it to measure stellar spectra and Doppler velocities, such as the spectroscopic binaries measured at the Steward Observatory 2.3 m Bok telescope.[16,17] The scanning delay requires time responsive detectors.

7.4 Comparison to Internally Dispersed Interferometers

The internally dispersed interferometer techniques called SHS[12,18] or heterodyning holographic spectroscopy[19] is related to an FTS, but where the delay range is recorded at once splayed spatially, and thus can use an integrating detector rather than being scanned over time [Fig. 18(d)].

Similar to the other hybrids, the photon-limited SNR is generally intermediate between purely dispersive and purely interferometric extremes, because there is some overlap of signal between pixels, which degrades SNR by a square root effect, as described in Eq. A40 of Ref. 19. The SHS technique is known for its very high resolution at high etendue of extended objects such as the atmospheric glow. Recording the delay range at once makes the instrument very rugged and well suited for aerospace platforms, such as described in Ref. 12, for measuring upper atmospheric wind by emission lines.

The one-dimensional (noncross-dispersed) internally dispersed interferometer can have a significantly reduced BW, because it produces a fringe comb, whose period varies strongly with wavenumber (due to the internal grating that changes the angle of interference), and thus, the fringe frequency can exceed the pixel pitch outside of a BW. Within this band, the resolution can be extremely high. (By contrast, the EDI has an almost uniform interferometer comb period. This allows a much larger BW, limited only by the native spectrograph.) However, newer cross-dispersed SHS have been demonstrated[18,20] that produce

a two-dimensional interferogram, and these have a much wider BW than the noncross-dispersed type.

7.5 Comparison to Super-resolution Techniques in 2-D Imaging

Related mathematical methods of enhancing image resolution by increasing its width in Fourier space have been developed in other fields. For example, the microscopy method of "structured illumination"[21] creates Moire patterns with spatial grids at various orientations.

We caution the reader not to confuse EDI with the "superresolution"[22] technique of photo enhancement that relies on the alias signal developed when the signal is undersampled and translated in subpixel displacements. Our technique does not use the alias signal and excludes it from the processed signal by filtering. We avoid using the alias signal because it is susceptible to irregular placement of the pixels, which can occur on a subpixel level, say, 0.1 pixel. This level may not be of concern to ordinary imaging but is significant to spectroscopy.

7.6 Relation to Amplitude Squeezed Light

Because the noise in the fringing and nonfringing signals is uncorrelated, when the two signals are combined, the net SNR can be better than the native nonfringing signal used alone. A related sub-shot-noise behavior has been previously observed by other researchers[23] (a topic called "amplitude-squeezed light"), using a conceptually similar experimental arrangement to EDI, but without the spectrograph. Namely, they have an interferometer with detectors on both (complimentary) outputs. This detects a photon both by summing the complementary outputs (the classical way) while also simultaneously subtracting the two complementary outputs; combining both signals produces a sub-shot-noise level of net SNR.

8 Concluding Remarks

We show that a single interferometer delay can be used to reduce the high frequency noise at the original resolution ("1× boost" case), and that except for delays much smaller than the native response peak half width, the fringing and nonfringing noises act uncorrelated and add in quadrature. This is due to the frequency shifting of the noise due to the heterodyning effect. We study the change between uncorrelated and partially correlated noise as the delay goes to zero.

We find a sum rule for the noise variance for multiple delays. The multiple delay EDI using a Gaussian distribution of exposure times has a noise-to-signal ratio (NSR) similar to a classical spectrograph with a proportionately reduced slitwidth to achieve the boost in the classical manner, but without the focal spot limitation and pixel spacing Nyquist limitations. That is, NSR $\sim \sqrt{\text{boost}}$. At low resolution boosts ($\sim 1\times$), the EDI has slightly smaller ($\sim 1.4\times$) noise than the conventional, and at higher than four boosts, the EDI has slightly larger ($\sim 1.4\times$) noise than conventional.

The $\sim 1.4\times$ better than conventional noise at low boost is due to combining fringing and nonfringing components while their noises are uncorrelated due to heterodyne shifting. The $\sim \sqrt{2}$ worse than conventional noise at high boosts is due to the factor 2 smaller height of the single delay fringing peak relative to the native, and the sum rule that spreads this SNR^2 over several delays.

The readout noise decreases as the square root of number of reads, motivating use of three or four reads instead of the 10 used

in TEDI. With the irregular phase steps stemming from using three or four phase step exposures at changing wavenumbers, the BW is still comfortably large ($\sim 2:1$) sufficient to handle, for example, the visible band (400 to 700 nm, 1.8:1).

8.1 Uncertainty Principle Followed

Some readers may find it nonintuitive that including an interferometer can boost resolution (or equivalently, decrease noise at constant resolution). Consider that the coherence length of the light passing through a high-resolution grating spectrograph is broadened more than passing through a low resolution spectrograph (imagine a single perfectly short input pulse). The output will consist of a train of pulses, one per grating groove. Thus, by including an interferometer of significant delay with a grating, the output will also have a broadened coherence length, comprising the convolution of the grating pulse train with a two-pulse impulse response (Fig. 19). Following the uncertainty principle, the increased coherence length is consistent with a spectral resolution increase.

This increase incoherence length causes greater ambiguity for time scales of τ/c or $0.05/3 \times 10^8$ or 166 picoseconds. This is not a problem for astronomy which typically measures much slower phenomena. The increased ambiguity applies to any method of increasing spectral resolution including those in conventional dispersive spectrographs.

Operating a high-resolution spectrograph system is then a business of producing very long coherence lengths, and doing it in a very controlled manner. The dithering of an EDI interferometer delay is very controlled, and the interferometer has only three degrees of freedom. By contrast, the grating has a multitude of degrees of freedom, at least one per grating groove, and many of these are uncontrolled by environmental insults.

8.2 Externally Dispersed Interferometry Can Benefit Adaptive Optics Spectrographs

The EDI can be useful in boosting the resolution and stability of an AO spectrograph, whose resolution is limited by the number of detector pixels. Ordinarily, to increase the resolution of an AO enhanced spectrograph (or classical spectrograph limited by FB), one increases the throw length, and the number of pixels, by a factor boost. This has the disadvantage of increasing the volume and weight of the instrument by a power law having an exponent between 2 and 3, with the expense growing nonlinearly as well.
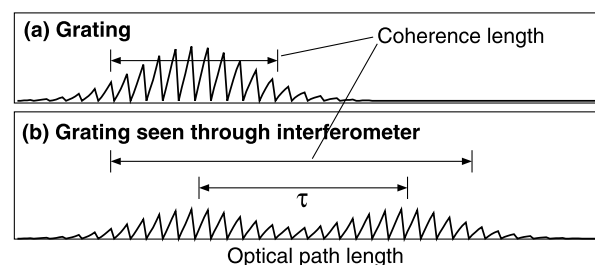

Fig. 19 Apparent increase in coherence length of a grating when viewed through an interferometer having delay τ. Any object viewed through a Michelson interferometer appears twice (with 50% intensity for each image), and with the second image delayed. The two images of the grating appear as a single grating with a longer coherence length. Since the spectral resolution is proportional to the net grating coherence length, the resolution increases.

Many airborne and spaceborne platforms have severe weight and volume constraints. Hence, using EDI with an AO spectrograph is an attractive means of achieving higher resolution and stability without exceeding the financial weight and volume limits.

8.3 Externally Dispersed Interferometry Can Benefit Integral Field Spectrographs

EDI can benefit integral field spectrographs that strive to produce spectra for each point on an image, because these systems are typically starved of pixels. Even if AO is used to produce a diffraction-limited focal spot, the spectral resolution is still limited by the paucity of pixels in the spatial dimensions. Consider an objective prism creating a rainbow for each object. One cannot spread the rainbows over very many pixels without danger of overlapping rainbows of adjacent objects.

The EDI can boost the spectral resolution by inserting an interferometer along the beam path. By taking a series of measurements at different delays and combining the results, the EDI boosts the final resolution to supersede what pixels allow in a single exposure.

Provided a fairly light efficient interferometer is used, we believe it would benefit spectrographs to have an EDI as a front end. The suppression of fixed pattern noise and the enhancement of the stability of the PSF through the sinusoidal fiducial comb are just as important of advantages as the resolution boost, especially for precision radial velocity, which requires an extremely stable PSF. New data analysis methods that can further improve the PSF stability, potentially up to 350×, by crossfading overlapped pairs of delays by reshaping the lineshape is an exciting new development (Sec. 10 of part 1).[2]

8.4 Interferometer Fiducial Comb Lowers Spectrograph Cost

The EDI enhances PSF stability through its sinusoidal fiducial comb, which is embedded with the input spectra and shifts along with it under a PSF shifting insult. Thus, the Moire (which depends on difference between spectrum and comb) is largely robust to a PSF shift.

By enhancing the robustness of the native spectrograph with the fiducial comb, the structural and optical tolerances of its design can be relaxed, saving cost and weight. It is possible that the bulky and heavy vacuum tank enclosing some spectrographs could be eliminated. Optical mounts can be made lighter. Thermal expansion can be less worrisome not requiring special materials. The diffraction grating and other optics could be optimized to maximize throughput rather than reduction of aberrations. Some lens elements or mirrors may be eliminated, increasing throughput and decreasing weight. Hence, the EDI presents a new and potentially useful leverage on the design trade study, which could potentially improve the SNR achieved at a given needed resolution.

Appendix A: Numerical Simulator Equations

The equations used for the numerical simulation of the EDI are below. The simulator begins at Eq. (1) with a computation in wavenumber space, using a hypothetical spectrum that is multiplied by a sinusoid and then blurred. Calculational pixels are 0.05 cm^{-1} wide. The blur is 40 pixels (2 cm^{-1}) FWHM at 7450 cm^{-1} for a native resolution of 3725.

The added random noise is of either two types, detector (magnitude independent of flux) or shot noise (magnitude scales

as square root of B_n prior to noise). We used a noise standard deviation amount of 3% in some simulations and 30% of continuum in others. The calculational pixels were 0.05 cm^{-1}.

The fringing (**W**) and nonfringing (B_{ord}) components are separated by idealized four step Eqs. (6) and (7). The variation of phase step with ν across the band is ignored here and treated elsewhere. This is justified since EDI works at a local level as small as an individual resolution element.

Since it is easy to forget a factor of 2 somewhere in the math chain, especially if sometimes we use a Fourier transform displaying both frequency branches and at other times a single branch, we confirm the correct relative sizes of the **W** and B components in the simulation results by inspecting the output of the purely sinusoidal test section of the input spectrum, 7420 to 7436 cm^{-1} in Fig. 7(a) with the expected magnitudes from the theory in frequency space indicated by red and green dots of panel (c).

Prior to heterodyning, reversal of some filtering is performed to kill off high frequencies that contain mostly noise and little or no signal. This can improve the photon-limited SNR by a factor of 1.4× because it prevents two frequency branches of noise, positive and negative, from combining into the final signal. At this point, the signal lies in the neighborhood of zero, so we optionally apply low pass filtering.

While rectangular lowpassing is the easiest to code and is the minimal amount done, it is optimal to apply a filter passband shape that has the same shape and magnitude as the expected signal response, which has a shape $\mathrm{psf}_0(\rho)$, which we model as a Gaussian. We call this "bell-shape weighting" to be more generic, denoted by $k(\rho)$. This is really meaningful in the region, where two different signals overlap (bass and treble, or trebles that belong to multiple delays), since for an isolated component, the EQ function would undo the effect of any weighting. The net noise changes slowly from the use of a nonideal shape or magnitude of weighting.

Gaussian shaped (bell-shaped) weightings are applied:

$$k_0(\rho) = \mathrm{psf}_0(\rho) = \mathrm{Gauss}(\rho), \tag{19}$$

or $k_0 = 1$ if no weighting is used.

$$w'(\rho) = 0.5 w(\rho) k_0(\rho), \tag{20}$$

$$b'_{\mathrm{ord}}(\rho) = b_{\mathrm{ord}}(\rho) k_0(\rho). \tag{21}$$

The factor 0.5 for weighting w is because its expected response is half that of the nonfringing component.

Reversal of heterodyning is more accurately computed in pixel or wavenumber space on $W(\nu)$, even though we discuss it as being $w(\rho)$ shifted in frequency space by an amount of delay τ:

$$W''(\nu) = e^{-i2\pi\nu\tau} W'(\nu). \tag{22}$$

The polarity of the exponent is chosen to restore the Moire signals to the high frequencies they originally had in the input spectrum S_0. A continuum in the input spectrum makes a sinusoidal comb in the measured **W**. Our convention is that we assign that a positive frequency. Hence, during heterodyning reversal, we apply a negative exponent to shift that down to zero frequency to recover the constant continuum.

The EQ is chosen to reshape the hump like shape of $(b_{\mathrm{ord}} + w)$:

$$\text{netEDI}(\rho) = [b'_{\text{ord}} + w(\rho)''] \text{EQ}(\rho) \tag{23}$$

into an ideal Gaussian of a user selected width (choosing the final resolution). The EQ is the ratio of desired shape divided by actual shape, which is

$$\text{EQ}(\rho) = \frac{\text{Gauss}(\rho, \text{FWHM} * \text{boost})}{(k_0 \text{psf}_0 + k_{\text{edi}} \text{psf}_{\text{edi}} + \epsilon)}, \tag{24}$$

where $\epsilon = 1e^{-10}$ to prevent divide by zero blow ups.

$$k_{\text{edi}} \text{psf}_{\text{edi}} = 0.5 k_0(\rho + \tau) \text{psf}_0(\rho + \tau)$$
$$+ 0.5 k_0(\rho - \tau) \text{psf}_0(\rho - \tau). \tag{25}$$

The factor 0.5 is because the fringing response peak is half of the native. This includes the contribution of the conjugate (or "twin") treble peak in the other (negative) frequency branch, which can bleed over into the positive branch for very small delays relative to the native psf_0 width.

Appendix B: Code for Externally Dispersed Interferometry Calculator

The EDI calculator is a faster way to visualize EDI behavior than the numerical simulator, and it lacks the statistical variations. It differs from the simulator by not using a specific input spectrum. It shows frequency response (ρ-space). It is based on equations that we have confirmed reproduce the numerical simulation results. Script written in Wavemetrics Igor data analysis application.

```
Functions below are in Fourier space, which is feature frequency space having variable rho
  (units cm).
=============== Gaussian Peak maker =======================
function GaussRsp(Res,aveWn,tau) // Calculates gaussian blurring vs rho, of grating
variable Res, aveWn, tau // Res is resolution, aveWn is in cm-1, tau is delay in cm
return exp(-(((x-tau)*1.133*aveWn)/(0.6006*Res))^2); // This is wavenumber savvy
end

================= Signals prior to EQ, and weighting ========
//<Bass signal, called hSigBassFinal>
hBass = GaussRsp(3725,7450,0); // Makes Gaussian peak centered at delay (3rd parameter), width
  from resolution in
    //first parameter, and average wavenumber in 2nd parameter.
//<Treble signal, called hSigTreb>
hTreble = 0.5*GaussRsp(3725,7450,tauG);
hTrebleTwin = 0.5*GaussRsp(3725,7450,-tauG); // This is the conjugate treble peak, on the other
  frequency branch,
    //that can bleed over when delay is small.

=============== Make Weighting: Bell, or none ==============
k0 = hBass // Make weighting have same bell shape as expected signal.
kedi = hTreble // Make weighting have same bell shape as expected signal.
kediTwin = hTrebleTwin // Make weighting have same bell shape as expected signal.
or if desiring no weighting, make all weights = 1

=============== Signals after weighting, prior to EQ =======
hBassFinal = hBass*k0 = hBass*hBass
HTreble2 = hTreble*kedi = kTreble*hTreble
HTrebleTwin2 = hTrebleTwin*kediTwin = hTrebleTwin*hTrebleTwin
hTrbBoth = HTreble2+HTrebleTwin2

//< Net EDI signal, called hFinal>
hAll = hBassFinal + hTrbBoth

=============== Make EQ multiplier ========================
// Currently the net signal has a hump and thus does not have a ideal Gaussian shape. The purpose
  of the EQ is to force
    // the final net signal into a Gaussian shape. The final width is chosen by user.
Boost = 1 // We are examining the particular case where final resolution is same as native.
hBassWider = GaussRsp(3725*Boost,7450,0) // This one differs from hBass by having resolution
  increased by Boost
hEq = hBassWider/(hAll+1e-10) // The 1e-10 useful for preventing blowups
```

```
=============== Signals after EQ =========================
hSigBassFinal = hBassFinal*hEq // Equalization
hSigTreb = hTrbBoth*hEq // Equalization
hFinal = hAll*hEq // Equalization

============== Noises =================================
//<Bass Noise, called hBassN>
hBassN = 1*k0 = 1*hBass // The noise (constant) times the bell weighting
If no bell weighting, then hBassN = 1
hBassN *= hEq // Equalization

//<Treble Noise, called hTrbBothN>
hTrbBothN = sqrt(kedi*hTreble + kediTwin*hTrebleTwin) = sqrt(hTreble^2 + hTrebleTwin^2)
 // Add conjugate from other
//frequency branch in quadrature, after applying weightings
If no bell weighting, then hTrbBothN = 1.414
hTrbBothN *= hEq // Equalization

//<Net EDI Noise, called hPurpleN>
//We will calculate both linear and quadrature versions, then mix them appropriate to the delay.
 Unless delay is very
    //small, quadrature is the usual result.
hAllN = sqrt(hBassN^2 + hTrbBothN^2) // <Sum in quadrature version>
hAllN *= hEq // Equalization
hAllNL = hBassN + hTrbBothN // <Linear version>
hAllNL *= hEq // Equalization

if (YesShotTypeNoise) // If Shot noise, then its somewhere between Quadrature and Linear
  (uncorr and corr).
    // Use Fig. 8(b) (noise vs delay), and MaxCorr*hBass(tauG) to estimate where in between the
      two asymptotic behaviors.
    // The peak is empirically found to fit a shape that is same as hBass, but only 70% high, //not
    the full 100% between the linear (upper) and quad (lower) behaviors. MaxCorr = sqrt(0.5) //
    Empirically found to be ~0.7, guessing its 0.707
    hPurpleN = hAllN*(1-MaxCorr*hBass(tauG)) + MaxCorr*hBass(tauG)*hAllNL
    else
    hPurpleN = hAllN // If detector noise, then purely quadrature
    endif

=============== SNR =================================
hPurpleSNR = hFinal/hPurpleN <For Net EDI, no readnoise case (photon limited)>
hPurpleSNR2 = hPurpleSNR/2 <For Net EDI, simulate four reads by dividing by sqrt(4)>
hTrebSNR = hSigTreb/hTrbBothN <For Treble component, fringing>
hConvSNR = hBass/1 <Bass is same as hBass because noise is uniform>
```

Acknowledgments

References

1. J. C. Wilson et al., "Mass producing an efficient NIR spectrograph," *Proc. SPIE* **5492**, 1295–1305 (2004).
2. D. J. Erskine et al., "High-resolution broadband spectroscopy using externally dispersed interferometry at the Hale telescope: part 1, data analysis and results," *J. Astron. Telesc. Instrum. Syst.* **2**(2), 025004 (2016).
3. D. J. Erskine et al., "High resolution broadband spectroscopy using an externally dispersed interferometer," *Astrophys. J. Lett.* **592**, L103–L106 (2003).
4. D. J. Erskine and J. Edelstein, "Interferometric resolution boosting for spectrographs," *Proc. SPIE* **5492**, 190–199 (2004).
5. 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).
6. 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).

7. J. Ge et al., "The first extrasolar planet discovered with a new-generation high-throughput Doppler instrument," *Astrophys. J.* **648**, 683–695 (2006).

8. 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).

9. D. J. Erskine et al., "Two-dimensional imaging velocity interferometry: data analysis techniques," *Rev. Sci. Instrum.* **83**(4), 043116 (2012).

10. 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).

11. S. Mahadevan et al., "An inexpensive field-widened monolithic Michelson interferometer for precision radial velocity measurements," *Publ. Astron. Soc. Pac.* **120**, 1001–1015 (2008).

12. J. M. Harlander et al., "Design and laboratory tests of a Doppler asymmetric spatial heterodyne (dash) interferometer for upper atmospheric wind and temperature observations," *Opt. Express* **18**(25), 26430–26440 (2010).

13. D. J. Erskine, *The WSPC Handbook of Astronomical Instrumentation*, Vol. **3**, Chapter Dispersed Interferometers, World Scientific Publishing Company, Singapore (2017).

14. D. J. Erskine, "Combined dispersive/interference spectroscopy for producing a vector spectrum," US Patent 6,351,307 (2002).

15. R. Beer, *Remote Sensing by Fourier Transform Spectrometry*, John Wiley & Sons, New York (1992).

16. B. B. Behr et al., "Stellar astrophysics with a dispersed Fourier transform spectrograph. I. Instrument description and orbits of single-lined spectroscopic binaries," *Astrophys. J.* **705**, 543–553 (2009).

17. B. B. Behr et al., "Stellar astrophysics with a dispersed fourier transform spectrograph. II. Orbits of double-lined spectroscopic binaries," *Astron. J.* **142**, 6 (2011).

18. J. Harlander, R. Reynolds, and F. Roesler, "Spatial heterodyne spectroscopy for the exploration of diffuse interstellar emission lines at far-ultraviolet wavelengths," *Astrophys. J.* **396**, 730 (1992).

19. N. Douglas, "Heterodyned holographic spectroscopy," *Publ. Astron. Soc. Pac.* **109**, 151 (1997).

20. A. Bodkin and A. Sheinis, "Multiband spatial heterodyne spectrometer and associated methods," US Patent 8,154,732 (2012).

21. M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).

22. G. Cristóbal et al., "Superresolution imaging: a survey of current techniques," *Proc. SPIE* **7074**, 70740C (2008).

23. Y.-Q. Li, D. Guzun, and M. Xiao, "Sub-shot-noise-limited optical heterodyne detection using an amplitude-squeezed local oscillator," *Phys. Rev. Lett.* **82**(26), 5225–5228 (1999).

David J. Erskine has been an experimental physicist at Lawrence Livermore National Laboratory since 1987 and has experience in femtosecond lasers, semiconductor physics, superconductivity, 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. Since 1998, he has collaborated with astronomers to innovate interferometric techniques for the Doppler planet search and high-resolution spectroscopy. He is a member of SPIE.

Edward H. Wishnow received his 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 midinfrared and visible.

Martin Sirk has over 30 years of 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, microchannel plate detectors and optics on six NASA missions (EUVE, FUSE, ORFEUS, CHIPS, SPEAR, ICON), and photographic plates at Lick Observatory.

Biographies for the other authors are not available.