Achieving High Absolute Accuracy for Group-Delay Measurements Using the Modulation Phase-Shift Technique

T. Dennis and P. A. Williams

***Abstract*—We have developed a modulation phase-shift (MPS) system for measuring relative group delay (RGD) in optical components with high absolute accuracy and simultaneously high temporal and wavelength resolution. Our 200-MHz system has a 3.2-pm wavelength resolution and has demonstrated a group-delay resolution of 0.072 ps for repeated measurements of an artifact based on a hydrogen-cyanide gas cell. The expanded uncertainty (2σ) is $\pm$0.46 ps for a single spectral measurement ($\sim$ 3.4-pm steps) of a narrow 20-ps group-delay feature of the artifact. To our knowledge, this is the first time that the sources of measurement uncertainty for this technique have been described and quantified. A method for predicting the group delay of the gas-cell artifact from measured absorption spectra is described, and an uncertainty analysis for the prediction method is also presented. The implementation required to achieve results of the highest accuracy for both measurements and predictions is discussed.**

***Index Terms*—Calibration, microwave photonics, optical communication, optical components, optical propagation in dispersive media, optical variables measurements, reference material, uncertainty.**

I. Introduction

FUTURE dense wavelength-division multiplexing systems pose a number of formidable metrology challenges as data rates increase and switching fabrics become more complex. Among the phenomena requiring high-resolution characterization is chromatic dispersion, which broadens optical data pulses through the wavelength-dependent variation in the refractive index of system elements. At higher data rates, characterization of relative group delay (RGD) due to chromatic dispersion in components becomes critical to system throughput. In optical fibers, broadband descriptions of the chromatic dispersion are sufficient. However, it is more challenging to characterize optical components such as narrowband filters for 40-Gbit/s data rates, since these systems may need subpicosecond RGD resolution in bandwidths on the order of tens of picometers. To help meet this industry challenge, the National Institute of Standards and Technology (NIST) has developed an RGD measurement technique targeted for metrology of high-data-rate components.

Our technique relies on refinements to the established modulation phase-shift (MPS) method for measuring RGD [1]–[3], yielding enhanced phase stability over the time interval of the measurement. This technique offers both high wavelength and temporal resolution [4], [5]. However, given the many sources of uncertainty in this measurement, the development of a calibration artifact with a theoretically predictable group delay is critical to the success of optical-component measurements [6], [7]. We present experimental and predicted results for the measurement of the RGD of a molecular-gas absorption line that has served as a stable and well-characterized standard. The estimation and validation of the uncertainty for our MPS system is described in detail.

Manuscript received October 19, 2004; revised June 23, 2005.

The authors are with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: tasshi@boulder.nist.gov; pwilliam@boulder.nist.gov).

Digital Object Identifier 10.1109/JLT.2005.856199

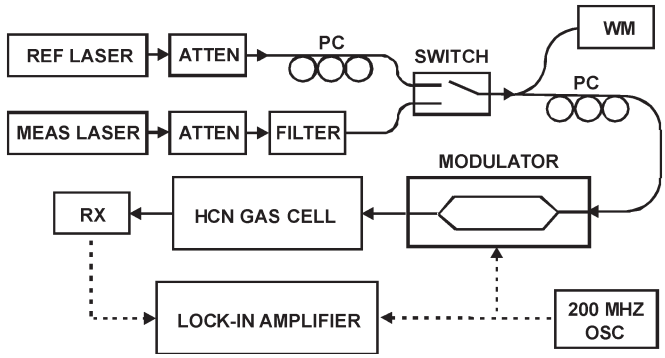

Fig. 1. MPS system for measuring the group delay of optical components. WM, wavelength meter; PC, polarization controller; OSC, crystal oscillator; RX, photoreceiver.

II. Experimental Apparatus

Fig. 1 shows the experimental apparatus we constructed for measuring RGD. The measurement laser was an extended-cavity diode laser having a single-mode tuning range of 100 nm centered at 1595 nm, with a measured 1-s linewidth of $\sim$ 5 MHz and a specified 50-ms linewidth of $<$ 300 kHz. The laser light was passed through a variable optical attenuator and a narrowband tracking filter. The reference laser was also of extended-cavity design with an emission wavelength that remained fixed throughout the measurements. A small portion of the light from the lasers was directed to a wavelength meter having subpicometer resolution and accuracy. The remaining light was amplitude modulated by a Mach–Zehnder modulator driven electrically by a crystal oscillator at 200 MHz. The amplitude-modulated signal was passed through the device under test, which was a fiber pigtailed gas cell. The cell contained the $H^{13}C^{14}N$ isotopic species of hydrogen cyanide (HCN) at a pressure of 13 kPa (100 Torr). The total optical

path through the gas was 22.5 cm using a three-pass geometry. An alternating-current (ac)-coupled photoreceiver was used to detect the modulated optical signal after the gas cell, and a lock-in amplifier referenced to the crystal oscillator measured the electrical amplitude and phase of the received signal.

During a measurement of group delay, the wavelength of the measurement laser was stepped, while the phase of the modulated signal from the receiver was recorded. Observed changes in arrival phase represented variations in the propagation time through the device, where 360° of phase represented one period at the modulation frequency. Therefore, the change in group delay $\Delta\tau$ for a measured change in phase $\Delta\varphi$ (in degrees) is given by

$$\Delta\tau = \frac{\Delta\varphi}{360 \cdot f} \tag{1}$$

where f is the modulation frequency.

To achieve measurements with the highest resolution and lowest uncertainty, the exact method of data collection was critically important. The variable optical attenuators were used to maintain constant received electrical power, which was necessary to counteract the power dependence of phase caused by electrical devices, such as the photoreceiver, amplifiers, and the lock-in amplifier. This is especially important when the device under test has large spectral variations in transmittance, such as those that occur in optical filters and molecular absorption lines. The tracking filter was used to remove amplified spontaneous emission (ASE) noise from the laser, which has been observed to cause phase errors in some measurement scenarios [8].

Temporal drift of the phase during a measurement can be particularly debilitating, because it is indistinguishable from the RGD structure of an optical component. The effects of system drift were removed from each value of phase by subtracting a subsequent reference phase measured with the fixed reference laser. Variations in reference phase represented the phase drift of the system, and recording its value in real time helped to remove drift from the RGD measurement. However, the reference phase must be recorded at a wavelength position having both sufficient transmitted power and small chromatic dispersion. If the chromatic dispersion at the reference wavelength is large, instabilities of the reference wavelength will translate into errors in the reference phase. A mechanical optical switch was used to rapidly alternate between the reference and measurement lasers.

We also minimized phase drift by using a modulator that was designed to operate at quadrature without a direct-current (dc) electrical bias, which tends to drift with time. The photoreceiver was ac coupled to avoid saturation caused by the detection and amplification of unmodulated (dc) light. The residual (background) dispersion in the fiber leads of the system was eliminated from the absorption-cell measurements by subtracting the RGD curve measured with the cell removed.

III. Group-Delay Artifact

The many sources of uncertainty in this method make the existence of a calibration artifact with a theoretically predictable group delay very important. In fact, it has been shown that the MPS method can give erroneous or misleading results when inappropriately used [9]. For example, increasing the modulation frequency to improve the temporal resolution can result in the distortion, and even the inversion, of finely structured RGD. In addition, the measured RGD near the edge of an optical filter can become distorted if the received power is not held constant. Neither Bragg gratings nor thin-film filters provide such an artifact, as their group-delay shape and ripple vary greatly with device and are not theoretically predictable. However, the group-delay profiles of molecular-gas absorption lines can be predicted from their absorption spectra, giving them the potential to be stable and well characterized absolute standards [6], [7]. HCN has about 50 strong absorption lines in the 1530–1560-nm region and a number of weaker lines [10]. Fig. 2 shows a transmittance profile of line P(16) of HCN measured with a lock-in detection system with wavelength steps of about 1.2 pm. The full-width-at-half-maximum of the line is about 50 pm with a depth of about 45%.

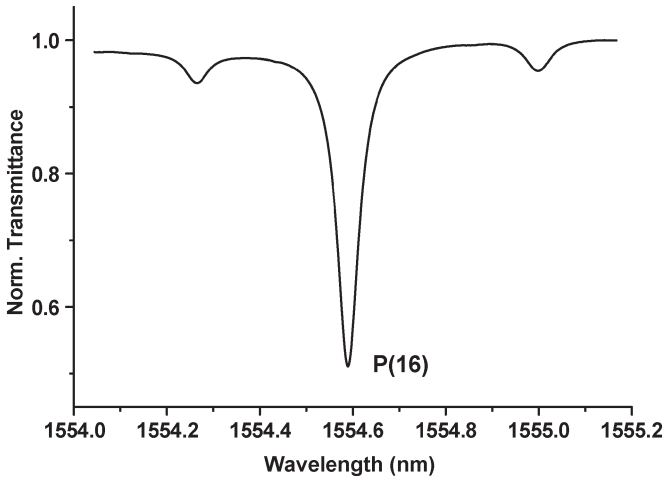

Fig. 2. Normalized transmittance profile of HCN line P(16) measured with a lock-in detection system and used to predict the RGD profile.

The theoretical motivation for using a molecular absorption line as a calibration reference is provided by the Kramers–Kronig relation [11], which enables the RGD $\tau(\lambda)$ to be predicted from a normalized transmittance profile $P(\lambda)$. The transmittance is normalized, such that the maximum measured transmittance is unity, and it represents a per-unit-length quantity. We begin by defining the imaginary component of the dielectric constant $K_{\mathrm{i}}(\lambda)$ for a low-pressure gas

$$K_{\mathrm{i}}(\lambda) = -\frac{\lambda}{\pi} \ln P(\lambda). \tag{2}$$

The Kramers–Kronig relations can then be used to calculate the corresponding real dielectric constant $K_{\mathrm{r}}(\lambda)$

$$K_{\mathrm{r}}(\lambda) = \frac{2}{\pi} \int_0^\infty \frac{K_{\mathrm{i}}(\lambda')\mathrm{d}\lambda'}{\lambda'^3 \left[\left(\frac{1}{\lambda'^2}\right) - \left(\frac{1}{\lambda^2}\right)\right]} + 1. \tag{3}$$

The refractive-index profile can be expressed in terms of $K_{\mathrm{i}}(\lambda)$ and $K_{\mathrm{r}}(\lambda)$ as

$$n(\lambda) = \sqrt{\left(\frac{1}{2}\right)\left[K_{\mathrm{r}}(\lambda) + \sqrt{K_{\mathrm{r}}^2(\lambda) + K_{\mathrm{i}}^2(\lambda)}\right]}. \quad (4)$$

Finally, by differentiation of $n(\lambda)$, the RGD $\tau(\lambda)$ can be calculated from

$$\tau(\lambda) = \frac{-\lambda L}{c}\frac{\mathrm{d}n(\lambda)}{\mathrm{d}\lambda} \quad (5)$$

where c is the speed of light in vacuum, and L is the physical length of the calibration reference. If these expressions are evaluated numerically, arbitrary profiles of measured spectra can be treated without assuming a functional form for the normalized transmittance.

Because the Kramers–Kronig relation is very sensitive to point-to-point fluctuations, equivalent to large transmittance slopes, digital filtering is often necessary to remove high-frequency noise originating from the measurement of absorption. We accomplished this using a low-pass least-squares finite-impulse response filter of order 60 to spectrally smooth the noise of our calculated RGD predictions. The filter characteristics were designed to remove the high-frequency noise without distorting or attenuating the shape of the RGD, particularly for critical features, such as the depth of a narrow absorption line. Visually, we ensured that the curve of the filtered data bisected the noise fluctuations of the unfiltered data.

IV. Measurements of Artifact RGD

Fig. 3 presents a group-delay profile of line P(16) of HCN measured using our MPS system. Ten measurements of the profile were recorded with average-wavelength steps of 3.4 pm and spectrally windowed and averaged to produce the result plotted as dots. The solid line is an average theoretical prediction calculated from ten measured normalized transmittance profiles of line P(16) using (2) through (5). As discussed above, Fig. 2 shows a sample transmittance profile. The agreement between the average measured and the average predicted RGD in Fig. 3 is excellent, even in the magnitude of the delay at the center of P(16) and the resolution of the $\sim$ 1 ps weak side features. The group-delay resolution given by the standard deviation of the difference between the averaged RGD measurement and the prediction over the $\sim$ 1-nm measurement bandwidth is 0.072 ps.

The reference phase was measured at 1554.1 nm, where the transmission of the cell is high and the RGD is flat. Without phase referencing, measurements of the absorption line had drifts of almost 2 ps over intervals as small as 0.1 nm. A small amount of background RGD due to the chromatic dispersion ($\sim$ 0.3 ps/nm) in the fiber leads, which connect the gas cell to the modulator and the photoreceiver, was measured and subtracted as a function of wavelength from each RGD measurement. This was especially important, since the Kramers–Kronig relation does not account for this broad, spectrally featureless background.

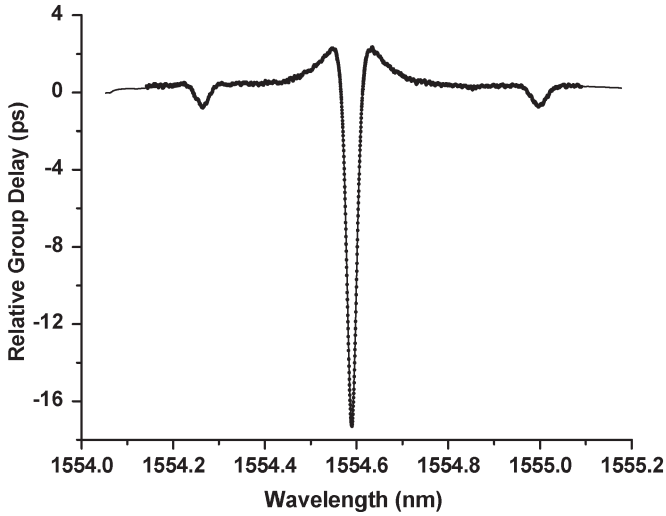

Fig. 3. Averaged measured group delay for line P(16) of hydrogen cyanide (dots), along with a theoretical prediction (solid line). The measured group-delay resolution is 0.072 ps.

V. Uncertainty Analysis

The goal of this study was to estimate the uncertainty of measurements made with our MPS system and to validate the accuracy of the estimate [12]. As was presented in the previous section, the validation of the measurement accuracy was achieved through measurements of the HCN cell, for which a theoretical RGD could be predicted. However, the RGD prediction was based on absorption measurements that contained uncertainty, and processing methods that introduced distortion. Therefore, the validation of the MPS uncertainty estimate (discussed at the end of this section) required that we also estimate the uncertainty in our theoretical prediction of RGD for the HCN cell.

The uncertainty estimate for our MPS measurement system was evaluated so as to be applicable to a single spectral scan of RGD at a wavelength step size of nominally 3.4 pm. This step size is approximately equal to the wavelength resolution of the system, as determined by the 400-MHz spacing (twice the modulation frequency) of the amplitude-modulation sidebands of the measurement signal at 1550 nm. Because many of the sources of uncertainty have been estimated from measurements of the P(16) line of HCN, the uncertainty is valid over a measurement range of 20 ps and is applicable to similarly sharp features. We describe below the uncertainty calculations for both the MPS measurement system and the RGD prediction for the P(16) line of HCN.

A. Sources of Uncertainty for Measured RGD

1) Average RGD Measurement: To reduce the point-to-point noise, we combined and averaged ten measurements of RGD, as described in Section IV. This enabled a direct comparison with the RGD prediction curve shown in Fig. 3 and the determination of any biases or slowly varying sources of measurement error. As previously stated, the standard deviation of the difference between the curves is 0.072 ps, which we use to estimate the random uncertainty of an averaged measurement.

We believe this uncertainty estimate is strongly influenced by distortions in the theoretical prediction, independent of the RGD measurement. Preliminary work shows that this standard deviation could be reduced to below ~ 0.05 ps by increasing the calculation bandwidth of the prediction (as described in Section V-B) to reduce the uncertainty of the RGD prediction curve.

2) Single-Measurement Consistency: By calculating the standard deviation of the discrepancy between the individual measurements and the predicted RGD curve, we could estimate the high-frequency (point-to-point) noise of a single measurement before averaging. The standard deviation was calculated as a function of wavelength (10-pm bin width, ~ 30 data points per bin) and showed no correlation with the RGD features of line P(16). The average of the calculated standard deviation for the ten measurements (0.17 ps) was taken as an estimate of the random uncertainty for a single measurement of RGD. This estimate is the largest component for the MPS system uncertainty and is attributed largely to the relative phase uncertainty of the lock-in amplifier.

3) Background Chromatic Dispersion: The group delay of the fiber leads was measured across the interval from 1545 to 1565 nm with the gas-cell artifact removed. A linear curve fit to data from six independent measurements was used to define a functional relationship for the background RGD with wavelength. This background curve was subtracted from each spectral measurement of RGD for the P(16) line of HCN. The sources of uncertainty for this background subtraction included a random component, determined from the six independent measurements, a temperature component, which we have studied for long lengths of dispersion-unshifted fiber, and a curve-fit component determined directly from the linear least-squares fitting. The total uncertainty due to the removal of the background RGD is estimated to be 0.005 ps.

4) Wavelength: Our wavelength meter has a wavelength resolution of 0.1 pm and was calibrated against our high-accuracy rubidium and methane wavelength standards [13], [14]. Absolute calibration was used to determine an offset adjustment for the wavelength scale and had negligible effect on MPS uncertainty. Consistency between the RGD measurements and the RGD prediction, based on measured absorption spectra, was achieved by using the same wavelength meter for all measurements. The random wavelength uncertainty caused by short-term drift during measurements and longer term drift between measurements has been accounted for in the first two components of uncertainty.

5) Modulator-Chirp Distortion: Modulator chirp can dramatically reduce the accuracy of RGD measurements by causing severe distortion [15]. To minimize this, we operated our modulator slightly off-quadrature, at a small bias determined by experiment. However, because the optimum bias point can drift with time, a small amount of distortion can still occur. The distortion can be seen as an asymmetric ripple in the plot of Fig. 4, which shows the residual difference between a single measurement of RGD and the average RGD prediction. The distortion is centered on line P(16) at 1554.59 nm and covers about 0.1 nm of optical bandwidth. To minimize the influence of noise, we determined the magnitude of the distortion from

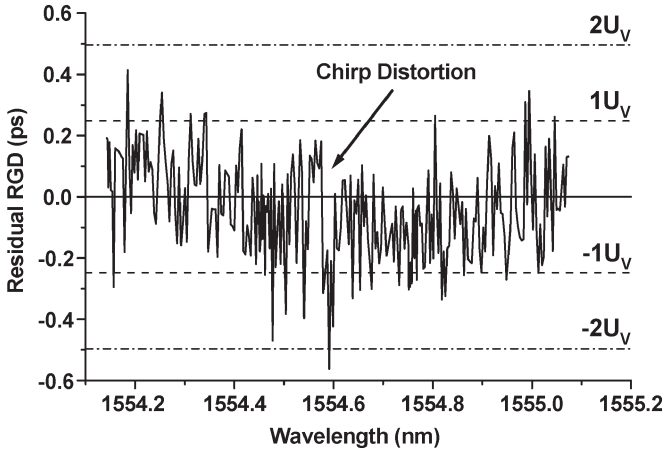

Fig. 4. Residual difference between a single spectral measurement of RGD and the average RGD prediction, shown with $1U_{\mathrm{V}}$ and $2U_{\mathrm{V}}$ uncertainty estimates. Both the curvature caused by a finite prediction bandwidth and the distortion due to modulator chirp are apparent.

the residual curve of the average RGD measurement and acquired a peak-to-peak value of 0.25 ps. We estimate the uncertainty component due to the residual modulator chirp to be 0.13 ps.

6) Averaging and Combining: We used a window average to combine the ten individual measurements of RGD for line P(16). The effective width of the moving average was approximately 4 pm, which compares closely to the 3.2-pm wavelength resolution of the MPS system at 200 MHz. This spectral averaging slightly reduces the depth of the sharp RGD feature at the peak absorption of line P(16). Comparing the averaged and individual measurement curves about this critical point, we estimate an uncertainty component of 0.058 ps due to spectral averaging.

7) Temperature and Pressure: We modeled the temperature dependence for the RGD of line P(16) using the square-root-of-temperature relationship for linewidth broadening [16], [17]. A sensitivity of -0.032 ps/°C at room temperature was estimated for the peak-to-peak depth of the RGD. The air temperature in our laboratory fluctuates with a standard deviation of 0.47 °C, as characterized at intervals of approximately 1 to 2 h. From this, we estimate the RGD uncertainty due to temperature to be 0.015 ps. The temperature of the cell is not expected to vary as much because of its larger thermal mass. While the RGD depends on the pressure of the gas through linewidth broadening, we did not observe variations over time that would indicate that the cell pressure had changed.

8) Frequency Stability and Accuracy: The change in group delay for an observed change in phase is calculated using (1), which depends on the modulation frequency. We measured the absolute frequency and stability of the crystal oscillator in our system with a counter incorporating an oven-controlled reference. The frequency was stable and measurable to at least five decimal places and therefore contributes negligible RGD uncertainty.

9) Polarization-Mode Dispersion (PMD) and ASE: PMD can cause random fluctuations in measured RGD, varying both with time and wavelength [2]. Between the ten artifact measurements, reorientation of the fiber leads at the input to

the gas cell showed no measurable change in RGD due to the 8.5 m of single-mode fiber in our system. This suggested that the system PMD was negligible, consistent with our estimate that the short length of single-mode fiber should have a PMD of about 5 fs. While we have shown that ASE can cause sizeable errors when measuring the dispersion of optical fiber [8], the error is negligible for measurements of narrow RGD features or MPS systems employing a laser-tracking filter. Therefore, the uncertainty due to ASE is also negligible.

10) Power Dependence of Phase: The programmable optical attenuators maintained constant received power on the lock-in amplifier to within 0.06 dB. Using a measured power dependence of the lock-in phase of $0.0084^\circ/\text{dB}$, the incomplete attenuation control caused an RGD uncertainty of 0.007 ps at 200 MHz. In addition, the attenuation control depended on the linearity of the received power measurement provided by the lock-in amplifier. For an estimated error in measured power of 0.05 dB, the RGD uncertainty is 0.006 ps. The quadrature sum of these two uncertainties gives a total RGD uncertainty of 0.010 ps, due to the power dependence of phase.

11) Phase Linearity: We conducted RGD measurements throughout the $\pm 180^\circ$ phase range of the lock-in amplifier without qualitatively observing discontinuities or inconsistencies in the phase linearity. However, we used a calibrated optical air gap to quantify the linearity of the lock-in amplifier, obtaining a maximum deviation of 0.4%. The 19.6-ps depth of the RGD feature for line P(16) represents a change in phase of 1.4° at a modulation frequency of 200 MHz. Therefore, the measurement of the RGD depth had, at most, an uncertainty of 0.0056° or 0.077 ps, which we used to define the limits of a uniform distribution. The RGD uncertainty component due to the deviation of the phase linearity is estimated to be 0.045 ps.

12) Amplitude-Modulation Distortion: Differential loss between the amplitude-modulation sidebands, caused by the wavelength-attenuation profile of the device under test and/or the MPS system, can cause an apparent phase shift. We investigated this source of error through analytic expressions [15] and numeric simulations and found that amplitude distortion could cause a maximum error in RGD of 0.011 ps when measuring line P(16) of HCN at 200 MHz. The maximum error occurs at two wavelengths symmetric about the center of the line, each where the transmittance slope is steepest. Because the differential loss is zero about the center of a symmetric absorption line, the measured RGD depth of line P(16) would not be affected. The RGD uncertainty component due to amplitude distortion is negligible, given its small maximum magnitude and its localization to a 60-pm spectral bandwidth. However, the error could become significant if an even sharper loss feature was measured or a higher modulation frequency was used.

B. Sources of Uncertainty for Predicted RGD

1) Average Prediction Consistency: The averaged prediction curve plotted as the solid line in Fig. 3 is the result of combining ten individual predictions, each derived from a separate measurement of normalized transmittance. To check the consistency, we collected a second set of ten individual predictions to form a second average prediction curve. The maximum difference between the two averaged prediction curves occurred where the transmittance slope of P(16) was largest because the Kramers–Kronig relation is very sensitive to slope. We assumed that the maximum difference defined the extent of a uniform distribution and estimated the random uncertainty for the average prediction curve to be 0.021 ps.

2) Wavelength: Being a random variation during the absorption measurements, the uncertainty due to wavelength is already contained in the uncertainty component for the average prediction consistency.

3) Digital Filtering: We applied a weak digital filter (see Section III) to each of the ten predicted RGD curves to perform spectral smoothing. To test for possible errors due to overfiltering, we identified the maximum deviation between each pair of filtered and unfiltered curves occurring across the central 20 pm of line P(16), where the filtering had the greatest impact. The largest of these maximum deviations was used to define a uniform distribution, from which we estimated a small uncertainty component of 0.009 ps for digital filtering.

4) Averaging and Combining: We used a window average to combine the ten individual predictions of RGD. The effective width of the moving average was approximately 1 pm. This spectral averaging slightly reduces the depth of the sharp RGD feature at the peak absorption of line P(16). Comparing the averaged and individual measurement curves about this critical point, we estimate an uncertainty component of 0.029 ps due to spectral averaging.

5) Temperature and Pressure: As discussed for the uncertainty of measured RGD, the component due to temperature was estimated to be 0.015 ps, while the pressure component was negligible.

6) Finite-Wavelength Range: The integral of the Kramers–Kronig relation (3) extends over all wavelengths, yet our prediction was based on measurements of normalized transmittance covering a finite bandwidth of 1 nm. This incurred a small broad distortion to the RGD of about 0.15 ps extending across the 1-nm bandwidth of the prediction. This appears as the slight curvature to the residual difference between measurement and prediction plotted in Fig. 4. This was discovered by comparison to a prediction based on a 3-nm-wide measurement, which encompassed both the P(15) line at 1553.756 nm and the P(17) line at 1555.435 nm. Increasing the spectral coverage can effectively eliminate the distortion, but at the expense of longer measurement time. We used half of the observed distortion to define the limits of a uniform distribution of uncertainty that applied to the entire 1-nm measurement interval. Therefore, we estimate the uncertainty component caused by using the 1-nm range to be 0.043 ps.

7) Gain-Stage Consistency and Linearity: The normalized transmittance was measured with a low-frequency lock-in detection system. Measuring the depth of line P(16) required one change in lock-in gain; however, we did not detect any discontinuities in either the measured absorption or the predicted RGD. The linearity of the detection system is critical to accurately measuring the maximum absorption depth of line P(16). Lacking a specification for the relative amplitude accuracy, we conservatively used the absolute accuracy of $\pm 0.2\%$ (typical)

TABLE I
SUMMARY OF UNCERTAINTY COMPONENTS FOR A SINGLE RGD MEASUREMENT

Source of Uncertainty	Uncertainty (in ps)
Average RGD Measurement	0.072
Single Measurement Consistency	0.17
Background Chromatic Dispersion	0.005
Modulator Chirp Distortion	0.13
Averaging and Combining	0.058
Temperature and Pressure	0.015
Power Dependence of Phase	0.010
Phase Linearity	0.045
Combined Standard Uncertainty (U_M) (RSS)	**0.23**
Expanded Uncertainty ($2U_M$)	**0.46**

TABLE II
SUMMARY OF UNCERTAINTY COMPONENTS FOR AN AVERAGE RGD PREDICTION

Source of Uncertainty	Uncertainty (in ps)
Average Prediction Consistency	0.021
Digital Filtering	0.009
Averaging and Combining	0.029
Temperature and Pressure	0.015
Finite Wavelength Range	0.043
Gain Consistency and Linearity	0.065
Combined Standard Uncertainty (U_P) (RSS)	**0.088**
Expanded Uncertainty ($2U_P$)	**0.18**

to estimate an RGD uncertainty of 0.065 ps from the measurement of amplitude.

C. Uncertainty Summary

The significant sources of uncertainty for our MPS measurement system are summarized in Table I, along with their estimates, while the significant sources of uncertainty for the averaged RGD prediction are summarized in Table II, along with their estimates. The combined standard uncertainty for each was obtained by combining the estimated components as a root-sum-of-squares (RSS). This gave a combined standard uncertainty of $U_M = 0.23$ ps for a single spectral measurement of RGD. The combined standard uncertainty for the predicted RGD was $U_P = 0.088$ ps for a calculation bandwidth of 1 nm but can easily be reduced to 0.077 ps for a bandwidth of 3 nm.

D. Uncertainty Validation

We validated our estimates for the combined standard uncertainty for both measured and predicted RGD simultaneously by re-examining our measurement data. Because the RGD measurement uncertainty was estimated for a single measurement spanning a 1-nm bandwidth about line P(16) of HCN with 3.4-pm average wavelength steps, we analyzed each of the ten individual measurements of RGD separately. Fig. 4 shows a sample residual difference between a single measurement of RGD and the average RGD prediction shown as a solid line in Fig. 3. Also shown on the plot are the $\pm 1U_V$ and $\pm 2U_V$ uncertainties, where U_V is equal to 0.25 ps and is the RSS combination of the standard uncertainties for measured U_M and predicted U_P RGD. The uncertainty for the validation U_V is only slightly larger than the uncertainty for a single measurement U_M, indicating that the validation is predominantly a test of the measurement uncertainty.

TABLE III
PERCENTAGE OF DATA POINTS WITHIN $\pm 1U_V$ AND $\pm 2U_V$ FOR INDIVIDUAL RGD MEASUREMENTS, WHERE THE VALIDATION UNCERTAINTY U_V IS THE RSS COMBINATION OF U_M AND U_P

	Data Set										
	1	2	3	4	5	6	7	8	9	10	Mean
$\pm 1U_V$ %	89.4	82.8	88.3	90.1	78.0	85.4	85.4	80.2	83.9	86.8	85.0
$\pm 2U_V$ %	99.6	99.3	99.6	98.9	98.9	99.3	98.9	99.6	98.9	99.6	99.3

For the sample residual difference of a single measurement shown in Fig. 4, 89.4% of the data points are within $\pm 1U_V$, while 99.6% are within $\pm 2U_V$. As shown in Table III, which presents the percentages for each of the ten RGD measurements, the percentages associated with Fig. 4 are some of the largest of the group. The mean percentage of data points within $\pm 1U_V$ for the ten measurements was 85.0%, while for $\pm 2U_V$, it was 99.3%. If the statistics for the fluctuations of the residual were normally distributed, we would expect the $\pm 1U_V$ and $\pm 2U_V$ uncertainties to define respective confidence intervals of $\sim$ 68% and $\sim$ 95%. However, after examining a histogram of the residual, we concluded that the actual distribution was narrower than a normal distribution. Likely contributing to this difference were the distortion due to chirp and the curvature due to the finite prediction bandwidth, as already discussed and evident in Fig. 4. As well, our combined standard uncertainties for both the measured and predicted RGD, as summarized in Tables I and II, are likely to be overestimates. We took a conservative approach to handling the estimates for spectrally localized sources of uncertainty, such as distortions from chirp or filtering that occurred at the center of line P(16), and applied them to the entire 1-nm measurement bandwidth.

VI. CONCLUSION

We have shown that the modulation phase-shift (MPS) technique can be implemented to measure optical group delay to very high absolute accuracy and simultaneously high temporal and wavelength resolution. We achieved an RGD resolution of 0.072 ps for repeated measurements spanning 1 nm across the spectrally narrow P(16) absorption line of HCN. A wavelength resolution of 3.2 pm was made possible by the low (200 MHz) modulation frequency of our system. From our uncertainty analysis, we conclude that the expanded uncertainty (2σ) of our MPS system is 0.46 ps for a single spectral measurement of RGD, where our combined standard uncertainty U_M was multiplied by a coverage factor of 2, yielding a confidence interval of approximately 95% [12]. The uncertainty is valid for a measurement spanning 1 nm and containing a narrow RGD feature of up to 20 ps in magnitude. Additionally, the spectral structure of the feature must not be finer than the 3.2-pm wavelength resolution of the system. Our expanded uncertainty depends largely on the relative phase resolution of the lock-in amplifier, as estimated by the uncertainty component for single-measurement consistency. This component is difficult to reduce, since it is determined mainly by the fundamental performance of the instrument and not by how the RGD data are collected. However, it can be reduced by averaging multiple measurements.

As indicated by (1), greater temporal resolution can be achieved by increasing the modulation frequency of the MPS system. Unfortunately, this also increases the wavelength separation of the modulation sidebands, which reduces the wavelength resolution of the system. For this reason, we have been careful to specify our temporal resolution together with the applicable wavelength resolution. Without this, specifications of temporal resolution are meaningless if measurements of RGD involve narrow or structured features that cannot be wavelength resolved.

The group delay of a gas absorption line provides a convenient and high-accuracy artifact for absolute calibration because measured absorption spectra can be used to make predictions. The 13-kPa (100 Torr) gas cell of HCN we constructed has proven to be a stable artifact with narrow features that are challenging to measure. We assigned an expanded uncertainty of 0.18 ps to the average prediction curve used in this calibration to the P(16) line of HCN. We anticipate that the uncertainty can easily be reduced by collecting absorption data and performing calculations over a broader spectral bandwidth. We also validated our uncertainty estimates by analyzing data without averaging, and concluded that the estimates for both measured and predicted RGD are conservative.

ACKNOWLEDGMENT

The authors are grateful to C. Wang for discussions and suggestions on the uncertainty analysis.

REFERENCES

[1] B. Costa, D. Mazzoni, M. Puleo, and E. Vezzoni, "Phase shift technique for the measurement of chromatic dispersion in optical fibers using LEDs," *IEEE J. Quantum Electron.*, vol. QE-18, no. 10, pp. 1509–1515, Oct. 1982.

[2] S. E. Mechels, J. B. Schlager, and D. L. Franzen, "Accurate measurements of the zero-dispersion wavelength in optical fibers," *J. Res. Nat. Inst. Stand. Technol.*, vol. 102, no. 3, pp. 333–347, May–Jun. 1997.

[3] *Chromatic Dispersion Measurement of Singlemode Optical Fibers by the Phase Shift Method*, EIA/TIA Standard Fiber Optic Test Procedure FOTP-169, 1992.

[4] G. Genty, T. Niemi, and H. Ludvigsen, "New method to improve the accuracy of group delay measurements using the phase-shift technique," *Opt. Commun.*, vol. 204, no. 1–6, pp. 119–126, Apr. 2002.

[5] R. Fortenberry, W. V. Sorin, and P. Hernday, "Improvement of group delay measurement accuracy using a two-frequency modulation phase-shift method," *IEEE Photon. Technol. Lett.*, vol. 15, no. 5, pp. 736–738, May 2003.

[6] T. Dennis and P. A. Williams, "Relative group delay measurements with 0.3 ps resolution: Toward 40 Gbit/s component metrology," in *Proc. Optical Fiber Communication (OFC)*, Anaheim, CA, 2002, pp. 254–256.

[7] A. Motamedi, B. Szafraniec, P. Robrish, and D. Baney, "Group delay reference artifact based on molecular gas absorption," in *Proc. Optical Fiber Communication (OFC)*, Anaheim, CA, 2001, pp. Th8C-1–Th8C-3.

[8] T. Dennis and P. A. Williams, "Chromatic dispersion measurement error caused by source amplified spontaneous emission," *IEEE Photon. Technol. Lett.*, vol. 16, no. 11, pp. 2532–2534, Nov. 2004.

[9] T. Niemi, M. Uusimaa, and H. Ludvigsen, "Limitations of phase-shift method in measuring dense group delay ripple of fiber Bragg gratings," *IEEE Photon. Technol. Lett.*, vol. 13, no. 12, pp. 1334–1336, Dec. 2001.

[10] H. Sasada and K. Yamada, "Calibration lines of HCN in the 1.5-μm region," *Appl. Opt.*, vol. 29, no. 24, pp. 3535–3547, Aug. 1990.

[11] J. R. Reitz, F. J. Milford, and R. W. Christy, *Foundations of Electromagnetic Theory*. Reading, MA: Addison-Wesley, 1980.

[12] B. N. Taylor and C. E. Kuyatt, "Guidelines for evaluating and expressing the uncertainty of NIST measurement results," Natl. Inst. Stand. Technol., Gaithersburg, MD, Tech. Note 1297, Sep. 1994.

[13] G. P. Barwood, P. Gill, and W. R. C. Rowley, "Frequency measurements on optically narrowed Rb-stabilised laser diodes at 780 nm and 795 nm," *Appl. Phys. B*, vol. 53, no. 3, pp. 142–147, Sep. 1991.

[14] T. Dennis, E. A. Curtis, C. W. Oates, L. Hollberg, and S. L. Gilbert, "Wavelength references for 1300-nm wavelength division multiplexing," *J. Lightw. Technol.*, vol. 20, no. 5, pp. 776–782, May 2002.

[15] T. Dennis and P. A. Williams, "High-accuracy optical group delay measurements and modulator chirp characterization," in *Proc. Int. Topical Meeting Microwave Photonics*, Ogunquit, ME, 2004, pp. 32–35.

[16] W. C. Swann and S. L. Gilbert, "Pressure-induced shift and broadening of 1510–1540-nm acetylene wavelength calibration lines," *J. Opt. Soc. Amer. B*, vol. 17, no. 7, pp. 1263–1270, Jul. 2000.

[17] W. Demtröder, *Laser Spectroscopy*, 2nd ed. Berlin, Germany: Springer-Verlag, 1996, pp. 67–82.

T. Dennis, photograph and biography not available at the time of publication.

P. A. Williams, photograph and biography not available at the time of publication.