Signal-to-noise ratio in chirped laser dispersion spectroscopy

Michal Nikodem,1 Damien Weidmann,2 Clinton Smith,1 and Gerard Wysocki1,*

1Electrical Engineering Department, Princeton University, Princeton, New Jersey 08544, USA
2Space Science & Technology Department, STFC Rutherford Appleton Laboratory, Harwell Oxford Campus, Didcot, Oxfordshire, UK
* gwysocki@princeton.edu

Abstract: Quantitative studies and experimental validation of noise sources occurring in chirped laser dispersion spectroscopy (CLaDS) are reported. Their impact on the signal-to-noise ratio (SNR) achievable with the CLaDS sensing method is analyzed through a noise model supported by experimental results. In particular the model shows that the SNR is optimal for a given value of the laser chirp rate. The experimental studies are conducted with a quantum cascade laser operating at 5.2 µm for the detection of nitric oxide. Optical fringing has been found to be a significant non-random source of noise and an effective reduction method that can improve the SNR is also discussed.

©2011 Optical Society of America

OCIS codes: (300.6390) Spectroscopy, molecular; (300.6310) Spectroscopy, heterodyne; (140.5965) Semiconductor lasers, quantum cascade.

References and links
1. Introduction

Tunable laser spectroscopy is an effective tool for molecular trace gas sensing. It provides excellent chemical selectivity, outstanding detection sensitivity and close to real time temporal resolution. Tunable laser spectroscopy is therefore relevant to a wide range of applications including environmental [1] and industrial monitoring [2], medical diagnostic [3] and security [4]. In a recent work, chirped laser dispersion spectroscopy (CLaDS) applied to trace gas detection has been introduced [5]. Unlike most of the tunable laser spectroscopy techniques that rely on molecular absorption of light, CLaDS detects the optical dispersion occurring in the vicinity of molecular transitions. The theoretical background and basic experiments demonstrating the measurement principles of CLaDS were presented in [5] and [6]. These studies were primarily focused on the physics of the generation of the CLaDS signals and its practical implications such as exceptional immunity to optical power changes, or large dynamic range in concentrations measurements. These advantages make CLaDS attractive for remote, long-distance, open-path sensing even in environments with large transmission fluctuations.

Experiments and theoretical models of CLaDS reported to-date have mainly focused on quantitative analysis of signals and on only qualitative description of associated noises. In this paper a detailed study of the noise sources impinging the CLaDS technique is presented. In the following section a noise model is constructed and validated by experimental observations. Then, the specific case of noise originating from parasitic optical etalon fringes is considered and an efficient suppression method is proposed. In the last section, the dependence of CLaDS noise on the power of the heterodyne signal is investigated to determine how changes in the received laser power affect CLaDS sensitivity.

2. Noise model of CLaDS

2.1 Experimental system

Interaction of light with matter close to resonance with an electronic, vibrational, or rotational transition of the irradiated medium results in absorption of the incident radiation and simultaneously causes dispersion of transmitted light wave. As described in [5], CLaDS is a method of molecular dispersion sensing using two-color interferometry and subsequent optical heterodyne detection. In the case of a molecular gas, as dispersion observed in the vicinity of ro-vibrational transitions is proportional to the molecular concentration, quantitative sensing can be carried out. CLaDS spectrum is detected as an instantaneous frequency of the heterodyne beatnote given by [5]:

$$f(\omega) = \frac{1}{2\pi} \left[ \Omega + \frac{S}{c} \Delta L \cdot \frac{c}{L_c} \cdot \frac{dn}{d\omega} \left| \frac{dn}{d\omega} \right| \right].$$

where \(S\) is the laser chirp rate, \(\Delta L\) is path length difference between the interferometer arms (shown in Fig. 1), \(L_c\) is the gas cell length and \(n\) is the refractive index of the medium. The frequency of the RF beat note, excluding the carrier \(\Omega\), provides information about the \(\Delta L\) and about the dispersion occurring in the optical path. The contribution originating from the physical path difference \(\Delta L\) can be suppressed by balancing the arms of the Mach-Zehnder interferometer. When the heterodyne beatnote is frequency demodulated at the carrier frequency \(\Omega\), the encoded molecular dispersion can be retrieved. Additionally the demodulated signal is enhanced by the laser chirp rate \(S\). This property is particularly important when fast frequency chirps are available, like for instance those of quantum cascade lasers (QCLs). However to estimate the ultimate detection limits not only the signal enhancement but also the system noise as function of the chirp rate must be analyzed.
The optical layout of the system used for noise analysis is shown in Fig. 1. The laser radiation is generated by a 5.2 μm distributed feedback (DFB) QCL provided by Alpes Lasers SA [7]. The laser beam is split into two frequency-shifted waves by a water-cooled acousto-optical modulator (AOM) driven at \( \Omega = 50 \text{ MHz} \). Both waves are recombined into a single beam in a Mach-Zehnder interferometer arrangement. The dual-frequency radiation subsequently interacts with a gas sample contained in a cell (12.5 cm long) and is detected by a fast photodetector (Vigo Systems, PVI-3TE-10.6). For an experimental analysis of the signal-to-noise (SNR) a gas sample of nitric oxide (NO) balanced with dry nitrogen was used. The R8.5 NO transition at 1906.73 cm\(^{-1}\) belonging to the fundamental \( \nu_1 \) ro-vibrational band was selected to conduct the study.

### 2.2 Noise model

To develop a noise model, various independent noise sources were analyzed. The first noise contribution considered originates from the frequency demodulation process. In FM detection, the random noise at the output of the demodulator depends on the carrier-to-noise (CNR) level (higher CNR gives lower noise) and, more importantly, is proportional to the square of the demodulation bandwidth [8]. In CLaDS, higher chirp rates require the dispersion spectrum to be acquired within shorter time. Hence the demodulation bandwidth needs to be increased proportionally with the chirp rate so that the FM demodulated spectra retain identical sampling (and therefore spectral) resolution. Maintaining a constant ratio of the demodulation bandwidth \( \Delta f \) to the chirp rate \( S \) is therefore a necessary condition to ensure meaningful comparison of signal-to-noise ratios at different chirp rates. With this condition and for a given CNR, the FM noise \( N_{FM} \) is expected to have a quadratic dependence with respect to the laser chirp rate \( S \) and also expected to be inversely proportional to the square root of the number of averaged spectral scans \( k \). With \( A \) being a constant that depends on multiple system parameters (e.g. the noise floor, CNR at the input of the demodulator, or molecular spectral feature that determines relationship between \( \Delta f \) and \( S \)) the FM noise can be expressed as:

\[
N_{FM} = A \cdot S^2 \cdot \frac{1}{\sqrt{k}}.
\]  

(2)

Since CLaDS signal is encoded in the phase of the electromagnetic field measured in the plane of the photodetector, any undesirable phase variations affecting either of the two frequency-shifted waves will contribute to the CLaDS signal. In most spectrometers based on highly coherent light sources, the dominant SNR limitation arises from spectral fringe patterns produced by interferences between uncontrolled stray light reflected or scattered from surfaces in the optical path. For example, despite antireflective coatings applied to all transmissive optical components, fringes formed by residual etalons between facets of the AOM were observed in the CLaDS instrument used in this study. Fabry-Perot etalon affects
both intensity and phase of light, thus it produces fringes in both absorption and dispersion spectra. This unwanted signal contribution is actually considered as a noise in the CLaDS spectrum and can be modeled as:

\[ N_{\text{fringe}} = B \cdot S, \]  

(3)

where \( B \) is a constant related to a root-mean-square (RMS) value of a parasitic fringe spectrum. The fringe noise component \( N_{\text{fringe}} \) is certainly not random. Thus, averaging has no effect in reducing the \( N_{\text{fringe}} \) contribution.

There are other sources of random noise that are independent of the chirp rate and they also contribute to the total noise. These include signal fluctuations caused by mechanical vibrations of the interferometer as well as a low-frequency noise of the electronic circuits. These contributions are accounted for in the model as a constant factor \( N_{\text{DC}} = C \) that also scales with \( k^{0.5} \). This contribution can be considered constant as long as their low frequency bandwidth is much smaller than the FM demodulation bandwidth in CLaDS.

By summing the three described noise sources, the CLaDS noise can be expressed by:

\[ N_{\text{total}} = \sqrt{N_{\text{FM}}^2 + N_{\text{DC}}^2 + N_{\text{fringe}}} \]  

(4)

with random noise contributions added in quadrature, whereas the fringe noise is added directly. Since the CLaDS signal scales linearly with \( S \) (see Eq. (1)), the SNR can be written as:

\[ \text{SNR}(S) = \frac{D \cdot S}{\sqrt{k \cdot \sqrt{A^2 \cdot S^4 + C^2 + B \cdot S}}}, \]  

(5)

where \( D \) corresponds to \( \frac{L_c}{c} \cdot \omega \cdot \left( \frac{dn}{d\omega} \bigg|_{\omega=\omega_m} - \frac{dn}{d\omega} \bigg|_{\omega=\omega_{m+\Omega}} \right) \) in Eq. (1) and is a constant representing a chirp-normalized CLaDS signal magnitude as a function of optical frequency. In the regime of random noise limited performance the functional form of \( \text{SNR}(S) \) exhibits a maximum at certain chirp rate around which the \( N_{\text{FM}} \) becomes dominating. However as \( N_{\text{fringe}} \) increases, this maximum flattens until the SNR becomes almost independent of the chirp rate. In practice an intermediate situation is most likely to occur and an optimum chirp rate value that maximizes the molecular detection performance should exist.

2.3 Experimental noise

The experimental investigation was undertaken with a gas sample of 1.1% of NO balanced with \( \text{N}_2 \) at a total pressure of 9 Torr. Laser chirp rates between 20 kHz/ns and 1325 kHz/ns were applied to the laser. The chirp rate was varied through a 10 mA peak-to-peak triangular laser current modulation at different frequencies ranging from 0.78 kHz to 100 kHz. As mentioned in the previous section, the ratio between the chirp rate and the demodulation bandwidth was kept constant throughout the experiment to maintain identical spectral sampling resolution. All measurements consisted of an average of \( k = 500 \) consecutive scans, corresponding to a total acquisition time of 640 ms for the lowest chirp rate and 5 ms for the fastest one. The noise was estimated as the standard deviation of the recorded trace away from the molecular transition, and the peak-to-trough amplitude of the dispersion profile was used as a measure of the signal. Plots of the signal, the noise, and the SNR measured as a function of the chirp rate \( S \) are shown in Fig. 2 (a), (b), and (c), respectively.
As predicted by Eq. (1) a linear dependence between signal amplitude and the chirp rate $S$ is clearly shown in Fig. 2(a). The noise dependence as a function of the chirp rate is also in good agreement with the noise model given in Eq. (4), which is shown in Fig. 2(b). This indeed results in the SNR that does exhibit a maximum for a given chirp rate, as shown in Fig. 2(c).

The validity of the noise model was further verified using different experimental conditions. First, the dependence of noise on the demodulation bandwidth was investigated. The chirp rate was fixed at 100 kHz/ns whilst the acquisition bandwidth $\Delta f$ was varied from 2.5 to 40 MHz. Since $N_{DC}$ is essentially independent of the bandwidth and $N_{fringe}$ is constant for a given $S$, the model in Eq. (4) can be used to express the noise as a function of $\Delta f$ as $N_{\text{total}}(\Delta f) = \sqrt{P_1^2 \cdot \Delta f^4 + P_2^2 + P_3}$. The function can be used for fitting of the experimental data. The constants $P_1$, $P_2$ and $P_3$ represent individual noise contributions for $N_{FM}$ (bandwidth-normalized), $N_{DC}$ and $N_{fringe}$ (at $S = 100$ kHz/ns) respectively.

Second, a dependence of the noise on the chirp rate $S$ was investigated while the demodulation bandwidth was kept constant at 20 MHz. The chirp rate was varied from 15 to 1325 kHz/ns. With $\Delta f$ kept constant only $N_{fringe}$ shows chirp rate dependence, which simplifies the Eq. (4) to a linear function $N_{\text{total}}(S) = B \cdot S + R$, where $B$ represents the fringe noise and $R$ is a constant representing the quadrature sum of $N_{DC}$ and $N_{FM}$ at $\Delta f = 20$ MHz. In both cases, experimental data were fitted by the proposed $N_{\text{total}}(\Delta f)$ and $N_{\text{total}}(S)$ functions showing coefficient of determination $R^2 > 0.99$, which additionally confirmed the validity of the developed model.
2.4 Noise budget

From the model and the experimentally determined constants $A$, $B$, $C$ and $D$, which characterize signal strength and individual noise contributions, the noise budget as a function of the chirp rate can be inferred and instrumental parameters can be optimized for sensitive molecular detection purposes.

The noise budget as a function of chirp rate shown in Fig. 4(a) was constructed using the data from Fig. 2(b) interpolated for one second acquisition time and with all noise contributions normalized to the total noise. The total noise as a function of chirp rate is shown in Fig. 4(b) and the SNR is shown in Fig. 4(c).

At small chirp rates (and small $\Delta f$ adjusted accordingly), the total noise is dominated by the random noise sources $N_{\text{DC}}$. This contribution becomes insignificant at faster chirp rates in part due to more time-effective averaging (as the chirp rate increases, more traces can be averaged within the one second acquisition time). The FM noise contribution also increases with the chirp rate as the $\Delta f/S$ ratio is kept constant. Quadratic dependence of $N_{\text{FM}}$ on the chirp rate is only partially compensated by the gain from the time averaging process, which can be roughly approximated as $S^{-0.5}$ dependence. Therefore as the chirp rate increases the FM noise quickly overcomes $N_{\text{DC}}$. More importantly, higher chirp rates favor the domination of the fringe noise in accordance with its linear dependence with $S$. 
The total noise dependence in an ideal situation of full fringe noise suppression is also shown in Fig. 4(b) and 4(c) as blue lines. Without $N_{\text{fringe}}$ the optimum chirp rate that maximizes SNR corresponds to a point at which $N_{FM}$ overcomes $N_{DC}$. Further increase in $S$ beyond this point would cause the noise to increase faster than the signal (noise $\propto S^{1.5}$ vs. signal $\propto S$). Improvement in detection sensitivity beyond the optimum chirp rate can only be achieved through longer averaging or, in some cases, by increase of a CNR of the heterodyne beatnote (see Section 4).

In the actual system presented here the total noise is strongly dominated by the non-random fringe noise contribution that cannot be suppressed by averaging or by the increase in CNR. As a result, in the fringe dominated regime ($S > 300$ kHz/ns) the total noise approximately grows as fast as the signal ($\propto S$) and the time-normalized SNR becomes almost chirp independent as shown by the red traces in Fig. 4(b) and 4(c).

This analysis indicates that significant sensitivity improvement can be achieved through reduction of parasitic fringing. As shown in Fig. 4(b), at the optimum chirp rate, more than one order of magnitude improvement in sensitivity can be achieved if the fringes are suppressed. Fringe-suppressed SNR reaches a theoretical maximum of $\sim 18900$ for an optimum chirp rate of $\sim 220$ kHz/ns. This corresponds to a theoretical minimum detection limit for NO of $73$ ppb-mHz$^{-0.5}$. Hence, an effective reduction of fringes poses the biggest challenge and is a very critical step in the system optimization. A method of mitigating the fringe contribution in this system is presented in the following section.

### 3. Reduction of optical fringes

The most straightforward and reliable approach to reduce fringes relies on careful optical design that utilizes reflective optics as well as efficient anti-reflection (AR) coatings on unavoidable transmissive parts of the optical setup. However, several other methods have been proposed including selection of optimal modulation frequency [9], two-tone frequency modulation [10–12], digital signal processing [13] or application of a Brewster-plate spoiler [14].

For the CLaDS spectrometer under study, the primary sources of optical fringing were identified to stem from the AOM crystal facets despite of being equipped with a standard AR coating. Thus an approach similar in principles to the scanning mirror method [15] has been implemented using thermal modulation of the optical length of the AOM crystal. Thermal cycling was achieved through variation of the flow of the AOM cooling water. The flow was periodically interrupted and restored with periods ranging from 5 to 20 seconds. Acquisition of CLaDS spectra was made synchronous with the flow alteration. During data averaging, the thermally induced frequency shift of the fringe pattern was effectively scrambling the fringes without influencing the molecular dispersion signal. The efficiency of the scrambling that resulted in fringe reduction is shown in Fig. 5. Both traces were recorded with a laser chirp rate of $420$ kHz/ns produced by a $25$ kHz triangular modulation of the laser current. The slow nature of the thermal process required long acquisition times (~few seconds). Hence, 10000 scans (40µs each) were acquired with ~1 ms interval between consecutive scans. The sample was a mixture of $850$ ppm of NO buffered by N$_2$, with a total pressure of 6 Torr.

Whether the cooling water flow is altered or not, the peak to peak amplitude of the fringes is ~1 kHz for 10 spectra averaged spectra (see blue traces in Fig. 5). Averaging of 10000 spectra yields no reduction of the fringe amplitude if the AOM cooling conditions remain steady (top red trace in Fig. 5). In contrast, with flow alteration (Fig. 5, bottom), the fringe amplitude is reduced by a factor of 16. The colored contour plots in Fig. 5 show the fringe stability over time without (top) and with (bottom) the flow alteration, respectively.
Fig. 5. Effect of fringe averaging on the overall noise reduction: (top) temperature of the AOM was stable during the measurement and no improvement is achieved when 10000 scans are averaged; (bottom) temperature of the AOM was changed during measurement which led to a significant fringe noise reduction after averaging of multiple scans.

The benefit of the fringe reduction method is analyzed by measuring noise as a function of the number of averaged spectra $k$. To obtain reference data, in which the fringe noise contribution is suppressed, a set of measurements with no laser chirp was acquired. Then, similar measurements were carried out while chirping the laser at 420 kHz/ns, with and without the thermally induced fringe scrambling. Results are shown in Fig. 6.

The reference data recorded without laser chirp shows the $k^{0.5}$ decay expected from the random nature of $N_{FM}$ and $N_{DC}$ contributions. Measurements taken without fringe scrambling show a small improvement due to averaging and the total noise is reduced by less than 20%. This decrease corresponds to a reduction of the random noise contributions, while contribution from fringe noise remains unchanged. Once the thermally induced fringe scrambling is turned on, noise decreases rapidly and just after 1000 scans most of the fringe noise is removed. Averaging of 10000 scans yields a total noise reduction by almost 90%. The limit achieved for this particular averaging time is about two times the noise floor measured with our fringe-free reference data, which suggests that the fringe scrambling is efficient and only a small contribution from the fringe noise is still present.

The reduction of fringe contribution to the total noise directly impacts the minimum detectable concentration of the CLaDS system. Considering 10000 averages and the $N_{fringe}$ contribution suppressed by 16 times, the total noise is reduced by ~8 times and the SNR improves from 54 to 452. This results in the NO detection limit of 235 ppb-m with an effective measurement time of 400 ms. Using the SNR model of Eq. (5), and accounting for the residual fringe noise observed with the fringe scrambling on, a detection limit of 184 ppb-m-Hz$^{0.5}$ can be estimated.
Fig. 6. Noise level as a function of number of spectra averaged. A fringe-free measurement (taken at $S = 0$) shows purely Gaussian nature (green squares). Blue squares and red triangles correspond to noise recorded with and without fringe scrambling, respectively (both at $S = 420$ kHz/ns).

4. CLaDS performance vs. carrier-to-noise ratio

CLaDS signal is encoded as changes in the frequency of the detected heterodyne beatnote, and as such are independent from amplitude variations. Immunity to variations of the received optical power has already been indicated in [5]. This effect is further investigated in this work with a focus on the dependence of output noise level on the CNR. Experiments were performed with similar conditions as described in the previous section, including the fringe reduction technique. The SNR was measured for 10000 traces averaged, and the CNR was varied through changing the total optical power reaching the detector that results in variation of the RF beatnote (carrier) power. Since the noise at the input of the FM demodulator was constant (detector and the front end electronics noise) changes of the carrier level are directly related the changes in CNR. The noise as a function of the carrier level is shown Fig. 7.

Fig. 7. Noise as a function of the power of the heterodyne beatnote shown in logarithmic (left) and linear scale (right). Measured data (dots) are fitted with linear spline. The threshold is visible, above which the contributions from fringe noise and FM-noise to the total noise are comparable, thus the impact of the carrier power on the total noise is reduced.

As a general trend, the noise increases as the carrier power decreases. However at different carrier levels the noise vs. CNR shows different sensitivity: 1) for low carrier levels
(below \(\sim-8\text{dBm}\)) the slope is about \(-3\text{dB/10dBm}\) and 2) for high carrier levels (above \(\sim-8\text{dBm}\)) the slope is about \(-1.5\text{dB/10dBm}\). This behavior is strictly related to the nature of the dominant noise in the system. Because the FM noise grows as CNR decreases, at low carrier levels \(N_{FM}\) becomes dominant. As the carrier level is increased, the \(N_{FM}\) decreases up to a point where noise contributions that are independent of the CNR (\(N_{DC}\) or, which is the case in this system, the residual fringe noise \(N_{fringe}\)) become comparable to \(N_{FM}\) and the impact of CNR on the total noise is reduced.

This suggests that the higher the carrier level, the more sensitive the CLaDS technique is. Yet, the noise dependence on CNR is small: in the FM-noise dominated region, noise is doubled when the carrier level decreases by one order of magnitude. If the fringe noise were to be overwhelmingly dominant, the CNR would not have any influence on the noise level. In practice, at an optimized chirp rate, contributions from FM and fringe noises will be comparable. Hence the CNR reduction will affect the noise level, although to a rather limited degree. For instance, the data shown in Fig. 7 (left) indicate that despite almost three order of magnitude decrease in the carrier power, the SNR decreases only \(\sim5\) times and the detection sensitivity still remains below \(1\text{ppm-m-Hz}^{-0.5}\).

5. Conclusions

A noise analysis in CLaDS has been performed: major noise sources have been fully identified and integrated into a noise model that has been validated experimentally. The noise characterization has allowed for the estimation of individual noise contributions to the instrument SNR. It was shown that, for a given experimental system an optimized value of the laser chirped rate that provides the best SNR exists.

Different ways of SNR improvement involving reduction of individual noise sources have been discussed. Noise which originates from spurious etalon in AOM was found to be dominant and an efficient method for fringe reduction was successfully implemented. Using averaging in combination with a thermally induced fringe scrambling approach, this noise contribution was reduced by a factor of 16. The noise component characteristic to the frequency demodulation process was found to be slightly dependent on the CNR. However in the presence of fringe noise this dependence becomes smaller as the carrier power increases and FM-demodulation noise is reduced. With the CLaDS instrument used in this study, bandwidth normalized minimum detection limit for nitric oxide was found to be 184 \(\text{ppb-m-Hz}^{-0.5}\), which is close to the ultimate theoretical fringe-noise-free limit of 74 \(\text{ppb-m-Hz}^{-0.5}\). This reduction in performance of about three times was found to be primarily related to the incomplete suppression of the fringe noise contribution.

The fundamental noise floor in CLaDS is defined by the acquisition bandwidth within which the frequency demodulation occurs. Therefore, further works to improve molecular detection capability using CLaDS will focus on reducing the demodulation bandwidth. Chirp modulation schemes allowing narrowband lock-in detection will be investigated to further improve CLaDS performance.

Acknowledgments

The work presented in this paper was supported by the NSF CAREER award CMMI-0954897, MIRTHE National Science Foundation Engineering Research Center, and by the 2010 Masao Horiba Award. Dr. Antoine Muller from Alpes Lasers SA is acknowledged for providing a laser for this study.