

Multi-harmonic detection in wavelength modulation spectroscopy systems

A. Hangauer · J. Chen · R. Strzoda · M.-C. Amann

Received: 5 February 2012 / Revised version: 30 March 2012 / Published online: 23 May 2012
© Springer-Verlag 2012

Abstract Two multi-harmonic detection methods for wavelength modulation spectroscopy (WMS) systems are presented and compared. The two possibilities discussed in this paper are: simultaneous curve fitting of multiple harmonic spectra, and reconstruction of the transmission from harmonic coefficients. The optimum number of harmonics is four and 25 harmonics, respectively. Compared with standard single-harmonic curve fitting, the methods give about a factor of 3 better performance than standard second-harmonic curve fitting. Concluding, multi-harmonic detection is better than single-harmonic detection and should be used if the system bandwidth is high enough to allow for proper detection of the higher harmonics.

1 Introduction

Tunable diode laser spectroscopy (TDLS) is a method where the absorption or transmission spectrum of a substance (usually a gas) is measured with a tunable diode laser for the purpose of analysis of the specific substance. The emission frequency of a spectrally single-mode laser is tuned, and the power of the light that passes through the sample is determined with a photodetector. The characteristics of the measured transmission spectrum are analyzed to determine the physical parameters, for example, concentration of specific

gas components (volume mixing ratio), pressure, or temperature.

For technical reasons a special method called wavelength modulation spectroscopy (WMS) [1] is frequently utilized to realize the TDLS measurement. Due to the additional sinusoidal wavelength or frequency modulation with repetition rates in the kHz range, harmonic spectra are obtained—instead of the transmission spectrum—which also contain the relevant information. It is a method for implementing an effective noise suppression and partially removing the influence of the laser power variation during frequency scanning. The light passes through the sample with transmission $T(\nu)$ and is detected, and the photocurrent is fed to a lock-in amplifier. The n -th harmonic component of the lock-in amplifier of the relative light power variation is called the harmonic coefficient $H_n = H_n(\nu; \nu_a)$ (Fig. 1). If these harmonic components are plotted against the slow sweep frequency ν , the harmonic spectra $H_n(\nu)$ are obtained (Fig. 2). The definition of the harmonic spectrum does not include system parameters (except for modulation amplitude). More details, mathematical definitions, and properties of WMS are outlined in Appendix A.

Obviously, in a WMS system the harmonic spectra have to be analyzed instead of the absorption spectrum itself. This is done either by curve fitting to an appropriate spectral model or a heuristic evaluation of the spectrum at a single frequency point. The single point evaluation is, however, only suitable for spectral models with a single parameter. Nevertheless, it is employed as detection of the maximum value of the second harmonic in a frequent number of applications due to its simple signal processing. For more realistic spectral models that include, e.g., a variable spectral baseline or laser emission frequency drift this is not suitable, and such drifts can not be compensated, which is the reason for employing curve fitting to measured spectra [3].

A. Hangauer (✉) · J. Chen · M.-C. Amann
Walter Schottky Institute, Technical University of Munich,
Am Coulombwall 3, 85748 Garching, Germany
e-mail: andreas.hangauer@wsi.tum.de

A. Hangauer · R. Strzoda
Siemens Corporate Research & Technologies, Siemens AG,
Otto-Hahn-Ring 6, 80200 Munich, Germany

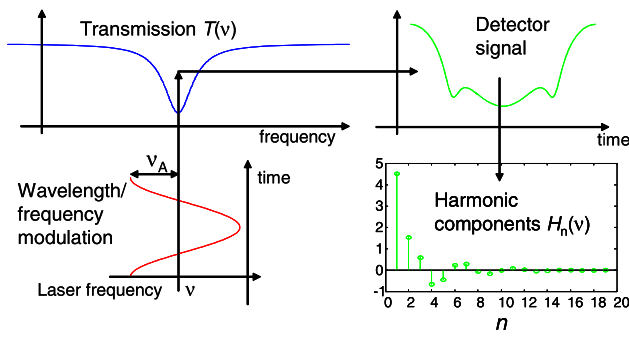


Fig. 1 Generation of the harmonic components in wavelength modulation spectroscopy. ν is the emission frequency of the slow sweep (not illustrated here) and ν_A the frequency modulation amplitude

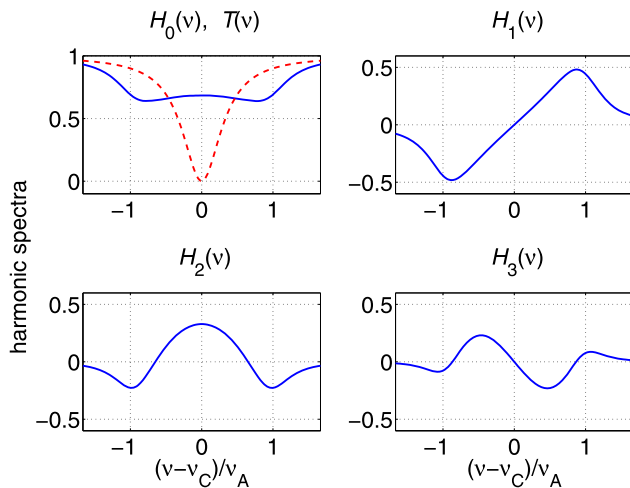


Fig. 2 The first four harmonic spectra (i.e., $H_n(\nu)$ with $n = 0, \dots, 3$) (blue, solid) for a Lorentzian absorption line (red, dashed) computed with theory from [2]. The wavelength modulation amplitude is ν_A , the absorption line center position is ν_C , and the half-width is $\nu_A/3$

As mentioned before, there are (theoretically) infinitely many harmonic spectra or harmonic components which all could be utilized to extract the wanted and extractable information or at least combined into one signal before data evaluation. For single-harmonic detection the signal energy that arrives at the photodetector is not fully utilized, whereas theoretically complete utilization of the signal energy should enable an improved sensing [4, Sect. 10.4]. Furthermore, there is no or only very little additional effort to detect the higher harmonic components if a digital lock-in technique is utilized. This *multi-harmonic detection* should therefore be a simple way to increase sensitivity or versatility for the WMS technique. However, it is not a standard approach in the literature. The analysis of suitable approaches will be the scope of this paper. Note that this multi-harmonic detection assumes a sufficiently wideband detection, which standard photodetectors always provide. But systems with an inherent resonant detection method like photoacoustic

spectroscopy may not benefit from multi-harmonic detection [5]. Also, optical systems like integrated cavity output spectroscopy (ICOS) [6] have limited bandwidth since the light buildup and decay in the optical cavity is not arbitrarily fast. This sets limits on the maximum number of harmonics that can be detected and/or the maximum modulation frequency.

Two different multi-harmonic detection methods have to be distinguished: the simultaneous curve fitting of harmonic spectra, or the evaluation of several harmonic components at a single spectral frequency. The harmonics $H_n(\nu_0)$ at a single frequency ν_0 contain all information about the shape of the absorption line (see Fig. 1), so in principle these should allow for extraction of all extractable parameters and not only for the peak absorbance with single-harmonic detection. This multi-harmonic detection at a single frequency was first proposed in Ref. [7], where an empirical approach that relates the ratios of the n -th harmonic components to the absorption line parameters, like half-width, center wavelength, or peak absorbance, is proposed. This method has the drawback that it requires an empirical calibration. This will be improved by the authors of this paper to an exact method where the full absorption line profile can be recovered. This method is exact and therefore allows for curve fitting to an arbitrary spectral model which is used for direct spectroscopy. Both methods are also suitable for saturated lines, i.e., thick optical sample conditions where the linearization of the Lambert-Beer law does not hold anymore. These methods will be presented, analyzed, and compared.

In the sections below the following questions will be answered:

- How should one implement multi-harmonic detection?
- How many harmonics are needed?
- Performance difference to single-harmonic detection?
- Performance difference to direct spectroscopy?

2 Multi-harmonic detection: reconstruction of the transmission

A reconstruction method that combines the harmonics at a single frequency ν_0 (i.e., no spectral scanning) can be derived as follows.

2.1 Theoretical background

After renaming ν to ν_0 and then substituting $2\pi ft$ with $\arccos((\nu - \nu_0)/\nu_A)$ on both sides of Eq. (6), we obtain a reconstruction formula for the transmission function from harmonic coefficients,

$$T(\nu) = \sum_{n=0}^{\infty} H_n(\nu_0) T_n((\nu - \nu_0)/\nu_A) \tag{1}$$

with $T_n(x) = \cos(n \arccos(x))$ the n -th Chebyshev polynomial [8, Chap. 22]. This reconstruction formula will serve as a multi-harmonic detection scheme. Practically, the right-hand side of Eq. (1) is evaluated with a finite number of measured harmonic coefficients $H_n(\nu_0)$ at the central laser emission frequency ν_0 (mostly close to the absorption line center). Equation (1) predicts that the obtained reconstructed data $T(\nu)$ must be the same as the transmission function, as if it had been measured by direct spectroscopy. This is also true if the laser changes its intensity during tuning. The signal model for the reconstruction method then has to include the proper baseline due to varying laser intensity.

Because the substitution made to derive Eq. (1) is only valid for $|\nu - \nu_0| \leq \nu_A$, this determines the range in which the reconstruction can be done. This is also clear from a physical point of view: since the frequency modulation only covers the range $\nu_0 - \nu_A$ to $\nu_0 + \nu_A$, actually no measurement is performed outside this spectral interval, and it cannot be expected to obtain information that is not measured. Practically, the method can be seen as an implementation of direct spectroscopy with sinusoidal frequency scanning instead of linear scanning.

2.2 Fitting model for reconstructed transmission and implementation

Since the frequency offset ν_0 and frequency modulation amplitude ν_A are generally unknown, it is convenient to use the normalized emission frequency coordinate $x = (\nu - \nu_0)/\nu_A$ for evaluation of the right-hand side of Eq. (1) and spectral fitting. For this the spectral model does not need to be changed; all determined frequency parameters like Gaussian or Lorentzian absorption line half-width or line position will be obtained in these normalized units (relative to scanning range, multiples of ν_A with a zero point at ν_0). Only if these obtained frequency parameters are compared with an absolute frequency, do the quantities ν_A and ν_0 need to be determined with a frequency calibration scheme (e.g., with a reference etalon or using the method described in Ref. [9]). If only a single Lorentzian line and a linear laser power characteristic are considered, the model has the following form:

$$\sum_{n=0}^N H_n(\nu_0) T_n(x) = \exp\left(\frac{-ax_\gamma^2}{x_\gamma^2 + (x - x_C)^2}\right)(y_o + y_s x) + \epsilon \quad (2)$$

The five unknown fit parameters are line position x_C , line half-width x_γ , both in normalized frequency coordinates ($x = (\nu - \nu_0)/\nu_A$), peak absorbance a and y_o and y_s for the y -offset and y -slope representing the laser amplitude

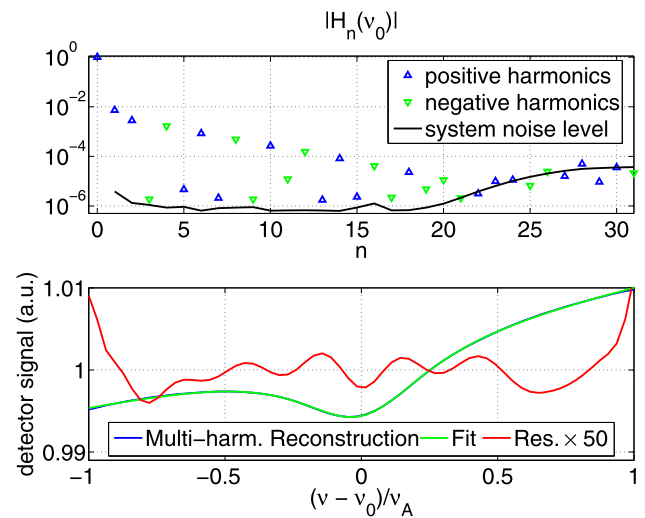


Fig. 3 The upper plot shows the harmonic coefficients $H_n(\nu_0)$ for the O₂ detection (different symbols for positive and negative values) and the corresponding noise level (black). The bottom plot shows the corresponding reconstructed transmission (using H_0, \dots, H_{20} with Eq. (2)) (blue) and the corresponding Lorentz fit (green) with residual (red)

modulation, respectively. The normalized frequency variable x lies between -1 to 1 and ϵ is a random variable that represents the noise on the reconstructed spectrum. The numerically stable evaluation of the sum of the left-hand side of Eq. (2) is preferably done using the Clenshaw algorithm [10, Sect. 3.11], and then the least-squares curve fit to the given model is carried out (for both see Fig. 3(bottom)). For the reconstruction and subsequent fit the variable x was discretized with 64 equidistant values from -1 to 1 . The actual number of points is irrelevant if it is high enough so that the reconstructed spectrum is adequately sampled.

2.3 Required number of harmonics

Due to the increasing oscillating behavior it is clear that each term in Eq. (1) adds more detail to the result with a magnitude corresponding to the amplitude of the harmonic coefficient H_n . This happens because the values of the Chebyshev polynomials all lie in the range of -1 to 1 . Since the harmonic coefficients typically show an exponential decay with order n , the sum in Eq. (1) can indeed be truncated when the desired accuracy is reached. The first few Chebyshev polynomials are shown in Fig. 4. The minimum number N of harmonics depends on the system noise level. The error by including the harmonic in the reconstruction is higher than the error by its omission if the noise amplitude in the harmonic is higher than the signal amplitude. The reconstruction thus should include all harmonics with a signal-to-noise ratio higher than one. Using a typical modulation amplitude

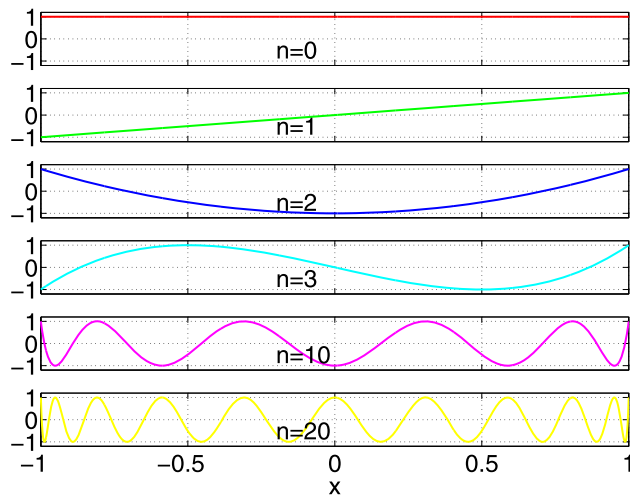


Fig. 4 The Chebyshev polynomials $T_0(x)$ to $T_3(x)$, $T_{10}(x)$, and $T_{20}(x)$

$\nu_A = 3\gamma$ (γ : half-width at half-maximum of the absorption line) and a system noise level of 1×10^{-5} to 1×10^{-6} a full reconstruction can be obtained with $N = 20$ to $N = 25$ harmonics, when the peak absorbance is around 1 %. This can be seen in Fig. 3(top), where the signal level of the harmonics drops below the noise level at approximately $n = 20$. Note that using more harmonics than necessary should not affect the overall precision of the parameters that the curve fits extract. The reason is that the contribution of the higher harmonics to the reconstructed transmission is of increasing detail (this is a property of the Chebyshev polynomials, cf. Fig. 4), so the fit will filter the added high frequency noise on the reconstructed transmission.

3 Multi-harmonic detection: fitting multiple spectra

As mentioned earlier, simultaneous curve fitting of multiple harmonic spectra is the other alternative to multi-harmonic detection.

3.1 Fitting model and implementation

Since the different harmonics all originate from a measurement of the same physical system (gas transmission, laser, and detector), a curve fit with a shared set of parameters to all harmonics has to be carried out. Note that this is different from the approach in Ref. [11], where several higher harmonics have been fitted, but separately with individual parameter sets and subsequent averaging. Note that this approach cannot be expected to have the highest precision, because a least-squares curve fit is the optimum method for estimation of parameters from noisy data. This holds exactly for linear models and is asymptotically true (for low noise) for nonlinear models.

The model used for fitting the unsaturated O_2 absorption line around 763 nm is the following:

$$\begin{aligned}
 H_1(\nu_i) &= -a \cdot S_1((x_i - x_C)/x_\gamma, x_A/x_\gamma) + y_{o,1} + \epsilon_1 \\
 H_2(\nu_i) &= -a \cdot S_2((x_i - x_C)/x_\gamma, x_A/x_\gamma) + y_{o,2} + \epsilon_2 \\
 H_3(\nu_i) &= -a \cdot S_3((x_i - x_C)/x_\gamma, x_A/x_\gamma) + \epsilon_3 \\
 &\vdots
 \end{aligned}
 \tag{3}$$

The $H_n(\nu_i)$ represent the measured harmonic spectra at frequency ν_i with $i = 1, \dots, N$ points per scan. The actual frequency points ν_i are unknown but assumed to be linearly increasing. The variable $x = \frac{\nu - (\nu_N + \nu_1)/2}{(\nu_N - \nu_1)/2}$ is defined as the normalized frequency scale. It runs from -1 to 1 and covers exactly one spectral scan. If the slow laser tuning is carried out linearly, the N spectral samplings correspond to the points $x_i = \frac{2(i-1)}{N-1} - 1$. The line center x_C is the index of the absorption line center, x_γ the line half-width, and x_A the frequency modulation amplitude, all in the relative units. $y_{o,1}$ and $y_{o,2}$ model the laser AM induced offset on the first and second harmonic spectrum. The laser AM effects are modeled in the same way as for single-harmonic detection. For a vertical-cavity surface-emitting laser (VCSEL) the simple approximation by variable offsets is sufficient (as the experiment shows). The exact model without approximations is described in Ref. [4].

The parameter a is the desired peak absorbance value and $S_n(x, m)$ the formula by Arndt [2, Eq. (12)], which is the harmonic spectrum of a peak normalized unsaturated Lorentzian function. The ϵ_n are random variables representing the measurement noise. The parameters describing the relationship between index x and ν (i.e., offset and proportionality factor) cannot be determined from the fit when only a single absorption line with unknown line width is scanned.

In the curve fit the six parameters $a, x_C, x_\gamma, x_A, y_{o,1}$, and $y_{o,2}$ are determined. Note that although the frequency parameters x_C, x_γ , and x_A are in nonstandard units, the model is formally the same as for absolute units. Ratios are independent of the units, i.e.,

$$\frac{\nu - \nu_C}{\gamma} = \frac{x - x_C}{x_\gamma} \quad \text{and} \quad \frac{\nu - \nu_C}{\nu_A} = \frac{x - x_C}{x_A}
 \tag{4}$$

Hence, $(\nu - \nu_C)/\nu_A$ can be chosen as the plot x -axis. In Fig. 5 the simultaneous fit of the first four harmonic spectra is shown.

3.2 Required number of harmonics

Since the amplitude of the harmonic spectra decreases with increasing order, it is expected that there is an optimum

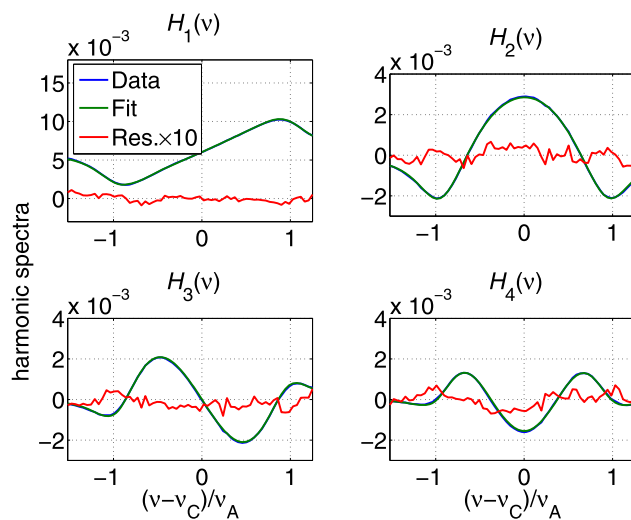


Fig. 5 Measurement (blue) and simultaneous curve fit to the first four harmonic spectra (green) and the residual (red)

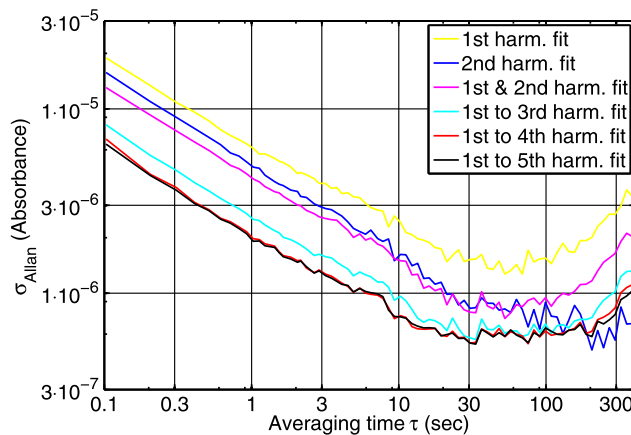


Fig. 6 The Allan plots for multi-harmonic detection with a variable number of spectra that are simultaneously fitted. The measurement data is the same for all curves. Fitting more than four harmonic spectra gives a saturation behavior, whereas including the third harmonic gives the highest improvement

number of harmonics for simultaneous curve fitting. Then using more harmonics than this optimum value will not further improve the noise on the extracted parameters.

It turned out experimentally that fitting of the first four harmonic spectra is optimum in the sense that using more does not give any further improvement (Fig. 6). This is also independent of the system noise level and peak absorbance value, as the relative improvement becomes negligible. This number will, however, depend on the modulation index ν_A/γ , because for higher modulation indices the signal energy is shifted to higher orders of the harmonic spectra. The same is true if a thick optical sample (i.e., absorbance >0.1) is measured.

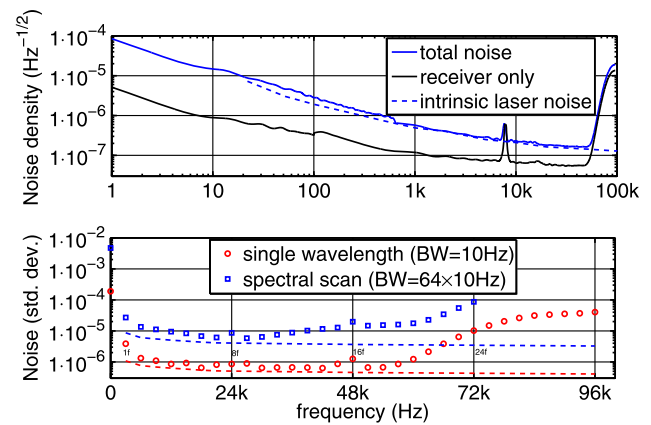


Fig. 7 Top: The total noise spectral density of the measurement system (blue) and the contribution of receiver (black) and intrinsic laser intensity noise (blue dashed). Bottom: the noise level (standard deviation) on the harmonic components for spectral scanning (blue) and single-point detection (red). The dashed lines indicate the noise level from the intrinsic laser noise

4 Experimental results and comparison of methods

4.1 Experimental setup and system noise analysis

For experimental verification both methods were tested using an O_2 absorption line probed with a 763 nm VCSEL. A 20 cm double pass gas cell exposed to laboratory air was used with neither temperature nor pressure stabilization. The line generates a peak absorbance of around 9×10^{-3} . A custom built electronics for laser and detector control as well as data processing served as a platform to record the harmonic spectra or harmonic coefficients, whereas the curve fitting was done off-line on a PC. The modulation frequency was set to 3 kHz and the sinusoidal laser modulation amplitude was set close to three times the absorption line width, so that the typical modulation index of around 3 was realized.

Regarding the ideality of the realization of the electronics, the noise on the harmonic spectra or harmonic coefficients is a factor 1.2 to 2 higher than the noise that would be expected from the intrinsic noise by the laser diode (Fig. 7(bottom)). This is a technical issue attributed to strong receiver noise above 55 kHz (Fig. 7(top)) and the frequency response of the applied filtering during detection of the harmonics. It indicates that in a better realization of the electronics the absolute noise performance could be a factor of 1.2 to 2.0 lower with the same laser. Comparison of the methods is, however, not affected.

In Fig. 8 the Allan plot of the two multi-harmonic detection methods, the corresponding single-harmonic detection, and the extrapolated direct spectroscopy detection are shown. For the latter a spectral noise density of $1.2 \times 3.5 \times 10^{-7}/\sqrt{\text{Hz}}$ was assumed (this is the receiver noise density at 3 kHz times a non-ideality factor 1.2). For conversion to noise density on the fitted absorbance a conversion factor of

3.4 was computed. It is obtained from the covariance matrix calculated with linearization of the spectral model for direct spectroscopy (right-hand side of Eq. (2)) [12].

These results are also summarized in Table 1 together with relevant system parameters. The column “scanning range” specifies the range of the slow frequency sweep of ν and hence is zero for the single spectral frequency detection.

4.2 Comparing both multi-harmonic detection methods

It can be seen that the reconstruction method has a slightly better noise performance than multi-harmonic curve fitting but is more sensitive to drift, as the relatively worse minimum of the Allan deviation shows. However, the behavior with respect to noise and drift performance of the reconstruction method is better than that for second-harmonic single-point detection. Although second-harmonic curve fit-

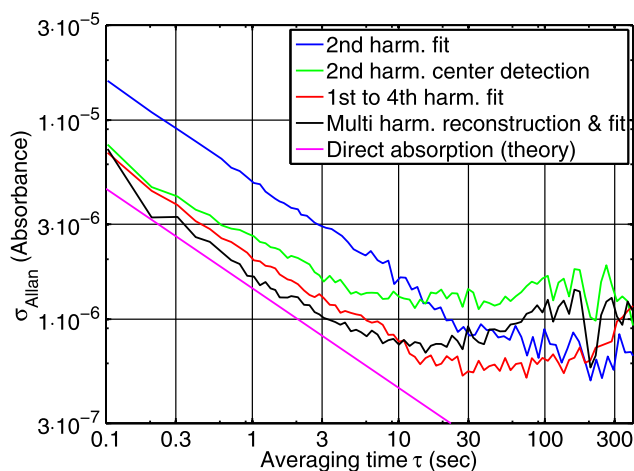


Fig. 8 Comparison between single-harmonic detection (single-point and spectral fitting) and the corresponding multi-harmonic detection (single-point reconstruction and simultaneous spectral fitting). The multi-harmonic detection methods lie close to what direct absorption spectroscopy could theoretically achieve (white noise only, no drifts)

ting shows a slightly better performance in terms of drift sensitivity than multi-harmonic detection, this is probably not a generic property of multi-harmonic curve fitting, because it is seemingly caused by the more drift affected contribution of the first harmonic in the multi-harmonic detection. This can be seen in Fig. 6 where the minimum of the Allan plot for second-harmonic detection is better compared to first-harmonic detection.

4.3 Comparing single- and multi-harmonic detection

If only white noise performance is considered (the region of the Allan plot with low averaging times where the slope is $-1/2$), both multi-harmonic detection methods give a factor of approximately 3 performance increase (2.7 for simultaneous curve fitting) to second-harmonic curve fitting.

The improvement is less compared to single-harmonic center point detection, which however is an unfair comparison because the numbers of determined parameters are different. For spectral scanning the noise on the harmonic spectra is higher than for single-harmonic center point detection, because a higher bandwidth around each harmonic frequency is required. Theoretically a \sqrt{N} fold increase of the noise standard deviation is expected (N : number of scanned spectral points). This is partly compensated again by the filtering of the curve fit, however, because the curve fit has to extract other parameters than absorbance, so the noise improvement is a certain factor less than \sqrt{N} . The reason does not lie in the curve fit itself, but is fundamental. The information content about peak absorbance is less in a scanned spectrum than for the single spectral frequency detection [12]. This explains that second-harmonic peak detection has a better white noise performance than second-harmonic curve fitting (cf. Fig. 8). The ability to determine more than one parameter and obtain better stability or calibration-free operation is payed off by a worse white noise performance. This is also indicated in Fig. 8, where the minimum of the Allan deviation is much worse for the single-harmonic center detection than for second-harmonic curve fitting, despite

Table 1 Overview and comparison of WMS detection methods. All measurements were taken with an identical experimental setup and identical settings

Description	Harmonics n	Scanning range	Modulation amplitude ν_A	Evaluation	$\sigma_{Allan} @ \tau = 1 \text{ s}$
1st harm. fit	1	$8.9\nu_\gamma$	$3.2\nu_\gamma$	Fit to Eq. (3)	6.1×10^{-6}
2nd harm. fit	2	$8.9\nu_\gamma$	$3.2\nu_\gamma$	Fit to Eq. (3)	4.8×10^{-6}
1st & 2nd harm. fit	1, 2	$8.9\nu_\gamma$	$3.2\nu_\gamma$	Fit to Eq. (3)	4.2×10^{-6}
2nd harm. center detection	2	0	$3.3\nu_\gamma$	Multiply with factor	2.6×10^{-6}
1st to 3rd harm. fit	1, 2, 3	$8.9\nu_\gamma$	$3.2\nu_\gamma$	Fit to Eq. (3)	2.5×10^{-6}
1st to 4th harm. fit	1, 2, 3, 4	$8.9\nu_\gamma$	$3.2\nu_\gamma$	Fit to Eq. (3)	2.0×10^{-6}
1st to 5th harm. fit	1, 2, 3, 4, 5	$8.9\nu_\gamma$	$3.2\nu_\gamma$	Fit to Eq. (3)	1.9×10^{-6}
Multi-harm. reconstruction & fit	0, ..., 20	0	$3.3\nu_\gamma$	Fit to Eq. (2)	1.6×10^{-6}

the better white noise performance. This is attributed to the higher sensitivity to fringes and laser wavelength instability, which are better compensated by the curve fit.

4.4 Comparing multi-harmonic detection and ideal direct spectroscopy

Both multi-harmonic detection methods have white noise performances in the order of what an estimated (white noise limited) realization of direct absorption spectroscopy with the sensor electronics would achieve (pink curve in Fig. 8). Theoretically, multi-harmonic WMS and direct spectroscopy should achieve both the same precision, because the overall frequency scanning range and relative distribution of time the laser spends on different spectral frequency regions during one scan are (approximately) the same (see Appendix B). The slightly lower white noise performance of the multi-harmonic curve fit may be explained by the missing zeroth-harmonic spectrum in the fit. It has been excluded, because the necessary weighting to equalize noise among all harmonic spectra (cf. Fig. 7) would greatly damp the zeroth harmonic, and so the benefit of its inclusion is negligible.

4.5 Differences in receiver and transmitter bandwidth requirement

Since the multi-harmonic detection methods theoretically should achieve the same precision than direct absorption spectroscopy operated at the same modulation or repetition frequency, the question is: Are there any differences between these two methods? One difference lies in the modulation bandwidth requirement on the transmitter side, for which multi-harmonic detection is advantageous. Direct absorption spectroscopy relies on the fact that the laser is able to emit fast and precise frequency sweeps, i.e., sawtooth or triangular-shaped emission frequency variations. Therefore the laser tuning bandwidth needs to be high, i.e., higher than the basic modulation frequency or repetition rate, since many higher harmonics have to be passed unattenuated through the transmitter including the tunable laser. For WMS only a sine wave frequency modulation is needed, so the bandwidth requirement is low (only a certain bandwidth around the modulation frequency is required). This is advantageous for lasers with low tuning bandwidth or a non-flat FM response [13].

Similar considerations can be made for the receiver side. The methods are summarized in Table 2, and a selection of the proper detection scheme can be made according to the bandwidth limitations of the system. One may ask what a fourth entry to the table might be (i.e., low receiver bandwidth but high transmitter bandwidth). It is a non-sinewave single-harmonic WMS (see, e.g., Refs. [14, 15]). Since this

Table 2 Comparison of bandwidth requirements for different methods. If a wide receiver bandwidth is available, either direct spectroscopy or multi-harmonic WMS should be selected, since both methods have similar and better performance than single-harmonic WMS

Method	Requirement on transmitter bandwidth	Requirement on receiver bandwidth
Direct absorption	wide	wide
Single-harmonic WMS	narrow	narrow
Multi-harmonic WMS	narrow	wide

realizes a very non-uniform emission frequency scanning coverage, the noise performance is not necessarily like direct spectroscopy with its uniform scanning; it depends on the system model.

5 Conclusion

It is shown that multi-harmonic detection gives an approximately factor of 3 noise improvement comparable to single-harmonic curve fitting. The overall performance is close to the theoretical performance of direct spectroscopy with its typically factor of 3 or more larger spectral amplitude for the same conditions. For simultaneous curve fitting the inclusion of the first to fourth harmonic spectrum is sufficient to achieve optimum performance. For the reconstruction of the transmission all harmonics above the noise level have to be included; this was 20 in our case. For systems with spectral scanning the simultaneous curve-fitting approach is probably more promising due to its lower drift sensitivity and the lower number of required harmonics.

Finally, we conclude that multi-harmonic detection can be regarded as a technical implementation method of TDLS with similar performance which would be expected from direct spectroscopy and the advantage that no requirement on the bandwidth of the laser tuning behavior is necessary.

The first author wishes to thank C. Smith, R. Lewicki, and M. Witinski for valuable discussions and for pointing out the limitations of multi-harmonic detection with QEPAS and cavity enhanced methods.

Appendix A: Wavelength modulation spectroscopy

Let ν denote the wavenumber or frequency of the central laser emission (without the sinusoidal modulation) which implements the slow (discrete) laser emission frequency sweep. The sweep is discrete, because ultimately the spectra are sampled at distinct frequency points and also ν has to be constant during at least one sinusoidal modulation period, so that the harmonic coefficients can be properly determined by

the lock-in amplifier. The instantaneous laser emission frequency $\nu_L(t)$ is given by

$$\nu_L(t) = \nu + \nu_a \cos(2\pi f_m t) \quad (5)$$

with ν_a the frequency modulation amplitude and f_m the modulation or repetition frequency (typically in the kHz range). The n -th harmonic output of the lock-in amplifier of the normalized light power variation $T(\nu_L(t))$ after passing through the sample with transmission $T(\nu)$ is called the harmonic coefficient $H_n = H_n(\nu; \nu_a)$ (Fig. 1). Mathematically, the Fourier series decomposition

$$T(\nu + \nu_a \cos(2\pi f_m t)) = \sum_{n=0}^{\infty} H_n(\nu; \nu_a) \cos(2\pi n f_m t) \quad (6)$$

is computed (by solving for $H_n(\nu; \nu_a)$). Note that the definition of harmonic spectra only requires knowledge of the normalized detector signal during one modulation period. The same is true for a practical realization of a digital lock-in. But for optimum noise performance all periods between two sampling points of ν should be averaged before Fourier decomposition is carried out. Also note that the Fourier decomposition is done in phase with the frequency modulation. The proper detection phase setting in a real lock-in amplifier is thus to compensate for the phase shift of the laser driver and the intrinsic phase shift between laser current and emission frequency [13]. Ideal harmonic spectra have no out-of-phase component. Out-of-phase components encountered in experimental systems may occur despite proper adjustment of the detection phase if there is additional laser amplitude modulation with an AM-FM phase shift. These out-of-phase components, however, contain no or only very little information about the gas spectrum, so only the in-phase component undergoes further signal processing.

Due to the inherent linearity of the signal processing, there is also an ideally linear relationship between the transmission spectrum and the corresponding harmonic spectra. It is linear in the sense that scaling and summation in the transmission also results in scaling and summation of the corresponding harmonic spectra. As a consequence, the relationship can be modeled as a convolution, and general closed-form expressions for the convolution kernel as well as for its Fourier transform can be derived [16]. The linearity of the relationship essentially means that, for unsaturated absorption lines, the individual lines simply add in both the transmission and the harmonic spectra. The concentration or peak absorbance scales the harmonic spectrum, whereas for orders n greater than zero additionally the large offset is removed. This is due to linearity of the relationship and the fact that the harmonic spectrum of a flat transmission is zero for higher orders. This offset removal property of WMS is usually considered as one of its advantages, because detection of a small signal on a large offset is not needed. Furthermore, it can be shown that the n -th harmonic spectrum has

the polynomial components of n -th degree in the transmission removed, i.e., the second harmonic spectrum is insensitive to linear slopes and the third harmonic insensitive to linear and quadratic components in the transmission, and so on. However, when the computation of the harmonic spectra is done by digital signal processing, one could also devise filtering methods to direct transmission spectra which have the same or possibly even better properties.

If the width of a line in the transmission changes, generally a complex change of the shape of the harmonic spectra is observed. Both the width and amplitude of the harmonic spectrum change. The overall shape of the harmonic spectrum generally depends on the modulation index, the ratio of line width and frequency modulation amplitude. The width of the harmonic spectrum is also generally much larger than the width of the absorption line, since the ‘‘broadening’’ is dominated by the frequency modulation amplitude, which is usually a factor of 2 to 3 larger than the absorption line half-width (see also Fig. 2). A more detailed and accurate description of the generation of harmonic spectra, which is not further relevant in this paper, can be found in Refs. [4, 16] or [17].

An interesting analogy between the Taylor series and harmonic spectra becomes evident from the reconstruction method in Sect. 2. In the case where all derivatives of the transmission $T(\nu)$ at a certain point ν_0 are known, then $T(\nu)$ can be reconstructed in a certain range around ν_0 . This is a consequence of the Taylor theorem:

$$T(\nu) = \sum_{n=0}^{\infty} \frac{d^n T(\nu_0)}{d\nu^n} \frac{(\nu - \nu_0)^n}{n!} \quad (7)$$

If we insert the asymptotic expression for harmonic spectra [17]

$$H_n(\nu; \nu_a) \rightarrow \frac{\epsilon_n (\nu_a/2)^n}{n!} \frac{d^n T(\nu)}{d\nu^n} \quad \text{for } \nu_a \rightarrow 0 \quad (8)$$

with $\epsilon_0 = 1$ and $\epsilon_n = 2$, for $n \neq 0$, we obtain

$$T(\nu) \approx \sum_{n=0}^{\infty} H_n(\nu_0; \nu_a) \frac{2^n}{\epsilon_n} \left(\frac{\nu - \nu_0}{\nu_a} \right)^n \quad (9)$$

However, this would only be valid for low values of ν_a where the signal-to-noise ratio is non-optimum. On the other hand the formula derived for the multi-harmonic detection scheme Eq. (1) has no such limitation, but has a very similar structure. For the convenience of the reader, Eq. (1) is stated again:

$$T(\nu) = \sum_{n=0}^{\infty} H_n(\nu_0; \nu_a) T_n \left(\frac{\nu - \nu_0}{\nu_a} \right) \quad (10)$$

The functions in Eq. (9) are $1, x, 2x^2, 4x^3, 8x^4, \dots$, whereas in the exact case the Chebyshev polynomials are given by

1, x , $2x^2 - 1$, $4x^3 - 3x$, $8x^4 - 8x^2 + 1, \dots$ with $x = (\nu - \nu_0)/\nu_A$ as a shorthand notation. By comparison we see that the approximate formula Eq. (9) only contains the leading coefficients of the Chebyshev polynomials in the exact formula Eq. (1).

Appendix B: Theoretical noise performance

The (theoretical) overall white noise performance of any tunable diode laser absorption spectroscopy (TDLAS) method only depends on the system model, the total spectral frequency coverage during one scan (“which wavelengths/frequencies”), and the relative distribution of time the laser spends on different frequency regions during one scan (“how long?”). The system model is independent of the specific detection method; it essentially describes the behavior of the black-box containing the optical system with laser and detector. This system has just an electrical input (laser control) and electrical output (detector current or preamplifier output). From Table 1 it can be seen that the covered wavelength range is approximately the same for direct spectroscopy and WMS. Hence, only the distribution of time (as a fraction of total time per scan) that the laser dwells on the individual frequencies of one spectral scan can influence noise performance. The specific variation in time of the frequency is not important, as long as the distribution of the dwelling time is the same. If “important” regions of the transmission (i.e., those which are more sensitive to changes in parameters we are interested in) are measured over longer time fractions than other frequency regions, we can expect a better quality of extraction of wanted parameters from the measurement. Vice versa, if less important regions of the transmission are sampled over relatively long fractions of time, the extraction will be of lower quality. For example, this may be illustrated in a very simplified model, where we scan over a single absorption line and are only interested in the peak absorbance. If in this case the measurement of the baseline consumes far more time than measurement of the absorption line peak, the performance will be non-optimum. The reason is that the peak absorbance is the difference between both values, and we should measure them with the

same accuracy, i.e., spend the same amount of time on both measurements. So even if it is fixed which wavelengths are to be sampled, the length of time that they are sampled also affects performance [12].

WMS with spectral frequency scanning and direct spectroscopy both realize an (approximately) uniform coverage of the wavelength during one scan. For WMS at a single point the absence of scanning gives a non-uniform frequency (sinusoidal) coverage over time. However, the deviation to uniform (linear) variation is not great, and numerical simulations have shown that the overall effect on noise performance is only on the order of 10 % (at least for the spectral model used in this paper).

References

1. J. Reid, D. Labrie, *Appl. Phys. B, Lasers Opt.* **26**, 203 (1981)
2. R. Arndt, *J. Appl. Phys.* **36**, 2522 (1965)
3. J. Liu, J. Jeffries, R. Hanson, *Appl. Phys. B, Lasers Opt.* **78**, 503 (2004)
4. P. Kluczynski, J. Gustafsson, Å.M. Lindberg, O. Axner, *Spectrochim. Acta, Part B, At. Spectrosc.* **56**, 1277 (2001)
5. A.A. Kosterev, Y. Bakhirkin, R.F. Curl, F.K. Tittel, *Opt. Lett.* **27**, 1902 (2002)
6. A. O’Keefe, J.J. Scherer, J.B. Paul, *Chem. Phys. Lett.* **307**, 343 (1999)
7. J.A. Silver, D.S. Bomse, Wavelength modulation spectroscopy with multiple harmonic detection. Patent US6356350 (1999)
8. M. Abramowitz, I.A. Stegun (eds.), *Handbook of Mathematical Functions*, 9th edn. (Dover, New York, 1970)
9. J. Chen, A. Hangauer, R. Strzoda, M.-C. Amann, *Appl. Phys. B, Lasers Opt.* **102**, 381 (2010)
10. L. Fox, I.B. Parker, *Chebyshev Polynomials in Numerical Analysis* (Oxford University Press, London, 1968)
11. T. Svensson, M. Andersson, L. Rippe, S. Svanberg, S. Andersson-Engels, J. Johansson, S. Folestad, *Appl. Phys. B, Lasers Opt.* **90**, 345 (2008)
12. J. Chen, A. Hangauer, R. Strzoda, M.C. Amann, *Appl. Phys. B, Lasers Opt.* **100**, 331 (2010)
13. A. Hangauer, J. Chen, R. Strzoda, M.-C. Amann, *IEEE J. Sel. Top. Quantum Electron.* **17**, 1584 (2011)
14. T. Iguchi, *J. Opt. Soc. Am. B, Opt. Phys.* **3**, 419 (1986)
15. J. Saarela, J. Toivonen, A. Manninen, T. Sorvajärvi, R. Hernberg, *Appl. Opt.* **48**, 743 (2009)
16. A. Hangauer, J. Chen, M.-C. Amann, *Appl. Phys. B* **90**, 249 (2008)
17. A.N. Dharamsi, *J. Phys. D, Appl. Phys.* **29**, 540 (1996)