

# Applications of wavelet transform to quantum cascade laser spectrometer for atmospheric trace gas measurements

Jingsong Li · Uwe Parchatka · Horst Fischer

Received: 23 March 2012/Revised: 9 July 2012/Published online: 16 September 2012  
© Springer-Verlag 2012

**Abstract** The application of wavelet transform-based digital filter to quantum cascade laser spectroscopy was investigated by its application to simulated spectra and experimental results obtained with our novel and compact QCL spectrometer, which offers the potential for high sensitive and selective measurements of trace gas concentration at high temporal resolution (less than 1 s) for species of interest in the atmosphere. With the application of wavelet digital filtering in post-signal processing, better measurement precision and higher detection sensitivity have been achieved, without reducing the fast temporal response. Details of wavelet transform theory applied to the spectroscopic data are presented. For comparison, the results of other commonly used filter techniques (i.e. Kalman filter, Wiener filter and moving average) are also tested. Finally, the experimental results show that the wavelet-based filter seems to be the superior choice for noise reduction, both for spectral signal processing and trace gas concentration signal evaluation.

## 1 Introduction

In atmospheric science, to gain reliable information from the measured trace gas data for improved understanding of atmospheric processes and transformations, to develop better climate models and to better understand their impact on global warming and climate change, it is of great importance, that the level of unwanted noise in scientific instruments should be suppressed as low as possible to

ensure sufficient precision and accuracy. Generally, for the typical white noise, this is achieved by signal averaging to gain a satisfactory signal-to-noise ratio (SNR). Thus, the noise can be reduced up to the limits of the spectrometer stability and the optimal averaging time can be determined by an Allan variance analysis [1]. Unfortunately, multi-signal averaging is a time-consuming method and it reduces the temporal resolution of the measurement.

Unfortunately, most optical instruments in particular in the IR-spectral range are limited in sensitivity by optical noise fringes superimposed on the recorded spectra and not by excess laser or detector noise. In order to reduce the effect of these fringes, many methods have been proposed, such as mechanical modulation or dithering of the etalon spacing, modified modulation schemes, background subtraction, and post-detection signal processing. From a practical standpoint, numeric filtering is easy to implement, since it requires no modifications or additions to the apparatus hardware and can be easily adapted to any experimental configuration. Mathematical filtering techniques for on-line noise reduction or off-line data processing of recorded spectra may be a better choice when temporal resolution is crucial. Traditional filtering methods in most cases rely on the identification of frequencies of noise contributions obtained in the stationary power spectrum. Many classic digital filter techniques and signal processing schemes have been developed and successfully applied to tunable diode laser absorption spectroscopy (TDLAS) for improving system performance [2, 3]. It is worth noting that one of the adaptive filtering techniques known as Kalman filter is still increasingly being employed in many research fields [3, 4]. The Kalman filter was thought of as an adaptive (or dynamic) bandwidth filter, because its frequency response is not fixed, but can adjust to changes in signal statistics and dynamic range during

---

J. Li (✉) · U. Parchatka · H. Fischer  
Department of Atmospheric Chemistry, Max Planck Institute  
for Chemistry, Hahn-Meitner-Weg 1, 55128 Mainz, Germany  
e-mail: Jingsong.li@mpic.de

operation. Thus, it offers an advantage over simple averaging and conventional low-pass filters, which have a fixed bandwidth and attenuate all fluctuations in the signal, including those outside the filter bandwidth. Indeed, the Kalman filter is an optimum estimation algorithm that was first used in aerospace modeling problems; this technique has proven to be extremely useful to estimate the trace gas concentrations in TDLAS measurements. However, one limitation of the Kalman filtering technique, as described by Leleux et al. [5], is the fact that it operates on concentration values only after a data processing algorithm has been applied to the original spectral signal. Indeed, the Kalman filter removes disturbances or faults from the signal using initialization and propagation of error covariance statistics, i.e. computes, and propagates the mean and the covariance matrix recursively for a linear system. In distributed systems, the computational expense of Kalman filter is thus dominated by the error covariance propagation step, which makes implementation of the Kalman filter impractical in large-scale models [6, 7].

Over the past decade, a new technique known as wavelet transform has been gradually developed using scaling and shifting properties to get better time and frequency resolution based on the Fourier transform and has been successfully employed in many fields, such as data compression, detecting features in images, and removing noise from signals. Wavelet transform has been proved as a powerful tool for signal processing mainly due to its multi-resolution characteristics, i.e. dividing the frequency contents of a signal into low and high sub-bands. Unlike the Fourier transform which considers only a single set of basis functions (sines or cosines), wavelet transforms use an infinite set of possible basis functions (i.e. mother wavelets or analyzing wavelets) with different properties. Thus, wavelet analysis provides immediate access to information that can be obscured to other time–frequency methods such as Fourier analysis. Discrete wavelet transform (DWT) is the only linear transform that can analyze non-stationary signals at varying resolutions by decomposing the signals into their frequency bands. In addition, DWT is a very fast algorithm with polynomial time and space complexity, which makes it more appealing. It is also worth mentioning that wavelet transform can be applied to any signal source if an appropriate basic wavelet is available, regardless of the measurement techniques. The wavelet transform tends to concentrate the signal energy into a relatively small number of coefficients with larger values. This energy-concentrating property makes the wavelet analysis appropriate for signal denoising, estimation and forecasting [8–11], and sometimes appears to be better suitable than the widely used Kalman filter [7, 12].

In this paper, we evaluate the application of the wavelet-based denoising technique to a quantum cascade laser

spectrometer for in situ and real-time atmospheric trace gas measurements. With the utilization of this algorithm to improve the spectral SNR and to minimize the dispersion of concentration values, higher detection sensitivity and better measurement precision can be achieved without reducing the fast temporal response. In order to assess the benefit of the wavelet filter, the efficiency of the wavelet filter was compared with other commonly used filter techniques, such as the Kalman filter, the Wiener filter and a moving average technique. The remainder of this paper is organized as follows: Sect. 2 gives a brief overview of the basic wavelet transform theory and the thresholding method used. Section 3 presents the procedure for determining the optimal wavelet parameters by simulation. Section 4 briefly describes the experimental instrument, the application wavelet to actual observed results, and the comparison with other filter techniques. Finally, Sect. 5 contains the conclusions and some outlook.

## 2 Wavelet transform

### 2.1 Wavelet theory and algorithm

Generally, the continuous wavelet transform (CWT) of a signal  $f(t)$  can be written as:

$$W_f(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t) \Psi\left(\frac{t-\tau}{s}\right) dt \quad (1)$$

where  $\tau$  and  $s$  are the so-called translation (or time location) factor and the scaling (or dilation) factor, respectively. The factor  $|s|^{-1/2}$  is for energy normalization across the different scales, whereas  $\Psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}} \Psi\left(\frac{t-\tau}{s}\right)$  is the so-called *continuous wavelet*, or “mother wavelet”. For each scale  $s$  and location  $\tau$ , the wavelet coefficient  $W_f(\tau, s)$  represents the information contained in  $f(t)$  at that scale and position. Thus, the original signal can be exactly reconstructed from the wavelet coefficients by inverse wavelet transform:

$$f(t) = \frac{1}{C_\Psi} \int_0^\infty \int_{-\infty}^{+\infty} W_f(\tau, s) \Psi_{\tau,s}(t) d\tau \frac{ds}{s^2} \quad (2)$$

where  $C_\Psi$  is defined as  $C_\Psi = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega} d\omega$ ,  $\Psi(\omega)$  is the Fourier transform of the mother wavelet.

If we assume  $s = s_0^j$  and  $\tau = k\tau_0 s_0^j$  ( $j, k \in \mathbb{Z}, s_0 \neq 0$ ), then the wavelet can be re-written as  $\Psi_{j,k}(t) = s_0^{-j/2} \Psi(s_0^{-j} t - k\tau_0)$ , which is commonly called *discrete wavelet*, where  $s_0 = 2$  and  $\tau_0 = 1$  are generally used. From the definition, one can see that the wavelet is a family of

functions derived from a basic function  $\Psi(t)$ . The DWT can be carried out with several different algorithms. A common pyramid algorithm of multi-resolution signal decomposition introduced by Mallat, for calculating DW coefficients that are widely used in signal processing, is graphically illustrated in Fig. 1. The algorithm consists of a series of successive decompositions of the signal (with length of  $2^n$ ) into two components: “detail coefficients  $D_j$ ” and “approximation coefficients  $A_j$ ” both with a reduced size of  $2^{n-j}$ , where  $j$  is the decomposition level. At each level, high and low-pass filters are applied to the input signal. The actual features of these filters depend on the wavelet basic function used. The high-pass filter at each scale filtering the high-frequency components is recorded as the wavelet coefficients. While the low-pass filter extracts the low-frequency components for the next scale before another set of high- and low-pass filters is employed. The procedure is repeated until at a prescribed level  $j$  is reached. Note that the maximum of  $j$  cannot exceed  $n$ . This way, the pieces of information represented by  $D_j$  and  $A_j$  are different and correspond to the high-frequency and the low-frequency part of  $A_{j-1}$ . They are always orthogonal to each other. The lengths of vectors  $D_j$  and  $A_j$  are the same and are half of that of  $A_{j-1}$ . More details about the algorithm can be found in the Mallat’s work [13].

Multi-resolution analysis provides wavelets with the capability of decomposing any signal into its contributions in different regions of time–frequency or time-scale space, where various features of the signals can be well studied. The basic idea of multi-resolution analysis is to write a function or the signal  $f(t)$ , as a limit of successive approximations at different approximation subspaces, while each of the approximations at subspaces is a

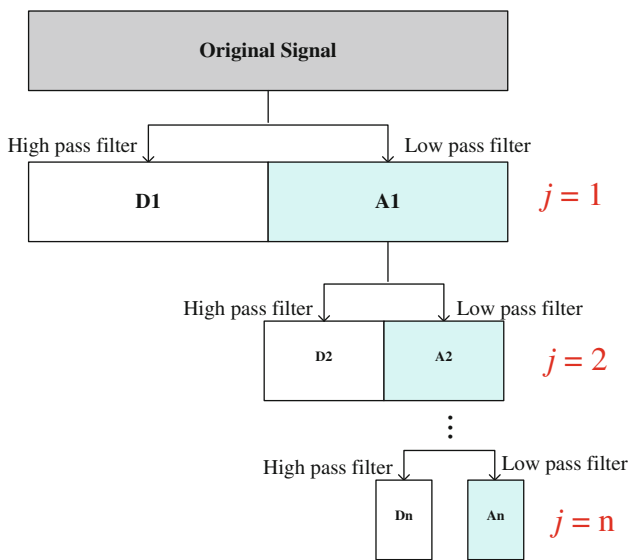


Fig. 1 Schematic diagram of the Mallat’s algorithm

smoother version of  $f(t)$ . According to the Mallat’s algorithm, in case of CWT, a signal can be decomposed with wavelets as follows [9]:

$$f(t) = \sum_{k=-\infty}^{\infty} A_{j,k} \varphi_{j,k}(t) + \sum_{j=1}^n \sum_{k=-\infty}^{\infty} D_{j,k} \Psi_{j,k}(t) \tag{3}$$

where  $\varphi_{j,k}(t)$  and  $\Psi_{j,k}(t)$  are called scaling functions and wavelet functions, respectively.  $A_{j,k}$  and  $D_{j,k}$  are the approximation coefficients and the detail coefficients mentioned in the Mallat’s algorithm. At each successive scale (or decomposition level), only high-frequency information (noises) is retained in the details, while the low-frequency information (signal features) is retained in the approximations. Subsequently, the standard denoising technique operating with thresholding policy is applied to the wavelet coefficients for removing the noises. The denoising procedure requires the estimation of the noise level from the detail coefficients. Finally, the de-noised signal can be reconstructed with the new estimated wavelet coefficients:

$$f^*(t) = \sum_{k=-\infty}^{\infty} A_{n_{opt},k} \varphi_{j,k}(t) + \sum_{j=1}^{n_{opt}} \sum_{k=-\infty}^{\infty} D_{j,k}^* \Psi_{j,k}(t) \tag{4}$$

where  $A_{n_{opt},k}$  are the approximation coefficients at the optimal decomposition level  $n_{opt}$ , and  $D_{j,k}^*$  are the detail coefficients that are retained. In practical applications, any measured variable signal can be assumed to be the sum of two components:

$$f(t) = f^*(t) + \zeta(t) \tag{5}$$

where  $f^*(t)$  is the ideal signal and  $\zeta(t)$  is the associated noise. Therefore, the basic idea to estimate  $f^*(t)$  (i.e. by denoising  $f(t)$  and extracting the true trend) using wavelets including the current work is as follows:

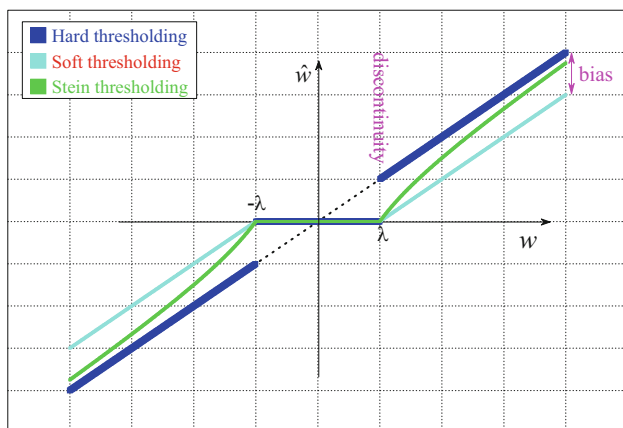
- (1) Decompose the signal using DWT into  $n$  levels ( $n$  depending on the signal length of  $2^n$ ) and obtain the empirical wavelet coefficients at each scale  $j$  ( $j = 1, \dots, n$ ).
- (2) Thresholding of the empirical wavelet coefficients using proper thresholding function, so that the estimated wavelet coefficients are obtained based on the selected threshold.
- (3) Reconstruct the processed signal from the thresholded wavelet coefficients using the inverse DWT.

### 2.2 Thresholding policy and threshold estimation

To perform the denoising algorithm described above, we must choose the appropriate transform, the thresholding function and the value of the threshold  $\lambda$  used by the thresholding function. Usually, two different thresholding

approaches are applied to signal denoising: hard thresholding or soft thresholding. The hard thresholding operation sets the wavelet coefficients smaller than the threshold to zero and keeps the values of the other wavelet coefficients. The soft thresholding operation consists of setting the wavelet coefficients smaller than the threshold to zero and shrinks the others toward zero. The hard thresholding operation keeps the amplitude constant before and after denoising, but might induce some Gibbs oscillation at the edges due to its discontinuity. The so-called soft thresholding function has the well-known and desirable properties of smoothness and adaptation, which might reduce the signal amplitude due to the constant presence of a bias at the wavelet coefficients  $w$  higher than the threshold ( $|w| > \lambda$ ). In view of these issues (depicted in the next section), another way to achieve a tradeoff between hard and soft thresholding is to use a soft-squared thresholding non-linearity, also named a Stein estimator, which has been used on the signal processing in this paper. The new estimated wavelet coefficients  $\hat{w}$  under different threshold estimators are graphically shown in Fig. 2.

The last parameter to specify is the value of the threshold. The choice of the threshold directly influences the effectiveness of the denoising algorithm. A too high threshold would result in too many wavelet packet decomposition coefficients being reset to zero, and thus losing too many details of the signal, while a too low threshold would not yield the expected denoising effect. Universal threshold estimation known as the most popular one was widely applied to estimate the threshold. It is based on the statistical properties of white Gaussian noise, but might lead to an overestimation of the noise level. In this study, an adaptive threshold selection rule of Stein's Unbiased Estimate of Risk (SURE) for signal denoising was used [14]. The aim of this estimate is to minimize the risk. Because the coefficients of the true signal are unknown, the true risk is also unknown. We derive the



**Fig. 2** Thresholding policy of wavelet coefficients

unbiased estimate of the true risk for generalized threshold functions, then, the SURE threshold value minimizes the unbiased risk estimate. The risk function based on the standard minimum mean square error (MSE) can be expressed as [15]:

$$R(\hat{w}, w) = \frac{1}{N} \sum_{j,k} (\hat{w}_{j,k} - w_{j,k})^2. \quad (6)$$

### 3 Simulation

The purpose of an optimal filter is to recover the de-noised signal without degrading the approximation degree between the real signal and the reconstructed signal. The key issue is which wavelet should be used for signal processing. Consequently, doing some theoretical simulation and test is necessary before applying this technique to actual observations. A computer program has been written in the numerical script language Python for the computations and signal simulations. In order to evaluate the performance of the denoising operation, large numbers of spectral simulations were performed. The criterion to quantitatively illustrate the effectiveness of the denoising operation is the SNR improvement, defined as follows:

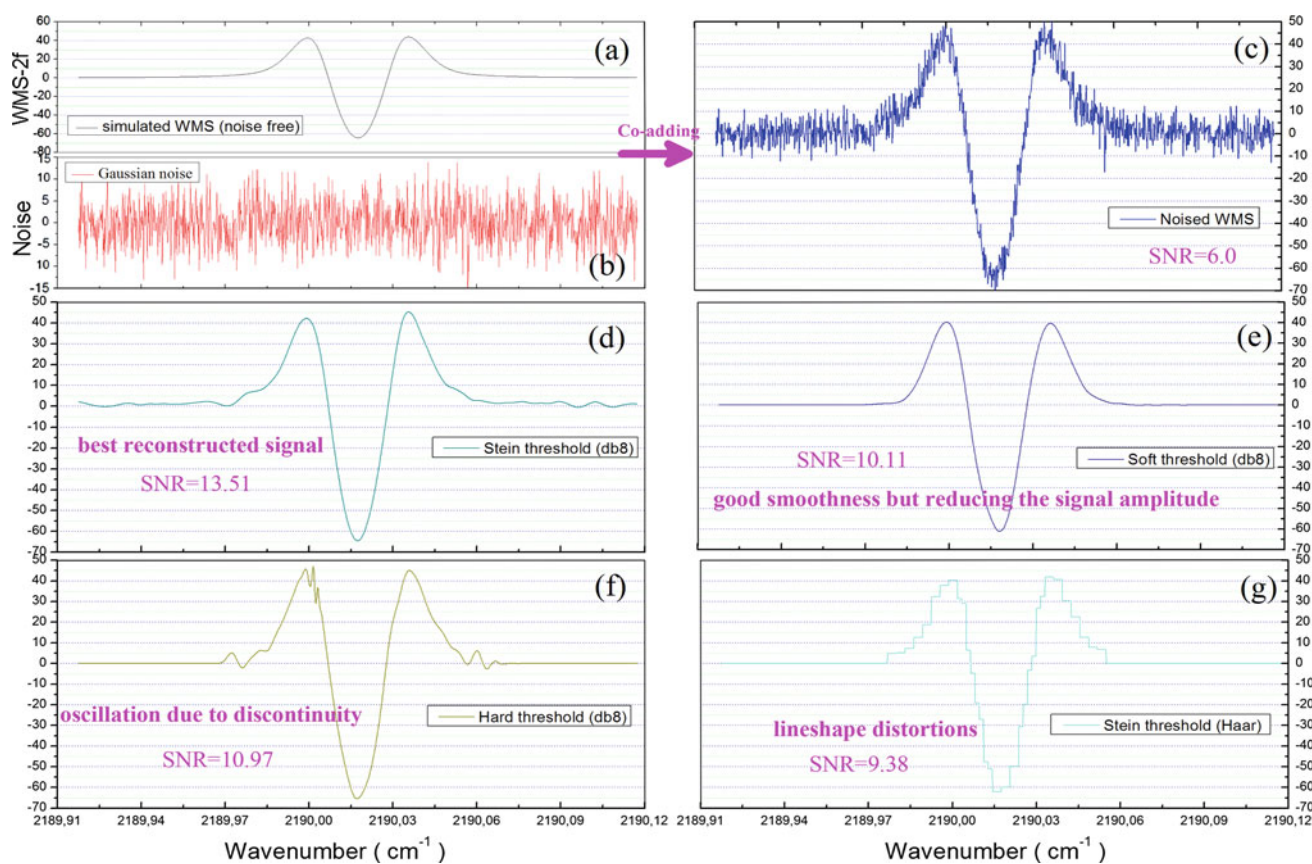
$$\begin{aligned} \text{SNR(dB)} &= 10 \log_{10} \left( \frac{\text{std}(\text{Signal}_{\text{noise-free}})}{\text{std}(\text{Signal}_{\text{noise-free}} - \text{Signal}_{\text{wavelet denoised}})} \right) \end{aligned} \quad (7)$$

where std means the standard deviation,  $\text{Signal}_{\text{noise-free}}$  and  $\text{Signal}_{\text{wavelet denoised}}$  are ideally simulated spectral and wavelet de-noised spectral signals, respectively.

To achieve a better filter performance, some parameters used for the wavelet filter should be optimized. The parameters mostly influencing the final signal denoising efficiency include the following five items: (1) wavelet type; (2) thresholding policy; (3) threshold; (4) method of decomposition and synthesis; (5) decomposition level. The key feature of the Mallat algorithm is how to choose a suitable wavelet basis and decomposition scale. We used the computer program to evaluate various kinds of parameter combinations.

From the viewpoint of practical applications, a CO absorption line near wavelength of  $2190.01 \text{ cm}^{-1}$  was very attractive for atmospheric CO measurements. Therefore, relative spectroscopic simulation of a wavelength modulated signal covering this spectral region was made (see Fig. 3). Related spectroscopic parameters are extracted from the HITRAN database [16], considering a set of given experimental conditions including temperature, pressure, gas concentration and optical path length. All the simulated spectra contain 1024 sample points. These simulated





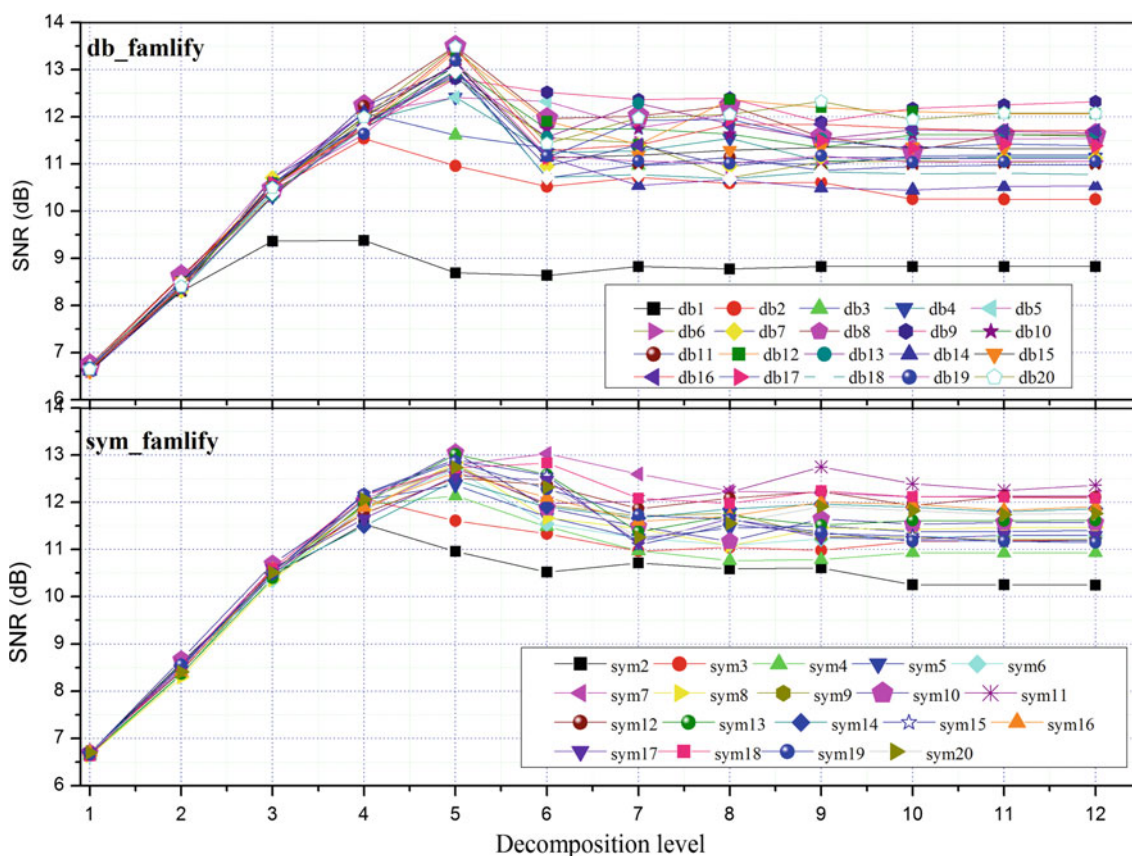
**Fig. 3** Spectral simulation (a) noise-free CO WMS-2f spectral signal, (b) Gaussian noise, (c) noised CO WMS-2f spectral signal, and wavelet filtered results at the same decomposition level of 5 (d) wavelet db8 with Stein thresholding, (e) wavelet db8 with soft

thresholding, (f) wavelet db8 with hard thresholding and (g) wavelet Haar with stein thresholding. The calculated SNR is also *inset* in the corresponding panel

spectra (2nd harmonic of WMS) were purposefully corrupted with Gaussian white noise to simulate an instrument-based signal. This type of noise interference often appears in optical spectroscopic instruments. To evaluate the performance of the wavelet-based signal denoising, a total of 76 wavelets belonging to seven different families of wavelets in the Python Toolbox were tested and compared according to the criterion mentioned above. Ultimately, we found that the Daubechies (db) and Symmlet (sym) wavelet families, except for the db1 and sym2 wavelets, show a relative better performance for spectral signal denoising. Similar to other filters, improper choice of wavelet parameters can cause distortions and artifacts in the constructed signal. For comparison, the optimal reconstructed signal using wavelet db8 and the improper choice of the wavelet basis (i.e. Haar wavelet) under the Stein thresholding policy, as well as sometimes the issues induced by soft and hard thresholding scheme mentioned previously, are also illustrated in Fig. 3.

Besides the wavelet basis and the thresholding function, selection of the analysis depth (i.e. the number of decomposition levels) also plays an important role in a suitable

denoising process. As an example, Fig. 4 shows both db and sym wavelet families at different decomposition levels with Stein thresholding policy applied. As one can see, the SNR of de-noised spectrum exhibits an approximate linear increase with the decomposition level, with a slight decrease after the optimal decomposition level has been reached. This trend is almost for all wavelets. Moreover, the figure shows that both db and sym wavelet families have approximately the same optimal decomposition level of 5 except for the db1 wavelet. Most of the later yields show almost equivalent performance with respect to noise remove comparable to levels 5 and 6 with the exception of wavelet sym2, sym3 and sym11, while in case of sym11, the SNR at level 5 is very close to that at the best level of 9 (the discrepancy is only  $2.82 \times 10^{-2}$ ). For clarity, the best SNR obtained for each wavelet at the optimal decomposition level are compiled in Table 1. In a further simulation test, by adjusting the noise intensity to the simulated noisy 2nd harmonic signal with varying SNR from 1 to 10 dB, we found that the optimal decomposition levels have slightly shifted between levels 5 and 7 following the variation of the SNR. In conclusion, db8 and



**Fig. 4** SNR as a function of decomposition level for both db and sym wavelet families applying to the noised CO WMS-2f spectral signal in Fig. 3

sym10 show a relative better performance at the optimal decomposition level for denoising the simulated noisy 2nd harmonic signal. Note that sometimes the difference of the best SNR obtained between the so-called optimal wavelet and others (such as db8 and db20, sym10 and sym15) is really very small, as can be seen from Table 1. Anyway, wavelet db8 and sym10 are selected as good candidates for denoising applications to the experimental spectra in this work.

By reviewing literatures and our experimental trials with many wavelet families, generally, we found that an ideal wavelet basis function must be symmetrical or anti-symmetrical, orthonormal complement, compactly supported and high number of vanishing moments for a given support width. However, it is not easy to find a wavelet with all these properties, being adaptive to any signal features. The basic idea is that the energy or information of the signal should be concentrated on as few coefficients as possible. This depends mostly upon the number of vanishing moments of the wavelet and the size of the support. Besides, the wavelet library should try to match the expected shape of the signal, i.e. have the largest correlation with the signal to be analyzed. It is worth noting that

wavelet denoising is a data-dependent process, thus, the optimal wavelet parameters, such as mother wavelet, threshold and decomposition level, need to be carefully adopted, by combining the consideration of the noised signal characteristics [17].

### 4 Application to actual observations

#### 4.1 Quantum cascade laser spectrometer

The design principles and construction details for our room temperature QCL spectrometer are discussed in Ref. [18], so we only briefly review them here. The instrument (see Fig. 5) combines a commercially available room temperature QCL (ALPES Laser) operated in continuous wave (cw) mode, a compact astigmatic Herriott cell (Aerodyne Research, Inc., Model AMAC-36), a short reference cell, a sophisticated optical and electronic system, and a portable computer-controlled system programmed with Labview software that incorporates the electronics for driving the QCL along with signal generation, data acquisition and on-line signal analysis and display.

The QCL spectrometer is devised for simultaneous measurement of sample absorption from the Herriott cell and a frequency-lock spectrum from the reference cell using two similar TE-cooled mercury cadmium telluride (MCT) infrared detectors (PVI-4TE-5, Vigo Systems). The laser used here can be tuned to the infrared frequencies between 2187.6 and 2202.1  $\text{cm}^{-1}$  over a temperature range of 243–283 K with a power range of 0.1–13.5 mW. This

**Table 1** The best SNR obtained for db and sym family wavelets and the corresponding optimal decomposition level

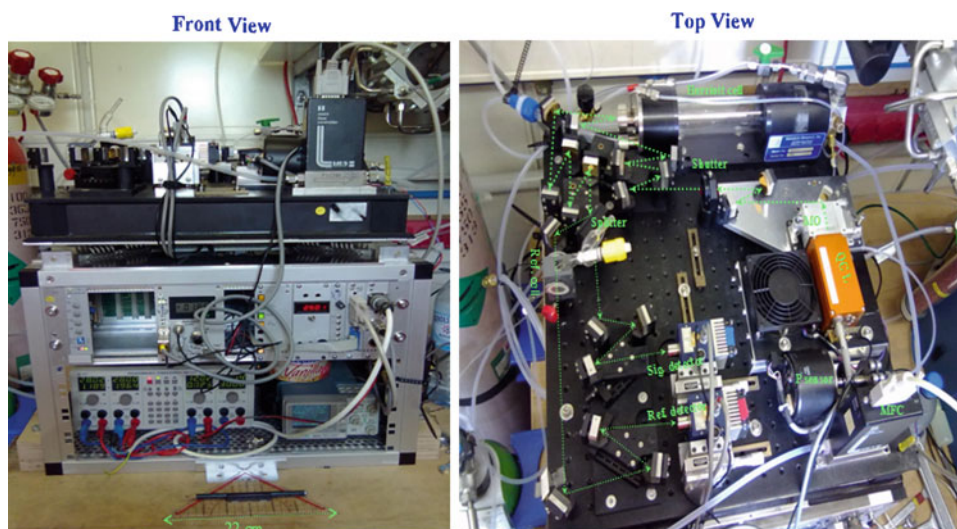
Wavelet order	Maximal SNR (dB)		Optimal decomposition level	
	db family	sym family	db family	sym family
1	9.38	–	4	–
2	11.54	11.54	4	4
3	12.08	12.08	4	4
4	12.42	12.14	5	5
5	12.41	12.36	5	5
6	12.83	12.70	5	5
7	12.88	13.03	5	6
8	13.51	12.84	5	5
9	12.81	12.77	5	5
10	13.07	12.92	5	5
11	13.05	12.75	5	9
12	13.39	12.51	5	5
13	12.93	13.02	5	5
14	13.10	12.47	5	5
15	13.13	12.91	5	5
16	13.45	12.64	5	5
17	12.97	12.57	5	5
18	12.97	12.84	5	6
19	13.19	12.86	5	5
20	13.49	12.74	5	5

wavelength range permits access to the R(11)–R(15) transitions in the CO fundamental vibration-rotation band. The CO absorption line R(12) near 2190  $\text{cm}^{-1}$  was chosen for the present experiments, because it exhibits the smallest overlap with absorption lines from other molecules, particularly  $\text{N}_2\text{O}$ . For this wavelength, the laser was operated at  $\sim 276$  K.

Wavelength modulation spectroscopy with 2nd harmonic detection technique was employed to improve the detection sensitivity. To acquire WMS signals, the combination of a low-frequency triangle ramp (12.5 Hz) and a high-frequency sinusoidal modulation (25 kHz) was supplied to the QCL as an addition to the injection current and the 2nd harmonic signals were demodulated using a digital lock-in amplifier programmed with Labview. A total number of 128 sampling points were acquired for each spectrum, averaged from 11 sequential laser scans with approximately 1-s integration time to improve the SNR. To determine the ambient CO concentration, a multi-linear least-square fitting routine was employed [19]. In this procedure, ambient air spectra are fitted to a known reference spectra obtained from a gas standard that is periodically added (typical at 1-h interval) to the sample cell.

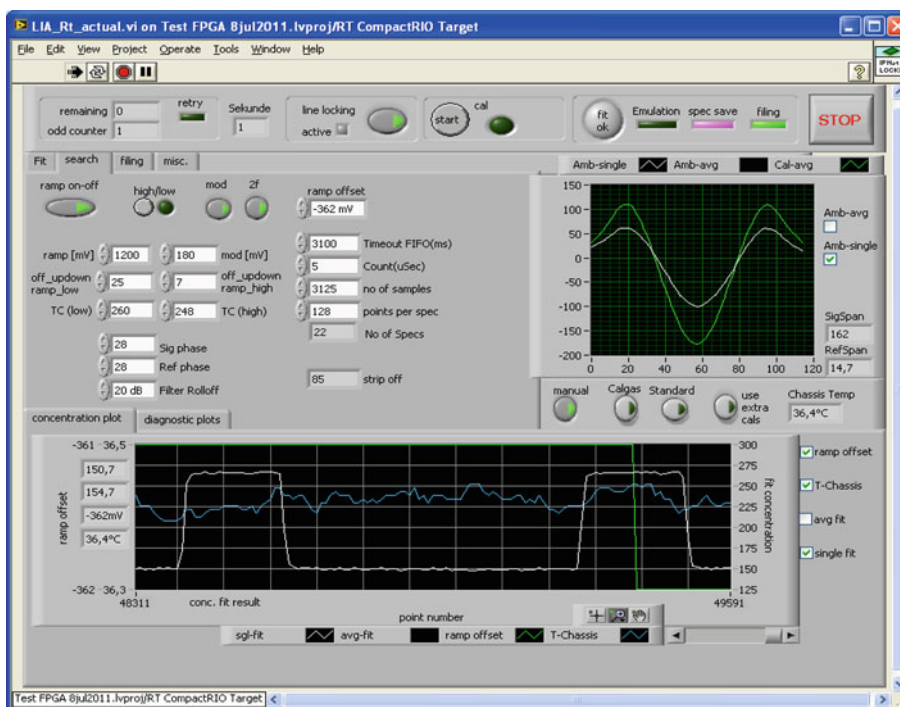
Data acquisition card, power supplies and the electronic elements controlled by National Instruments (NI) field-programmable gate array (FPGA) board are mounted in a standard 19-inch rack housed below the main optical breadboard. Control of all electronics, signal acquisition and real-time signal processing is completely implemented through a LabView-based graphical user interface software program using a laptop linked via a local area network (shown in Fig. 6). Further details of the optical setups, electronics design and sampling line have been described in previous publications due to the use of identical or modified components [20–23].

**Fig. 5** Photograph of the prototype QCL spectrometer





**Fig. 6** Front panel interface of QCLS control system based on the Labview software

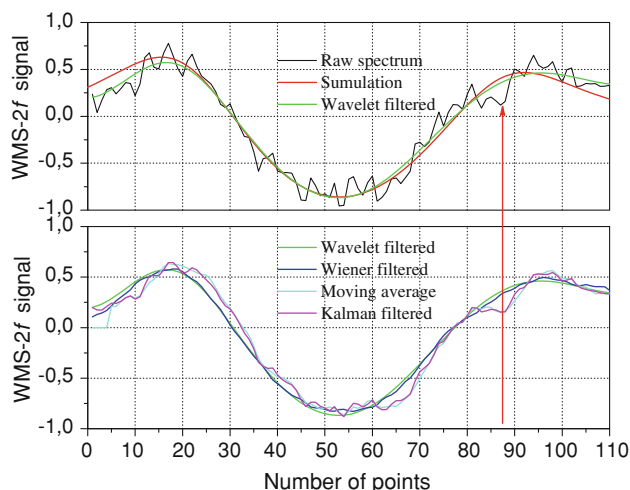


## 4.2 Results and discussion

It is well known, that all signals obtained from any analytical apparatus are often corrupted by noise obtained in the processes of data acquisition and transmission. The noise will degrade the accuracy and precision of the data, and it also affects the detection limit of the instrumental technique. Signal denoising is therefore highly necessary to improve the instrument sensitivity, detection limit, measurement precision and accuracy, in particular at low ambient concentration level. The goal of signal denoising is to remove the noise while retaining as much as possible the important signal features. In this study, we applied the wavelet-based denoising technique to our QCL spectrometer in post-signal processing to achieve two objectives: improving the SNR of the raw spectra and diminishing the dispersion of the final concentration values, in order to improve the detection limit, measurement precision and accuracy.

### 4.2.1 Application of wavelet to QCL spectral signal

We evaluate the wavelet-based denoising technique by applying it to actually recorded spectra. A series of WMS-2f spectra recorded for very lower CO concentrations (several ppbv) were measured by diluting a primary standard with a certified mixing ratio (Scott Marrin Specialty Gases, Inc., CA) with CO-free zero air obtained from a home-made CO scrubber, consisting of SOFNOCAT 423 (Molecular Products, Thaxted, Essex, UK), a hydrophilic



**Fig. 7** Typical example of experimentally determined CO 2nd harmonic spectrum of R(12) transition at  $2190.01 \text{ cm}^{-1}$  recorded with QCL spectrometer and corresponding simulated spectrum (*upper panel*), as well as the results by applying different filter techniques (*lower panel*) (for details of comparison, see text)

CO oxidizing agent, that scrubs CO from ambient air. Spectra from this CO-free zero air are also used to determine the instrument background and can be subtracted from the sample spectra to improve the system performance. According to the discussion of the simulated spectra mentioned above, wavelet db8 was used to study the improvement of the detection limit after denoising. Figure 7 presents an experimental spectral signal after background subtraction with a CO concentration of



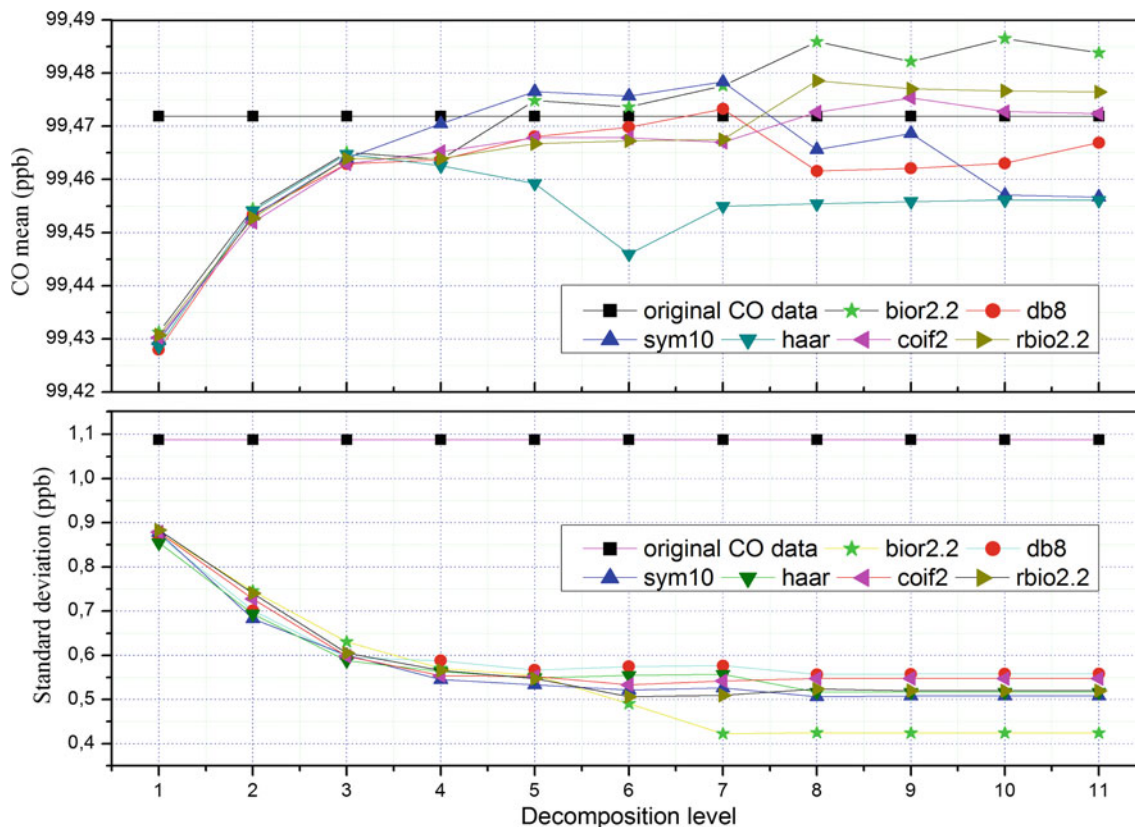
3.69 ppb in air. For comparison, the simulated signal according to the harmonic theory is also given. Note that the slight asymmetric lineshape in both wings is potentially due to the residual amplitude modulation effect, which has been considered during the spectral simulation.

As can be seen, after the application of the wavelet filter, the noisy signal becomes clearer and smoother. Also note, that the wavelet de-noised signal keeps its original spectral features without any distortion relative to the simulated signal. Only the SNR is significantly enhanced. In addition, the performance of the wavelet filter was compared to other classic filter techniques, such as Wiener filter, moving average and Kalman filter. The results show similar behavior for Kalman filter and moving average, as shown in the figure (lower panel). As noted in Sect. 1, the Kalman filter is generally used for forecasting trace gas concentrations but is not well suitable for application to denoising of absorption spectra. Therefore, it is not surprising that the performance of the Kalman filter to improve the detection sensitivity is not perfect here. As for the moving average, because of the limit of total sampling points in the spectral signal, the optimal number of points in the moving average window was set to 5 points. A larger moving window will degrade the approximation degree between the raw and the

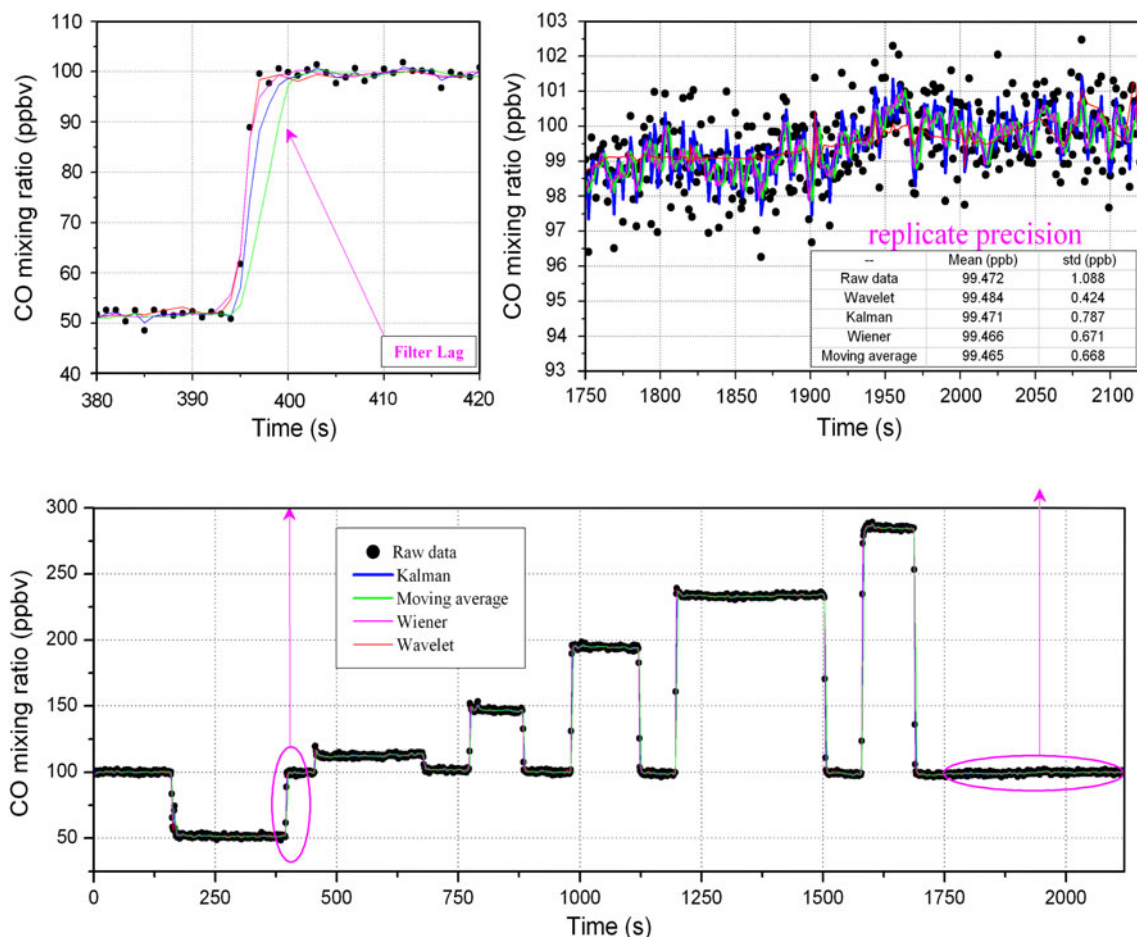
filtered signal. The Wiener filter presents a better performance compared to Kalman filter and moving average. However, it is obvious that the wavelet filter shows the best result and keeps the maximum approximation degree. It greatly improves the spectral SNR yielding a detection sensitivity enhancement by a factor of at least a factor of 3.

#### 4.2.2 Application of wavelet filter to QCL concentration signal

In TDLAS, a high measurement precision can in principle be achieved with longer integration times for signal averaging. The optimal stability time before instrument drifts influence the averaging can be determined by the Allan variance technique [1]. However, the multi-signal averaging technique is time consuming and thus unsuitable for some special applications like eddy correlation flux measurement, which requires a high temporal resolution. In field deployments, the measurement precision will additionally depend upon vibrations and the thermal stability of the instrument, as well as the duty cycle of additions of zero air or calibration gas to compensate for instrument drift. A more meaningful evaluation of the instrument performance is the so-called replicate precision, which can be



**Fig. 8** Comparison of mean CO concentration (*top panel*) and standard deviation (std) (*bottom panel*) obtained by applying different wavelet as a function of decomposition level to the time series of CO concentration signal



**Fig. 9** Continuous measurements of CO from a series of standard tanks with 1-s sampling rate under lab conditions and comparison of various filter results (lower panel). A typical zoom-in on the CO concentration measurements between 380 and 420 s for clear

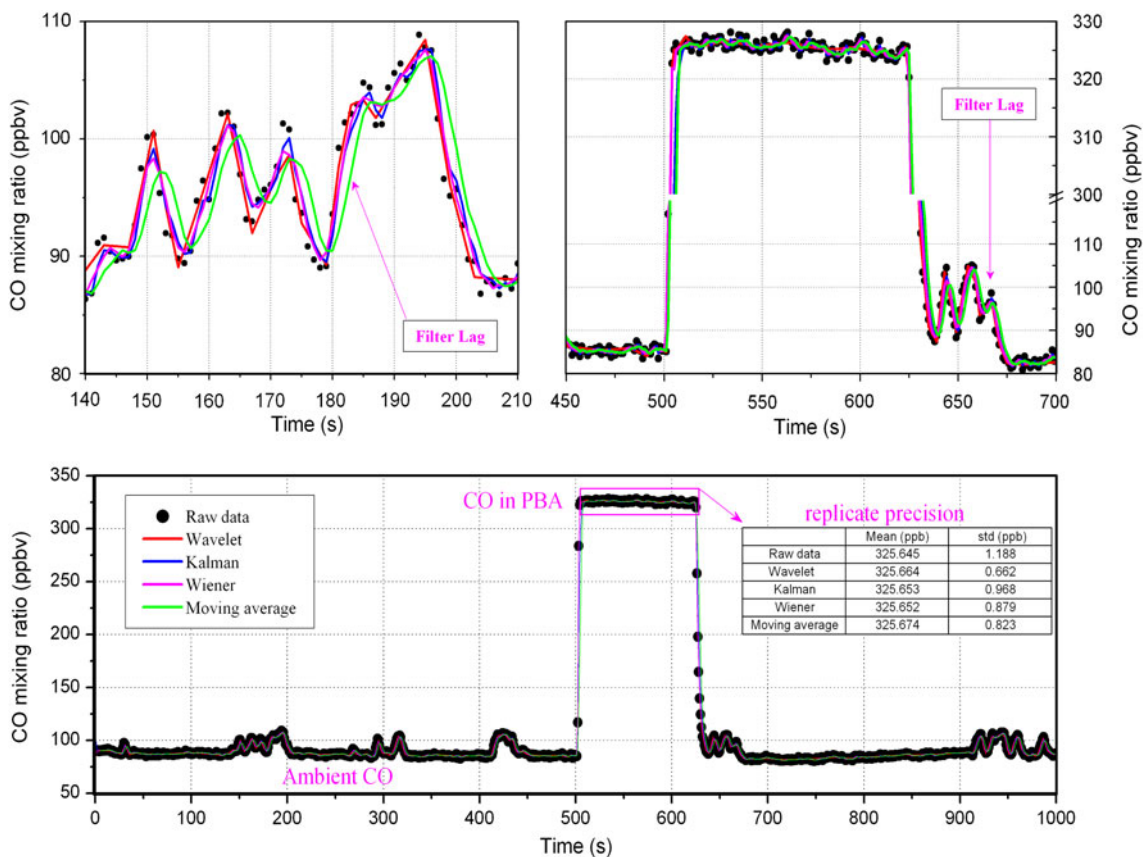
comparison of filter lag effect (left upper panel), and replicate precisions calculated from data between 1750 and 2120 s are shown in the right upper panel

obtained from the standard deviation of replicate measurements of a reference sample with constant concentration.

Indeed, the time series of gas concentration data can also be considered a noisy signal, with noises induced by instrument noises (optically and electrically), instability and environmental factors that the wavelet denoising applied previously to the raw spectral signals could not eliminate completely. In the next evaluation test, the wavelet filter was further applied to minimize the dispersion of the gas concentration values, in order to improve the instrument measurement precision. As discussed above, the wavelet basis should be carefully adopted depending upon the signal characteristics for optimal data denoising. However, in case of practical gas concentration data, any simulation is no longer possible. In order to obtain an ideal wavelet basis for this kind of application, a similar procedure with the test of wavelets to spectral signals in the simulation section has also been used for evaluating various mother wavelets on the experimental time series of gas

concentration data. Herein, the selection criterion is the measurement precision, which was defined as the standard deviation (std) for a range of CO concentrations which were rather constant.

For clarity, we only present results obtained with some typical wavelets from different wavelet families, as shown in Fig. 8. The raw experimental CO data used for the wavelet test and the data range (between 1750 and 2120 s) selected for calculating the measurement precisions are shown in Fig. 9 (data details see next paragraph). From this figure, we can see that the mean CO value after the application of wavelet filtering is approaching the “real value” of about 100 ( $\pm 10\%$ ) ppb (manufacturer provided), while the std is decreasing with the increasing of decomposition level (i.e. measurement precisions is improving), and retains almost constant after the optimal decomposition level (between 6 and 8 depending on the wavelet basis) is reached. This varying trend is similar to the case of applying wavelets to the spectral data as discussed previously. However, here

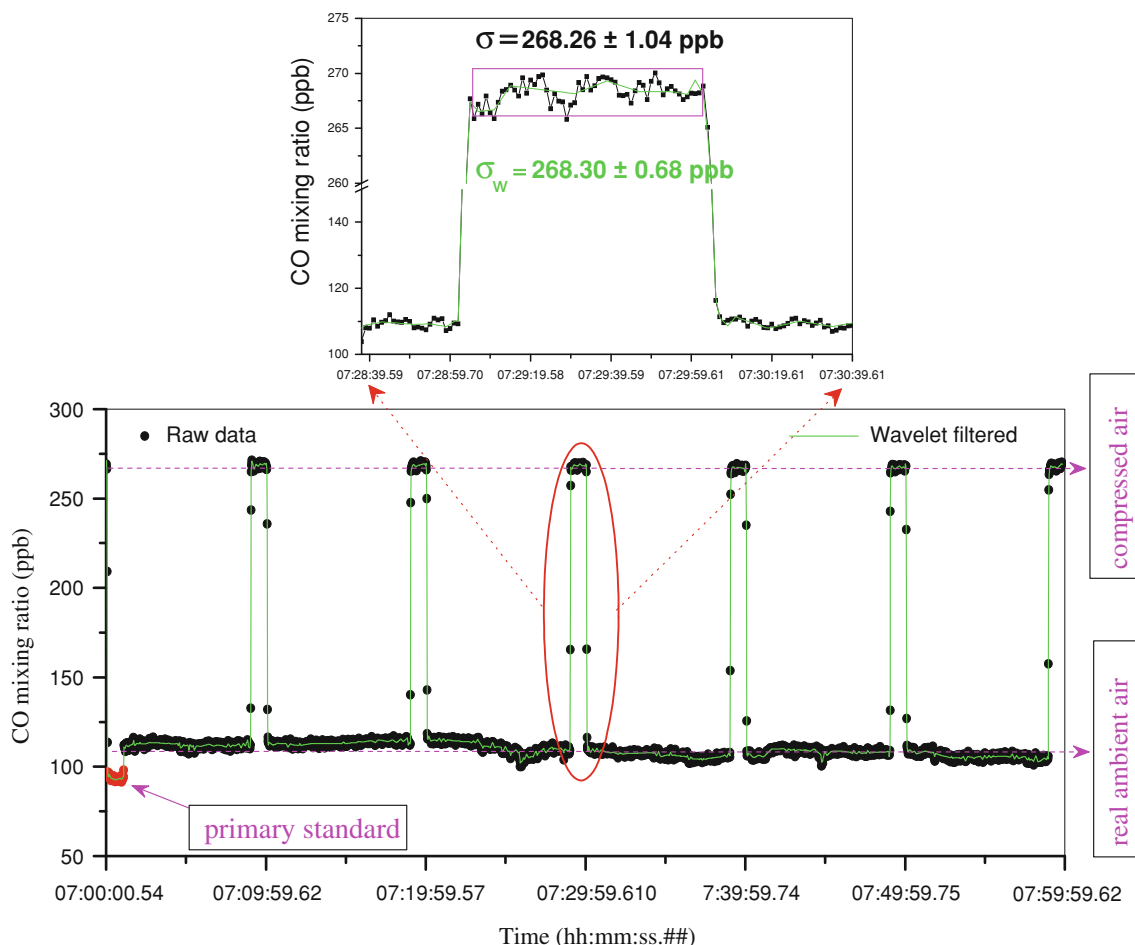


**Fig. 10** Pre-field CO concentration measurements in Mainz for a 17-min data collection. A typical zoom-in on the CO concentration measurement data for clear comparison (*both upper panels*), and replicate precisions are *inset (lower panel)*

wavelet bior2.2 shows a better performance at the optimal decomposition level for the time series of concentration data. This may be due to the different noise characteristics and the correlation between mother wavelet and the analyzed signal, since wavelet denoising is a data-dependent process. In addition, from Fig. 8, we can see that the difference in the std between the optimal decomposition level and the maximum is very small. Taking wavelet bior2.2 as an example, the discrepancy between level 7 and 11 is only  $1.6 \times 10^{-3}$ . Therefore, the maximal decomposition level  $n$  (depending on the signal length of  $2^n$ ) and Stein thresholding scheme were widely used in the following studies.

On the basis of analysis above, a subsequent experiment involved applying the best wavelet candidate (i.e. bior2.2) to various concentration data with different features recorded under different experimental conditions. Similar to the analysis of experimental spectral data, a comparison with other widely used filter techniques was also investigated. First, the performance of the wavelet filter was studied over a significant dynamic concentration range by allowing CO mixing ratios to change by a factor of about 6, between 50 and 300 ppb, as shown in Fig. 9. Another distinct feature of this dataset is the rapid changes between

different CO concentration levels, which are useful to evaluate the adaptability of the filter techniques. This experiment included 2,120 concentration measurements obtained with a 1-Hz sampling rate under laboratory conditions, using a series of primary standard tanks (Scott Marrin Specialty Gases, Inc., CA). From Fig. 9, we can see that the “T” shape concentration increase becomes smoother after application of the wavelet filter compared to the raw data. This indicates that small fluctuations due to optical fringe instability, laser frequency or laser power drifts are efficiently reduced through the adaptive filter. However, the true information of the concentration change is conserved and the features of the sensor’s response to variations are still reflected in the filtered data. Replicate precision evaluated from the raw concentration data and various filtered results are shown in the inset in the right upper panel. These results show similar behavior. By comparing the replicate precisions, we can see that wavelet filter provides a better performance than other filter techniques. Moreover, it indicates that a slight delay is introduced by the 5-point moving average (left upper panel). The wavelet filter is not sensitive to this kind of quick changing effects.



**Fig. 11** In-field CO concentration measurements at the Taunus Observatory in Kleiner Feldberg for a typical 1-h interval and the application of wavelet filter (*lower panel*). *Upper panel* presents a

zoom-in on an approximate 2-min interval for clarity, and an approximate 1-min replicate precision evaluated from the compressed air

A pre-field test of the QCL spectrometer was performed at the outside of MPI-C main building in Mainz, as described in Ref. [18]. Figure 10 shows an example of applying wavelet filter to the pre-field results. The high CO mixing ratios are periodically sampled from pressurized bottled air (PBA, 325 ppbv traced to a primary gas standard), which was used to evaluate the system performance. The lower trace near 100 ppbv shows the real-time ambient CO mixing ratios. As can be seen, the observed atmospheric CO concentrations vary a lot, due to changes in meteorological parameters, for example, wind speed and direction, or local emissions. It is mandatory, that the application of any digital filtering conserves the characteristic of the high temporal resolution. From the figure it is clear, that all filter techniques show a similar behavior, except the time lags introduced by the 5-point moving average. In general, the wavelet filter provides a better measurement precision. A long-term (66 h) precision of 1.41 ppbv was deduced from the original data of the

pre-field measurements, improving to a better precision of 0.88 ppbv after utilization of the wavelet filtering.

From these three figures above, one can see that wavelet filter can not only be used for spectral signal denoising but is also effective for concentration estimating. The results show that the wavelet filter has a superior performance in improving detection sensitivity and increasing measurement precision compared to either Kalman filter, Wiener filter or moving average, both visually and quantitatively.

Primary analysis from a recent field campaign at the Taunus Observatory on the summit of the Kleiner Feldberg [24] confirms that this completely TE-cooled system is capable of long term, unattended and continuous operation at RT without cryogenic cooling of either laser or detector. In a similar measuring procedure as shown in Fig. 11, the primary CO standard is used for instrumental calibration, while the continuous measurements of the commercial compressed air are used to evaluate the instrument performance (precision, stability). Finally, the field measurement



precision of 1.55 ppb improves to 0.74 ppb after applying the wavelet filter technique during a nearly 40-day measurement period. Data analysis is still ongoing and will be discussed in a further publication.

## 5 Conclusions and outlook

We presented an application of wavelet transform-based denoising technique to quantum cascade laser spectrometer for in situ and real-time atmospheric trace gas measurements. With the utilization of the wavelet digital filter technique in post-signal processing, better measurement precision and higher detection sensitivity have been achieved without reducing the fast temporal response. A procedure to optimize parameters used in signal denoising in the wavelet domain was presented. In order to assess the benefit of the wavelet filter, we compared the wavelet-based filter to other commonly used digital filter techniques, i.e. Kalman filter, Wiener filter and moving average, by application to actual observations. Of the four methods studied, wavelet filtering demonstrated a higher ability to considerably improve detection sensitivity and increase measurement precision and accuracy. The method implemented here largely depends on wavelet parameters, for example, wavelet type, thresholding policy, threshold estimation and decomposition level, etc. To achieve a better performance in practical applications, wavelet parameters should be optimized according to the characteristics of signals to be de-noised.

Wavelets are becoming an increasingly important analysis tool for signal processing. Wavelets can effectively extract both time and frequency-like information from a time-varying signal. Given the powerful and flexible multi-resolution decomposition, the linear and non-linear processing of signals in the wavelet transform domain offer new methods for infrared laser spectra denoising and trace gas concentration evaluation. The adaptive Kalman filter is a popular tool in the discipline of environmental science for predicting the concentration of trace gases in real time due to its capability of making a more accurate prediction by minimum-variance estimations. However, the Kalman filter is usually applied with some prior assumption on the variances of the process noise and the measurement noise, which are difficult to be obtained in practice. Both wavelets and Kalman filters can handle non-stationary signals. Therefore, the wavelet coefficients can be modeled as the state variables of Kalman filters, using the recursive Kalman filter algorithm to best estimate the wavelet coefficients. Thus, a more promising technique based on incorporating both wavelet transform and Kalman filter (i.e. wavelet-based Kalman filter) can be more effectively applied on trace gas sensors. Recently, the so-called wavelet-based

multi-model Kalman filter has been successfully applied in the field of hydrology for food forecasting and electric power systems [25, and references therein]. Future research work will focus on this proposed method, to implement this method to our QCLS system for various field applications.

**Acknowledgment** The authors gratefully acknowledge the anonymous reviewers and editors for their valuable comments and suggestions to improve the quality of the paper. The authors would like to thank Dieter Scharffe for providing CO scrubber and primary standards for instrumental calibration.

## References

1. P. Werle, R. Mücke, F. Slemr, *Appl. Phys. B* **57**, 131 (1993)
2. P. Werle, B. Scheumann, J. Schandl, *Opt. Eng.* **33**, 3093 (1994)
3. H. Riris, C.B. Carlisle, R. Warren, *Appl. Opt.* **33**, 5506 (1994)
4. T. Wu, W. Chen, E. Kerstel, E. Fertein, X. Gao, J. Koeth, K. Rößner, D. Brückner, *Opt. Lett.* **35**, 634 (2010)
5. D.P. Leleux, R. Claps, W. Chen, F.K. Tittel, T.L. Harman, *Appl. Phys. B* **74**, 85 (2002)
6. A. Tangborn, S.Q. Zhang, *Appl. Numer. Math.* **33**, 307 (2000)
7. S. Postalcioğlu, K. Erkan, E.D. Bolat, in *Proceedings of the International Conference on Neural Networks and Brain*, vol 2 (Beijing, 2005), pp. 951–954
8. B.K. Alsberg, A.M. Woodward, M.K. Winson, J. Rowland, D.B. Kell, *Analyst* **122**, 645 (1997)
9. R.V. Tona, C. Benqlilou, A. Espuña, L. Puigjaner, *Ind. Eng. Chem. Res.* **44**, 4323 (2005)
10. S. Yin, W. Wang, *Proc. SPIE* **6367**, 63670T (2006)
11. S. Yousefi, I. Weinreich, D. Reinartz, *Chaos Solitons Fract.* **25**, 265 (2005)
12. J. Chun, J. Chun, *Signal Process* **80**, 441 (2000)
13. S.G. Mallat, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 674 (1989)
14. D.L. Donoho, I.M. Johnstone, *J. American Statist. Ass.* **90**, 1200 (1995)
15. X.-P. Zhang, M.D. Desai, *IEEE Signal Process Lett.* **5**, 265 (1998)
16. L.S. Rothman, I.E. Gordon, A. Barbe, D.C. Benner, P.F. Bernath, M. Birk, V. Boudon, L.R. Brown, A. Campargue, J.P. Champion, *J. Quant. Spectrosc. Radiat. Transf.* **110**, 533 (2009)
17. L. Pasti, B. Walczak, D.L. Massart, P. Reschiglian, *Chemom. Intell. Lab. Syst.* **48**, 21 (1999)
18. J.S. Li, U. Parchatka, R. Königstedt, H. Fischer, *Opt. Express* **20**, 7590 (2012)
19. P. Werle, P. Mazzinghi, F. D'Amato, M. De Rosa, K. Maurer, F. Slemr, *Spectrochim. Acta A* **60**, 1685 (2004)
20. R. Kormann, H. Fischer, F.G. Wienhold, *Proc. SPIE* **3758**, 162 (1999)
21. F.G. Wienhold, H. Fischer, P. Hoor, V. Wagner, R. Königstedt, G.W. Harris, J. Anders, R. Grisar, M. Knothe, W.J. Riedel, F.-J. Lübken, T. Schilling, *Appl. Phys. B* **67**, 411 (1998)
22. R. Kormann, R. Königstedt, U. Parchatka, J. Lelieveld, H. Fischer, *Rev. Sci. Instrum.* **76**, 075102 (2005)
23. C.L. Schiller, H. Bozem, C. Gurk, U. Parchatka, R. Königstedt, G.W. Harris, J. Lelieveld, H. Fischer, *Appl. Phys. B* **92**, 419 (2008)
24. J.N. Crowley, G. Schuster, N. Pouvesle, U. Parchatka, H. Fischer, B. Bonn, H. Bingemer, J. Lelieveld, *Atmos. Chem. Phys.* **10**, 2795 (2010)
25. Ch-M Chou, R.-Y. Wang, *Hydrol. Process* **18**, 987 (2004)