Theoretical investigation of applicability and limitations of advanced noise reduction methods for wavelength modulation spectroscopy

Lenard L. Röder[1] · **Horst Fischer**[1]

Abstract

In this study, the applicability and limitations of several statistical and mathematical methods for noise reduction in wavelength modulation spectroscopy are analyzed. Background noise is simulated for different frequencies and frequency confinement. The noise is added to an absorption line of varying amplitude. The noise reduction methods (NRMs) are applied to the simulated signals and their performances are analyzed and compared. All NRMs show great increase in signal to noise ratio (SNR) while keeping bias low under certain conditions of the simulated signal. For each NRM the subspace of best performance is summarized and highlighted. Little to no overlap is found between these subspaces. Therefore, the optimal NRM strongly depends on measurement conditions and NRM quality cannot be compared in a general context.

1 Introduction

Tunable diode laser or quantum cascade laser absorption spectroscopy (TDLAS/QCLAS) using wavelength modulation spectroscopy (WMS) have become major tools for molecule detection in various disciplines due to their high flexibility, time resolution, sensitivity, precision and selectivity. Main applications are atmospheric measurements of trace gases and combustion process controls in industry [1–3]. Advancements in laser technology like the quantum cascade laser (QCL) [4] and the implementation of multi pass absorption cells [5, 6] have lead to major improvements of the technique in the last decades. The main limitations for precision and minimum detectable mixing ratios are laser disturbances, power or frequency drifts and optical noise patterns [7–9]. These noise patterns often have an underlying dominant frequency and a varying phase. The frequency can lie in the same region as the absorption signal, depending on the optical setup.

Many new data analysis techniques for noise reduction have been adapted to absorption spectroscopy experiments during the last decades and were reviewed several times in a qualitative manner [9, 10]. These methods range from signal analysis and denoising via time–frequency transformations [11, 12] to adaptive statistical methods [13, 14]. However, most recent publications of state-of-the-art experiments still rely on basic data analysis established in the last century [3, 15]. This may originate from the specific design of these analysis methods for their target experiment; thus, leaving questions unanswered about the experimental requirements and limitations. This paper attempts to compare a few prominent and also lesser known noise reduction methods (NRM) via simulation of a wide array of possible noise shapes. The goal to reach is a quantitative statement about the experimental conditions necessary for the given NRM to increase the signal-to-noise-ratio (SNR) while simultaneously maintaining a low bias.

This paper is organized as follows: In Sect. 2, an overview of the analyzed NRMs is given and small modifications to ensure compatibility are explained. In Sect. 3, the simulation procedure is shown. Finally the simulation results are analyzed and interpreted.

2 Noise reduction methods

First, the theory of WMS is described in short. According to the Lambert–Beer law the transmitted intensity *I* through an absorbing medium for small optical density can be approximated by

✉ Lenard L. Röder
lenard.roeder@mpic.de

1 Department of Atmospheric Chemistry, Max Planck Institute for Chemistry, Hahn-Meitner-Weg 1, 55128 Mainz, Germany
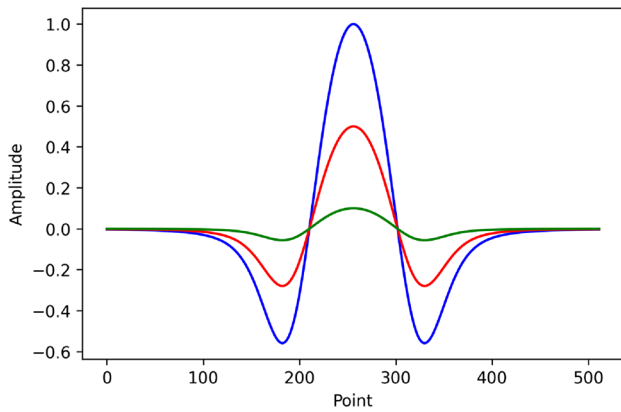
Fig. 1 Reference 2*f* wavelength modulated spectrum of Voigt profile of carbon monoxide transition at at wave number 2190.02 cm^{-1} for a pressure of 50hPa, a temperature of 293K and modulation depth 2.2· HWHM, scaled to a maximum amplitude of 1 (blue), 0.5 (red) and 0.1 (green)

$$I(\omega) \approx I_0(1 - \alpha(\omega)), \tag{1}$$

where I_0 is the transmitted intensity without absorption and $\alpha(\omega)$ is the absorption profile. α approaches a Gaussian profile for low pressure and a Lorentzian or Cauchy profile for high pressure. In intermediate cases the profile can be expressed as a convolution of both, a so called Voigt profile:

$$\alpha_V(\omega) = \alpha_G(\omega) * \alpha_L(\omega)$$

$$\sim \exp\left(-\frac{(\omega - \omega_0)^2}{2\sigma^2}\right) * \frac{1}{1 + \left(\frac{\omega - \omega_0}{\gamma}\right)^2}. \tag{2}$$

In WMS the frequency ω is modulated with a modulation depth a and modulation frequency Ω and the signal is subsequently demodulated using an analog or digital lock-in amplifier set to an integer multiple of the modulation frequency $n\Omega$. Therefore, the retrieved signal S is proportional to the nth harmonic H_n of the cosine series of the absorption profile:

$$S \sim H_n(\omega) = \int \alpha(\omega + a\cos(\Omega t))\cos(n\Omega t)\mathrm{d}t. \tag{3}$$

More detailed derivations have been given several times in the literature and can be found, for example, in [2, 7, 10]. In the course of this paper the second harmonic is used as most methods considered in this study were originally designed for second harmonic detection. It is depicted in Fig. 1. In principle this study can be conducted using other harmonics, normalization schemes [3] or direct absorption, as well.

2.1 Wiener filter

Werle et al. [13] have described the implementation of an adaptive finite impulse response filter (FIR), the Wiener filter. It is calculated for each measurement spectrum X by a least squares algorithm, minimizing the mean square error between the measurement spectrum and a synthesized version by convolution of the filter with a calibration spectrum C. Thus, knowledge about the target signal is used to optimize the FIR:

$$X_i = \sum_{k=-m}^{m} \gamma_k C_{i+k}. \tag{4}$$

A free parameter of this method is the filter size m. The appropriate value has to be chosen by a measurement of known mixing ratio.

2.2 Singular value decomposition

A simple statistical approach was proposed by Mappe-Fogaing et al. [14] that assumes a stable and dominant absorption signal compromised by an unstable background. By performing a singular value decomposition (SVD) on a batch of measured spectra the signal is split into eigenvectors. Under the right circumstances some of the eigenvectors contain signal information, while the rest contains only noise.

The singular value decomposition decomposes a matrix X with dimension $N \times M$ into the basis of normalized wavelets V, corresponding singular values Σ and the propagation matrix U:

$$X = U\Sigma V^T. \tag{5}$$

In their paper a decent initial signal to noise ratio is given so the signal information is stored in the first eigenvectors. They choose the appropriate effective dimension of the data set empirically. Naturally this becomes a very difficult task when SNR is low. To adapt to the simulation conducted in this paper, Pearson correlation coefficients between each normalized wavelet and a reference spectrum are calculated and only the first m eigenvectors, sorted by their correlation coefficients, are considered during reconstruction. This procedure extends the empirical approach from the original paper but requires optimization of the free parameter m. The dimension has to be chosen during the calibration phase.

2.3 Discrete wavelet transform

Denoising techniques based on the discrete wavelet transform (DWT) have become widely used in signal processing, feature extraction and image denoising [16, 17]. The continuous wavelet transform is a time–frequency transformation similar to the Fourier transform [18]. It is defined by a convolution of the signal x with modified versions of a special function called *mother wavelet ψ*, that is scaled by the parameter a and translated by the parameter b:

$$\mathcal{W}(a,b) = \int \psi^* \left(\frac{t-b}{a} \right) x(t) \mathrm{d}t. \tag{6}$$

Employing the redundancy of this procedure, Daubechies [19] carefully designed special wavelets and evaluated only discrete scales. This way the transform no longer has a two-dimensional result but has the same dimension as the input. With Mallats algorithm this procedure now becomes an iterative convolution with discrete filter banks to decompose the signal into approximation and detail coefficients [20]. Based on this representation of the original signal several studies have shown applications for removal of trends and suppression of noise for a given signal, also for TDLAS [10, 21, 22]. A study by Li et al. [11] has successfully applied the noise reduction based on Steins unbiased risk estimator (SURE) [23] with the non-negative Garrote threshold [24] on WMS. In their study they compared several wavelets and maximum decomposition levels and retrieved the optimal choice for their experimental setup. Hence, the choice of wavelet and maximum decomposition level is a free parameter and has to be chosen during calibration.

2.4 Fourier domain analysis

Another approach utilizing time–frequency transformation has been proposed by Hartmann et al. [12]. They take advantage of the linear decrease of the logarithmic power spectrum of a Cauchy profile and interpolate discrete disturbances to this slope caused by sinusoidal background, effectively removing the fringe patterns. They emphasize, however, that the technique required fine tuning of signal and noise for this procedure to work that easily.

A number of adjustments have to be made for this method to be used fully automatically on second harmonic Voigt profiles. The absorbance of a measured signal is calculated by performing a masked linear fit of the power spectrum of the measured signal to the power spectrum of a second harmonic Voigt, acquiring the parameters from a reference spectrum. Alternatively it can be fit directly to the power spectrum of the measured reference. The mask is a free parameter with left and right boundary index.

Depending on the background structure the background is more or less confined in the frequency domain. This way the mask would be able to ignore points in the power spectrum influenced by the background and only consider noise free parts of the signal power spectrum.

Regarding the emphasis on the special requirements by the authors it can be seen here that improper choice of the mask or a background that is not at all confined to a small bandwidth can lead to unforeseen consequences if this method is applied blindly.

2.5 Empirical mode decomposition

The empirical mode decomposition (EMD) is an adaptive decomposition scheme that splits a signal into its intrinsic mode functions (IMF) with the same number of zero crossings as extreme points [25]. Meng et al. [26] have applied Savitzky–Golay filters to the IMFs and reconstructed the signal using the Pearson correlation coefficients of the filtered IMF and the original signal as a weight. They have shown in their study that this method called EMD-FCR (Filter-Correlation-Reconstrunction) outperforms Wiener filters and DWT noise reduction for their particular experimental setup. A drawback of this method is the high computational effort. In addition, the order and window size of the Savitzky–Golay filter have to be optimized during the calibration phase as they are free parameters.

The FCR part was modified slightly for this simulation, since the Pearson correlation with the original signal enhances noise instead of suppressing it in cases of very low SNR. As a modification the correlation coefficients are evaluated against a reference spectrum instead, similar to the SVD method.

2.6 Artificial neural network

A feed-forward neural network consists of layers of affine transformations that are separated by a non-linear function called activation. This can be expressed by a matrix multiplication of the input x_i with a weight Matrix W_i and addition of a bias vector b_i, where i is the layer number. The result of this operation is then pointwise evaluated using the activation function σ:

$$y_i = \sigma(W_i x_i + b_i). \tag{7}$$

During the training phase the output is compared to the target value via a metric called loss function. The weight matrix parameters $w_{i,jk}$ and biases $b_{i,k}$ are optimized by propagation of the gradient of the loss backwards through the network. Many sophisticated improvements to model design, loss functions, optimization, learning and

prevention of overfitting have been made in the last few years. This paper will not cover any more details on this topic, since this is not the main focus of this study. Here we refer to reviews and books about this field [27, 28].

A simple artificial neural network has been proposed by Nicely et al. [29] as a proof of concept for optical fringe reduction. The model design of the authors is kept very simple as they emphasize their goal being only to proof the concept of applicability. Clearly the model is not optimally chosen for the given task and does not learn very fast.

For the simulation in this study eight identical networks were trained on different subspaces of the target simulation and the choice of one specific network is set as a free parameter to be optimized during calibration. Each model is trained on synthetic data for 250 epochs with 2^{14} iterations each.

3 Simulation

For the simulation a measurement scheme is chosen that mimics real experimental conditions and allows for optimization of the free parameters of the NRMs. The measurement scheme requires three different kinds of input data:

- *Reference* A noise-free absorption spectrum. In experiments this could be either a very high concentration measurement or a simulation using the HITRAN database [30]. However, deviations of the used spectrum from the real underlying noise-free spectrum of the experiment can lead to undesired effects. In the simulation the reference is used to initialize some of the NRMs and also to calculate the absorbance.
- *Calibration* A noisy spectrum with known mixing ratio. This could be a measurement used for instrument characterization or a calibration during field experiments.

This data will be used to optimize free parameters of the NRMs for the given background signal.

- *Ambient* Target noisy spectrum to be denoised.

As reference a Voigt profile is calculated using data from the HITRAN database [30]. The chosen absorption profile is an idealized absorption spectrum of carbon monoxide at wave number 2190.02 cm^{-1} for a pressure of 50hPa and a temperature of 293K [31]. The number of points is set to $2^9 = 512$, the line is centered at 2^8 and the scale is adjusted so that the Gaussian width σ caused by Doppler Broadening is set to 16. The modulation depth is set to 2.2·HWHM according to the studies of Reid and Labrie [32]. The resulting $2f$ modulated spectrum is numerically calculated using (3) and is shown in Fig. 1 for several amplitudes.

The noisy part of the spectrum $\tilde{n}_{fc}$ is generated as a function of two parameters: frequency f and noise complexity c. The Fourier power spectrum is set to a Gaussian of mean f and width $f \cdot c$ and the positive angles are sampled from a uniform distribution between $-\pi$ and π. The negative frequency angles are set to the negation of the positive angles to acquire a purely real signal after inverse Fourier transform. The complexity parameter c is a measure of the contribution to the noise by neighboring frequencies. For high values of c and f the noise becomes almost white, e.g., $c > 4$ and $f > 512$.

The resulting time signal is then normalized to $\sigma = 1$. Afterwards the reference spectrum is weighted with the third simulation parameter, the amplitude s, and added to the signal. Examples with $s = 0$ are shown in Fig. 2.

$$\tilde{n}_{fc} = \mathcal{F}_{f'}^{-1}(t)\left[\exp\left(\frac{-(f'-f)^2}{2f^2c^2}\right) \cdot e^{i\phi}\right]$$

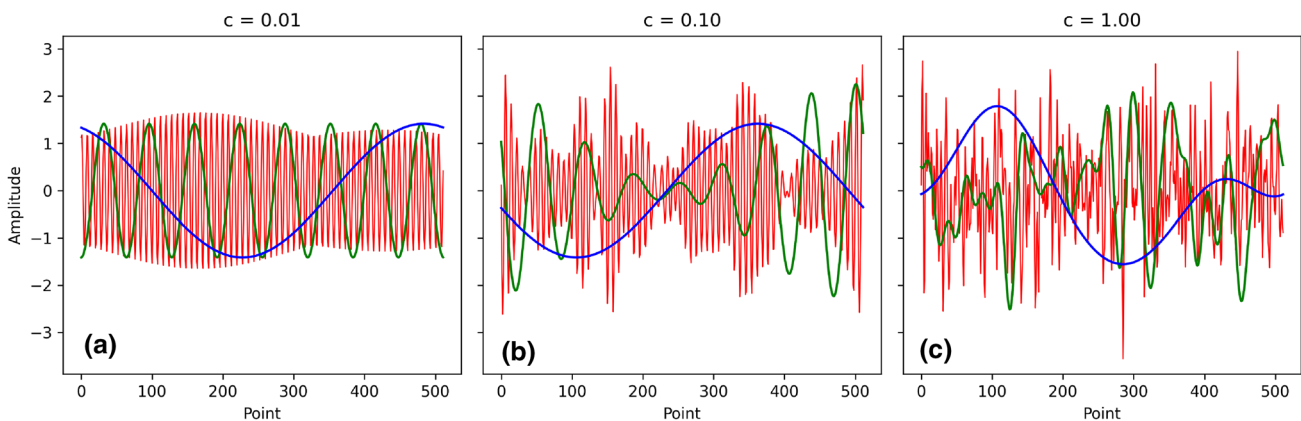$$\text{with } \phi \in \mathcal{U}(-\pi, \pi), \tag{8}$$

Fig. 2 Example spectra for different noise parameters. Complexity c is set to 0.01, 0.1 and 1 in plot **a**– **c**, respectively. In all plots the blue line, green line and red line corresponds to a frequency of 2^0, 2^3 and 2^6, respectively

$$n_{sfc} = \left(\frac{\tilde{n}_{fc}}{\sqrt{\mathrm{Var}\left[\tilde{n}_{fc}\right]}} \right) + s \cdot x. \tag{9}$$

Now, the three parameters are discretized along logarithmic scales to form a 64^3 cube consisting of $2^{18} = 262,144$ individual triplets. The signal amplitude, frequency and complexity range from 0.01 to 100, 2^0 to $M/2 = 2^8$ and 0.01 to 1, respectively. For each triplet 32 calibration spectra are generated for free parameter optimization and 1024 ambient spectra are simulated. A simple grid search is used to optimize the free parameters. In possible future simulations more sophisticated approaches can be applied to speed up the procedure, e.g., some publications propose algorithms to efficiently choose the optimum window of a Savitzky–Golay filter which is used in the EMD-FCR [33, 34]. The NRM is applied to the N ambient spectra using the optimized parameters. The absorbance a_i is retrieved by a final linear fit to the reference spectrum. These absorbances are then compared to the true amplitude s and the mean squared residual (MSR) and also the mean residual squared (MRS) are returned. The MSR and MRS are defined as follows:

$$\mathrm{MSR} = \frac{1}{N} \sum_i \left(s - a_i \right)^2, \tag{10}$$

$$\mathrm{MRS} = \left(s - \frac{1}{N} \sum_i a_i \right)^2. \tag{11}$$

This procedure is repeated for each NRM and also run once without noise reduction as a baseline. This is important, since a least squares linear fit already shows great noise reduction properties under certain background conditions.

4 Simulation results

For more clarity the resulting values from the simulation MSR and MRS can be expressed by two different measures:

$$\Delta S = 10 \log_{10} \left(\frac{\mathrm{MSR}_0}{\mathrm{MSR}} \right), \tag{12}$$

$$\beta = \frac{\sqrt{\mathrm{MRS}}}{s} \cdot 100\%. \tag{13}$$

MSR_0 is the baseline MSR without noise reduction. Now ΔS describes the increase in SNR in the unit of decibels (dB) and the bias β describes the relative deviation of the sample mean from the true value in percent. Note that in extreme

cases the SNR increase can be negative and the bias can reach values above 100%. These performance measures are displayed by individual animations, where each frame shows an image plot of the fc-plane at a certain signal amplitude s which is increased from frame to frame from 0.01 to 100. The animations can be found in the electronic supplementary material. As an example the Fourier Domain Analysis method (FDA) is evaluated explicitly in the following subsection, as it shows a variety of different characteristics. Then the performances of all NRMs is summarized and compared.

4.1 Case study: Fourier domain analysis

The resulting SNR increase and corresponding bias are shown in Fig. 3 for several signal amplitudes s. In the first image at $s = 0.01$ the SNR increase gradually grows from 30dB at highest frequency to about 60dB at a frequency of 2^5 for complexity values below 0.2. There ΔS evolves into a stripe pattern of several overlapping discrete lines. Here ΔS ranges up to over 70dB. Below frequency 2^3 the SNR increase of the stripes drops slightly and the pattern is overlaid with a smooth 40dB region that tapers down at higher complexity. Finally ΔS increases again at lowest frequency.

Meanwhile, β is mostly close to zero for low complexities. The bias in the overlying smooth region at low frequency ranges from 20 to 50%, however. The bias is greatly enhanced at high complexity, especially at low frequency, where it reaches values above 100%.

Obviously at high complexity the part of the Fourier spectrum containing the signal information is dominated by noise. The bias is reduced for low complexity but still significant for low frequencies. Here the noise interferes with the first peak of the Fourier power spectrum. At frequencies above 2^3 the algorithm effectively ignores parts of the power spectrum that differ from the underlying signal. The stripy pattern corresponds to discrete changes to the boundary indices of the mask chosen in the calibration phase. The gradual decrease for higher frequencies probably does not originate from the NRM but from the baseline performance of the least squares linear fit, which acts as a low pass filter and, therefore, performs better on higher frequency noise.

In the second plot at $s = 0.22$ the SNR increase has shrunk uniformly, while the described patterns stay the same. This also follows from the baseline performance increasing with higher signal, hence lowering the ratio between absolute SNR and baseline SNR. The bias starts to tend to zero at high frequency and high complexity. The higher frequencies have lower impact on the method, since all signal information is contained in the lower frequency components.

As displayed in the next image at $s = 4.64$ this process continues. The SNR increase is now constrained to the 2^2

Fig. 3 Sequence of image plots of SNR increase (left) and bias (right) of the FDA algorithm in the *fc*-plane for different signal amplitudes, increasing from top to bottom. Each plot sequence has a shared color map ranging from 0 to 70dB and 0 to 100%, respectively. White pixels in the SNR increase plot indicate points, where the bias is 100% or above. Depending on the experimental requirements the method should only be applied when high SNR increase is achieved while maintaining a low bias
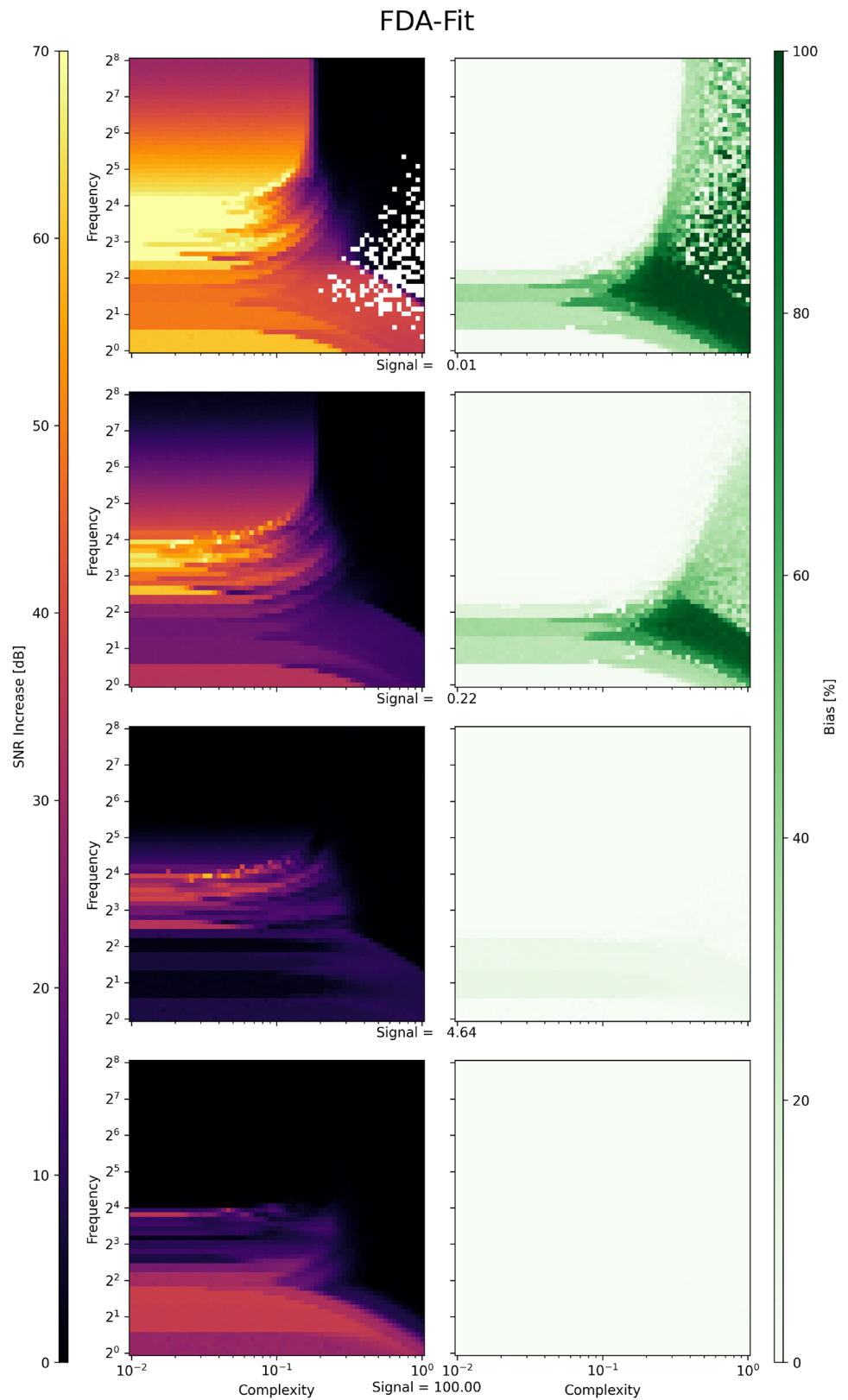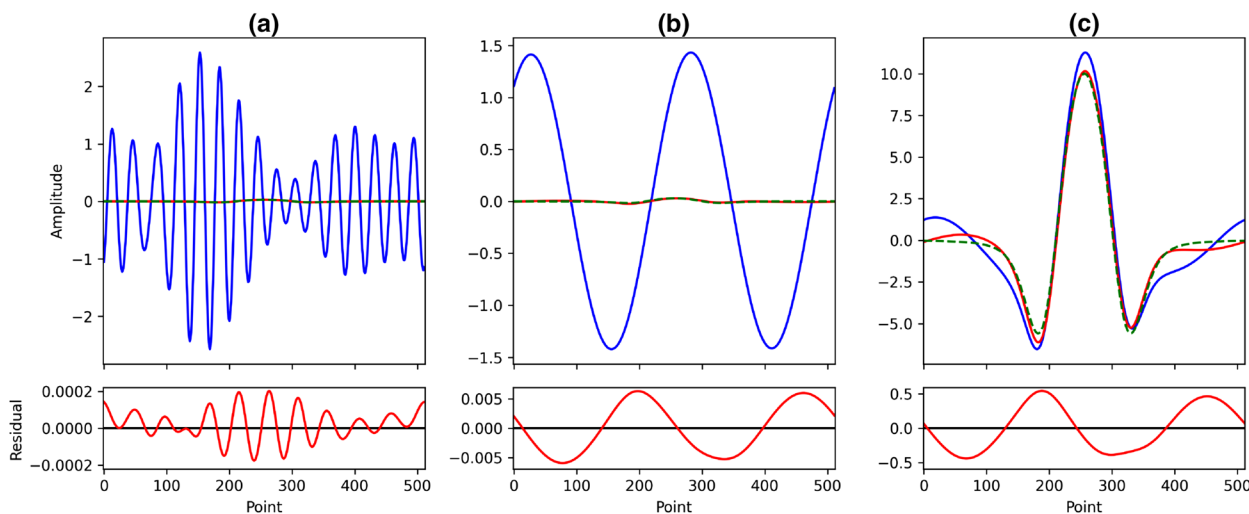
Fig. 4 Examples of simulated signal (dashed green), noise (solid blue) and result of the FDA NRM (solid red) for three different variable settings. The lower subplots each show the residual (reconstruction minus true spectrum). The chosen values of the variables s, f, and c and the resulting SNR increase ΔS and β are given as follows: **a**

$s = 0.03$, $f = 16$, $c = 0.1$, $\Delta S = 47$dB, $\beta = 4.7 \cdot 10^{-3}\%$; **b** $s = 0.03$, $f = 2$, $c = 0.1$, $\Delta S = 39$dB, $\beta = 32\%$; **c** $s = 10$, $f = 2$, $c = 0.1$, $\Delta S = 19$dB, $\beta = 0.34\%$. As mentioned in the text ΔS and β strongly depend on variable setting

-2^4 frequency region, while β is close to zero almost everywhere. Only the overlying smooth region at low frequency and low complexity shows bias values up to 10%. At the same time the SNR increase starts to rise at this region.

The final image at $s = 100$ supports this new characteristic. While bias and SNR increase have gone down to zero almost everywhere, SNR increase grows to about 40dB in the former biased region at low frequency and low complexity. A cause could be an increased capability to

Fig. 5 Highlights of the best performances of the analyzed NRMs in the fc-plane. The mean SNR increase and bias is given for the rectangular regions or the line trajectory, respectively: Wiener: $\overline{\Delta S} = 7.3$ dB, $\overline{\beta} = 8.7 \cdot 10^{-4}\%$; SVD: high frequency region: $\overline{\Delta S} = 24.9$ dB, $\overline{\beta} = 2.0\%$; low frequency region: $\overline{\Delta S} = 40.9$dB, $\overline{\beta} = 5.6\%$; DWT-SURE: $\overline{\Delta S} = 10.6$ dB, $\overline{\beta} = 1.2 \cdot 10^{-5}\%$; FDA: $\overline{\Delta S} = 38.2$dB, $\overline{\beta} = 7.1 \cdot 10^{-3}\%$; EMD-FCR: $\overline{\Delta S} = 14.3$dB, $\overline{\beta} = 4.9\%$; ANN: $\overline{\Delta S} = 15.0$ dB, $\overline{\beta} = 1.0\%$. Rectangles filled with a stripe pattern indicate negative values of ΔS
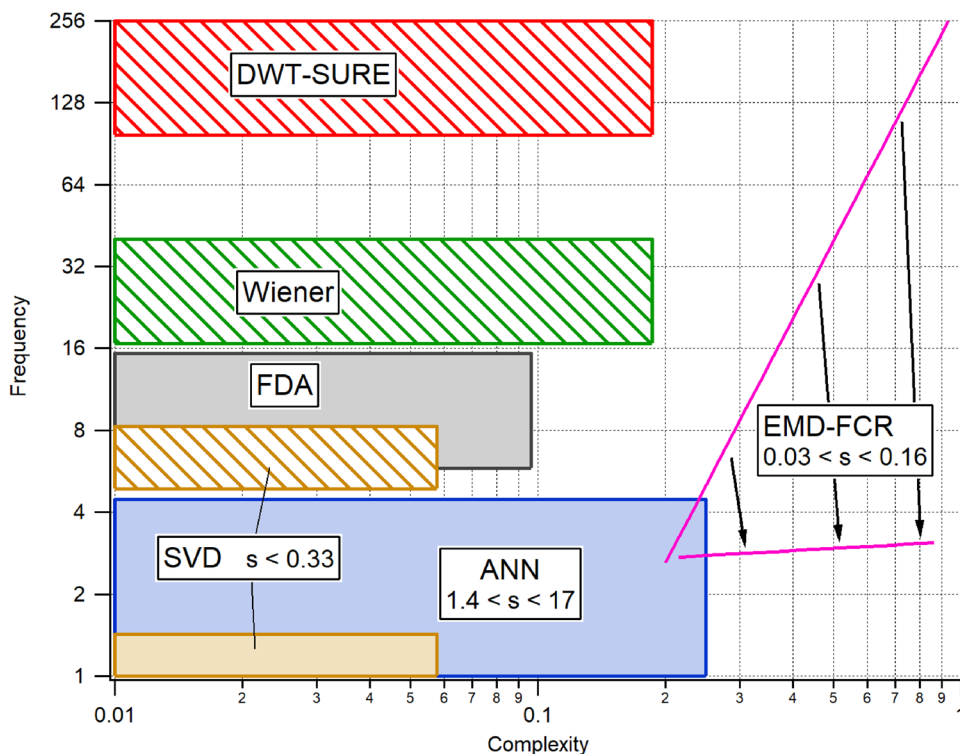
reconstruct the low frequency part of the power spectrum from the higher frequency features due to the high signal. This implies that up to $s \approx 5$ only the first peak in the power spectrum may have been used by the algorithm.

As a summary the FDA method performs extraordinary well at complexity below 0.2 and in the frequency range 2^3 –2^5 and also at lower frequencies for high input SNR. The described features are visualized in Fig. 4. Here the method is applied to example spectra for three different settings.

4.2 Performance summary

Interpreting the information from the animations given in the electronic supplementary material, the other NRMs are evaluated in a similar manner as the case study from the previous section. These results will not be discussed in such detail at this point but are instead summarized to limit the extent of this paper.

The Wiener filter achieves an SNR increase up to 19dB for frequency between 2^4 and 2^6 and complexity below 0.2, independent of signal amplitude s. Similar to the gradual decline of ΔS for higher frequencies due to baseline performance, this method performs better for lower frequencies inside the described region. The bias differs from zero significantly only at very low frequency or high complexity.

The NRM based on SVD performs best at signals below 1, complexity below 0.08 and two different frequency ranges. At frequencies higher than 2^3 the method shows no significant performance gain compared to the baseline, between 1.4 and 2^2 β grows to values of 60%. In this intermediate region the noise is too close to the signal to be distinguished.

The Discrete Wavelet Transform with Stein Threshold shows noisy SNR increase independent from input signal s for complexity below 0.2. Some broad bands emerge at frequencies equal to a power of two above 2^5, where ΔS reaches values above 90dB. In this region the bias is always close to zero. The noisy SNR increase corresponds to a noisy choice of wavelet family and order in the calibration phase.

The method called EMD-FCR shows high bias for almost all parameter combinations. Obviously the output is lowered drastically if the noise free spectrum does not match with an IMF, since the filtered IMFs are multiplied by the correlation coefficient. In the simulation even at high signals s the absorption spectrum is split into two or more IMFs very often. Here a thresholding scheme instead of a weighted sum could lead to better performance. However, at signal amplitude between 0.02 and 0.2 a narrow region of low bias and ΔS above 10dB propagates along a curved path through the fc–plane. Here it seems the EMD

separates signal from background and maps the absorption signal into a single IMF.

The final NRM analyzed is the combined set of simple artificial neural networks (ANN). Although the base model is primitively designed and was not optimized for this special task, the results show an average SNR increase of 15dB and average bias below 1% for signal amplitude between 1.4 and 17, frequency below 2^2 and complexity below 0.26. The performance coincides with a particular network choice during calibration phase.

Figure 5 shows an overview of the best performance for each NRM in the fc-plane. The corresponding mean values for ΔS and β are given in the figure caption.

4.3 Remarks

Now, some flaws of the simulation will be mentioned which influence the results of this paper.

- The reference spectrum and added signal is assumed to be noise-free and stable over time. Distortions and drifts in real experiments can influence the performance of some of the NRMs. Prominent effects that can cause distortions in experimental conditions that have not been considered are residual amplitude modulation [2], the phase shift between wavelength modulation and intensity modulation [35] and the non-linearity of (1).
- Although the simulated noise covers a big range of noise characteristics, it is only an idealization of real background structures. Only fringe-like noises with a dominant frequency have been considered. Thus, the performance of some NRMs may vary depending on the background structure.
- Since the number of simulated ambient spectra is finite, this paper only gives statistical estimates of expected results.
- Since the number of simulated calibration spectra is finite, the grid search not always chooses the optimal free parameters. This could result in an underestimation of performance.

5 Conclusion

There is no best noise reduction method for wavelength modulation spectroscopy. Parameters like the dominant noise frequency, noise power spectrum confinement and input signal strength can vastly influence the performance of some methods and lead to distortions of the output. A proper algorithm has to be carefully chosen according to the experimental setup. The results from this paper can help

predict possible performances given the noise parameters for a particular setup. Each analyzed method showed different behavior for the simulation parameters. Although some NRMs were altered slightly to fit the simulation conditions, all of them were able to outperform the least squares linear fit by at least 10dB without introducing large biases. The best performances were highlighted and little to no overlap of these best performance ranges were found. This result could explain the odd phenomenon that the methods outperformed other established techniques in the studies in which they were proposed, sometimes leading to apparent contradictions between different studies.

While a quantitative comparison for a large space of noisy input shapes was given, many questions arose or were left unanswered. There are infinitely many more noise shapes to be tested. The effect of distortions to the noise free spectrum were ignored. The choice of free parameters and small adaptions to the methods could be improved. Finally there is still a long way to be able to identify and evaluate the best noise reduction method under experimental conditions, but this paper tries to take a first big step.

A follow-up study is planned to further analyze and quantify the results given in this paper under several experimental conditions. The methods will be tested for two different absorption spectrometers (different noise patterns) and several different trace gases (varying initial signal to noise ratios).

References

1. P. Werle, F. Slemr, K. Maurer, R. Kormann, R. Mücke, B. Jänker, Opt. Lasers Eng. **37**(2–3), 101 (2002)
2. J. Hodgkinson, R.P. Tatam, Meas. Sci. Technol. **24**(1), 012004 (2012)
3. C.S. Goldenstein, R.M. Spearrin, J.B. Jeffries, R.K. Hanson, Prog. Energy Combust. Sci. **60**, 132 (2017)
4. J. Faist, F. Capasso, D.L. Sivco, C. Sirtori, A.L. Hutchinson, A.Y. Cho, Science **264**(5158), 553 (1994). https://doi.org/10.1126/science.264.5158.553
5. J.U. White. Optical system providing a long optical path (1957)
6. D. Herriott, H. Kogelnik, R. Kompfner, Appl. Opt. **3**(4), 523 (1964)
7. P. Werle, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. **54**(2), 197 (1998)
8. P. Werle, R. Mücke, F. Slemr, Appl. Phys. B **57**(2), 131 (1993)
9. Z. Wang, P. Fu, X. Chao, Appl. Sci. **9**(13), 2723 (2019)
10. J. Li, B. Yu, W. Zhao, W. Chen, Appl. Spectrosc. Rev. **49**(8), 666 (2014)
11. J. Li, U. Parchatka, H. Fischer, Appl. Phys. B **108**(4), 951 (2012)
12. A. Hartmann, R. Strzoda, R. Schrobenhauser, R. Weigel, Appl. Phys. B **115**(2), 263 (2014)
13. P.W. Werle, B. Scheumann, J. Schandl, Opt. Eng. **33**(9), 3093 (1994)
14. I. Mappe-Fogaing, L. Joly, G. Durry, N. Dumelié, T. Decarpenterie, J. Cousin, Appl. Phys. B **108**(4), 933 (2012)
15. D. Richter, P. Weibring, J.G. Walega, A. Fried, S.M. Spuler, M.S. Taubman, Appl. Phys. B **119**(1), 119 (2015)
16. A.K. Leung, F. Chau, J. Gao, Chemomet. Intell. Lab. Syst. **43**(1–2), 165 (1998)
17. P.S. Addison, Physiol. Meas. **26**(5), R155 (2005)
18. A. Grossmann, J. Morlet, SIAM J. Math. Anal. **15**(4), 723 (1984)
19. I. Daubechies, Commun. Pure Appl. Math. **41**(7), 909 (1988). https://doi.org/10.1002/cpa.3160410705
20. S.G. Mallat, IEEE Trans. Acoust. Speech Signal Process. **37**(12), 2091 (1989)
21. B.K. Alsberg, A.M. Woodward, M.K. Winson, J. Rowland, D.B. Kell, Analyst **122**(7), 645 (1997)
22. I. Mappe-Fogaing, L. Joly, G. Durry, N. Dumelié, T. Decarpenterie, J. Cousin, B. Parvitte, V. Zéninari, Appl. Spectrosc. **66**(6), 700 (2012)
23. D.L. Donoho, I.M. Johnstone, J. Am. Stat. Assoc. **90**(432), 1200 (1995)
24. H.Y. Gao, J. Comput. Graph. Stat. **7**(4), 469 (1998)
25. N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **454**(1971), 903 (1998)
26. Y. Meng, T. Liu, K. Liu, J. Jiang, R. Wang, T. Wang, H. Hu, IEEE Photonics J. **6**(6), 1 (2014)
27. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, Cambridge, 2016)
28. Y. LeCun, Y. Bengio, G. Hinton, Nature **521**(7553), 436 (2015)
29. J. Nicely, T. Hanisco, H. Riris, J. Quant. Spectrosc. Radiat. Transf. **211**, 115 (2018)
30. I. Gordon, L. Rothman, C. Hill, R. Kochanov, Y. Tan, P. Bernath, M. Birk, V. Boudon, A. Campargue, K. Chance, B. Drouin, J.M. Flaud, R. Gamache, J. Hodges, D. Jacquemart, V. Perevalov, A. Perrin, K. Shine, M.A. Smith, J. Tennyson, G. Toon, H. Tran, V. Tyuterev, A. Barbe, A. Császár, V. Devi, T. Furtenbacher, J. Harrison, J.M. Hartmann, A. Jolly, T. Johnson, T. Karman, I. Kleiner, A. Kyuberis, J. Loos, O. Lyulin, S. Massie, S. Mikhailenko, N. Moazzen-Ahmadi, H. Müller, O. Naumenko, A. Nikitin, O. Polyansky, M. Rey, M. Rotger, S. Sharpe, K. Sung, E. Starikova, S. Tashkun, J.V. Auwera, G. Wagner, J. Wilzewski, P. Wcisło, S. Yu, E. Zak, J. Quant. Spectrosc. Radiat. Transf. **203**, 3 (2017). https://doi.org/10.1016/j.jqsrt.2017.06.038 (HITRAN2016 Special Issue)
31. Q. Zou, P. Varanasi, J. Quant. Spectrosc. Radiat. Transf. **75**(1), 63 (2002)

32. J. Reid, D. Labrie, Appl. Phys. B **26**(3), 203 (1981). https://doi.org/10.1007/BF00692448

33. G. Vivó-Truyols, P.J. Schoenmakers, Anal. Chem. **78**(13), 4598 (2006)

34. M. Sadeghi, F. Behnia, R. Amiri, IEEE Trans. Instrum. Meas. **69**(8), 5418 (2020)

35. S. Schilt, L. Thévenaz, Infrared Phys. Technol. **48**(2), 154 (2006)