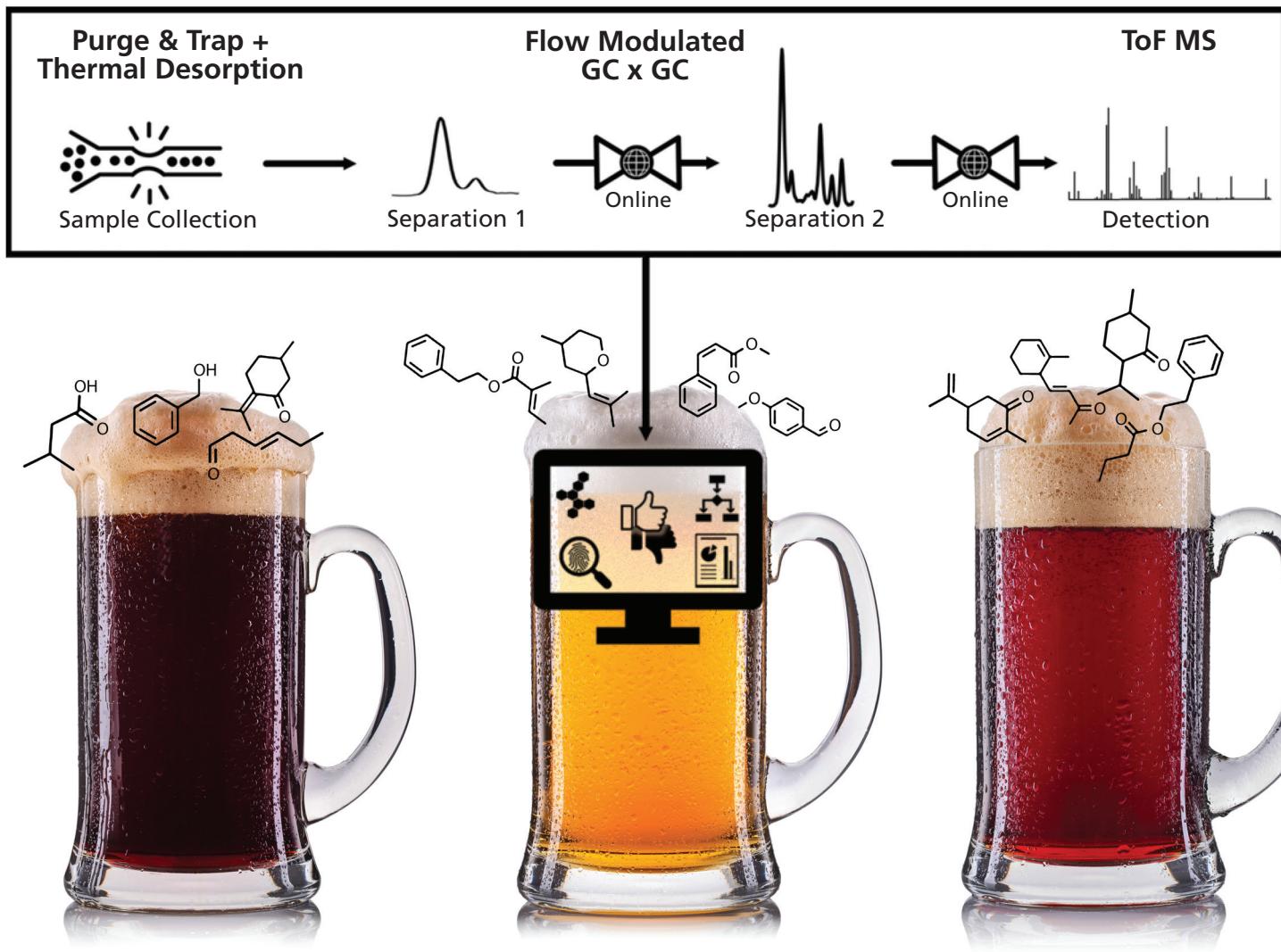


J
S
S

JOURNAL OF SEPARATION SCIENCE

9-10 | 20



Special Issue
Emerging Thought Leaders in Separation Science

Guest Editor
Elia Psillakis

www.jss-journal.com

WILEY-VCH

REVIEW

Wavelet transforms in separation science for denoising and peak overlap detection

Muhammad Farooq Wahab¹  | Thomas C. O'Haver²

¹Department of Chemistry and Biochemistry, University of Texas at Arlington, Arlington, TX, USA

²Department of Chemistry and Biochemistry, University of Maryland at College Park, College Park, MD, USA

Correspondence

Muhammad Farooq Wahab,
Department of Chemistry & Biochemistry, University of Texas at Arlington, Arlington, TX 76019, USA.
Email: mfwahab@uta.edu,
mawahab@ualberta.ca

Wavelet transform is a versatile time-frequency analysis technique, which allows localization of useful signals in time or space and separates them from noise. The detector output from any analytical instrument is mathematically equivalent to a digital image. Signals obtained in chemical separations that vary in time (e.g., high-performance liquid chromatography) or space (e.g., planar chromatography) are amenable to wavelet analysis. This article gives an overview of wavelet analysis, and graphically explains all the relevant concepts. Continuous wavelet transform and discrete wavelet transform concepts are pictorially explained along with their chromatographic applications. An example is shown for qualitative peak overlap detection in a noisy chromatogram using continuous wavelet transform. The concept of signal decomposition, denoising, and then signal reconstruction is graphically discussed for discrete wavelet transform. All the digital filters in chromatographic instruments used today potentially broaden and distort narrow peaks. Finally, a low signal-to-noise ratio chromatogram is denoised using the procedure. Significant gains (>tenfold) in signal-to-noise ratio are shown with wavelet analysis. Peaks that were not initially visible were recovered with good accuracy. Since discrete wavelet transform denoising analysis applies to any detector used in separation science, researchers should strongly consider using wavelets for their research.

KEY WORDS

digital filters, noise removal, peak overlap, signal processing, wavelet transforms

1 | INTRODUCTION

Separation scientists are continuously challenged by the complexity of *in vivo* and *in vitro* chemical systems. These problems are currently tackled by exploiting highly selective detection technologies, surface chemistries, and obtaining high peak capacities by using small particles or narrow open tubular columns [1, 2]. Another less trodden path for solving separation problems is signal processing for enhancing peak to peak resolution and noise reduction [3-7].

Article Related Abbreviations: CWT, continuous wavelet transform; DWT, discrete wavelet transform; WT, wavelet transform.

Chromatographers have not yet openly accepted the idea of peak processing techniques unlike spectroscopists who routinely use mathematical tools to process their data. The apprehension is understandable from the fact that the pharmaceutical industry, which is the primary user of chromatography, is subject to rigorous regulations. Any applied mathematical peak processing techniques must be foolproof if the chromatography community has to accept it wholeheartedly. In contrast, the raw spectroscopic data are rarely presented, even for legal purposes, but instead it goes through a series of transformations to make it interpretable. For example, the raw nuclear magnetic resonance data (free induction decay curve in time) cannot be interpreted by any chemist until and unless they see the Fourier transform (FT) of these

data. Similarly, in vibrational spectroscopy, the interferogram is transformed into the familiar IR spectrum by FT. The electron spin resonance spectrum is displayed as a first derivative.

A question then arises on exploring the utilities of newer or better signal processing techniques in chromatography in order to visualize the same data in another domain and obtain “hidden” information from it. Simple visualization of the data with simple chromatographic figures of merit cannot provide conclusive information about the peak purity or overlapping peaks. Similarly, noise can bury some useful signal such as a small peak in a chromatogram. This could be true for biological samples, which can have thousands of components in disproportionate amounts [2]. Hence, signal processing for separation science aims to reduce the number of experiments to obtain the same level of information, which were obtained by carrying out an extensive set of experiments. For instance, two partially overlapping peaks can be mathematically resolved, and the area can be recovered without doing an actual baseline separation by the screening of various columns and several mobile phases [4, 5, 8, 9]. Similarly, a small peak area can be recovered from a noisy chromatogram by repeating the experiment with another detector setting.

FT has reigned supreme for almost a century in every domain of analytical chemistry. It can be predicted that the next logical stage is to use wavelets for the analysis of time varying signals such as those seen in separation science, as another powerful tool with unique capabilities. Herein, we explore and review the utility of wavelet transforms (WTs) in analyzing the chromatographic signals, and use them as a data visualization tool for extracting hidden information or recovering signals from noise. The term wavelet was introduced in the 1940s in analyzing earthquake signals by Ricker [10]. The French mathematicians called them as *ondelette* (from onde, a wave, -ette, small) [11]. The logical reason for this name will become clear in the next section.

Wavelet analysis entered analytical chemistry in 1992 for the evaluation of peak intensities in flow-injection analysis in a noisy environment of enthalpimetric measurements [12]. Most separation techniques are flow based and have several noise sources such as the pump pulsations, back-pressure regulators, solvent mixing, lamp noise, drift, and the fundamental shot noise. This is coupled with the severe problem of highly undesirable partial or complete peak overlap. Wavelet analysis can assist separation scientists in improving the S/N and can help in the detection of hidden peaks, as illustrated in the application sections. The article shows that continuous wavelets and discrete wavelet transforms (DWTs) are powerful approaches toward thinking and visualizing analytical chemistry problems.

2 | OVERVIEW OF WAVELETS

2.1 | Continuous wavelet transform

The purpose of any mathematical transform is to enhance certain features of the original data, make it more amenable to interpretation, or make certain calculations easier. Some operations in the frequency domain are much easier to accomplish than in the time domain. Convolution that requires calculation of integrals in time domain becomes a simple multiplication in the Fourier domain. One major drawback of mathematical manipulation in the frequency domain is the loss of all the time information and generation of a high level of noise during multiplication or division in the frequency domain [13]. A continuous wavelet transform (CWT) is very similar to a FT. The key difference is that instead of using infinitely extended sines and cosines, WT uses small wavelike functions $\psi(t)$, which are localized in space or time, hence the name wavelet (Figure 1) [14]. Unlike FT, wavelet analysis has a vast choice of functions for analyzing the signal such as Haar wavelet, Mexican hat wavelet, Coiflet family of wavelets, Daubechies family of wavelets, and Symlets among others as shown in Figure 1 [14, 15]. A significant advantage of the WT is that it allows to localize a given signal into a 2D time–frequency plot, detect discontinuities, edges, and/or signal overlap. The analyzing wavelet function $\psi(t)$ must meet these mathematical requirements: (i) the integral of the wavelet function must be zero, that is, the area of a wavelet function is zero and (ii) the square of the wavelet function should integrate to unity [14].

The schematics of a WT is shown in Figure 2. Mathematically, a WT is the convolution of the given signal with the wavelet function. Since the wavelet is localized in space or time, a wavelet function has to be translated by steps of b in order to cover the entire signal. Once a given wavelet has been entirely translated over the whole signal, the same wavelet function can now be compressed or dilated by scale a . This process is always called dilation in the wavelet literature to imply dilation or compression [15]. As an illustration, consider a simulated chromatogram with random and periodic noise (as typically seen in some supercritical fluid chromatographs).

In Figure 2, we have chosen a “Mexican hat” wavelet because the Mexican hat wavelet is related to the second derivative of a Gaussian. It may “probe” Gaussian-like signals well. Qualitatively, we can understand the wavelet transformation as follows. The original or so-called mother wavelet is placed at the beginning of the signal with a scale $a = 1$. This mother wavelet is multiplied and integrated over the available time range (Figure 2A). Now the wavelet is translated by b time units over the signal, and the transform values are calculated for the scale $a = 1$. The mother wavelet is continued to

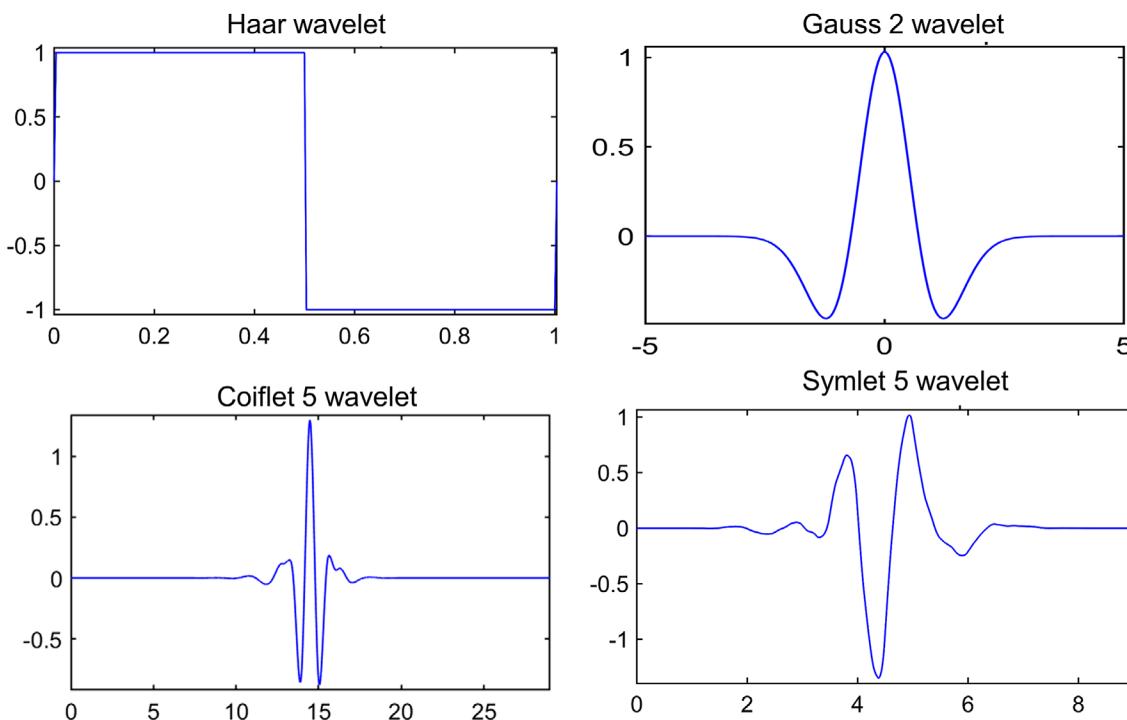
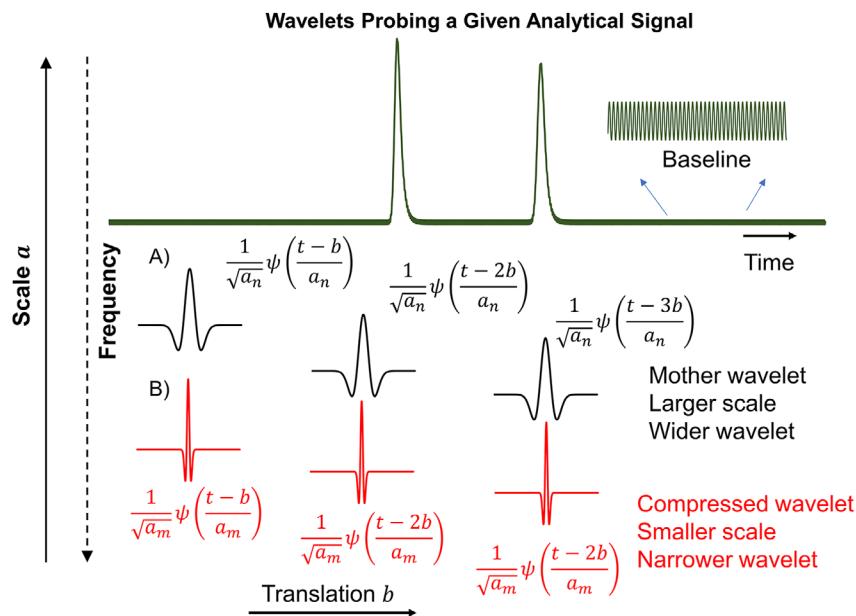


FIGURE 1 Graphical representation of some commonly used wavelets used in the analysis of signals. Gaus2 wavelet is obtained from the second derivative of a Gaussian function. Some wavelets do not have an analytical expression such as Coiflet and Symlet wavelets



be translated, and its transform calculated until we reach the end of the signal. Next, the scale a is changed (increased or decreased), and all the previous integration steps are repeated for all translation values. One row of a given scale value a fills a single row of the 2D time-scale plot for all possible translations. At a given local position during translation, if the wavelet matches very well with the local signal, then a large value of transform is observed. Figure 2B shows an example of a compressed mother wavelet, which is also

FIGURE 2 A schematic diagram explaining the concept of continuous wavelet analysis on a chromatogram. The figure shows the concept of scale and translation using a Mexican hat wavelet. The Mexican hat wavelet is equivalent to the second derivative of a Gaussian. (A) A given wavelet with a fixed scale a is translated in b steps eventually sweeping across the entire signal length. This generates a row of data consisting of coefficient values for all possible translations at a given scale. (B) The mother wavelet can be compressed and the entire procedure repeated to fill another row of coefficients for all possible translations as shown in Figure 3 covering the entire time range

shifted in steps of b . If this compressed wavelet locally matches with the signal, then a substantial transform value is obtained. As shown in Figure 2, one expects larger values of the transform, where the peaks exist, because the wavelet matches the shape of peak-like signals. Thus, it is aptly stated that a wavelet is a “mathematical microscope,” [16] where the translation value b signifies its position on time series being studied, and the parameter a is the desired magnification at a given position b . Mathematically, CWT of a signal $s(t)$ is

represented as

$$C(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (1)$$

The symbol ψ^* indicates the complex conjugate. If the wavelet function is real, the complex conjugate is the same as the original wavelet function. Here, a is a scaling parameter, which is related to the frequency of the scaled mother wavelet $\psi^*(\frac{t-b}{a})$ [14]. The larger the scale values, the larger is the time width of the wavelet. A wider wavelet in the time domain can capture a smaller range of lower frequencies, and a narrow wavelet can probe a large window of frequencies. This property arises from the property of Fourier transformation, where an expanded wavelet in the time domain has a narrow energy spectrum in the frequency domain and vice versa [14] (see Figure 2).

Figure 3 compares the concept of FT and WT. A chromatograph with random noise and 10 Hz noise is shown (also see the inset of Figure 2). These types of random white noise superimposed with regular pulsations can arise in supercritical fluid chromatography due to a backpressure regulator for carbon dioxide. Its vibrations at times can generate a periodic noise of 5 or 10 Hz in the signal [17]. A FT of the data in Figure 3A data shows a spike at 10 Hz and then a peak at low frequencies. The low frequencies with wiggles correspond to the two peaks of the chromatogram and the random fluctuations in the baseline correspond to random noise. Note the x -axis, where we have entirely lost the time information (Figure 3B) as to where the useful signal and where the noise are. On the other hand, a CWT of the same data shows two useful signals are located at 0.4 and 0.6 min, and there is a continuous noise of the same frequencies in the entire chromatogram. Note that high-frequency noise appears at low scales (shown in green in Figure 3C). There is a qualitative inverse relationship of frequencies and scales a [14].

2.2 | Discrete wavelet transforms

In Eq. (1), every possible value of scale and translation could be analyzed. If alternatively discrete values of scaling and translations are used, then we can write

$$\begin{aligned} a &= a_0^j \\ b &= k b_0 a_0^j, \end{aligned} \quad (2)$$

where $a_0 > 1$, $b_0 \neq 0$, j, k are integers. Hence the DWT can be written as

$$C(j, k) = \frac{1}{\sqrt{a_0^j}} \int_{-\infty}^{\infty} s(t) \psi^* \left(\frac{t - k b_0 a_0^j}{a_0^j} \right) dt \quad (3)$$

If $a_0 = 2$ and $b_0 = 1$, a power of 2 logarithmic scalings of both the dilation and translation steps, then this choice is called dyadic grid scaling in the wavelet analysis literature [14]. Then Eq. (3) simplifies to

$$C(j, k) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} s(t) \psi^* \left(\frac{t - k 2^j}{2^j} \right) dt, \quad (4)$$

which shows that the scale varies as powers of two and the translations are integral multiples of scale [14]. The larger the level j , the wider is the wavelet and corresponds to coarser resolution. Smaller values of j represent a narrower wavelet and hence finer resolution [18].

In practice, DWT is done with the help of a multiresolution algorithm with orthogonal wavelets [19]. It is crucial to keep in mind that signal decomposition by DWT produces a family of level-wise organized decompositions, which is unlike all other common transforms used in analytical chemistry. The DWT operates on data whose length is an integral power of 2 as seen in Eq. (4) with the help of filters of fixed length [20]. In reality, most data length is not a power of two and in that case software, such as Matrix Laboratory MATLABTM, use built-in data extension mode to cover this requirement. The data are then passed through a pair of high-pass (H) and low-pass (L) decomposition wavelet filters. The scaling option, which was available in CWT, is now achieved by shrinking the length of data instead of compressing or dilating the wavelet. Keeping length of the wavelet filters the same, and the size of the data is dyadically reduced because the wavelet scale changes by the power of 2 rather than continuously [15]. There are several wavelets families available in the DWTs such as Haar, Symlet, Coiflet, and Daubechies.

This DWT decomposition process is schematically shown in Figure 4. At level 1, the output of the high-pass filter captures high frequencies of the signal and yields the *detail coefficients* (cD_1). The low pass operation on the data captures low frequency information and yield the *approximation coefficients* (cA_1) for the given signal. Now, this process is further continued to the next desired levels as follows. At level 2 DWT of the signal, the approximate coefficients are further passed through the same high-pass and low-pass filter after down-sampling by 2, giving us cA_2 and cD_2 . Note that this process is akin to zooming out of the signal, because at level 2, we are already operating on the approximate coefficients of the signal or in other words looking at a “coarser magnification” of the signal. At level 3, the cA_2 is further decomposed by the same high- and low-pass wavelet filter to give cA_3 and cD_3 . The cA_3 is an even a coarser version of cA_2 , and cD_3 has captured the higher frequencies present in the cA_2 . The details of the exact step-by-step process are outlined in an excellent reference on “Conceptual Wavelets” by Fugal for a data set of eight points [15].

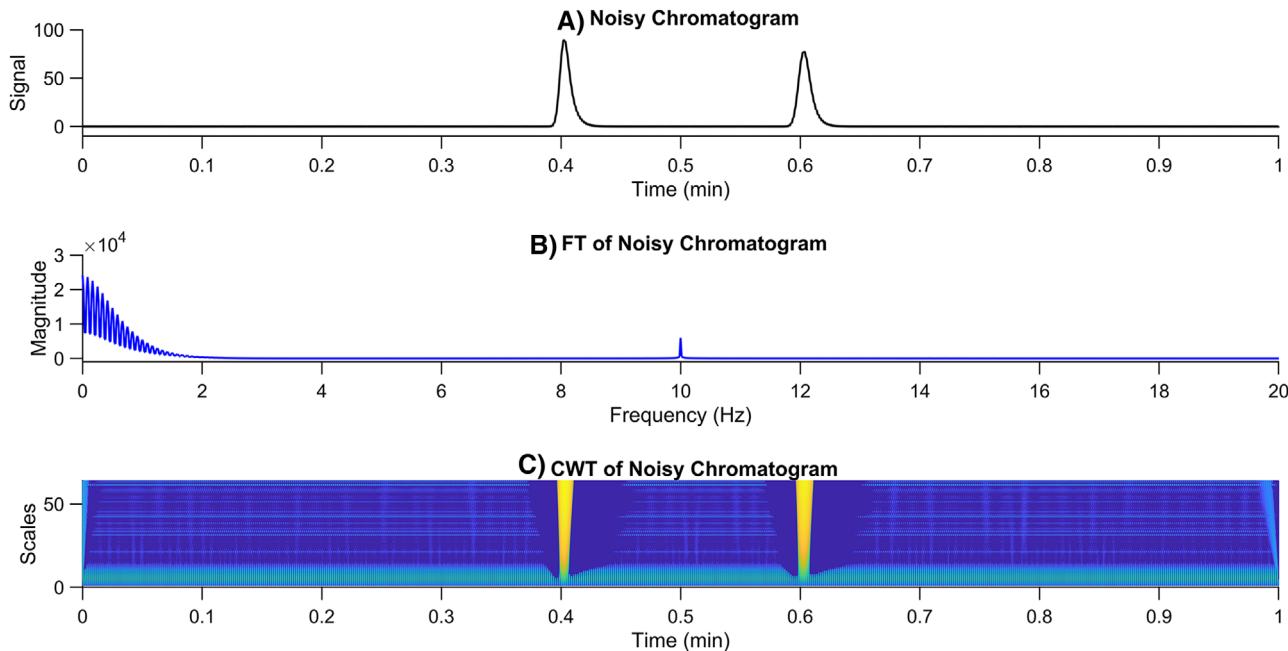


FIGURE 3 Comparison of Fourier transform of the same chromatogram shown in Figure 2 with continuous wavelet transforms using a Mexican hat wavelet. (A) The data sampling rate is 200.8 Hz. The chromatogram has random noise (standard deviation of = 0.005) and 10 Hz superimposed noise. (B) The relevant frequencies are found below 20 Hz including a 10 Hz spike and low-frequency component corresponding to the two Gaussian peaks. The Fourier scale extends to 100.4 Hz (half the sampling rate). (C) A continuous wavelet transform shows the exact time location of useful signal and shows that the random and 10 Hz is present everywhere uniformly and shown in green

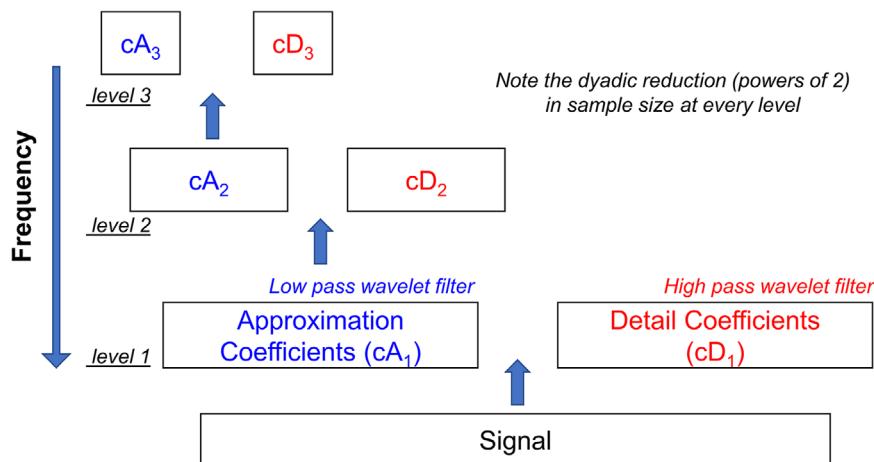


Figure 5 visually shows the DWT decomposition, on the same signal in Figure 3A, at four levels using Coiflet-5 wavelet filters. The noisy chromatogram with an additional 10 Hz noise is decomposed at four levels. As discussed above, a high-pass Coiflet filter on the raw data ($N = 12\,050$ points) yields about $N/2$ cD_1 (after downsampling by 2). Similarly, a low-pass Coiflet filter on the raw data (12 050 points) yields about $N/2$ cA_1 (not shown in the figure). These detail coefficients cD_1 at level 1 and capture all the high frequency components of the signal (Figure 5). Further decomposition of the approximation coefficients cA_1 at level 2 with high- and low-pass filters gives cD_2 and cA_2 , with $N/4$ points. The

FIGURE 4 A schematic explanation of the concept of the discrete wavelet transforms for the decomposition of a signal in a dyadic fashion (powers of 2). The signal is decomposed into approximate coefficients and detail coefficients at each level with the help of low pass and high pass decomposition filters. Downsampling by 2 is also done at each level

level 2 detail coefficients, cD_2 , look at the “zoomed out” signal, that is, they capture lower frequencies of the signal. Next, if cA_2 is decomposed analogously and yields cD_3 and cA_3 with $N/8$ points at level 3. Note the low frequencies captured by cD_3 . Further decomposition of cA_3 by high- and low-pass Coiflet wavelet filters finally give us cD_4 and cA_4 with $N/16$ points. Figure 5 shows both cD_4 and cA_4 . The approximation coefficients at level 4 are a highly smoothed and downsampled by $N/16$ approximate coefficients of the original version of the signal [21]. Supporting Information Figure S1 shows the decomposition Coiflet 5 filters.

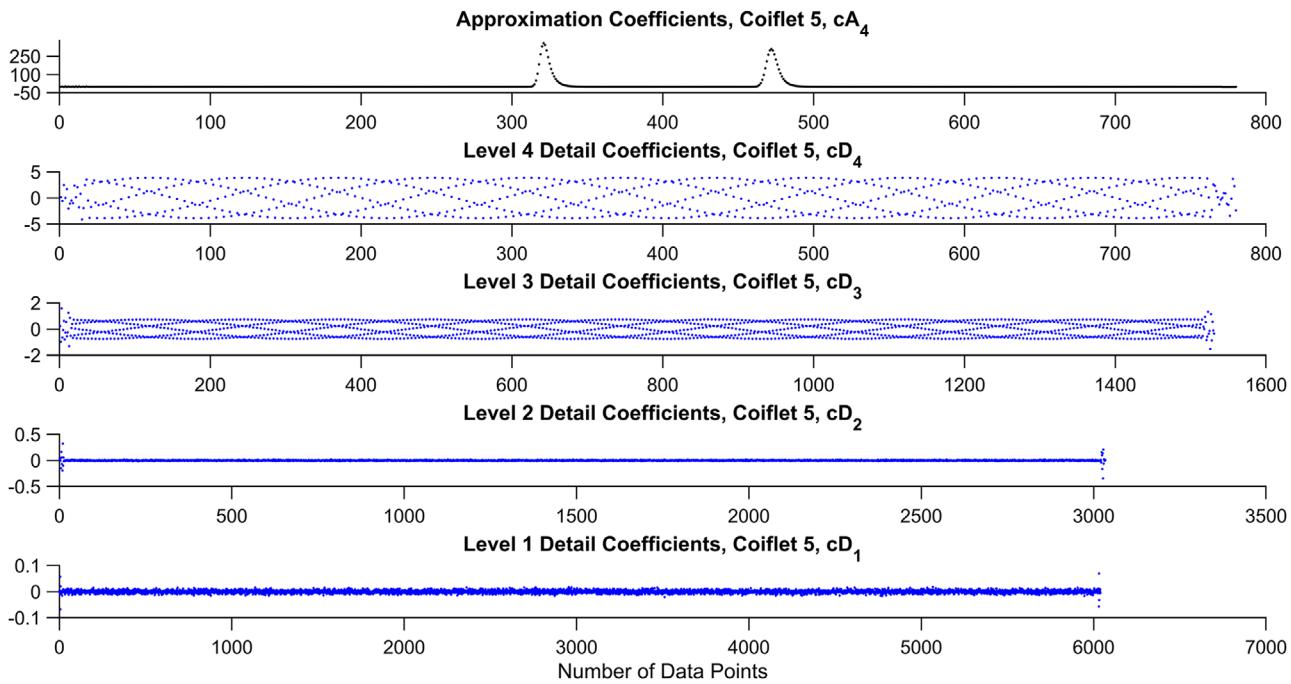


FIGURE 5 Explanation of the concept of discrete wavelet decomposition of the same signal shown in Figure 3 into approximate and detail coefficients using a coiflet 5 (coif5) decomposition wavelet filter. Note how the approximate coefficients capture the low-frequency components of the signal and the detail coefficients different ranges of the frequency of the signal. At level 1, the detail coefficients capture the highest frequency noise (50.2–100 Hz), level 2 (50.3–25 Hz), level 3 (25.2–12.5 Hz), level 4 (6.26–12.6 Hz), and approximation (0–6.27 Hz) as obtained from multiresolution analyzer toolbox in MATLAB

At this stage for denoising applications, the detail coefficients (cD s) at various levels can be eliminated, shrunk, or thresholded before reconstructing the signal to provide an ideally noise-free signal. If we set all the detail coefficients (high frequency) to zero, we would lose all the frequency information and sharp features of the signal. Instead of setting the detail coefficients to zero, there are several methods for thresholding the detail coefficients. For the sake of brevity, the wavelet literature mentions two types of thresholding: (i) hard thresholding and (ii) soft thresholding. In hard thresholding, all coefficients below a threshold value T on a given level are set to zero. In soft thresholding, the detail coefficients are shrunk by T and the values below T are set to zero [20]. An excellent review of DWT denoising is available for chemistry applications and along with a comparison of the thresholding methods [22]. The main strength of DWT is that it allows multiscale decomposition, which is not possible in conventional filtering and, independent thresholding at various levels. It also provides time interval-dependent denoising, which means that if some regions of chromatogram have a poor S/N or spikes, then one can remove noise from that region and leave other regions intact where there is less noise [23]. This is also not possible by conventional means.

The question arises that once we have decomposed the signal at various levels, and did thresholding to dampen or eliminate noise, how do we get the modified signal back?

The inverse DWT allows us to reconstruct or synthesize the signal from approximation and detail coefficients by upsampling and running reconstruction filters on the coefficients [15]. The size of the reconstructed signal is the same as the original data [9]. Supporting Information Figure S1 shows the reconstruction Coiflet 2 filters. The output of reconstruction filters on the approximation coefficients and detail coefficients yield *approximations As* and *details Ds* of the signal at various levels. The reconstructed signal is then simply the sum of approximations and details:

$$\begin{aligned} \text{Reconstructed Signal} &= A_1 + D_1 = A_2 + D_2 + D_1 \\ &= A_j + D_j + D_{j-1} + \dots + D_1 \quad (5) \end{aligned}$$

Equation (5) shows that if the DWT decomposition was performed at level 1, the reconstructed signal is equal to the sum of $A_1 + D_1$. If a level 2 DWT was performed, the reconstructed signal is the sum of $A_2 + D_2 + D_1$. One can immediately see the advantage of multilevel decomposition that it allows a great deal of flexibility in analyzing and manipulating signal especially in noise removal applications. Figure 6 illustrates the signal reconstruction from the approximation and details. The details D_1 , D_2 , D_3 , and D_4 were obtained from detail coefficients cD_1 , cD_2 , cD_3 , and cD_4 (after upsampling and applying reconstruction filters). The final size of the details is the same as the signal size. After DWT decom-

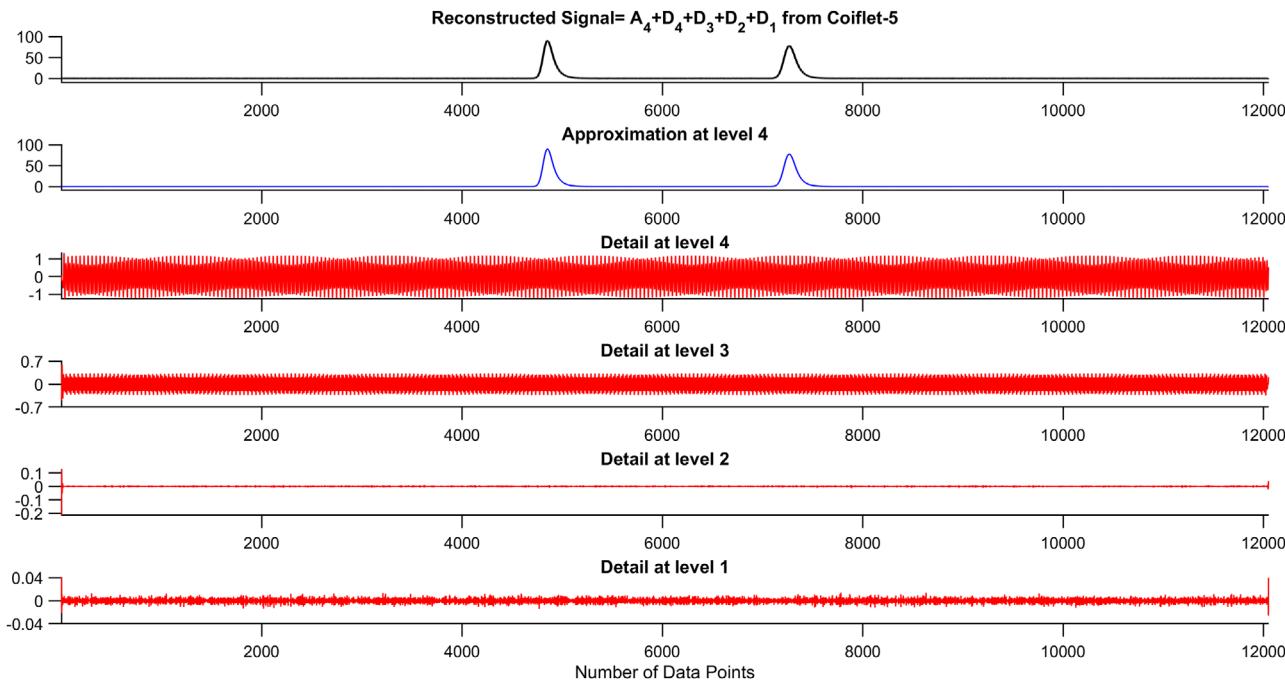


FIGURE 6 Explanation of the concept of inverse discrete wavelet transform or reconstruction of the signals after DWT decomposition. By applying reconstruction filters of the Coiflet 5 wavelet on the upsampled approximate and detail coefficients, the approximation and details are obtained. The final signal is simply the sum of approximation at the given level and the sum of all the details

position, the original signal at level 4 can be reconstructed by adding approximations and details $A_4 + D_4 + D_3 + D_2 + D_1$. If we look at Figure 6 carefully, the approximation A_4 is already a highly smoothed chromatogram. Readers will recognize that the whole wavelet denoising process is somewhat analogous to denoising by Fourier transformation, where the signal is first transformed in the frequency domain, then unwanted frequencies are removed, and then an inverse transform is performed to obtain a lower noise-level signal. The noise removal section will demonstrate the major advantages of using DWT for noise application as compared to traditional filters used in chromatography. A question is where should we stop the signal decomposition? The rule of thumb is that when we begin to lose 70% of the energy of the signal, we can stop the decomposition.

3 | MATERIALS AND METHODS

3.1 | Simulations

Simulations of the Gaussian and bidirectional exponentially modified Gaussians were done in MATLAB R2019a. Peak characterization was done in PeakFit v4.12 (SeaSolve Software, USA). All continuous and discrete wavelets transforms are available in MATLAB library of Wavelet Toolbox 5.2. The CWT was performed in the graphical user interface of Wavelet Analyzer. All the useful commands and complete codes for chromatographic applications are provided in the Supporting Information.

4 | APPLICATIONS OF WAVELET ANALYSIS IN SEPARATIONS

4.1 | CWTs for detecting peak overlaps

In image processing, CWT is used as a powerful tool to detect discontinuities/edges and this is related to an attractive property of certain wavelet functions. After the transform, the CWT coefficients are useful for detecting peak overlap. If the wavelet has n vanishing moments, the CWT can be thought of as an n th-order multiscale differential operator on a given signal. Chromatographers are already familiar with statistical moments for chromatographic peaks, where the zeroth moment of a peak represents its area, the first moment represents the centroid, and the second moment represents the peak variance [24]. If we are interested in a second derivative, then we should choose a wavelet whose zeroth and first moment vanish. The details behind derivative calculations by CWT can be found here and in Mallat's work [25, 26].

What is the primary advantage of using CWT versus simple numerical derivatives to detect peak overlap? The leading utility of CWT wavelet, with n vanishing moments, is that the n th derivative calculated by the particular wavelet is already smoothed. It is significant advantage of finding hidden peaks with CWT, where that noise is already suppressed. It is very well known that taking a derivative degrades S/N. Even under optimum detector settings, derivatives significantly enhance noise especially at high sampling rates. The strength of CWT

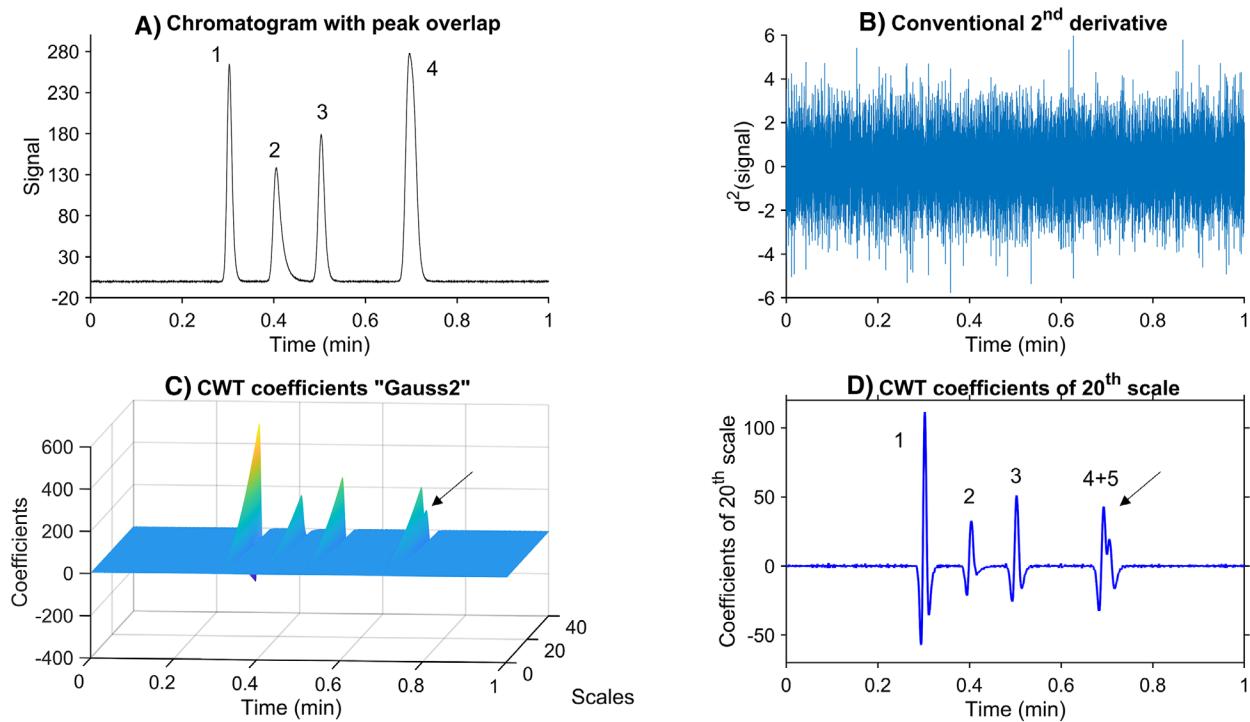


FIGURE 7 Detection of peak overlap with continuous wavelet transform. (A) A signal with four peaks is shown where one of the peaks is a double component. (B) Taking the conventional second differential shows severe degradation of the S/N. Heavy smoothing will be needed in order to see the shape of the second derivative. (C) The same signal is decomposed by Gauss 2 wavelet (gaus2) up to 40 scales. A three-dimensional plot is obtained. (D) A slice of the 3D plot at scale 20 shows a hidden peak underneath peak 4

is that the users have plenty of choices for wavelets, such as Coiflet, biorthogonal, or Gaussian wavelets and so on with zero moments. For example, WT with gaus1, gaus2, gaus3, and gaus4 will generate multiscale first, second, third, and fourth derivatives of the signal, respectively.

Figure 6 illustrates how CWT can be qualitatively used for detecting hidden peaks in a chromatogram with noise. In the previous work, it was demonstrated how even derivatives could help in detecting overlapping signals of analytes [5]. In those calculations, a large point moving average window had to be applied before interpreting any results because the order of noise increment is as follows: first derivative < second derivative < third, and so on [27]. By the time we reach the second derivative, most differentiated chromatograms are saturated with noise because the derivatives amplify high frequencies, which is noise, but not the low-frequency signal. Thus, the S/N after taking a derivative degrades because of an increase in the noise. Hence this attractive powerful technique sometimes becomes almost useless. Second, applying a filter such as the moving average filter changes the shape of the derivative.

In Figure 7A, there are four apparent peaks in the chromatogram and it is not possible to decide which one is a singlet, overlapping doublet, or multiplet. The peaks are asymmetric as well. Calculating the second numerical derivative (Figure 7B), by numerical differentiation, results in a highly degraded S/N, where it is impossible to see any peak

overlap without applying a secondary heavy smoothing of the data. The noise has not been smoothed in the derivative calculation. In Figure 7C, a CWT with a Gaussian WT with two vanishing moments (gaus2) was applied on the chromatogram at 40 scales starting from 1 to 40 with unit intervals. This results in a 3D multiscale second derivatives of the chromatographic signal. If a slice from this figure is taken, say at the 20th scale, Figure 7D is obtained. It is immediately apparent that the fourth peak in the chromatogram is actually a composite of two peaks. This application of detecting edges and discontinuities in image processing is thus readily applicable in processing chromatographic data. Table 1 shows the outline for using CWT for detecting peak overlap.

4.2 | Denoising applications of DWT and comparison other digital filters used in chromatographic instruments

The purpose of any ideal noise removal, or so-called denoising of the signal in the wavelet literature, is to extract the actual signal from the collected data, that is,

$$\text{Observed Signal} = \text{True signal} + \sigma(\text{Gaussian noise}, \mu = 0, \text{std. dev.} = 1) + \text{baseline} \quad (6)$$

One can start with the assumption that the observed signal is a sum of “true signal” plus random Gaussian noise as shown

TABLE 1 General scheme for using continuous and discrete wavelet transforms

General outline for peak overlap detection with CWT

- (A) CWT provides a huge flexibility in the choice of wavelets. Names of continuous wavelets that can be used for detecting peak shoulders or overlaps: Daubechies family (1–9), Symlet family (2–8), Coiflet family (1–5), and Gauss family (1–8). Gaussian wavelet transform gives the most accurate estimation of frequency components localized in time. Gauss wavelet 1 will have an output similar to the shape of the first derivative. Gauss-2 shows an output similar to second derivative. In MATLAB, for the CWT the scales can be varied continuously from 0 to 4000 (with a step size > 0).
- (b) Analyze a real chromatographic signal with a given wavelet, say Gauss 2 at a scale of 1. If the output is noisy, analyze the signal from scales of 1 to 4. For noisy chromatograms, a scale range of 1–100 can be evaluated. Larger the scale size, higher is the magnitude, higher is the smoothing but less sharp the features of the n th derivative.
- (c) Generate a 3D plot of time, scale, and wavelet coefficients. Look at the output of individual scales.
- (d) Compare the chromatogram with the CWT output at a chosen scale starting from 1. If the CWT plot shows splitting or shoulder for a peak which appeared to a single peak, it is a strong sign that this is not pure peak and another component is most likely present in it.

General denoising procedure by DWT

- (a) Since DWT is done on multiresolution analysis, not all wavelets available in CWT are usable in DWT because of orthogonality requirements. Not all wavelets are orthogonal. Some common wavelets employed in DWT denoising are Haar wavelet, Symlet family, Coiflet family, and Daubechies family, among others. The user has a great flexibility in choosing the wavelet and any DWT wavelet can be tried.
- (b) From the given number of data points in the chromatogram, determine the maximum level of decomposition possible for a chosen wavelet. The maximum level is the largest level for which one coefficient is correct as the minimum requirement. The MATLAB command “wmaxlev” allows the maximum level of allowed decomposition.
- (c) Decompose the signal up to the allowed level. Look at the detail coefficients and do level by level thresholding of coefficients. There are several thresholding methods (Bayes, Universal, SURE, etc.) with an option of soft or hard thresholding. Hard thresholding can potentially generate small occasional spikes and soft thresholding output is smooth.
- (d) After thresholding the detail coefficients, reconstruct the denoised signal by adding the approximation and details (obtained from modified detail coefficients). If a certain wavelet does not give desired level of denoising, another wavelet can be tried.

in Eq. (6) with a mean, μ , of 0 and unit standard deviation multiplied by the standard deviation, σ , which scales the noise level.

Most state of the art HPLC/UHPLC software use Savitsky–Golay, digital RC-filter, and various varieties of moving average filters such as simple moving average, Gaussian weighted moving average, and Hamming weighted moving average [28]. From an analytical chemist's perspective, these traditional filters used in chromatography always lead to the broadening and deformation of the peak, albeit without affecting peak area. The conservation of peaks with an increase in peak width decreases the chromatographic resolution and lowers the peak height and sensitivity. It would be ideal to have an approach that can increase the signal-to-noise without distorting the signal. The goal here is to show that DWT is highly promising for denoising chromatographic data. As explained in Section 2.2, DWT denoising uses a level by level decomposition, modification of detail coefficients, followed by signal reconstruction, unlike traditional filters that rely on moving averages or fitting of polynomials.

One of the best visual ways to “see” the distorting effect of these filters is to run these filters on square waves instead of chromatographic peaks. The generation of a square wave can be easily understood as a light source that is rapidly turning on and off as shown previously [28]. These square wave patterns on an HPLC can be readily achieved by attaching a light-emitting diode on an HPLC detector flow cell as shown previously [28]; when the light is off, the “absorbance” is high, and when the light is on, the absorbance is “low.” Since the manufacturers do not explicitly state the noise suppressing filters, we simulate several commonly used filters and show these effects on a perfect square wave in the absence of noise. Figure 8 shows that the digital RC filter, which is used by a major HPLC manufacturer, distorts the square wave in two ways: first it introduces a tail and second it shifts the retention time to a longer time [29]. The next filter is the Gaussian weighted moving average with an averaging of 11 and 81 points. If we take a large number of averages, the square wave broadens but its retention time remains the same and the area is conserved. This problem would manifest itself as a broadened peak, that is, poor chromatographic resolution. Similar behavior is shown by the Hamming filter used by a major HPLC/UHPLC manufacturer. It does not change the retention time but broadens the peak of the Gaussian weighted moving average. The Savitsky–Golay is among the most popular filters. It works by fitting a polynomial in the chosen number of points. The user chooses a polynomial of order, 1, 2, 3 and so on. Herein, an order 3 polynomial was chosen. Figure 8 clearly shows that Savitsky–Golay filter generates highly undesirable side lobes near the steep rise and fall of the square wave. On the other hand, the DWT denoising process correctly reconstructs the square wave. A Haar wavelet was chosen with a level 6 decomposition. The detail coefficients

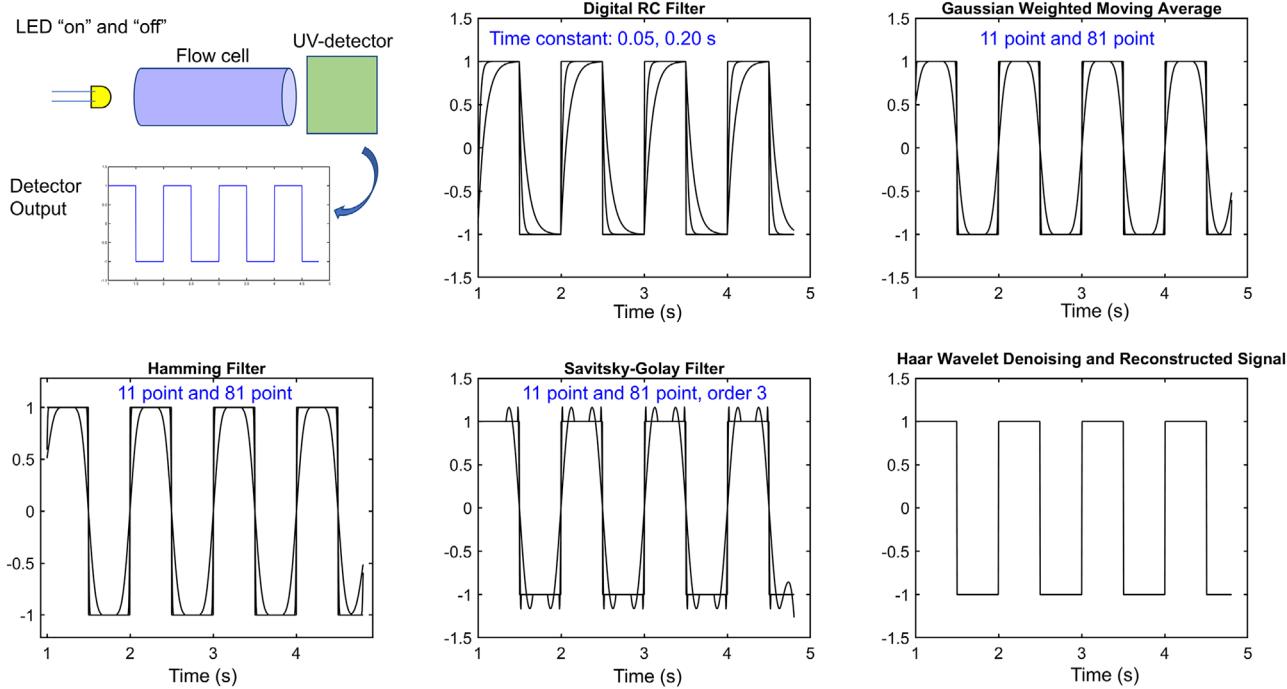


FIGURE 8 Comparison of conventional noise smoothing techniques used in HPLCs/UHPLCs with the DWT denoising process. The distortion of each square wave is evident with more smoothing; on the other hand, the DWT denoising maintains the original signal shape. See Supporting Information for the effect of the digital filters on noisy square waves

(cDs) were thresholded in MATLAB with a built-in function (see Supporting Information for the code). The reconstructed signal in Figure 8 is then the sum of approximation and details $A_6 + D_6 + D_5 + D_4 + D_3 + D_2 + D_1$. The procedure can be repeated on a noisy square wave and the lowest root mean square (rms) is found to be lowest with the DWT denoising procedure. A general outline for DWT denoising is presented in Table 1.

The advantages of wavelet denoising can be shown on a very low/S/N chromatogram, with highly asymmetric peaks, as shown in Figure 9A. The standard deviation of the noise is ± 0.6 units. The asymmetry at 10% ranges from 1.0 to 3.6. In such cases, running conventional filters, such as moving averages, may not be beneficial. It is clear that peak 6 is not easily detectable. Note that this is an extreme example with small heights and areas where one might be interested in trace impurities and their peak areas. It is highly challenging to integrate or even extract areas of peaks 3 and 6. Using DWT denoising techniques, a level 7 decomposition of coif5 wavelet denoises and reconstructs the chromatogram (Figure 9B), which is almost the same as the original ideal noiseless signal (Figure 9C). The recovered heights of peaks 1 to 6 are 4.308, 6.051, 1.219, 4.865, 5.682, and 0.668, which are very close to the original heights (4.864, 6.201, 1.372, 5.176, 5.842, and 0.818 respectively). The heights were obtained from PeakFit by considering the peaks as exponentially modified Gaussians. Chromatographers are often interested in peak areas for calibration purposes. The recovered areas were

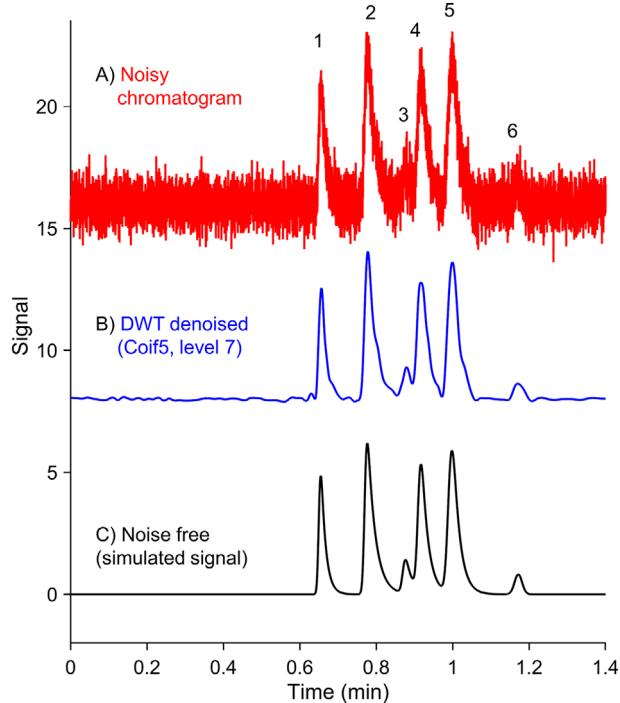


FIGURE 9 A denoising example on a very low S/N chromatogram with (A) asymmetric peaks, and small magnitudes of heights and areas of peaks 1–6. (B) A DWT coiflet 5 and level 7 decomposition was done. After thresholding, the detail coefficients using the command are described in the Supporting Information, and the signal was reconstructed. (C) Ideal noise-free chromatogram

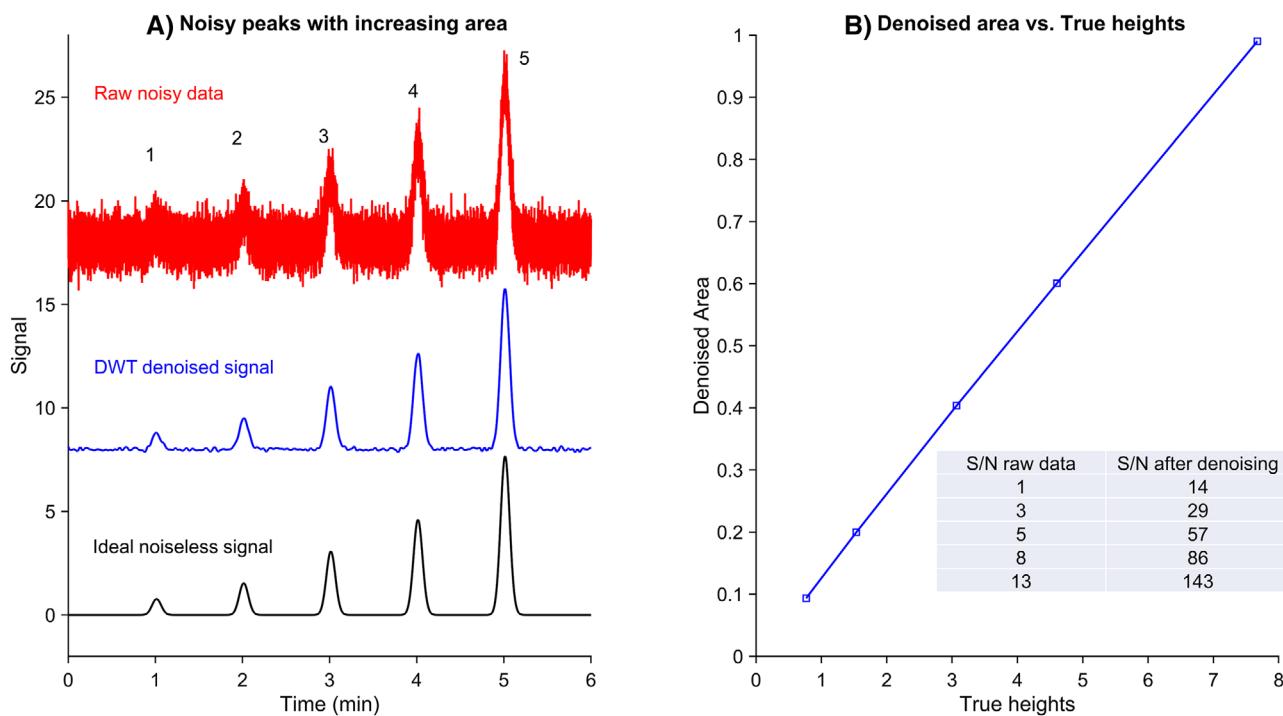


FIGURE 10 S/N gains with DWT denoising. (A) Top: Simulation of a raw noisy data representing injection of the same analyte at increasing concentration; middle: smoothed by Coiflet-5 wavelet; bottom: represent ideal noise-free signal. (B) Resulting calibration curve after denoising the raw data and measuring the recovered areas

0.109, 0.185, 0.040, 0.161, 0.197, and 0.022, whereas the true areas were 0.100, 0.191, 0.040, 0.158, 0.198, and 0.020 for peaks 1–6, respectively. The peak 6 was initially invisible in the original chromatogram.

The effect of DWT denoising can be assessed in terms of the gains in S/N and the recovery of peak areas and heights for quantitation. Figure 10A shows the simulation of an analyte peak, with increasing concentration. The lowest S/N for concentration “1” is 1. By conventional definition, this is below the detection limit; similarly, the concentration “2” is barely at the detection limit. Applying the same denoising protocol, as above, the S/N of the recovered injection is 14. This is a 14-fold increase in the S/N, which means that even five times lower concentration can still be analyzed. Figure 10B shows a linear relationship ($R^2 = 0.9998$) between true heights of the peaks and recovered areas, showing the remarkable power of DWT in denoising applications.

A significant advantage of the wavelets is that the smoothing parameters (widths, time constants, etc.) have to be optimized by the user for different signals in traditional smoothing process, but not for the wavelets. Wavelets may appear initially complicated internally, but they are eventually simpler and rewarding to use with some efforts.

5 | CONCLUDING REMARKS

Wavelet transforms have found wide applications in almost every branch of science of physics, analytical chemistry, and medical imaging. The purpose of this article was to introduce the main concepts of WT to separation scientists and encourage its use in advanced applications. Two types of wavelets (continuous and discrete) were introduced, which have different properties and applications. The CWT is used for detecting edges and discontinuities in image analysis. The coefficients of certain CWTs are related to n th derivatives and the “shape” of the derivative depend on the scale as well. The main advantage of CWT is that a smoothing function is built into the peak overlap detection. Next, a DWT was introduced, which finds primary applications in denoising of analytical signals. DWT is accomplished by a multiresolution algorithm that decomposes a signal into approximate and detail coefficients. A comparison of standard filters on square waves shows that all types of weighted moving averages and Savitsky–Golay filters distort sharp signals in their own unique ways. DWT works successfully in extremely noisy chromatograms and enhances S/N. These denoising procedures will help in lowering the LOD in challenging separations.

ACKNOWLEDGEMENTS

The authors thank Dr. Edward Voigtman (University of Massachusetts–Amherst) for a helpful feedback on this manuscript.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

ORCID

Muhammad Farooq Wahab 

<https://orcid.org/0000-0003-4455-2184>

REFERENCES

1. Kaplitz, A. S., Kresge, G. A., Selover, B., Horvat, L., Franklin, E. G., Godinho, J. M., Grinias, K. M., Foster, S. W., Davis, J. J., Grinias, J. P., High-throughput and ultrafast liquid chromatography. *Anal. Chem.* 2020, **92**, 67–84.
2. Xiang, P., Yang, Y., Zhao, Z., Chen, M., Liu, S., Ultrafast gradient separation with narrow open tubular liquid chromatography. *Anal. Chem.* 2019, **91**, 10738–10743.
3. Villalanti, D. C., Burke, M., Phillips, J., Frequency modulated correlation chromatography. *Anal. Chem.* 1979, **51**, 2222–2225.
4. Wahab, M. F., Gritti, F., O'Haver, T. C., Hellinghausen, G., Armstrong, D. W., Power law approach as a convenient protocol for improving peak shapes and recovering areas from partially resolved peaks. *Chromatographia* 2019, **82**, 211–220.
5. Wahab, M. F., O'Haver, T. C., Gritti, F., Hellinghausen, G., Armstrong, D. W., Increasing chromatographic resolution of analytical signals using derivative enhancement approach. *Talanta* 2019, **192**, 492–499.
6. Wright, N. A., Villalanti, D. C., Burke, M. F., Fourier transform deconvolution of instrument and column band broadening in liquid chromatography. *Anal. Chem.* 1982, **54**, 1735–1738.
7. Dasgupta, P. K., Chen, Y., Serrano, C. A., Guiochon, G., Liu, H., Fairchild, J. N., Shalliker, R. A., Black box linearization for greater linear dynamic range: The effect of power transforms on the representation of data. *Anal. Chem.* 2010, **82**, 10143–10150.
8. Wahab, M. F., Berthod, A., Armstrong, D. W., Extending the power transform approach for recovering areas of overlapping peaks. *J. Sep. Sci.* 2019, **42**, 3604–3610.
9. Goodman, K. J., Brenna, J. T., Curve fitting for restoration of accuracy for overlapping peaks in gas chromatography/combustion isotope ratio mass spectrometry. *Anal. Chem.* 1994, **66**, 1294–1301.
10. Ricker, N., The form and nature of seismic waves and the structure of seismograms. *Geophysics* 1940, **5**, 348–366.
11. Luisa, B. G., Wavelets: A Tutorial in Theory and Applications, Academic Press Inc., San Diego, CA 1992.
12. Bos, M., Hoogendam, E., Wavelet transform for the evaluation of peak intensities in flow-injection analysis. *Anal. Chim. Acta* 1992, **267**, 73–80.
13. Kauppinen, J. K., Moffatt, D. J., Cameron, D. G., Mantsch, H. H., Noise in Fourier self-deconvolution. *Appl. Opt.* 1981, **20**, 1866–1879.
14. Addison, P. S., The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, 2nd ed. CRC Press, Boca Raton, FL 2017.
15. Fugal, D. L., Conceptual Wavelets in Digital Signal Processing: An In-depth, Practical Approach for the Non-mathematician, Space & Signals Technical Pub., San Diego, CA 2009.
16. Nanavati, S. P., Panigrahi, P. K., Wavelet transform. *Resonance* 2004, **9**, 50–64.
17. Barhate, C. L., Wahab, M. F., Tognarelli, D., Berger, T. A., Armstrong, D. W., Instrumental idiosyncrasies affecting the performance of ultrafast chiral and achiral sub/supercritical fluid chromatography. *Anal. Chem.* 2016, **88**, 8664–8672.
18. Chau, F.-T., Liang, Y.-Z., Gao, J., Shao, X.-G., Chemometrics: From Basics to Wavelet Transform, John Wiley & Sons, Hoboken, NJ 2004.
19. Mallat, S. G., A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 1989, **11**, 674–693. <https://doi.org/10.1109/34.192463>
20. Barclay, V., Bonner, R., Hamilton, I., Application of wavelet transforms to experimental spectra: Smoothing, denoising, and data set compression. *Anal. Chem.* 1997, **69**, 78–90.
21. Chen, P., Lu, Y., Harrington, P. B., Application of linear and non-linear discrete wavelet transforms to MALDI-MS measurements of bacteria for classification. *Anal. Chem.* 2008, **80**, 7218–7225.
22. Cai, C. S., Harrington, P. D., Different discrete wavelet transforms applied to denoising analytical data. *J. Chem. Info. Comput. Sci.* 1998, **38**, 1161–1170.
23. Ehrentreich, F., Sümmchen, L., Spike removal and denoising of Raman spectra by wavelet transform methods. *Anal. Chem.* 2001, **73**, 4364–4373.
24. Misra, S., Wahab, M. F., Patel, D. C., Armstrong, D. W., The utility of statistical moments in chromatography using trapezoidal and Simpson's rules of peak integration. *J. Sep. Sci.* 2019, **42**, 1644–1657.
25. Mallat, S., A Wavelet Tour of Signal Processing, Elsevier Science, Amsterdam, The Netherlands 1999.
26. Shao, X., Ma, C., A general approach to derivative calculation using wavelet transform. *Chemom. Intell. Lab. Syst.* 2003, **69**, 157–165.
27. O'Haver, T., Begley, T., Signal-to-noise ratio in higher order derivative spectrometry. *Anal. Chem.* 1981, **53**, 1876–1878.
28. Wahab, M. F., Dasgupta, P. K., Kadjo, A. F., Armstrong, D. W., Sampling frequency, response times and embedded signal filtration in fast, high efficiency liquid chromatography: A tutorial. *Anal. Chim. Acta* 2016, **907**, 31–44.
29. Butterworth, G., Effect of detector response time on the distortion of spectral peaks of Gaussian shape. *J. Phys. E-Sci. Instrum.* 1968, **1**, 1165.

AUTHOR BIOGRAPHY



M. Farooq Wahab is a Research Engineering Scientist at the University of Texas at Arlington. After completing his PhD from the University of Alberta, Canada, he worked as a post-doctoral fellow with Daniel W.

Armstrong and assisted him in his start-up company. Research interests: separation science, new detection technologies, and signal processing. He serves on the editorial advisory board of *Chromatographia*. Wahab was selected among “Top 40 Under 40 Analytical Chemists” by the Analytical Scientist Magazine (UK). In Pittcon 2019, he received the “Young Investigator Award” from the Chinese American Chromatography Association.

How to cite this article: Wahab MF, O’Haver TC. Wavelet transforms in separation science for denoising and peak overlap detection. *J Sep Sci.* 2020;43:1998–2010. <https://doi.org/10.1002/jssc.202000013>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.