Practical Methods for Noise Removal: Applications to Spikes, Nonstationary Quasi-Periodic Noise, and Baseline Drift

Delphine Feuerstein, Kim H. Parker, and Martyn G. Boutelle*

Department of Bioengineering, Imperial College London, London, U.K.

A new approach to signal processing of analytical time-domain data is presented. It consists in identifying the types of noise, characterizing them, and subsequently subtracting them from the otherwise unprocessed data set. The algorithms have been successfully applied to three classes of noise commonly found in analytical signals: spikes, ripples, and baseline drift. Traditional filters have been used as an intermediary step to detect and remove spikes in the signal with 96.8% success. Adaptive ensemble average subtraction has been developed to remove nonstationary ripples that have similar time scales as the signal of interest. This method increased the signal-to-noise ratio by up to 250% and led to minimal distortion of the signal, unlike conventional Fourier filters. Finally the removal of baseline drift has been achieved by subtraction of a mathematical model for the baseline. These three methods are generic, computationally fast, and applicable to a wide range of analytical techniques. Full Matlab codes and examples are included as Supporting Information.

Experimental data are usually complicated by noise that may be due to interfering physical or chemical processes, instrumental noise, environmental noise, or any number of causes which result in spurious fluctuations of the signal generated by the detector. The main goal of the analytical chemist is therefore to devise techniques that provide a maximum signal-to-noise ratio (SNR). Efforts generally focus on maximizing the signal by improving the sensitivity, precision, selectivity, and limit of detection of the system. While the production of signal is well understood, the production of noise is not and solutions to minimize the sources of noise are not well-defined. Some recommendations include cleanliness to reduce the chemical noise, efficient shielding, and grounding to protect from environmental and instrumental noise.[1] However, in real world conditions, analytical systems are frequently operated very close to their limits of detection, where the signal-to-noise ratio (SNR) is small. This means that an otherwise small uncertainty in the background noise can induce a large error in the evaluation of the signal of interest.

One approach to this is a hardware approach where a second "sentinel" sensor, that responds to interfering signals but not the analytical signal, is placed alongside the primary sensor. Modern operational amplifiers can then efficiently reject common-mode noise at all frequencies within a wide operational bandwidth. This has been used successfully for implanted electrochemical biosensors.[2,3] Such an approach is very useful but depends on the availability of a suitable sentinel sensor that can be combined with the analytical system. Alternative approaches to increase the SNR of analytical signals have been based on digital signal processing methods. The most widely used smoothing filters among analytical chemists include least-squares polynomial fitting popularized by Savitzky and Golay in 1964,[4] and a wide range of Fourier-based filters.[5] Historically, a solution to noise has been very heavy damping or filtering. Such an approach relies on a characteristic time scale for the signal of interest, t_{signal}, being so long that a simple low-pass filter effectively removed all types of noise. However, with the move of analytical instrumentation to "real-time" or rapidly responding approaches such as capillary electrophoresis, lab-on-a-chip, microbore columns, and flow injection analysis, this is no longer tenable. The problem increasingly becomes the selection of the optimal smoothing parameters (polynomial degree and width of the moving data window for Savitzky−Golay filters and spectral cutoff for Fourier filters) that will achieve the best compromise between noise reduction and signal fidelity. Recent works have reported methods to select the *optimal* smoothing parameters, based on the maximization of the Shannon information entropy,[6] or statistical tests, including Pearsons's correlation[7] and *F* distribution[8,9] on the residuals.

Unlike these methods that aim at crushing the noise within the signal and assume that the noise has a Gaussian distribiution of zero mean,[6,8−12] this paper describes a range of practical solutions to improve the SNR by recognizing and removing real

* Author to whom correspondence should be addressed. E-mail: m.boutelle@imperial.ac.uk.

(1) Morgan, D. M.; Weber, S. G. *Anal. Chem.* **1984**, *56*, 2560–2567.
(2) Boutelle, M. G.; Stanford, C.; Fillenz, M.; Albery, W. J.; Bartlett, P. N. *Neurosci. Lett.* **1986**, *72*, 283.
(3) Kulagina, N. V.; Shankar, L.; Michael, A. C. *Anal. Chem.* **1999**, *71*, 5093–5100.
(4) Savitzky, A.; Golay, M. J. E. *Anal. Chem.* **1964**, 36.
(5) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*; Cambridge University Press: Cambridge, U.K., 1996.
(6) Larivee, R. J.; Brown, S. D. *Anal. Chem.* **1992**, *64*, 2036–2041.
(7) Browne, M.; Mayer, N.; Cutmore, T. R. H. *Digital Signal Process.* **2007**, *17*, 69–75.
(8) Barak, P. *Anal. Chem.* **1995**, *67*, 2758–2762.
(9) Jakubowska, M.; Kubiak, W. W. *Anal. Chim. Acta* **2004**, *512*, 241–250.
(10) Battistoni, C.; Mattogno, G.; Righini, G. *J. Electron Spectrosc. Relat. Phenom.* **1995**, *74*, 159–166.
(11) Lee, O.; Wade, A. P.; Dumont, G. A. *Anal. Chem.* **1994**, *66*, 4507–4513.
(12) Bromba, M. U. A.; Ziegler, H. *Anal. Chem.* **1984**, *56*, 2052–2058.

10.1021/ac900161x CCC: $40.75

world noise from the otherwise unprocessed signal. We here address three different types of noise that are frequently found in experimental data. The general approach consists in using prior knowledge about the noise to fit different noise types and subtract them from the otherwise unprocessed signal. We can thereby achieve both maximal noise reduction and minimal signal distortions. The three noise types amenable to this approach are classified according to their characteristic time scales relative to that of the signal of interest, t_{signal}: (1) Spikes are short signals with time scales (t_{spike}) very much shorter than the signal of interest, $t_{spike} \ll t_{signal}$. They can be larger or smaller in magnitude than the signal of interest. We will here distinguish "stereotypical spikes" that have a reproducible pattern, with a readily identifiable and constant time scale from other sorts of spikes that can include simple and complex spikes (i.e., clusters of simple spikes) of varying time scale. Stereotypical spikes can be inherent to the experimental process, such as sample injection in an analysis stream, while other spikes can be caused by various sources, such as noise on the power line, imperfection of electronic circuits, and environmental conditions. (2) Ripples are quasi-periodic signals that frequently have a similar time scale t_{ripple} to the signal of interest. Ripples can have any magnitude, distribution, and shape and can be nonstationary. They can arise, for example, from piston pump flow fluctuations, thermostatic valves, or cardiac and respiratory induced fluctutations in clinically related measurements. (3) Baseline drifts are signals that vary at a time scale that is large compared to the signal of interest: $t_{drift} \gg t_{signal}$. In particular it is assumed that the gradient of the drift is much less than the gradient of the signal of interest.

The characteristic time scale of the signal of interest t_{signal} is data-dependent. It can in general be defined as the rate limiting factor in the process, either 90% response time of the detector, the typical time scale of the measured phenomenon, or the full width at half-maximum (fwhm) for chromatographic peaks.

Three generic methods to remove these three classes of noise are described here, and their applications to a flow injection analysis assay is discussed as an example. They are relevant to a wide range of analytical time-domain signals where similar problems with noise are found.

EXPERIMENTAL SECTION

Data Processing. All algorithms discussed in the following sections were written using Matlab 7.2 and run on a Power Macintosh dual G5 computer (Apple). The codes and full details about the algorithms are given in the Supporting Information.

The programs are written as Matlab functions and can be directly implemented on the data provided: (1) the function *despiking1* removes stereotypical spikes; (2) the function *despiking2* removes nonstereotypical spikes; (3) the function *derippling* removes nonstationary ripples and uses the children functions *findtemplate*, *synchronisetemplate*, and *updatetemplate*; (4) the function *detrending* removes baseline drift and the exponential fitting is performed in *expfit*.

The original data used to generate the figures in this paper and some additional signals are available in the Supporting Information: (1) *Stereotyp1.mat* and *Stereotyp2.mat* are examples of stereotypical spikes; (2) *NonStereotyp1.mat* and *NonStereotyp2.mat* are examples of nonstereotypical spikes; (3) *Ripples1.mat*, *Ripples2.mat*, *Ripples3.mat* are examples of ripples; (4) *Trend.mat* is an example of trends. The window lengths used for the despiking method (stereotypical spikes) and the training data sets utilized for the derippling method are included in the Matlab data files.

Experimental Data Collection. The denoising algorithms can be applied to any type of experimental data. As an illustration in this paper, they are applied to data recorded using rapid sampling microdialysis (rsMD). This technique allows online measurements of the concentrations of glucose and lactate in the brains of patients with head injury. It basically consists in enzyme recognition of the analytes followed by a mediated flow injection analysis (FIA). A full description of the system can be found in refs 13 and 14.

RESULTS AND DISCUSSION

Despiking Method. *Algorithms.* Spikes often mask details of the true data, which can lead to misidentification of the signal of interest and subsequently significant error in the quantification of the signal. To avoid these false readings, it is essential to remove the spikes reliably without introducing distortion into the rest of the curve.

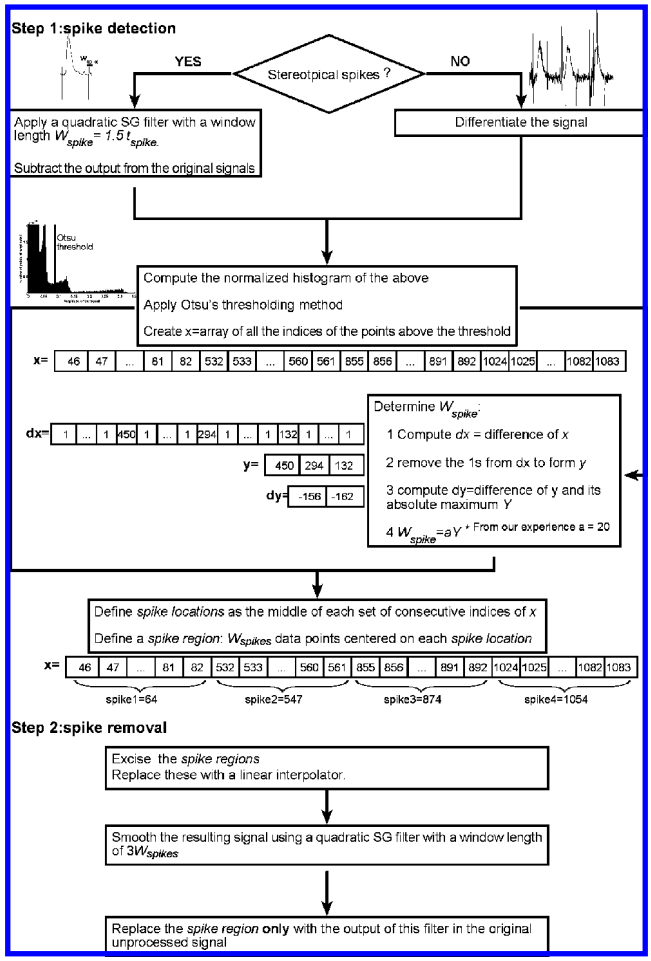
Most of the methods used for noise reduction are not effective for spike removal. Methods such as Savitzky−Golay (SG) smoothing[15] or adaptive degree polynomial filter[8] fail in the presence of spikes, as a result of trying to fit a polynomial function across discontinuities. Similar problems occur when using Fourier based algorithms, such as Butterworth or Chebyshev filters,[16] because the filters produce ripples at sharp edges. In fact, these methods result in the propagation of the spike to the preceding and following points and change the signal shape undesirably.

We here propose an alternative method that does not use the filtered data as the denoised signal but as an intermediate step for the localization of the spikes within the unprocessed raw data set. It is a two-step procedure that achieves both maximal removal of spikes and minimal distortions of the signal of interest: (1) detection of the position and width of a spike or spike train and (2) removal of the spikes in the original unprocessed signal by excision, linear interpolation, and smoothing of the identified spike region.

The main difficulty is the reliable identification of spikes. Manual detection of spikes is an easy operation, but automatic detection of short pulses is more complicated, especially when the data are complicated by other forms of random noise. Several different types of spike detection methods are described in the literature, usually based on outlier detection algorithms, as the spike may be treated as an outlier from the correct experimental signal. These methods are generally based on examining higher moments, in practice skewness and kurtosis.[17] Nonlinear smoothing algorithms have also been applied to Raman spectra.[18,19] We

(13) Parkin, M.; Hopwood, S.; Jones, D.; Hashemi, P.; Landolt, H.; Fabricius, M.; Lauritzen, M.; Boutelle, M.; Strong, A. *J. Cereb. Blood Flow Metab.* **2005**, *25*, 402–413.
(14) Hopwood, S.; Parkin, M.; Bezzina, E.; Boutelle, M.; Strong, A. *J. Cereb. Blood Flow Metab.* **2005**, *25*, 391–401.
(15) Madden, H. H. *Anal. Chem.* **1978**, *50*, 1383–1386.
(16) Horlick, G. *Anal. Chem.* **1972**, *44*, 943–947.
(17) Huber, P. J. *Ann. Math. Stat.* **1972**, *43*, 1041–1067.
(18) Hill, W.; Rogalla, D. *Anal. Chem.* **1992**, *64*, 2575–2579.
(19) Phillips, G. R.; Harris, J. M. *Anal. Chem.* **1990**, *62*, 2351–2357.

Scheme 1. Algorithm 1: Despiking



here propose two alternative methods that are easy to implement and fast to compute. They both exploit the main property of spikes: their short time scale, t_{spike}, as compared to the data. The algorithm is presented in detail in the flowchart "Algorithm 1: Despiking" in Scheme 1, and the Matlab codes are given in Supporting Information.

The first step is the detection of the spikes in the data. Two methods have been developed. The first one is applicable to stereotypical spikes whose time scale t_{spike} can be directly estimated from the data set. It is based on locating the spikes by the subtraction of a quadratic SG filtered curve from the original data set. Using the known t_{spike}, one can significantly attenuate the spikes using a window width $W_{spike} \geqslant t_{spike}$, typically $W_{spike} = 1.5t_{spike}$. However, the use of such a window also leads to significant distortions of smaller scale signal features, such as the very top of a chromatographic peak, thus reducing the corresponding peak height. Here we subtract the filtered trace from the original trace to obtain an intermediate signal in which only regions of fast dynamics, in particular spikes, and smoothing errors remain. The second method is applicable to any sort of spikes. It is based on the derivative of the signal. As spikes are rapidly varying signals relative to the signal of interest, their slopes are generally larger than those of the signal of interest. So when the derivative of the signal is taken, large components will be ascribed to spikes in the original data. In both methods the spikes are subsequently detected using a simple thresholding detection in the difference signal (step 1a) or the derivative signal (step 1b). We have used Otsu's algorithm[20] to determine the optimum threshold as this is simple and fast to compute. With the use of a histogram of the difference signal, it calculates the threshold that best separates two classes of data points: spikes and other forms of random noise in our case. The optimum threshold is defined as the one that minimize the within-class variance.

The second step is the removal of the detected spikes. Each identified spike is first excised from the original signal, and the excised section is replaced by a linear interpolation. The resulting signal is then smoothed using a traditional quadratic SG filter with a window length $W = 3W_{spike}$. The output is then used to replace *only* the spikes in the original unprocessed signal (step 2).

Application and Discussion. The first detection method is computationally faster than the second one but requires setting the window length W_{spike} using *a priori* knowledge of t_{spike}. The minimal window length W_{spike} should be $W_{spike} = 1.5t_{spike}$. If a smaller window length is used, some spikes will be missed in the subtraction step. When using a longer window length, a somewhat larger region of the data set is affected by the spike removal step (step 2). For data where the signal of interest is composed of peaks, the window length W_{spike} should be less than 0.7 fwhm of the narrowest peak of interest. This causes minimal distortion of the peaks by an SG filter according to ref 21. In cases where t_{spike} is not easily determined, then the second method is preferred.

The polynomial order of the SG filter used in the spike detection step (step 1a) is not critical as signal integrity is not necessary in the spike identification step. It is preferable to select the order of the polynomial that will best approximate the signal of interest for the second SG filter of width W_2 used in the spike removal step (step 2). Adaptive degree polynomial methods can be used for this purpose.[8] For chromatographic data, a quadratic filter is usually suitable, although the choice is also not very critical since it affects only a very limited segment of the final signal.

This despiking procedure has been successfully applied to clinical data recorded with rsMD. An illustrative case is given in Figure 1 for both stereotypical spikes (Figure 1a) and nonstereotypical spikes (Figure 1b).

In both cases, the original signal (upper trace) is distorted by the presence of spikes. In the case of nonstereotypical spikes in particular, the spikes are quite diverse with different magnitudes, both positive and negative. The output of the despiking methods is shown in the lower traces. A total of 100% of the spikes are removed in the case of stereotypical spikes, while 93% of the nonstereotypical spikes are removed. Even in the worst-case scenario, where a spike occurs at exactly the same time as a peak, the algorithm performs well. This can be seen in Figure 1b in the peak at approximately 50 s. In practice, out of the 10 clinical data records processed this way, which represents approximately 19 days of data, i.e., 10^8 data points with about 10^4 spikes, the despiking method eliminated 90% of the nonstereotypical spikes and 98% of the stereotypical spikes. Overall, 96.8% of all the spikes in our data were removed. From an analytical perspective, the despiking procedure enables reliable automatic

(20) Otsu, N. *IEEE Trans. Syst., Man, Cybernet.* **1979**, *9*, 62–66.
(21) Willson, P. D.; Edwards, T. H. *Appl. Spectrosc. Rev.* **1976**, *12*, 1–81.

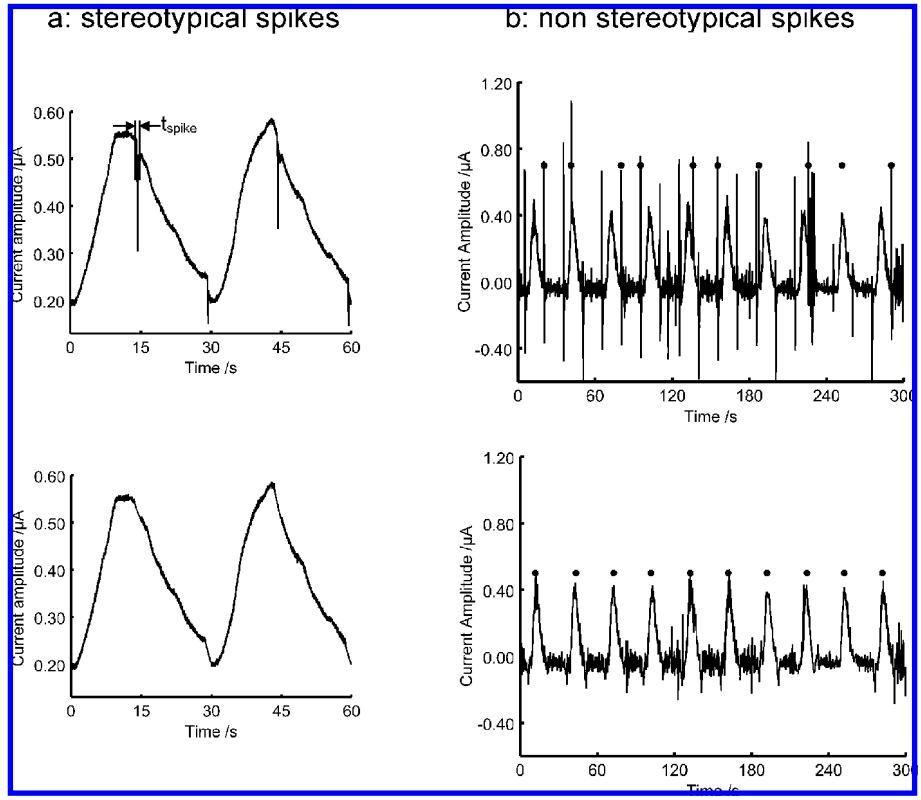


Figure 1. Application of the despiking method to real clinical data. (a) Stereotypical spikes: upper trace, example of data sets with stereotypical spikes. The characteristic time scale of the spikes, t_{spike}, can readily be determined by inspection of the data: t_{spike}= 0.835 s or 167 data points (sampling at 200 Hz). Using an SG filter with a window length $W_{spike} = 1.5t_{spike} = 1.25$ s or 251 data points, enables 100% spike detection. These are then removed from the original unprocessed signal. Lower trace: output of the despiking technique for stereotypical spikes. Only the spike region surrounding a detected spike has been smoothed while the rest of the signal has been left intact. (b) Nonstereotypical spikes: upper trace, example of data with nonstereotypical spikes. Within 5 min of data, there are more than 20 spikes, with different widths, amplitudes, and patterns. The dots above the trace show the output of a local maximum detection algorithm (from Chart 5.6, ADInstruments). It fails to detect the analyte peaks, confusing them with the spikes. Lower trace: output of the despiking method for nonstereotypical spikes applied to the data above. A total of 93% of the spikes have been successfully removed. The remaining spikes have much lower amplitudes. As a result, the same local maximum detection technique successfully read 100% of the analyte peaks and hence was able to evaluate their respective peak heights for further analysis.

estimation of the peak heights using an automatic local maximum detection algorithm embedded in the data acquisition software (Chart 5.6, ADInstruments). In the original data, most of the analyte peaks detected by the algorithm are in fact spikes (upper trace). In the despiked data (lower trace), all of the analyte peaks are identified correctly. This proved invaluable for further clinical interpretation of the data collected with rsMD.[22]

Comparison to Other Methods. Another spike removal technique is the moving median filter. In this method, the central value of an interval sliding along the curve is substituted by the median of the data in the interval. Originally described by Tukey,[23] it has been applied to remove stereotypical spikes. However, it leads to considerable distortions of the curve shape, giving flat-topped truncated peaks,[24] which has restricted its use in practice. We nevertheless tested it in the despiking method for stereotypical spikes (available in the Supporting Information). We replaced the SG filter used for the spike detection step with a moving median filter with a window length half of W_{spike} to effectively remove them. Then, we subtracted this filtered data from the original data set and removed the identified spikes as described previously. It gave very similar results to our method, with 98% spike removal and less than 5% signal distortion. However, it was computationally less efficient, usually 10 times slower than SG filters, which agrees with Stone's observations.[24]

A more recent algorithm designed for spike removal,[25] based on a gross error statistical test, has been successfully used on voltammetric data, giving satisfactory spike removal while not distorting the curve. However, like the median filter, it requires sorting the data into either ascending or descending order, which makes it computationally slower than the despiking method presented here. In comparison to other despiking methods we have tried, the main advantage of our method is its computational speed, which makes it appropriate for real-time online processing as data are collected.

Derippling Method. *Algorithm.* Ripples are defined as quasi-periodic fluctuations. They are neither random nor sinuslike and

(22) Feuerstein, D.; Manning, A.; Hashemi, P.; Bhatia, R.; Tolias, C.; Fabricius, M.; Parker, K.; Strong, A.; Boutelle, M. In *Monitoring Molecules in Neuroscience: Proceedings of the 12th International Conference on in Vivo Methods 2008*; Phillips, P. E. M., Sandberg, S. G., Ahn, S., Phillips, A. G., Eds.; Vancouver, Canada, 2008; pp 39−42.
(23) Tukey, J. W. *Exploratory Data Analysis*; Addison-Wesley: Reading, MA, 1977.
(24) Stone, D. C. *Can. J. Chem.* **1995**, *73*, 1573–1581.
(25) Jakubowska, M.; Kubiak, W. W. *Electroanalysis* **2005**, *17*, 1687–1694.

can be nonstationary, changing in morphology and distribution with time. As a further complication, they can have a characteristic time scale t_{ripple} very similar to that of the signal of interest. In the case of a periodic signal, for example, chromatographic peaks, the similar time scale between the signal and ripple can lead to aliasing of the ripples on the peaks of interest. This affects the baseline and, over time, all parts of the peak. This in turn leads to slow variations in measures of the peak, such as height and peak area. To limit this problem, it is important to remove the ripples without compromising the signal of interest.

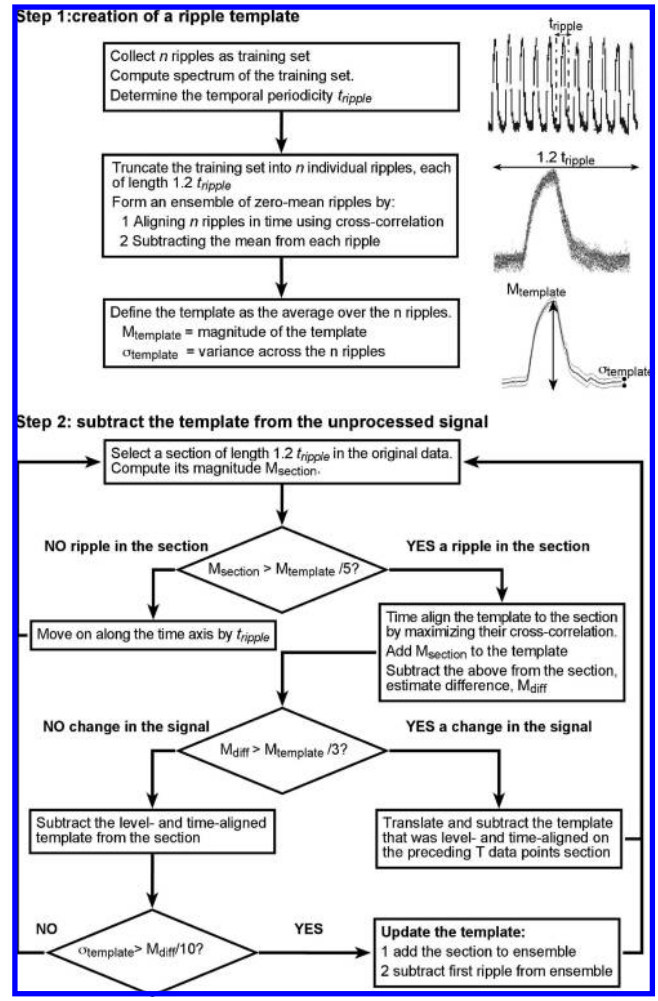
The most commonly used technique to remove periodic noise is based on Fourier filters.[26] This method works reasonably well when the typical time scale of the signal of interest is long, such as the broad analyte peaks in the case of traditional liquid chromatography.[27] However, with the introduction of faster chromatographic methods, the separation between noise and signal in the spectral domain is not so clear. In this case, it is very difficult to determine the optimal spectral cutoff that will remove noise without compromising signal integrity.[16] In any case, Fourier filters will inevitably fail when noise and signal have the same time scales so that they overlap in the frequency domain. In such cases any Fourier filter will lead to dramatic distortion of the signal of interest, especially flattening and broadening of the analyte peaks, and therefore cannot be used in practice. An example of the effect of a low-pass Fourier filter is given in Figure 2a (light gray trace on the right panel). Another main drawback of Fourier filters and other linear filters in general is their limited success with nonstationary signals.

An alternative method using the quasi-periodicity of the ripples is proposed here. It consists in estimating the ripples with a template (obtained by averaging a training data set) that is a good representation of the noise. This template is then subtracted from the original unprocessed signal. Careful alignment of the template onto the signal is required to avoid distortion of the signal of interest. A flowchart with the details of the algorithm can be found in "Algorithm 2: Deripplling" in Scheme 2.

In case of a change in the signal of interest, such as a chromatographic peak, occurring at the same time as a ripple, it is assumed that the ripples are locally repetitive and the previous aligned ripple is subtracted. This procedure should ensure both that there is no contribution of the signal of interest when fitting the underlying ripple and that it gives the best chance of effectively zeroing the noise on the signal of interest.

Importantly, after each subtraction step, the template shape is updated to account for any changes in the patterns of the ripples. The decision to update the template with new ripples is determined by the variance of the template. This is calculated during the subtraction step. If this variance is much smaller than the amplitude of the ripples (typically 10 times smaller than the amplitude of the ripples), then the template is still adequate for subtraction. If the variance is above this critical value, then the template needs to adapt to the new morphology of the ripples. In this case, new ripples are added to the ensemble as the algorithm moves along the trace and the "oldest" ripples are removed to keep the size of the ensemble constant. If, after 20 updates of the ensemble, the variance is consistently more than half the size of the ripples, the ripples are not stable enough for this method.

Scheme 2. Algorithm 2: Deripplling



Application and Discussion. The computation of the representative template requires a training data set with n typical ripples. For this purpose, it is preferable to have a section of data with no signal of interest before the actual recording starts, but the algorithm can be trained using already recorded data as well. Such training data can be found in the Matlab data files given in the Supporting Information. The number of ripples needed to have a good representation of the noise will depend on the initial SNR and the regularity of the ripples. From our experience, n of the order of 20 suffices even for poor SNR (less than 3). If the data set has spikes, the despiking method should be applied first to avoid the presence of spikes within the training set, which would otherwise prevent the detection of the periodicity of the ripples.

The template is slightly longer than t_{ripple}. This is to ensure there is a sufficient number of data points overlapping at the "junction" between two consecutive subtraction steps. From our experience, a template length of $1.2t_{ripple}$ is sufficient.

Continual updating of the template as the subtraction occurs ensures that any nonstationarities in the ripples are accommodated. The only assumption here is that the ripples superimposed on a signal change have a local consistency and a similar pattern to the ripple just preceding this change.

The deripplling method has been applied to 10 clinical traces and achieved significant enhancement of the signal-to-noise ratio

(26) Paley, R.; Wiener, N. *Fourier Transforms in the Complex Domain*; American Mathematical Society: Providence, RI, 1934.
(27) Kaiser, J. F.; Reed, W. A. *Rev. Sci. Instrum.* **1977**, *48*, 1447–1457.

(SNR), thus considerably reducing the error when quantifying the signals. Some examples are presented in Figure 2 (right panel). The baseline is effectively zeroed by the subtraction step. All parts of the peak are restored and can be fully fitted by a double-Gaussian peak function as described in ref 28. The SNR, calculated as the ratio of peak height over the standard deviation of the noise, is globally improved by a factor of 2.5 (Figure 2a) to 3.5 (Figure 2b). This does not depend on the initial SNR level, and the method performs equally well with very low initial SNR (Figure 2b). The improvement on the SNR has a direct impact on the analytical assay. It significantly reduces the error when estimating the magnitude of the peaks. When a calibration with $n = 5$ peaks per concentration is performed, the relative standard deviation within the peaks (calculated as the standard deviation of the peak heights over their average absolute amplitude for each concentration) is reduced by up to 136%. As a result, there is less than 2% variation instead of 5.4% when the analyte levels reach 250 μM. This means that the limit of detection of the assay improved from 70 to 25 μM (Figure 2c).

Comparison to Other Methods. As previously mentioned, the deripplling method works well even for nonstationary signals where traditional linear filters usually fail. Some nonlinear methods have been developed to overcome the problems associated with nonstationarity. The most common of these methods is based on wavelets thresholding. The wavelet schemes involve projecting the time-domain signal into a wavelet domain where the basis functions are wavelets (Daubechies or Haar for instance).[5] Contrary to sine and cosine (the Fourier basis functions), the wavelet functions are dually localized in both time and frequency domains, so that wavelet filters can be used in the time domain. Smoothing here relies on the fact that the energy of a signal will often be concentrated in a few coefficients in the wavelet domain while the energy of noise is spread among all coefficients.[29] Wavelet noise filters are constructed by calculating the wavelet transfer for a signal and then applying an algorithm that determines which wavelet coefficients should be modified (usually by being set to zero). Wavelet denoising is very efficient in cases where the wavelet basis functions best fit noise and signal of interest. As the bases functions are fixed, they do not necessarily match the wide variety of real-world analytical signals. Unlike the wavelet approach, the deripplling method developed here is derived from the recorded noise itself. There is no assumption about the shape and distribution of the signal, which makes this method much more versatile and applicable to a much wider range of signals.

Another approach that is also completely data driven is empirical mode decomposition (EMD). In this scheme, the modes are derived from the signal itself, based on the sequential extraction of energy associated with various intrinsic time scales of the signal starting from finer temporal scales to coarser ones.[30] The problem is that the time scale of the ripples can be the same as that of the signal of interest, so that the ripples will be in the same mode as the signal itself and the EMD approach will thereby fail in separating ripples from signal. The approach used in this paper avoids this pitfall.

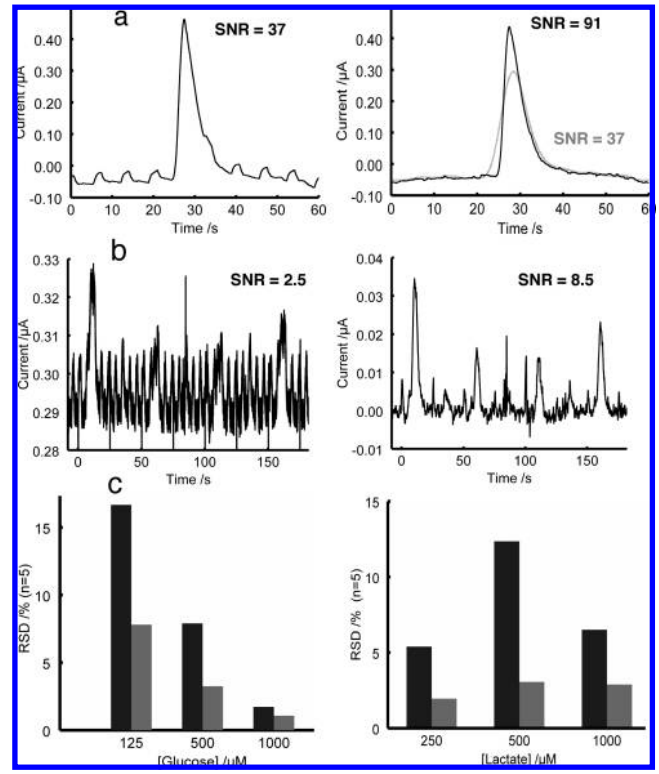


Figure 2. The deripplling method significantly improves the signal-to-noise ratio. (a) Left panel, original unprocessed signal, with a moderate SNR. Right panel: outputs of a low-pass Fourier filter (cutoff of 0.1 Hz) (gray line in the background) and of the deripplling algorithm (dark line in the foreground). The Fourier filter effectively attenuate the ripples but also significantly distorts the peak, making it smaller and broader. There is no improvement in the signal-to-noise ratio. The deripplling method reduces the ripples and does not affect the peak shape and height, hence improving the SNR by 250%. (b) Left panel: original unprocessed signal, with a poor SNR. Right panel: output of the deripplling algorithm. The SNR is improved by 350%, the peaks are identifiable, and a clear fall and rise in peak heights is seen. (c) Bar plots of the relative standard deviations for the raw signals (dark bars) and the derippled signal (light bars) when using three different standard concentrations for glucose (left panel) and lactate (right panel). The variations on the peak height estimations in the derippled data are reduced in the range 38−136%.

The main advantage of our method is that, once sufficient training data set has been recorded (which can be prior to the actual measurement), deripplling can be performed in real time while the wavelet filtering and empirical mode decomposition are postacquisition approaches that transform the entire data set as a complete block. Like the despiking method, the deripplling algorithm is computationally fast: it takes approximately 50 s to process a trace with 10^6 data points that included about 10^3 ripples. It is therefore also suitable for online signal processing.

Detrending. *Algorithm.* For completeness, the same approach has been used to remove baseline drift, which is noise that has a characteristic time scale that is longer than that of the signal of interest, commonly referred to as low-frequency baseline drift. This has been a main issue in traditional chromatography and a large number of digital methods have been developed for baseline correction. These include Fourier-based filters, whereby high-pass filters can remove the background drift assuming that the frequency components of this drift are in a lower range than those of the signal of interest. It was first used in spectroscopy by Atakan

(28) Li, J. *Anal. Chem.* **1997**, *69*, 4452–4462.
(29) Barclay, V. J.; Bonner, R. F.; Hamilton, I. P. *Anal. Chem.* **1997**, *69*, 78.
(30) Boudraa, A. O.; Cexus, J.; Saidi, Z. *Int. J. Signal Process.* **2004**, *1*, 33.

et al.[31] The problem is the selection of the filter parameters: if the frequency cutoff is too low, some residual baseline drift will remain in the baseline-corrected spectra; if the spectral cutoff is too high, some of the signal peak frequency components will be removed along with the baseline and the signal peaks will be reduced. Another very common approach is based on differentiation. The slopes of the background are in fact generally lower than those of signal peaks, so taking the first derivative will discriminate against slowly varying background components irrespective of their absolute values.[32] However, such methods lead to peak distortions with 90° phase shifts and attenuation of the signal.[33] To limit signal distortion, other methods have been devised, based on a SG filter subtraction[34] or median filter subtraction.[33] With both of these methods, the peaks of interest are smoothed out using a large window for an SG filter or a moving median filter. The output of this filter is therefore the baseline and its drift, without the peaks of interest. When this filtered data is subtracted from the original signal, it results in peaks on a flat baseline. The limitation is to determine the right window length that will preserve the broader peaks while removing the baseline from the sharper peaks, essentially the same problem as in high-pass filtering. Many other techniques have been developed: using wavelet filtering, Gaussian filtering, artificial neuronal networks, etc. A good review can be found in ref 32.

Here, we propose a strategy that is similar to the homomorphic procedure described in 2001 by Michel et al.[35] It is a two step process and consists of (1) fitting the noise with a mathematical function and (2) subtracting this model from the original data.

A huge variety of models can be used, depending on the nature of the drift, including polynomials, exponentials, or power laws. Detrending is particularly applicable to chromatography derived methods, including FIA, capillary electrophoresis, separation techniques using gradients of phases, where sections of data with no injection of the sample and only baseline drift alternate with sections containing the analyte peaks.

Application and Discussion. This method has effectively been applied to clinical data recorded with rsMD using an exponential approximation of the baseline drift ($I = B - Ae^{-(t)/(T)}$ where I is the drift signal and A, B, and T are constants). The coefficients A, B, and T were determined using an unconstrained nonlinear optimization technique based on the Nedler−Mead simplex technique.[36] An example of the results obtained with the detrending method is shown in Figure 3. The peaks are readily identifiable and a baseline can be defined to determine the amplitudes of the peaks and quantify the analyte concentrations after detrending.

It is worth noting that the model coefficients were recomputed for each section of baseline drift preceding a section with analyte peaks. This allows for nonstationarities in the baseline drift.

In this particular example, the despiking and detrending methods were combined. The first step of the despiking procedure, i.e., the spike identification step, was used to avoid any spikes in the estimation of the baseline drift. That would have led to nonconvergence of the Nedler−Mead method.

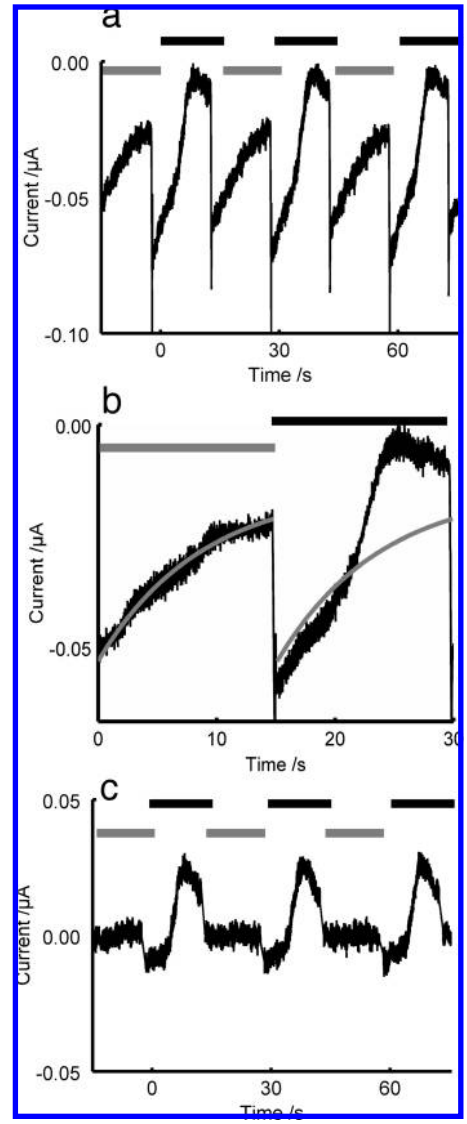


Figure 3. Detrending method: (a) Raw data consisting of alternating 15 s of only baseline drift (indicated by the gray block bars) followed by 15 s of baseline drift with an analyte peak (indicated by the black block bars). Injection spikes can also be seen every 15 s. (b) Close-up of 30 s from the trace in part a. The first 15 s are fitted to an exponential model: $I = B - Ae^{-(t)/(T)}$ (solid curve). This mathematical function is subsequently subtracted from the "no injection" interval and from the following 15 s when the analyte is injected. The exponential function is re-evaluated each time there is a "no injection" segment with baseline drift only. (c) Output of this detrending procedure followed by the despiking method on the data set shown in part a. The peaks are identifiable and their amplitudes quantifiable automatically. Derippling could also be applied as the restored flat baseline show evidence of quasi-periodic ripples.

This *ad hoc* method can be generalized to other sets of data, using different modeling functions. It can also be combined with other methods presented here to remove other types of noises. In our case, the detrending procedure has allowed the quantification of high value clinical traces that could not possibly be analyzed without removal of the drift. However, other methods for baseline correction may be preferable. This will depend on factors such as the nature of the measurement and noise, the magnitude of the signal-to-noise ratio, the

(31) Atakan, A. K.; Blass, W. E.; Jennings, D. E. *Appl. Spectrosc.* **1980**, *34*, 369.
(32) Schulze, G.; Jirasek, A.; Yu, M. M. L.; Lim, A.; Turner, R. F. B.; Blades, M. W. *Appl. Spectrosc.* **2005**, *59*, 545–574.
(33) Moore, A. W.; Jorgenson, H. W. *Anal. Chem.* **1993**, *65*, 188–191.
(34) Carbonneau, R.; Bolduc, E.; Marmet, P. *Can. J. Chem.* **1973**, *51*, 505–509.
(35) Michel, J.; Bonnet, N. *Ultramicroscopy* **2001**, *88*, 231–242.
(36) Nedler, J. A.; Mead, R. *Comput. J.* **1965**, *7*, 308–313.

computational and programming resources available, and the time scale for the required calculations.

CONCLUSIONS

We have here presented three different methods to remove three particular types of noise that are encountered in many analytical techniques: spikes, quasi-periodic nonstationary ripples, and baseline drift. All of these techniques are based on identifying specific patterns of the noise types and then subtracting the fitted noise from the original unprocessed signal. Key to the success of the methods presented here is the correct identification of these noise types based on their time scale relative to the analytical signal. These filtering methods effectively improve the signal-to-noise ratio with minimal distortion of the signal of interest. They can be combined successively or simultaneously for denoising, permitting reliable subsequent automatic quantification of the data. Some parameters, including the length of the SG filter for the despiking procedure, the periodicity of the ripples, and the mathematical function describing the drift, have to be derived using *a priori* knowledge of the noise. In many cases, they can be easily determined from the recorded noise itself. As these signal processing tools are fully data-driven with no assumptions about the signal of interest, they can be applicable to a wide variety of time domain analytical signals.

ACKNOWLEDGMENT

We would like to acknowledge King's College Hospital for making the clinical monitoring with our assay possible. In particular, we would like to thank Professor Anthony Strong and Dr. Andrew Manning for their support and help.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review January 22, 2009. Accepted April 30, 2009.

AC900161X