A Wavelet-based Optimal Filtering Method for Adaptive Detection: Application to Metallic Magnetic Calorimeters

B. Censier, M. Rodrigues and M. Loidl

CEA, LIST, Laboratoire National Henri Becquerel (LNE-LNHB), F-91191 Gif-sur-Yvette, France

Abstract. Optimal filtering allows the maximization of signal-over-noise ratio for the improvement of both energy threshold and resolution. Nevertheless, its effective efficiency depends on the estimation of signal and noise spectra. In practice, these are often estimated by averaging over a set of carefully chosen data. In case of time-varying noise, adaptive non-linear algorithms can be used if the shape of the signal is known. However, their convergence is not guaranteed, especially with 1/f-type noise. In this paper, a new method is presented for adaptive noise whitening and template signal estimation. First, the noise is continuously characterized in the wavelet domain, where the signal is decomposed over a set of scales, corresponding to band-pass filters. Both time resolution and decorrelation properties of the wavelet transform allow an accurate and robust estimation of the noise structure, even if pulses or correlated noise are present. The whitening step then amounts to a normalization of each scale by the estimated noise variance. A matched filter is then applied on the whitened signal. The required signal template is constructed from a single event, denoised by a filtering technique called wavelet thresholding. As an example, application to metallic magnetic calorimeter data is presented. The method reaches the precision of conventional optimal filtering, further allowing noise monitoring, adaptive threshold and improving the energy resolution of up to 8% in some cases.

Keywords: wavelet, optimal filtering, adaptive filtering, matched filter, metallic magnetic calorimeter.
PACS: Replace this text with PACS numbers; choose from this list: http://www.aip.org/pacs/index.html

INTRODUCTION

We propose a wavelet-based method for adaptive noise measurement on a signal containing pulses and for template signal construction. The main idea is to take profit of the discrimination features of signal and noise in the time/scale plan. We first briefly recall the main properties of wavelet transform. The method is presented in a second section, followed by some example applications on metallic magnetic calorimeter data in the last section.

WAVELET TRANSFORM

A wavelet transform is a decomposition on a set of time localized oscillating functions basis $\psi_{j,k}$, instead of sine and cosine functions basis used for Fourier transform. The j index stands for time while k stands for scale. The wavelet transform of a signal is thus composed of several time series, each of them being the signal content at a given scale, i.e. in a given frequency band with a signal processing point of view (see figures 1 and 2). While the Fourier transform of a signal has a constant frequency resolution and a complete loss of time resolution, its wavelet transform keeps both in a somewhat optimized way. In analogy with the Heisenberg's uncertainty principle, large time scales have a better frequency resolution and a worse time resolution than small time scales. Besides this general property, a wavelet functions basis $\psi_{j,k}$ must match several mathematical constraints. Among them let us cite orthogonality, built-in scaling (all basis functions are translation and dilatation of each other), Fourier compact support (all basis functions can be seen as band-pass impulse responses) and number of vanishing moments[1]. Having p vanishing moments means that wavelet-coefficients for p^{th} order polynomial will be zero. This property is strongly connected to the decorrelation properties of wavelet, allowing a signal to be represented with few non-zero wavelet coefficients, and an approximate whitening of long range correlated noise at each scale[2]. These features are extensively used for image denoising and compression: wavelet coefficients at each scale are set to zero when being lower than a given threshold, and

the inverse transform gives the denoised image. This "wavelet thresholding" technique is often seen as a nearly optimal smoothing method. It can be used on a decimated discrete wavelet transform that allows the computing time to be as low as the fast Fourier transform one. This is due to the use of a dyadic scale grid, where scale k is twice the scale k-1, and where the signal at scale k is decimated k times[9]. Although this does not imply any information loss, the main drawback is the loss of time invariance that can generate non-linear artifacts on the thresholded signal[7]. This problem can be overcome using a non-decimated wavelet transform on a dyadic grid, at the cost of computing time. Wavelet thresholding is thus a very efficient, almost non-parametric tool for smoothing an unknown shape buried in noise, but it is non-linear and often ill-adapted to measurements requiring precise determination of specific parameters such as pulses amplitude.

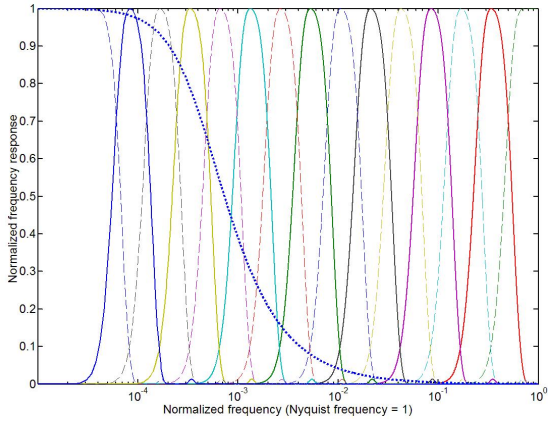


FIGURE 1. Frequency responses for a filter bank implementation of a Daubechies 8 wavelet transform (thick solid lines and dashed lines). The thick dotted line is the Fourier spectrum of a typical pulse. Note the logarithmic scale of the frequency axis.

WAVELET-BASED METHOD

Let us first recall the definition of a matched filter. It is constructed from a template pulse Fourier spectrum S(f) and a noise power spectral density N(f) and is theoretically optimal in the sense of maximizing the instantaneous signal-over-noise ratio (SNR in the following)[3]. Its transfer function H(f) is given by:

$$H(f) = K\frac{S^*(f)}{N(f)} \qquad (1)$$

where K is normalization constant. This can be interpreted in practice by a whitening filter (the N(f) denominator) followed by a correlation of the signal with the time reversed template pulse (the $S^*(f)$ numerator). The main advantage about implementing a matched filter using wavelet is the possibility of being adaptive by evaluating continuously the structure of the noise. An example application and a discussion about the advantages of an adaptive implementation of the matched filter in the wavelet domain can be found in [5]. In our implementation though, only the estimation of N(f) and the whitening process is done in the wavelet domain, while the correlation is done afterward back in the time domain. This choice is mainly driven by the reduction of computing time as will be explained further.

Thanks to the decorrelation properties of wavelets, the wavelet transform of a Gaussian correlated noise at each scale k is a centered, nearly white Gaussian noise [2] with a variance σ_k^2, which is also the noise power at scale k. Once this variance is calculated at each scale, the whitening step amounts to a simple normalization of scale k by σ_k^2.

The question is: how to compute these variances? We can once again use the decorrelation properties of wavelets. As stated above, the wavelet transform of a pulse has only few non-zero coefficients, contrary to the wavelet transform of the noise (see figure 2). We can then estimate σ_k by computing the median absolute deviation (MAD) which is often used as a robust estimator for wavelet thresholding[4][5]. The quality of the estimation depends in some way on the ratio of the number of non-zero wavelet coefficients over the total number of samples, thus on the number of pulses in the data. Note that the duration of the pulses is not playing a significant role here: the duration of a wavelet transformed pulse at a given scale k mostly depends on k. The filter used to derived the wavelet transform at scale k has indeed a finite support (it is a band-pass filter), and its bandwidth sets a corresponding time scale.
The main limitation is thus concerning the largest scales (or lowest frequencies): for a given number of pulses in the data, the larger is the scale k, the larger is the proportion of non-zero wavelet coefficients, and the larger is the bias to be expected on the estimation of σ_k. As a reference, we know an efficient matched filter requires template spectra approaching the real spectra at the percent level.

In order to check the validity domain of this noise estimation technique, simulated signals including exponentially decaying pulses, white noise and low frequency $1/f^2$ noise have been analyzed. The decay time constant have been set to 15 ms, the (exponential) rise time constant to 10µs, and the simulated data sets are 2^{20} points long (about 10s with a 100kHz sampling frequency). These parameters, as well as those of the

noise, have been chosen to be close to the real data analyzed further in the last section of this article. We tested two SNRs of 100 and 1000. Simulated pulses all have the same amplitude and their occurrence times are drawn from a Poisson law. The FWMH obtained on the corresponding peak in the amplitude histogram have been used as an indicator of efficiency. Up to about 50 pulses per acquisition (5 events/s), the measured FWMH was less than 5% close to the theoretical FWHM of baseline noise. At higher event rates, the resolution is degraded because of an over-estimation of the noise power at large scales. It should however be noted that pile-up events become predominant at these rates, which also limits the achievable resolution in practice, and calls for a shortening of the pulses decay constant. By shortening this decay, the energy content of a pulse is spread over a larger part of the frequency spectrum toward high frequencies, so that the relative importance of low frequencies (and of large scales) is diminished.

If we use a decimated wavelet transform, which is not time invariant, the filtering of a pulse at a given time will be different than that of the very same pulse at another time, which is redhibitory for the precision of pulse amplitude measurement. On the other hand, the use of a non-decimated transform may dramatically increase the computing time. A good compromise can been found in [7], where the authors present an optimized, faster algorithm for time invariant wavelet transform. Its main drawback is the loss of time ordering of the wavelet coefficients at each scale. Nevertheless, this not a concern for us since the computation of MAD does not depend on time ordering. This prevents us from computing the correlation step in the wavelet domain and exploring all the possibilities of the time/scale plan. While this latter point is quite obvious, the former one may not be an advantage as stated in [8].

Finally, we need to construct a template pulse. This is done by wavelet thresholding a single pulse with sufficient SNR (typically >100). The use of a time invariant transform is also crucial here, since it ensures the quality of the filtering.

To Summarize, here is the different steps of the method:

- Time invariant wavelet transform of a signal with pulses and noise
- Estimation of the σ_k by MAD
- Wavelet thresholding with the computed σ_k's of a high SNR pulse: definition of the template pulse
- Whitening of each scale (normalization by σ_k)
- Inverse wavelet transform back to the time domain
- convolution of the whitened signal by the pulse template

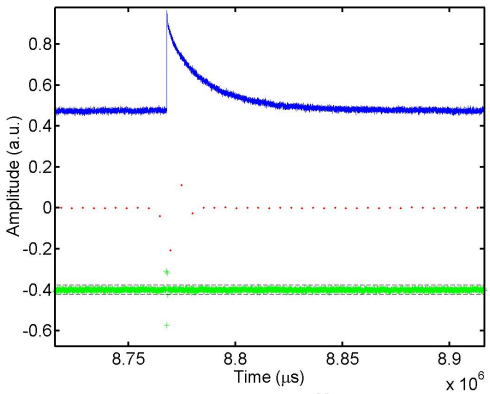


FIGURE 2. 5.9 keV event from a ^{55}Fe source (top) and its corresponding wavelet coefficients at scale 9 (middle, '.') and 4 (bottom, '+', with an offset of -0.4 for a clearer view). Scale 9 corresponds to a frequency band ≈ 100Hz to 200Hz and scale 4 ≈3.1 kHz to 6.2 kHz. Wavelet coefficients are decimated according to their scale: if the original data has 2^N points, scale k will have 2^{N-k} points. The dashed lines are the negative and positive thresholds derived from σ_4. The type of wavelet used is Daubechies 8.

EXAMPLE APPLICATION

Two data sets have been analyzed from two different detectors with two different radioactive sources, namely ^{55}Fe and ^{133}Ba. These runs have a SNR of the order of 100 and 1000 respectively. Data streams are digitalized continuously at a 100 kHz sampling rate and stored on a hard disk. They are then analyzed off-line, being divided into contiguous segments of 2^{20} samples (≈10s) that are treated one by one. The counting rate is about 0.4 event/s for both sources. The analysis have been carried out under GNU Octave[10], an open source equivalent to Matlab, with the free Wavelab802 toolbox[11] including the time invariant wavelet transform algorithm[7] discussed above.

The K normalization constant of equation (1) has been chosen to get a unity standard deviation σ of the residual noise in the filtered signal [6]. By doing so, we can set an adaptive threshold in σ units. Threshold has been set to 6σ.

Three different analyses have been tested: conventional non adaptive matched filter, adaptive wavelet-based matched filter, and linear fit with a template pulse. Results from a linear template fit are usually close from a matched filter ones and may sometimes be more robust against noise variations. Two different wavelets have been tested: Haar and

TABLE 1. FWHM of three lines of the ^{133}Ba energy spectrum. X lines FWHM have been corrected from natural width.

Analysis method/line name	Ka1 (30.974 keV)	K-M3 (34.987 keV)	γ (53.168 keV)
Non adaptive matched filter	52.3 eV	51.5 eV	55 eV
Linear fit with template	51.5 eV	50.6 eV	54.6 eV
Wavelet based matched filter	46.7 eV	41 eV	44 eV

Daubechies 8. Although the latter has more vanishing moments, no significant differences have been seen on the results, but this behavior may not be the same for every types of noise.

The ^{55}Fe run shows an increasing drift of pulses amplitude due to a progressive lowering of the temperature. The FWHM of the 5,9 keV peak is 55 eV for the matched filter, 50 eV for the template fit, and 46 eV for the wavelet-based method. Given the chosen theshold and considering a gaussian noise, we expect about 30 noise events on the total run, which is compatible with the ~20 events counted under 500 eV. Besides the monotonic amplitude drift, a complex drift of the noise structure is observed, with a different behavior at low and high frequencies. The energy threshold is varying by about 2 eV on the whole run.

The ^{133}Ba run shows even more important variations of amplitude, with steep jumps and a smoother, non monotonic amplitude drift. Energy resolutions are summed up in table 1. The time varying FWHM of baseline noise as measured by the wavelet algorithm is varying by about 4eV during the run. The number of noise events is 5 to 10 times the expected one. This could be due to events interacting inside the wafer support of the detector or in the sensor itself. A preliminary analysis of the wavelet coefficients of the pulses tends to confirm this hypothesis: two populations of faster (sensor interaction) and slower (wafer interaction) pulses can be discriminated in the lower part of the energy spectrum.

CONCLUSION

The structure of wavelet transform is particularly adapted to noise measurements when pulses are present. The use of MAD allows statistics accumulation about the noise independently in each frequency band and on an adapted time scale. Application to metallic magnetic calorimeter shows a clear improvement in energy resolution compared to non-adaptive methods, and allows a deeper insight into the time/frequency evolution of the noise structure. Besides being adaptive about noise variations, the method could also be made adaptive to pulse shape variations, since wavelet thresholding allows defining a template on a single event. Further developments could also include pulse shape discrimination by wavelet coefficients or scale-dependent detection of pile-up events. In this perspective, an implementation of non decimated wavelet transforms on FPGA would be particularly convenient in terms of computing time and would allow a real time and almost non parametric filtering of experimental signals.

REFERENCES

1. I. Daubechies, "*Ten lectures on wavelets*". *2nd ed.* Cambridge university press (1992).
2. P. Flandrin, "*Wavelet analysis and synthesis of fractional brownian motion*", *IEEE transactions on information theory*, **38**(2), 910-917 (1992).
3. D. Mc. Cammon in "*Cryogenic particle detection*", Topics in applied physics, Vol. 99, edited by C. Enss.
4. P. J. Huber, "*Robust statistics*", John Wiley and Sons, New York, 1981.
5. I. M. Johnstone and B. W. Silverman, "*Wavelet threshold estimators for data with correlated noise*", *J. Royal Statist. Soc.*, **B59**(2), 319-351 (1997).
6. J. M. Jenkins, "*The impact of solar-like variability on the detectability of transiting terrestrial planets*", *Astroph. Journal*, **575**, 493-505 (2002).
7. R. Coifman and D. Donoho, "Translation invariant denoising", Technical report 475, Dept of Statistics, Stanford University (May 1995).
8. O. Lemoine and P. Abry, "*Redundancy in a frame of wavelets-based matched-filter*", *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 257-260 (1996).
9. S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", IEEE transactions on pattern analysis and machine intelligence, **11**, 674-693 (1989).
10. http://www.gnu.org/software/octave/
11. J. Buckheit and D. Donoho, "*Wavelab and reproducible research*," in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, Eds. Berlin, Springer-Verlag, 1995, pp. 55–81.