# On the Correlation Property of Multiscaling Coefficients for Signal Denoising

Eddie C.Y. Ma, Daniel P.K. Lun and Tai-Chiu Hsung
Centre for Signal Processing
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University
enpklun@polyu.edu.hk

***Abstract*—The discrete multiwavelet transform (DMWT) enables a signal to be analyzed in a multiresolution and multidimensional way. While the generated multiwavelet coefficients are vectors in nature, it has been generally understood that correlation exists between the vector elements. This feature has been adopted particularly in image coding applications to allow efficient design of VQ codebook. For a multiresolution analysis, the multiwavelet coefficients are generated from the multiscaling coefficients of the upper level. In this paper, we show that many multiwavelet systems cannot give correlated multiscaling vector elements, as different from the multiwavelet vector elements. But for those that can give correlated multiscaling vector elements, they can provide much information to assist in identifying the "blank" regions in a noisy signal. A new denoising algorithm is then proposed based on this feature and is particularly useful for sparse source signals.**

Index Terms— ***Multiwavelet, wavelets, denoising, cross correlations.***

## I. Introduction

Multiwavelet, which can be treated as an extension to the scalar wavelet, has drawn much attention in recent years [1]-[5]. Much effort has been made in the studies of its kernel design [1][5], prefilter design [4], and applications [2][3]. Multiwavelet is different from the traditional scalar wavelet in that it consists of a set of scaling functions namely, multiscaling functions, which jointly form a Riesz basis for $V_0$. It also consists of a set of wavelet functions namely, multiwavelet functions, which jointly form an orthonormal basis of $L^2(\boldsymbol{R})$. The number of scaling functions used in a multiwavelet system is indicated by its multiplicity. Similar to wavelet, multiwavelet construction is also associated with a multiresolution analysis, but of multiplicity r, where r is the number of scaling functions used in the system. The two-scale difference equations for multiscaling and multiwavelet functions can be derived as follows [1]:

$$\Phi(x) = \sum_{k\in z} \mathbf{H}_k \Phi(2x-k); \;\; \Psi(x) = \sum_{k\in z} \mathbf{G}_k \Phi(2x-k) \qquad (1)$$

$\boldsymbol{\Phi}(x)$ and $\boldsymbol{\Psi}(x)$ in (1) represent the multiscaling function $\boldsymbol{\Phi} = (\phi_1, \phi_2 \ldots \phi_r)^T$ and the multiwavelet function $\boldsymbol{\Psi} = (\psi_1, \psi_2 \ldots \psi_r)^T$ respectively. $\boldsymbol{\Psi}$ belongs to the spaces $W_j$ with the following relationship:

$$V_{j+1} = V_j \oplus W_j \qquad (2)$$

$\mathbf{H}_k$ and $\mathbf{G}_k$ are the so-called multifilters which link up the multiscaling and multiwavelet functions of different resolutions. As different from the scalar wavelet case, $\mathbf{H}_k$ and $\mathbf{G}_k$ are matrix-based FIR filters. It means that every coefficient of the filter is an $r$ x $r$ matrix, where $r$ is the multiplicity. Taking the Fourier transform of (1), we have the following relationship:

$$\hat{\Phi}(w) = \boldsymbol{H}(w/2)\hat{\Phi}(w/2)\,; \hat{\Psi}(w) = \mathbf{G}(w/2)\hat{\Phi}(w/2) \qquad (3)$$

where $\boldsymbol{H}(w) = \frac{1}{2}\sum_k H_k e^{-jwk}$ , $\boldsymbol{G}(w) = \frac{1}{2}\sum_k G_k e^{-jwk}$ . The increased free parameters allow us to design orthogonal and symmetric multiwavelet filter kernels, which can never be achieved by the traditional scalar wavelets [1][5]. When the multiwavelet is used to process scalar signals, the input signals need to be vectorized first before multiresolution analysis can be performed. To maintain the desirable properties of the multiwavelet, a prefilter is often used to serve the purpose [4]. Figure 1 shows the discrete multiwavelet transform (DMWT) with prefilter.

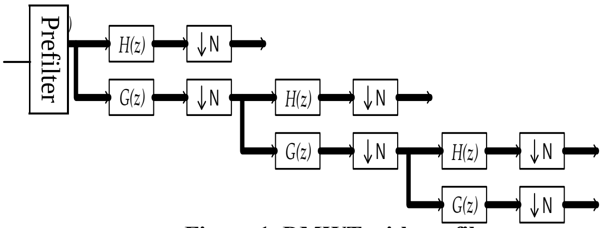


**Figure 1. DMWT with prefilter**

## II. Correlation Property of Multiscaling Coefficients

Multiwavelet uses multiple scaling functions to jointly analyze and synthesize a signal. It is obvious that the frequency responses of different scaling functions should have some differences. Depending on the type of multiwavelet, the amount of differences between the scaling functions is also different. For instance, the multiscaling functions may not be all low-pass. And in fact most of the multiwavelets as reported in the literature have one or more band-pass scaling functions. The multiscaling coefficients generated by these band-pass scaling functions obviously will be much different from the low-pass ones. For example, Figure 2 shows the overall transfer functions (the transfer function including the prefilter) of the first level low-pass branch of the DMWT. In the figure, the overall transfer functions when using the DGHM [1] and CL [5] multiwavelets of multiplicity 2 with orthogonal second-order prefilter are shown. For the CL multiwavelet, one of the multiscaling functions is band-pass while the other is low-pass. Hence the overall transfer function of the two channels is quite different as shown in Figure 2. For the DGHM multiwavelet, the overall transfer functions are similar. Hence we can expect the multiscaling coefficients generated by using the DGHM multiwavelet should have a higher correlation among the vector elements than the CL multiwavelet.

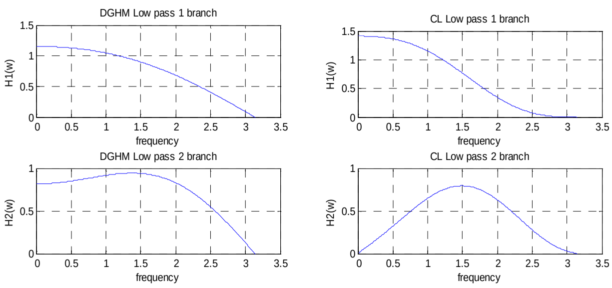


**Figure 2. Overall transfer functions of the first level low pass branch of DMWT using DGHM (left column) and CL Multiwavelets (right column) with prefilter**

## III. DMWT of Signal and Noise

Simulations were performed to study the differences between the multiscaling coefficients of signal and white Gaussian noise. We are particularly interested in the correlation among the vector elements of their multiscaling coefficients. The testing signal "bumps" with sample size 4096 was selected in our experiments. The DGHM multiwavelets and orthogonal second-order prefilter were adopted for the implementation of the DMWT. Figure 3 show some of the simulation results. It is found that, for the signal "bumps", the multiscaling coefficients of the two channels are similar up to a scale difference at least for the first few levels of DMWT. It is not the case for noise. Its multiscaling coefficients have much dissimilarity among the two channels starting from the prefilter output.

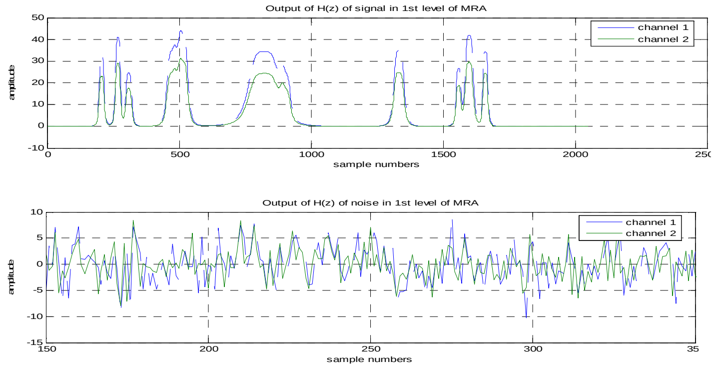


**Figure 3. Level 1 multiscaling coefficients of signal (upper) and noise (lower).**

The result can be explained from a statistical point of view. Since the prefilter is an orthogonal filter bank, the statistical property of noise will not change after the prefiltering. So if we measure the correlation of noise among the two outputs of the prefilter, they should be zero in theory. It is not the case for signal since it is not a random process. The prefilter outputs can still have high correlation depending on the prefilter design. The same argument can also be applied to the subsequent multiwavelet transform since it is also orthogonal. The white Gaussian property of noise will be preserved. In addition, we have seen in Figure 2 that the overall transfer functions of the DGHM multifilters are similar. It helps to maintain the correlation among the vector elements of the multiscaling coefficients of signal. It is only when the level is high then the correlation of signal coefficients may be reduced due to the repeated application of the decimation operator. Certainly the above argument does not apply to those multiwavelets that have great difference in their multiscaling functions, such as the CL multiwavelet.

## IV. Adaptive Multiwavelet Denoising

Based on the studies on the dissimilarity between the multiscaling coefficients of signal and noise, a new adaptive multiwavelet denoising algorithm is proposed. For simplicity, DGHM multiwavelet of multiplicity 2 is used in the algorithm. It is noted that the algorithm can be extended to other multiwavelets and multiplicities only if their low-pass multifilters have similar overall transfer functions at least for the first few levels.

The algorithm starts with the estimation of the local cross correlation among the vector elements of each multiscaling coefficient as follows:

$$Rxorr_j(i) = \left( \sum_{n=0}^{2N-1} c_{j,i-(n+N),1} c_{j,i-(n+N),2} \right) M \quad (4)$$

where $N$ is the number of adjacent samples that will be taken account for the local cross correlation computation, $M$ is an empirical constant which is related to the variance of noise and $c_{j,k,l}$ represents the multiscaling coefficient of sample number $k$ channel $l$ at level $j$. Using the above equation, the estimated local cross correlation at the first level low pass branch of the DMWT of the signal "bumps" and noise are shown in Figure 4. Note that whenever there are signal coefficients, the local cross correlation *Rxorr* will have a very high value. In contrary, *Rxorr* becomes small for the regions containing only noise coefficients.

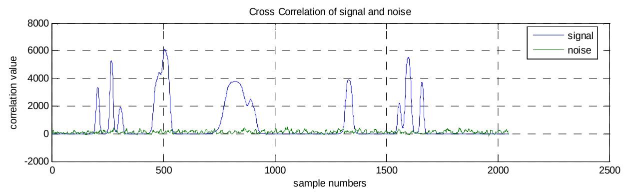


**Figure 4. Cross correlation of the first level multiscaling coefficients of signal and noise**

Hence for a noisy signal that contains both signal and noise, we expect its multiscaling coefficients will exhibit a similar behavior as above at least for the first few levels. The noise multiscaling coefficients are then suppressed by using the following equation:

$$\hat{\mathbf{c}}_{j,i} = \frac{\left| Rxorr_j(i) \right|}{1 + \left| Rxorr_j(i) \right|} \mathbf{c}_{j,i} \quad (5)$$

In (5), $\mathbf{c}_{j,i}$ is the $i^{\text{ih}}$ multiscaling coefficient at level $j$. As discussed above, $Rxorr_j(i)$ will be small if the coefficient $i$ at level $j$ contains only noise energy. In this case, a factor tends to 0 will be multiplied to $c_{j,i}$ to suppress its magnitude as indicated in (5). On the other hand, $Rxorr_j(i)$ will be large if the coefficient $i$ at level $j$ contains signal energy. A factor tends to 1 will be multiplied to $c_{j,i}$ and hence the signal energy will be retained. The processed multiscaling coefficients will be used for generating the next level of multiscaling and multiwavelet coefficients. The above procedure will be repeatedly applied to the next level of multiscaling coefficients if the signal coefficients still exhibit a high correlation among the vector elements.

Note that (5) can only suppress the noise in the regions where signal is not found. Hence it is particularly useful to those signals that are extremely sparse. Many practical signals have such characteristic. Examples include the electrocardiogram (ECG) signals and the activation patterns in functional magnetic resonance imaging (fMRI) [6]. For those regions which are mixed with signal and noise, the multivariate shrinkage [2][3] can be applied to the multiwavelet coefficients to reduce the noise energy. The proposed adaptive multiwavelet denoising algorithm is summarized as follows

1. Perform the first level of DMWT with prefilter.
2. Compute the cross correlation function (4) for each multiscaling coefficient and apply (5) to suppress the noise coefficients.
3. Apply multivariate shrinkage to the multiwavelet coefficients to further reduce the noise energy.
4. By using the denoised multiscaling coefficients, generate the next level of multiscaling and multiwavelet coefficients.
5. Repeat step 2 for the new level if the multiscaling coefficients exhibit a high correlation among the vector elements.
6. Repeat step 3 for the new level of multiwavelet coefficients to reduce the noise energy.
7. Repeat step 4 – 6 until all levels of multiscaling and multiwavelet coefficients are denoised.
8. Reconstruct the signal by using the inverse DMWT.

## V. Simulation and Results

The performance of the proposed adaptive multiwavelet denoising algorithm is demonstrated in this section. The testing signal "bumps" with sample size 4096 was selected in the simulation. The signal is contaminated with zero mean white Gaussian noise at different SNR 5, 10, and 15dB. The noisy signal is then transformed by the DMWT using the DGHM multiwavelet of multiplicity 2 and the orthogonal second-order prefilter. Seven levels of multiscaling and multiwavelet coefficients are generated. While all levels of multiwavelet coefficients are denoised using the multivariate shrinkage, only the first 4 levels of multiscaling coefficients are denoised using the proposed algorithm. The parameter $N$ of (4) is set to 5 and $M$ is set to 0.05, 0.125 and 0.5 for the noisy signal with SNR 5, 10 and 15dB, respectively. The performance of the proposed algorithm is compared with four other wavelet-based denoising methods, namely multivariate shrinkage [2][3], wavelet SureShrink (hybrid and rigorous) [7], and spatially adaptive shrinkage [8]. 50 Monte Carlo simulations were performed for the noisy signals of different SNR in order to obtain accurate results.

As it is shown in Table 1, the proposed adaptive multiwavelet denoising algorithm outperforms the other 4 consistently. The testing system "bumps" is the kind of signal that is sparse in nature. The SureShrink (or the rigorous mode SureShrink, the name used in Matlab) cannot perform well due to the insufficient signal data to construct the SURE profile. The hybrid mode SureShrink works better as it allows switching back to using universal threshold when signal data are not enough. However, the universal threshold cannot adapt to the varying characteristic of the signal hence the performance can only be average. The worse performance of the spatially adaptive shrinkage is expected. “Bumps” is a kind of sparse signal that its wavelet coefficients can hardly be described by GGD with $\beta$ between 0.5 to 1, which is the criterion for the spatially adaptive shrinkage algorithm to be optimal [8]. The proposed algorithm in fact is the combination of the multivariate shrinkage and the denoising of the multiscaling coefficients. Hence the performance should be better than using the multivariate shrinkage alone. The amount of improvement depends on the sparsity of the signal. Figure 5 further shows the denoised signals using different algorithms. As seen in the figure, the proposed algorithm gives a visually more pleasant denoised signal.

| SNR (dB) | Multi-variate Shrink | **Proposed** | SureShrink (Hybrid/ Rigorous) | Spatially Adaptive Shrinkage |
|---|---|---|---|---|
| 5 | 18.267 | **18.843** | 18.086 / 17.755 | 9.970 |
| 10 | 22.541 | **23.906** | 22.464 / 22.112 | 21.132 |
| 15 | 26.643 | **27.627** | 26.061 / 25.940 | 27.081 |

**Table 1. Signal to error ratio (dB) of the enhanced signal using different denoising algorithms**

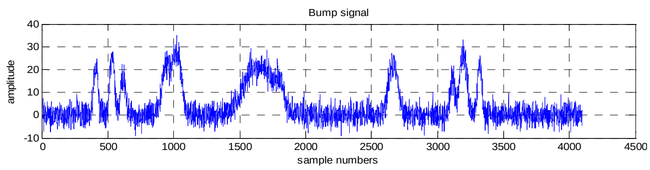

**Figure 5(a). Noisy “bumps” signal with SNR = 10dB**

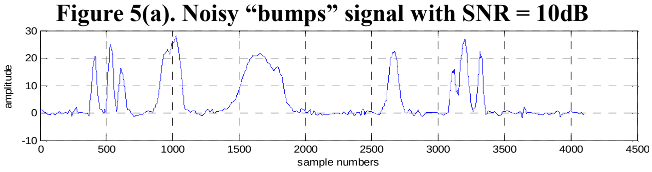

**Figure 5(b). Denoised signal using multivariate shrinkage**

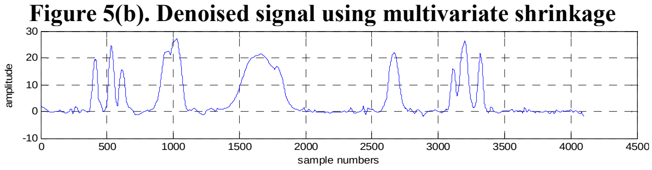

**Figure 5(c). Denoised signal using SureShrink (hybrid)**

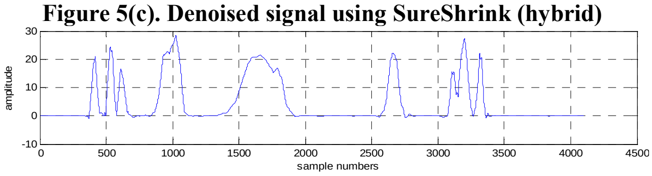

**Figure 5(d). Denoised signal using the proposed approach**

## VI. CONCLUSION

An adaptive denoising algorithm based on the local cross correlation between the vector elements of the multiscaling coefficients is proposed. It suppresses the noise in the multiscaling coefficients before they are further decomposed to generate the next level of multiscaling and multiwavelet coefficients. Multivariate shrinkage is also applied to all multiwavelet coefficients to further reduce the noise energy. Simulation shows that the proposed algorithm gives good result as compared with the traditional multiwavelet-based or wavelet-based denoising methods particularly when the signal is sparse.

## ACKNOWLEDGMENT

This work is supported by the Hong Kong Polytechnic University (grant no. G-U412).

## REFERENCES

[1] J. S. Geronimo, D. P Hardin and P. R. Massopust, “Fractal functions and wavelet expansions based on several scaling functions”, Journal of Approximation Theory, vol.78, no.3, pp.373-401, Sept 1994.

[2] T. R. Downie and B. W. Silverman, “The discrete multiple wavelet transform and thresholding methods”, IEEE Trans. Signal Proc., vol.45, no.9, pp.2558-2561, Sept. 1998.

[3] Tai-Chiu Hsung and Daniel Pak-Kong Lun, “Optimal thresholds for multiwavelet shrinkage”, Electronics Letters, vol.39, no.5, pp.473-474, March 2003.

[4] X. G. Xia, J. Geronimo, D. P. Hardin and B. W. Suter, “Design of prefilters for discrete multiwavelet transforms”, IEEE Trans. Signal Proc., vol.44, no.1, pp.25-35, Jan. 1996.

[5] Charles K. Chui and Jian-ao Lian, “A study of orthonormal multi-wavelets”, Applied Numerical Mathematics, vol.20, pp.273-298, 1996.

[6] Alle Meije Wink and Jos B.T. M. Roerdink, "Denoising functional MR images: A comparison of wavelet denoising and gaussian smoothing", IEEE Trans. Med. Imag., vol.23, no.3, pp.374-387, Mar. 2004.

[7] D. L. Donoho, I. M. Johnstone, “Adapting to unknown smoothness via wavelet shrinkage”, J. American Statistic Association, vol.90, pp.1200-1224, 1995.

[8] S. Grace Chang and Martin Vetterli, “Spatial adaptive wavelet thresholding for image denoising”, Proc., IEEE Int. Conf. Image Proc., (ICIP-97), vol.2, pp.374-377, 1997.