MULTISCALE ANALYSIS BY MEANS OF DISCRETE MOLLIFICATION FOR ECG NOISE REDUCTION

ANÁLISIS MULTIESCALA POR MOLIFICACIÓN DISCRETA PARA LA REDUCCIÓN DE RUIDO EN SEÑALES ECG

JUAN PULGARÍN-GIRALDO
Grupo de Investigación en Ingeniería Biomédica G-BIO, Universidad Autónoma de Occidente, Cali, jdpulgarin@uao.edu.co

CARLOS ACOSTA-MEDINA
Departamento de Matemáticas y Estadística, Universidad Nacional de Colombia, Manizales, cdacostam@unal.edu.co

GERMÁN CASTELLANOS-DOMÍNGUEZ
Grupo de Control y Procesamiento Digital de Señales, Universidad Nacional de Colombia, Manizales, gcastell9@gmail.com

ABSTRACT: Multiscale analysis and computation is a rapidly evolving area of research that have had a fundamental impact on computational science and applied mathematics and have influenced the way we view the relation between mathematics and science. Even though multiscale problems have been longly studied in mathematics, such techniques suffer of the ill-posedness nature of the problem. In the solution of several ill-posed problems, discrete mollification has been used for regularization. In this paper, we propose a new technique (procedure) for multiscale analysis by using discrete mollification. The multiscale scheme is based on numerical linear algebra results combined with the mollification method applied to the Mallat algorithm. The new technique has a simple theory, an efficient implementation and compares fairly well with classical wavelet transform procedures. An application on electrocardiographic signals contaminated with typical non-white noise is considered.

KEYWORDS: ECG, GCV, multiscale analysis, mollification, non-white noise, regularization, thresholding.

RESUMEN: El análisis multiescala es un área de gran actividad investigativa con fuerte impacto en computación científica y matemática aplicada, ocupando un lugar de privilegio en la forma como se entiende la relación entre la matemática y las demás ciencias. Aunque el estudio matemático de problemas multiescala está bastante documentado, estas técnicas heredan en el entorno discreto la naturaleza mal condicionada del problema. La molificación discreta ha sido empleada con éxito en la solución numérica de diversos problemas mal condicionados. Este árticulo propone una nueva técnica de análisis multiescala, basada en molificación discreta. El procedimiento aquí propuesto usa resultados de algebra lineal numérica para implementar molificación discreta en el algoritmo de Mallat. La nueva técnica tiene una teoría simple, una implementación eficiente y proporciona resultados de calidad comparable a técnicas multiescala clásicas tipo onditas (wavelet). El proceso es aplicado en señales electrocardiográficas contaminadas con ruido no blanco, para fines de comparación.

PALABRAS CLAVE: ECG, GCV, análisis multiescala, molificación, ruido no blanco, regularización, umbralización.

1.   MOLLIFICATION

The mollification method is a filtering procedure based on convolution, which is appropriate for the regularization of ill-posed problems and for the stabilization of explicit schemes in the solution of partial differential equations. As a regularization method for ill-posed problems, the method is well documented, for instance in [1], [2] and [3]. For the definition of mollification, the implementation of numerical boundary conditions for discrete mollification and the automatic selection of mollification parameters, we recommend [1] and [4].

1.1   Abstract Setting

Let $\delta > 0$, $p > 0$ and

$$A_{p\delta} = \left(\int_{-p/\delta}^{p/\delta} \exp(-s^2) ds \right)^{-1}. \qquad (1)$$

We work with the following truncated Gaussian kernel:

$$\kappa_{\delta p}(t) = \begin{cases} A_{p\delta}\delta^{-1}\exp(-t^2/\delta^2), & |t| \le p \\ 0, & |t| > p. \end{cases} \qquad (2)$$

This kernel satisfies: $\kappa_{\delta p} \ge 0$, $\kappa_{\delta p} \in C^\infty(-p, p)$, $\kappa_{\delta p}$ is zero outside $[-p, p]$ and $\int_{\mathbb{R}} \kappa_{\delta p} = 1$.

Given $f : \mathbb{R} \to \mathbb{R}$ locally integrable, we define its $\delta p-mollification$, denoted $J_{\delta p}f$, as the convolution of f with the kernel $\kappa_{\delta p}$.

That is,

$$J_{\delta p}f(t) = \left(\kappa_{\delta p} * f\right)(t)$$
$$= \int_{-p}^{p} \kappa_{\delta p}(-s)f(t+s)ds. \qquad (3)$$

1.2   Unbounded Discrete Domain

Definition   1:   Let $X = \{x_j : x_j = x_0 + jh, \ j \in \mathbb{Z}\}$ *be a discrete domain with* x_0 *and* h *given real numbers and* $h > 0$. *Let* $G : X \to \mathbb{R}$ *be a function defined by* $G(x_j) = y_j$. *Set*

$$S_j = (x_{j-1} + x_j)/2, \quad j \in \mathbb{Z}$$
$$I_j = [S_j, S_{j+1}), \qquad j \in \mathbb{Z} \qquad (4)$$
$$f(t) = \sum_{j=-\infty}^{\infty} y_j \chi_j(t), \quad t \in \mathbb{R},$$

with χ_j *the characteristic function of* I_j. *Then for* $\delta > 0$ *and* η *a given non-negative integer, we define the* $\delta\eta-mollification$ *of* G *as the* $\delta p-mollification$ *of* f *with*

$$p = (\eta + 1/2)h, \qquad (5)$$

that is,

$$J_{\delta\eta}G(x) = J_{\delta p}f(x). \qquad (6)$$

We are particularly interested in the value of $J_{\delta\eta}G$ at the points in X. Let

$$t_j = (j - 1/2)h, \ j \in \mathbb{Z}. \qquad (7)$$

Then we can write

$$J_{\delta\eta}G(x_j) = J_{\delta p}f(x_j)$$
$$= \sum_{i=-\eta}^{\eta} \int_{t_i}^{t_{i+1}} \kappa_{\delta p}(-s)f(x_j + s)ds. \qquad (8)$$

Furthermore,
$t_i < s < t_{i+1}$ if and only if $S_{i+j} < x_j + s < S_{i+j+1}$.

Thus,

$$J_{\delta\eta}G(x_j) = \sum_{i=-\eta}^{\eta} w_i y_{j+i}, \qquad (9)$$

where

$$w_i = \int_{t_i}^{t_{i+1}} \kappa_{\delta p}(-s)ds. \qquad (10)$$

Notice that w satisfies

$$\sum_{i=-\eta}^{\eta} w_i = 1. \qquad (11)$$

Theorem 2: *Let g be a function defined on $\mathbb{R}$ with fourth derivative $g^{(4)}$ continuous and bounded in $\mathbb{R}$. Let G be its discrete version defined on X. If G^{ε} is another discrete function defined on X and such that*

$$\left| G^{\varepsilon}(x_j) - G(x_j) \right| \le \varepsilon, \text{ for } x_j \in X, \qquad (12)$$

then, there exists a constant C such that

$$
\begin{aligned}
\left| J_{\delta\eta} G^{\varepsilon}(x_j) - J_{\delta\eta} G(x_j) \right| &\le \varepsilon, \\
\left| J_{\delta\eta} G(x_j) - g(x_j) \right| &\le Ch^2.
\end{aligned} \qquad (13)
$$

Furthermore, if g is smooth enough, there exists a constant C such that

$$
\begin{aligned}
\left| D_{+} J_{\delta\eta} G(x_j) - g'(x_j) \right| &\le Ch, \\
\left| D_{0} J_{\delta\eta} G(x_j) - g'(x_j) \right| &\le Ch^2, \\
\left| D_{-} D_{+} J_{\delta\eta} G(x_j) - g''(x_j) \right| &\le Ch^2,
\end{aligned} \qquad (14)
$$

where D_{+}, D_{-} and D_{0} are the forward, backward and central finite difference operators respectively.

For a proof, see [4].

1.3 Discrete Mollification

For working with finite length data some kind of boundary condition has to be assumed, [5], [6]. The most common and documented are Zero Padding (Dirichlet), Scaled version (Non-homogeneous Dirichlet), Even Reflection (Neumann) and Periodic. All these options lead to linear operators with Toeplitz, Scaled Toeplitz, Toeplitz plus Hankel and Circulant matrix representation, respectively.

Under periodic boundary conditions the FFT allows us to know the actual spectral action of the mollification kernel on the data, [6]. In the case of even reflection the Discrete Cosine Transform (DCT) results useful, [5].

2. MULTISCALE ANALYSIS (MSA)

2.1 Multiscale Decomposition Tree

The clue in the multiscale analysis is the Decomposition Tree, [7], [8] (Fig. 1).
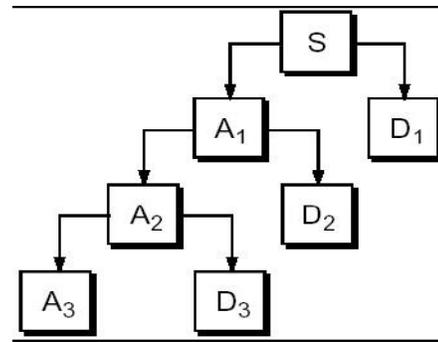
Figure 1. Multiscale decomposition tree

The $A_i's$ are filtered versions of S at different scales or resolutions. The $D_i's$ are the complementary details for obtaining the original source. The idea is that the approximations $A_i's$ give us an overview of the signal S and the details $D_i's$ give us its specials features. The approximations are usually obtained from successive applications of lowpass filters, and the details from associated highpass filters.

2.2 The Idea

A decomposition tree could be set up by using discrete mollification at different resolutions ($\delta's$) as approximations and residuals as details (Fig. 2).
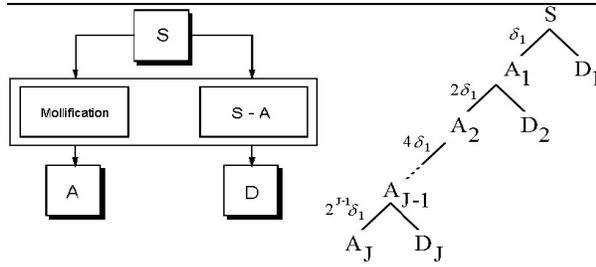
Figure 2. Multiscale decomposition tree using discrete mollification

2.3 Discrete MSA

Weinan E and Bjorn Engquist in [9] state:
"Given such a variety of multiscale methods in many different applications, it is natural to ask whether a general framework can be constructed.

The **general framework** should ideally:
- unify existing methods,
- give guidelines on how to design new methods and improve existing ones,
- provide a mathematical theory for stability and accuracy of these methods."

Definition 3: *A 3-upla of collections* $(\beta, \uparrow 2, \downarrow 2)$ *is called a DMSA if* $\beta = \{\beta_k\}_{k \in \mathbb{N}}, \uparrow 2 = \{\uparrow 2_k\}_{k \in \mathbb{N}}, \downarrow 2 = \{\downarrow 2_k\}_{k \in \mathbb{N}}$ *satisfy*

1) $\beta_k = \left\{ V_1^k, V_2^k, \cdots, V_k^k, W_1^k, W_2^k, \cdots, W_{k-1}^k \right\}$ *is a set of subspaces of* $\mathbb{R}^{2^k}$.

2) $V_1^k \subseteq V_2^k \subseteq \cdots \subseteq V_k^k = \mathbb{R}^{2^k}$.

3) $V_j^k = V_{j-1}^k \oplus W_{j-1}^k, \quad 2 \le j \le k$

4) $\dim V_j^k = 2^j$.

5) $\uparrow 2_j : \mathbb{R}^{2^j} \to \mathbb{R}^{2^{j+1}}$ *is an upsampling operator and* $\downarrow 2_j : \mathbb{R}^{2^j} \to \mathbb{R}^{2^{j-1}}$ *is a downsampling operator.*

6) *if* $v \in V_j^k$ *and* $j < k$, *then* $v^{\downarrow 2} \in V_j^{k-1}$, $v^{\uparrow 2} \in V_j^{k+1}$ *and* $\left(v^{\downarrow 2} \right)^{\uparrow 2} = v$.

Example 4: *The decomposition generated by orthogonal DWTs is a DMSA.*

Definition 5: *A vector* $S \in \mathbb{R}^{2^k}$ *is said to be decomposed to* J *levels if it has been written in*

the form
$$S = D_1 + D_2 + \cdots + D_J + A_J, \qquad (15)$$
where $D_j \in W_{k-j}$ *and* $A_J \in V_{k-J}$.

3. DMSA BY MOLLIFICATION

3.1 Kernel Adjusting

For $\delta > 0$, consider the Gaussian kernel
$$\kappa^\delta(x) = \delta^{-1} \exp(-x^2 / \delta^2), \quad x \in \mathbb{R}. \quad (16)$$

Its Fourier transform is
$$\hat{\kappa}^\delta(\omega) = \sqrt{\pi} \exp(-\delta^2 \omega^2 / 4), \quad \omega = 2\pi f.$$

$$(17)$$

Then
$$\hat{\kappa}^{2\delta}(\omega) = \sqrt{\pi} \exp(-4\delta^2 \omega^2 / 4)$$
$$= \sqrt{\pi} \exp(-\delta^2 (2\omega)^2 / 4) = \hat{\kappa}^\delta(2\omega). \quad (18)$$

So, a dyadic dilation in the kernel's parameter δ generates a dyadic contraction in its spectrum.

Furthermore, under periodic boundary conditions, by taking

$$h = \frac{2}{2\eta + 1} \quad \text{and} \quad \delta_1 = \frac{16}{\pi(2\eta + 1)}, \qquad (19)$$

the discrete mollification process cuts the upper half of the frequency components off.

Then the following decomposition scheme will produce a dyadic spectral decomposition of a vector in $\mathbb{R}^{2^k}$ (Fig. 3).
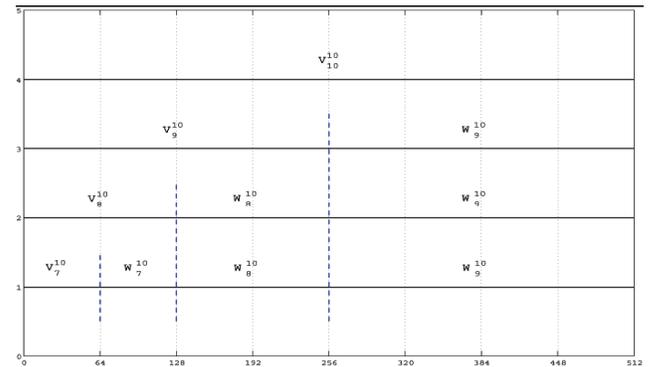
Figure 3. Dyadic spectral decomposition

To this procedure we could associate a sub-band DMSA in the frequency domain. The problem is that, because of the shape of the Gaussian kernel's spectrum, the residuals (i.e. details) contain low-frequency information of the vector and consequently $D_j \notin W_{k-j}$. So, our scheme is not a perfect DMSA.

3.2   Mallat's Algorithm

In a perfect DMSA we can apply the Mallat's Algorithm for obtaining the MSA decomposition of a vector, [7], [8]. However, if this algorithm is directly applied to our scheme we will not obtain a perfect recovery structure.

In our case, we propose the following algorithm (Fig. 4), whose action takes place in the frequency domain and consists on applying the downsampling operator to suppress the entries associated to the upper half of the spectrum.
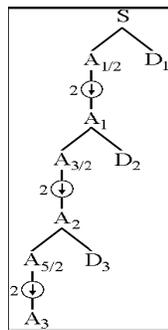
Figure 4. Decomposition tree (dyadic case) using discrete mollification

3.3   Non-dyadic case

For non dyadic vectors the DMSA by Discrete Mollification is implemented by increasing the $\delta's$ values (Fig. 2).

4.   REGULARIZATION BY THRESHOLDING

4.1   Introduction

A measured vector S^ε, contaminated with additive noise, can be modeled as

$$S_i^\varepsilon = s(x_i) + \varepsilon_i, \text{ for } i = 1, 2, \cdots, 2^k, \quad (20)$$

where $s(x_i)$ represents the exact data function and ε_i the corresponding added noise, usually a random variable. Let

$$S_i = s(x_i), \text{ for } i = 1, 2, \cdots, 2^k. \quad (21)$$

Now we apply J levels of DMSA to S^ε to obtain a decomposition of the form

$$S^\varepsilon = A_J^\varepsilon + D_J^\varepsilon + \cdots + D_2^\varepsilon + D_1^\varepsilon. \quad (22)$$

We expect A_J^ε to be somewhat similar to S and the details to depend in some way on the noise. At this point we do some assumptions

- The magnitude of the noise is small compared with the signal S (SNR>10dB [10]).
- The approximation procedure is good enough, so the residuals in the details are expected to be zero or close to zero at most of the points.

Under these hypotheses we make zero the detail components with magnitude under a certain **threshold** value, because they are suspicious of being noise. The entries with magnitude above this value are reduced in magnitude, for continuity purposes (s*oft-thresholding*). With this procedure we obtain a modified version of the measured data which is expected to be smooth and to preserve the most important features of the signal, [7], [10].

A different threshold value is selected for each level, because each level of detail deals with different frequency bands and probably different noise behavior.

Theorem 6 (Numerical Convergence): *Let S,* $S^\varepsilon \in \mathbb{R}^{2^k}$ *such that*

$$\left\| S^\varepsilon - S \right\|_\infty \le \varepsilon, \quad (23)$$

If S_λ^ε and S_λ denote the results of reconstructing after J levels of DMSA by mollification and soft-thresholding with threshold values $\lambda = [\lambda_1, \lambda_2, \cdots, \lambda_J]^T$ applied to

S^ε and S respectively, then

$$\left\| S_\lambda^\varepsilon - S_\lambda \right\|_\infty \leq \varepsilon, \ (Stability)$$
$$\left\| S_\lambda - S \right\|_\infty \leq J \max_{1 \leq j \leq J} \lambda_j, \ (Consistency) \quad (24)$$

Consequently, it holds the convergence

$$\left\| S_\lambda^\varepsilon - S \right\|_\infty \leq \varepsilon + J \max_{1 \leq j \leq J} \lambda_j. \quad (25)$$

4.2   Threshold value selection

As usual in regularization methods, the most difficult and relevant task is the automatic selection of regularization parameters that allow an optimal balance between the errors of regularization and perturbation, [10], [11]. The procedure selected for this task is Generalized Cross Validation (GCV), which offers:
- Efficiency
- No information about the magnitude or variance of the noise is required
- Asymptotic optimality.

The GCV function in this case is

$$\lambda_j = \arg\min_\lambda \frac{\frac{1}{N}\left\| D_{j,\lambda}^\varepsilon - D_j^\varepsilon \right\|_2}{\left(\frac{N_0(\lambda)}{N} \right)}, \quad (26)$$

where $N_0(\lambda)$ denotes de number of components in D_j^ε with magnitude under the threshold value λ.

5.   EXPERIMENTAL SETUP

5.1   ECG Database

The experiments were performed on 55 registers extracted from the MIT-BIH arrhythmia database and corrupted, at a 6dB signal-to-noise ratio, with three different types of synthesized noise: electromyographic interference, 60Hz powerline interference and electrosurgical noise [12],[13].
This test shows how DMSA by discrete mollification works in noise reduction in comparison with another popular multiscale procedure: wavelet. Four levels of decomposition were used for both methods. The wavelet of choice was "db4".

5.2   Scoring criterion

For each register, the scouring routine obtains absolute error between original and filtered signal for both methods. The result on the *i-th* register is denoted by e_i.

Afterwards, mean value, variance, maximum and minimum values are computed as follows:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} e_i \qquad Var = \frac{1}{n}\sum_{i=1}^{n}(e_i - \mu)$$
$$Max = \max_{1 \leq i \leq n} |e_i| \qquad Min = \min_{1 \leq i \leq n} |e_i| \quad (27)$$

6.   RESULTS

Table 1. Absolute error in ECG signals after regularization by thresholding. Four levels, wavelet of choice: 'db4'

	μ	*Var*	*Max*	*Min*
Original	0.1390	0.0106	0.4542	0.0002
DMSA by mollification	0.0334	0.0008	0.1871	0.0002
DMSA by wavelet	0.0353	0.0008	0.1582	0.0000

Both methods are aceptable for non-white noise reduction in ECG signals, but a discriminat factor could be the maximus error. Because of that, DMSA by mollification still shows a little pikes reduction on ECG signals.
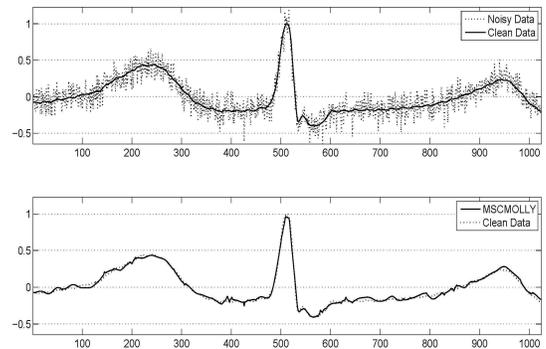
Figure 5. Result of applying DMSA by Discrete Mollification for filtering ECG signals

7.   CONCLUSION

The results show how the new procedure is capable of making an automatic filtering preserving not only the low frecuencies but also the main high frecuency features of the signal.

Additionally, this work offers a fully discrete mathematical analysis of the tool. This could be useful in the convergence analysis of algorithms using DMSA for the solution of multiscale problems.

All this imply that DMSA by mollification could be applied to a wide range of problems with strong multiscale interaction. Additional work on this matter has to be done.

8. ACKNOWLEDGMENT

REFERENCES

[1] MURIO D. A. Mollification and space marching, In: Inverse Engineering Handbook (ed. K. Woodbury), CRC Press, 2002.

[2]   MURIO D. A., The mollification method and the numerical solution of ill-posed problems, John Wiley, New York, 1993.

[3] MURIO D. A., C. E. MEJÍA, AND   S. ZHAN, Discrete mollification and automatic numerical differentiation, Computers Math. Applic., 35, 1–16, 1998.

[4] ACOSTA C. D. AND C.E. MEJÍA, Stabilization of explicit methods for convection diffusion equations by discrete mollification, Computers Math. Applic., 55, 368 – 380, 2008.

[5]   NG M. K., CHAN R. H., AND TANG W., A fast algorithm for deblurring models with Neumann boundary conditions, SIAM J. Sci. Comput., 21, 851–866, 1999.

[6]   NAGY J. G., Iterative technics for the solution of Toeplitz systems, Siam News, 28, 1–16, 1995.

[7]   MATHWORKS, Wavelet toolbox: User's guide - version 2.1 for use with Matlab, The MathWorks Inc., Natick, 2001.

[8]   CHUI C., An introduction to wavelets, Academic Press, 1992.

[9]   E W. AND ENGQUIST B., Multiscale modeling and computation, Notices Amer. Math. Soc., 50, 1062–1070, 2003.

[10]   JANSEN M., Noise reduction by wavelet thresholding, Springer, 2001.

[11] HANSEN P. C., Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion, SIAM, 1997.

[12] FRIESEN G. M., JANNETT T. C., JADALLAH M. A., YATES S. L., QUINT S. R. AND NAGLE H. T., A comparison of the noise sentitivity of nine qrs detection algorithms, IEEE Transactions on Biomedical Engineering, 37, 85–98, 1990.

[13] OROZCO R., PÉREZ M., LORENZO J. V., GRAU R. AND RAMOS R., Evaluation of qrs morphological classifiers in the presence of noise, Computers and Biomedical Research, 30, 200–210, 1997.