

Avaliação de desempenho de wavelet shrinkage pela esparsidade dos coeficientes

Vinícius Ávila da Silva e Evandro Ottoni Teatini Salles

Resumo—A técnica conhecida como *wavelet shrinkage* se consagrou como um método de baixo custo computacional e desempenho satisfatório em *denoising* de sinais. Este método, porém, não é totalmente adaptativo, pois depende de parâmetros como a escolha de uma *wavelet* e escala de aplicação. Este artigo visa avaliar a relação entre o desempenho do método e a esparsidade dos coeficientes *wavelet*. Foi observado que o ganho, em dB, da SNR dos sinais depende grandemente dos parâmetros citados, e que a esparsidade pode ser uma ferramenta para adaptar os parâmetros a cada conjunto de amostras.

Palavras-Chave—*Denoising, Wavelet Shrinkage, Esparsidade.*

Abstract—The wavelet shrinkage approach has been considered a method of low computational cost and reasonable performance in signal denoising. However, this method is not entirely adaptive, because of its dependence on parameters such as the wavelet function and scale. This paper aims to evaluate the relation between the performance of the method and the sparsity of the wavelet coefficients. It was observed that the signal SNR increase, in db, relies greatly on the aforementioned parameters, and that the sparsity can be a tool to adapt the parameters to each sample set.

Keywords—*Denoising, Wavelet Shrinkage, Sparsity.*

I. INTRODUÇÃO

O problema de sinais contaminados por ruído é comum em diversas aplicações e pode reduzir a eficiência de sistemas diversos, como reconhecimento de voz, telemedicina e transmissão de dados. Um algoritmo de *denoising* é aquele que busca remover ruído, erros e perturbações de um sinal [1], de forma a preservar o sinal original. Semelhantemente, a partir de N amostras de um sinal f contaminado por um ruído e ,

$$y[i] = f[i] + e[i] \quad i = 1, \dots, N, \quad (1)$$

deseja-se estimar um vetor \hat{f} com o menor erro possível em relação ao sinal original f . O desempenho do algoritmo pode, então, ser aferido pelo erro médio quadrático,

$$MSE = \frac{\sum_{i=1}^N (f[i] - \hat{f}[i])^2}{N}, \quad (2)$$

e pelo relação sinal-ruído,

$$SNR = 10 \log \left(\frac{VAR(\hat{f})}{VAR(\hat{f} - f)} \right). \quad (3)$$

O método de *wavelet shrinkage*, como descrito em [2], foi desenvolvido na década de 90 e é de fácil implementação e baixo custo computacional. Este método consiste na aplicação de limiares aos coeficientes *wavelet* do sinal contaminado. Assim, pode ser descrito como duas etapas sucessivas: transformação e

limiarização. Esta, baseada no estimador não-polarizado de risco de Stein (SURE) [3], não depende de nenhum parâmetro a ser definido por usuário, ou seja, é adaptativa. A etapa de transformação, por sua vez, decorre da escolha de uma *wavelet* e uma escala de aplicação, de forma que o desempenho final do algoritmo depende grandemente destes parâmetros.

Este artigo visa avaliar a esparsidade dos coeficientes *wavelet* de um sinal contaminado como ferramenta para adaptar os parâmetros de um *denoiser* a um determinado conjunto de amostras. De outra forma, deseja-se relacionar o desempenho do algoritmo (MSE e SNR) à esparsidade no domínio *wavelet*.

O presente artigo está dividido da seguinte forma. A seção II descreve o método de *wavelet shrinkage* e a técnica de limiarização baseada em SURE. Posteriormente, na seção III, é descrita a medida de Hoyer como métrica de esparsidade e como esta se reduz à norma L1 para transformações ortogonais. O método para avaliar a relação entre as métricas de desempenho (MSE e SNR) e a esparsidade (norma L1) é descrito na seção IV, bem como os resultados da avaliação. Exemplos de utilização do método são expostos também na seção IV, enquanto a conclusão do trabalho é apresentada na seção V.

II. WAVELET SHRINKAGE

A transformação *wavelet* discreta (DWT) de um sinal y com $N = 2^j$ amostras pode ser escrita na forma matricial

$$w = \mathcal{W}y, \quad (4)$$

na qual \mathcal{W} é a matriz de transformação e w são os coeficientes *wavelet* resultantes. A matriz \mathcal{W} é quadrada e depende da chamada *wavelet* mãe, a função na qual o sinal y será decomposto. Para o caso em que esta função é periódica, a matriz de transformação é ortogonal. Assim, a reconstrução do sinal é simples e definida por

$$y = \mathcal{W}^T w. \quad (5)$$

Segundo [4], um sinal qualquer é melhor representado por uma função *wavelet* semelhante ao sinal. A representação é dita melhor quando poucos coeficientes são não-nulos, ou seja, quando a representação é esparsa. O ruído branco, devido a sua característica aleatória, não é bem representado por nenhuma função e , para *wavelets* ortogonais, o ruído branco é transformado em ruído branco [2], ou seja, em amostras não correlacionadas e com média zero.

Assim, se a representação de um sinal qualquer for esparsa, a distinção entre os coeficientes do sinal e os coeficientes do ruído é perceptível, conforme Fig. 1. O sinal dado como

Vinícius Ávila da Silva e Evandro Ottoni Teatini Salles, Centro Tecnológico, Universidade Federal do Espírito Santo (UFES), Vitória-ES, Brasil, E-mails: vinicius_ads@hotmail.com, evandro@ele.ufes.br. Este trabalho foi parcialmente financiado pelo CNPq.

exemplo é o *blocks* [2] e, devido a sua natureza quadrada, a transformação utilizada foi a Haar.

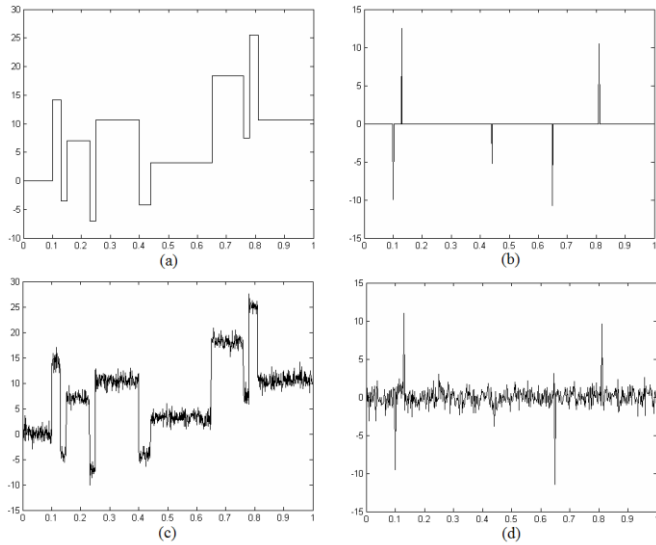


Fig. 1. Exemplo de sinal e seus coeficientes *wavelet*. (a) *Blocks* no domínio do tempo, (b) coeficientes de *blocks*, (c) *blocks* contaminado por ruído branco, (d) coeficientes de *blocks* contaminado por ruído branco.

É possível perceber que, como a representação do sinal é esparsa, os coeficientes do sinal possuem amplitude superior aos coeficientes resultantes do ruído. Logo, se um limiar apropriado t for definido, novos coeficientes w' definidos por

$$w'[i] = \begin{cases} w[i] - t & |w[i]| \geq t \\ 0 & |w[i]| < t \end{cases} \quad (6)$$

podem ser usados para reconstruir o sinal com SNR superior à condição anterior.

O limiar é definido de forma adaptativa pelo SURE, que estima de forma não-polarizada o erro quadrático de um estimador polarizado. O ponto de partida, portanto, deve ser um estimador polarizado para o limiar t . Frequentemente utilizado, o limiar universal [5] pode ser definido por:

$$t = \sigma \sqrt{2 \log N}, \quad (7)$$

em que σ é o desvio padrão do ruído no domínio *wavelet*, que é desconhecido e pode ser estimado pelo desvio mediano absoluto (MAD)

$$\sigma_{MAD} = \text{mediana}(|w[i] - \text{mediana}(w)|). \quad (8)$$

Desta forma, o erro quadrático do limiar t é estimado por

$$\text{SURE}(t, w) = N - 2 \cdot \#\{i: |w[i]| \leq t\} + \sum_{i=1}^N (\min(|w[i]|, t)^2), \quad (9)$$

na qual $\#\{i: |w[i]| \leq t\}$ indica uma soma conectada entre o vetor de índices e os componentes *wavelet* com módulo inferior ou igual a t .

Um novo limiar t_s pode, então, ser definido de forma a minimizar o risco de Stein, ou seja

$$t_s = \underset{0 \leq t \leq \sigma \sqrt{2 \log N}}{\text{argmin}} \text{SURE}(t, w). \quad (10)$$

Este método de limiarização é adaptativo e de simples implementação [2], e por isso é utilizado largamente mesmo em

denoisers que utilizam transformações diferentes da *wavelet* [6].

III. ESPARSIDADE

O método descrito na seção anterior é, se comparado a métodos modernos mais complexos como [6], relativamente simples e de baixo custo computacional. Porém a etapa de transformação não é totalmente adaptativa e deixa a cargo do usuário definir uma *wavelet* mãe e uma escala para decompor o sinal. Como observado em [2], o MSE do sinal recuperado possui grande variação em função desses parâmetros, porém em situações reais (sinal original desconhecido) o MSE não está disponível para permitir a adaptabilidade do método.

Como sugerido em [4] e [5], a representação de um sinal é considerada boa quando for esparsa, e esta é a característica que permite a aplicação de *wavelet shrinkage* (Fig. 1). A esparsidade pode ser facilmente calculada e está disponível em situações reais. Em [7], diversas medidas de esparsidade são comparadas e duas métricas são consideradas superiores: a medida de Hoyer e o índice de Gini. Como o índice de Gini depende da ordenação do vetor a ser aferido, neste trabalho optou-se, por simplicidade, pelo emprego da medida de Hoyer. Dado o vetor de coeficientes *wavelet* w , sua esparsidade no sentido de Hoyer é medida por

$$H(w) = \frac{(\sqrt{N} \frac{\|w\|_1}{\|w\|_2})}{(\sqrt{N}-1)}, \quad (11)$$

na qual $\|w\|_1$ e $\|w\|_2$ indicam, respectivamente, a norma L1 e a norma L2 dos coeficientes *wavelet*. Se o vetor de coeficientes possuir apenas uma componente não-nula, as normas L1 e L2 são iguais, logo $H(w) = 1$. Para qualquer outro caso, a norma L1 é superior à norma L2, assim $H(w) < 1$. Portanto a medida de Hoyer representa a esparsidade de um vetor com um número real de valor máximo 1 (esparsidade máxima). A Fig. 2 exemplifica a medida de Hoyer para duas transformações aplicadas ao sinal *blocks*.

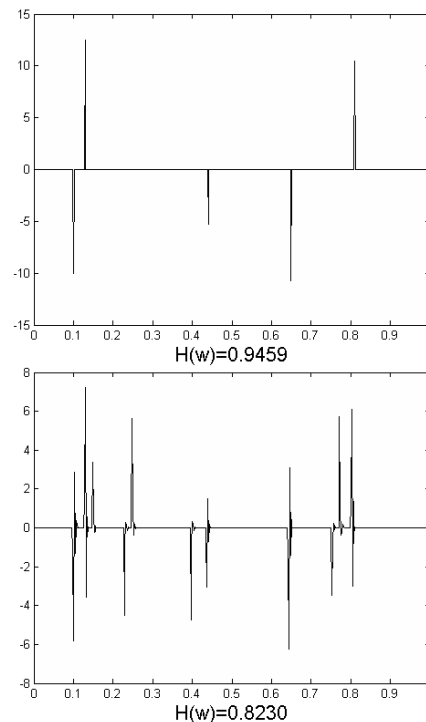


Fig. 2. Transformações aplicadas ao sinal *blocks*. (a) Haar, (b) db7.

É possível observar que a representação do sinal *blocks* com a *wavelet* Haar possui apenas cinco componentes não-nulas. Isto se reflete na medida de Hoyer, que é próxima de um (0,9459). A representação pela *wavelet* db7 possui um número maior de componentes, por isso $H(w)$ é menor que o caso anterior. Pode-se dizer, então, que a função Haar representa o sinal *blocks* melhor que a função db7.

Para calcular a medida de Hoyer, a norma L2 dos coeficientes *wavelet* pode ser determinada por

$$\|w\|_2 = \sqrt{w^T w} = \sqrt{x^T W^T W x}, \tag{12}$$

e pode ser reduzida, para transformações ortogonais, a

$$\|w\|_2 = \sqrt{x^T x} = \|x\|_2. \tag{13}$$

A medida de Hoyer, portanto, pode ser escrita como

$$H(w) = \frac{\sqrt{N}}{\sqrt{N}-1} - \frac{\|w\|_1}{\|x\|_2(\sqrt{N}-1)}. \tag{14}$$

Como N e $\|x\|_2$ são constantes dado um conjunto de amostras, $H(w)$ varia exclusivamente com $\|w\|_1$. De outra forma, pode-se dizer que para qualquer transformação ortogonal, a representação mais esparsa no sentido de Hoyer é aquela cuja norma L1 dos coeficientes é mínima. Logo, dado um conjunto de amostras, a esparsidade pode ser facilmente aferida.

IV. TESTES E RESULTADOS

Para avaliar a relação entre a esparsidade da representação de um sinal e o desempenho do *denoiser* foram utilizados quatro sinais, conforme [2]. Os sinais estão expostos na Fig. 3. Para cada sinal, uma versão com ruído foi gerada (Fig. 4) de forma a ter SNR de 40dB e foi aplicado o método de *wavelet shrinkage* nas 10 primeiras escalas de 22 transformações ortogonais (Apêndice I). Portanto, para cada sinal foram geradas 220 versões após a aplicação do *denoiser*, cada uma com um determinado desempenho (MSE e SNR) e esparsidade (Norma L1).

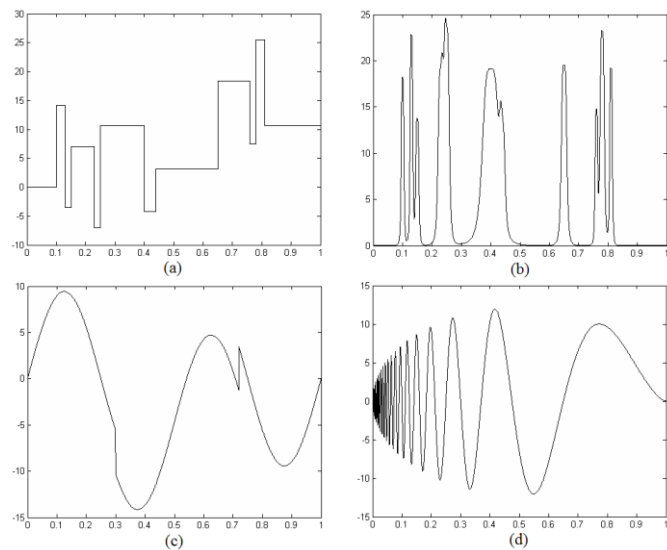


Fig. 3 Sinais de dinâmica variável sem ruído. (a) *Blocks*, (b) *bumps*, (c) *heavy sine*, (d) *doppler*.

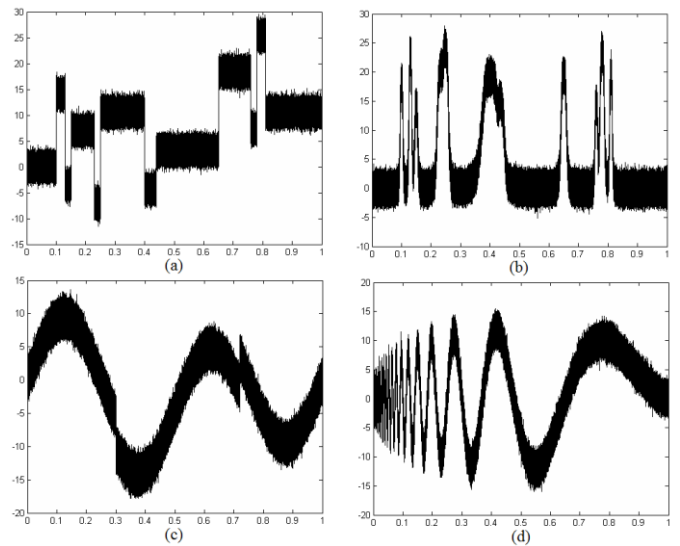


Fig. 4 Sinais contaminados por ruído branco (SNR=40dB). (a) *Blocks*, (b) *bumps*, (c) *heavy sine*, (d) *doppler*.

Como produto do teste descrito acima, os gráficos abaixo (Fig. 5-8) possuem 220 pontos para cada sinal e representam os valores de MSE e SNR em função da norma L1.

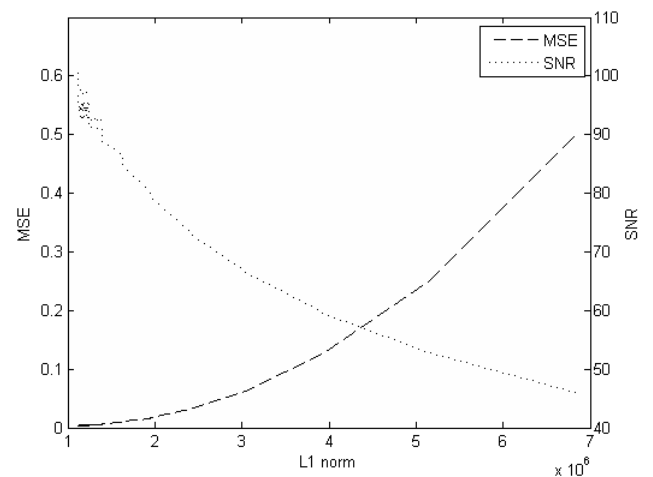


Fig. 5 Desempenho do método (MSE e SNR) em função da esparsidade (norma L1) para o sinal *blocks*. O gráfico possui 220 pontos para cada função.

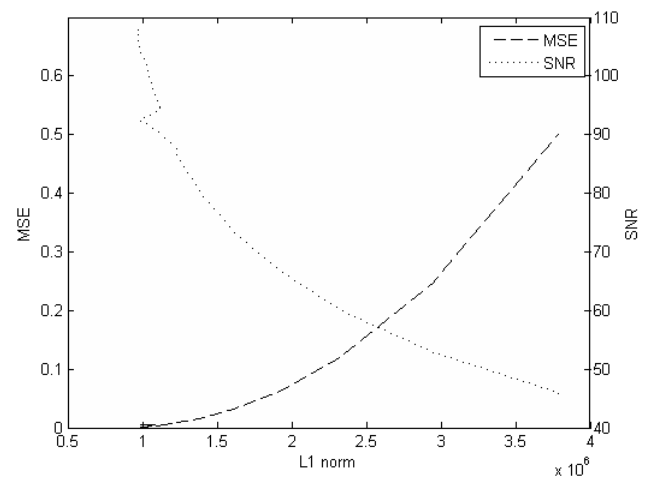


Fig. 6 Desempenho do método (MSE e SNR) em função da esparsidade (norma L1) para o sinal *bumps*. O gráfico possui 220 pontos para cada função.

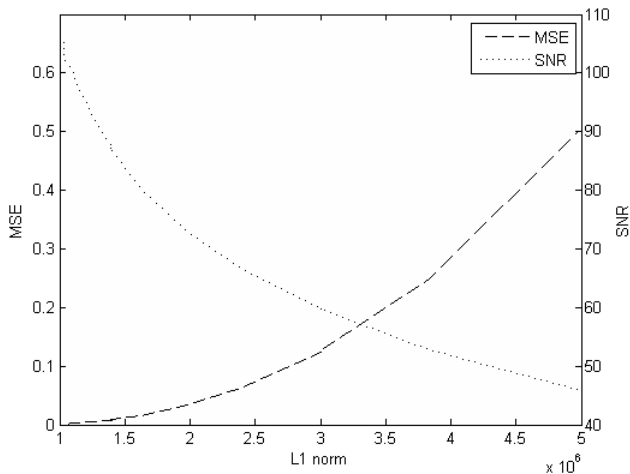


Fig. 7 Desempenho do método (MSE e SNR) em função da esparsidade (norma L1) para o sinal *heavy sine*. O gráfico possui 220 pontos para cada função.

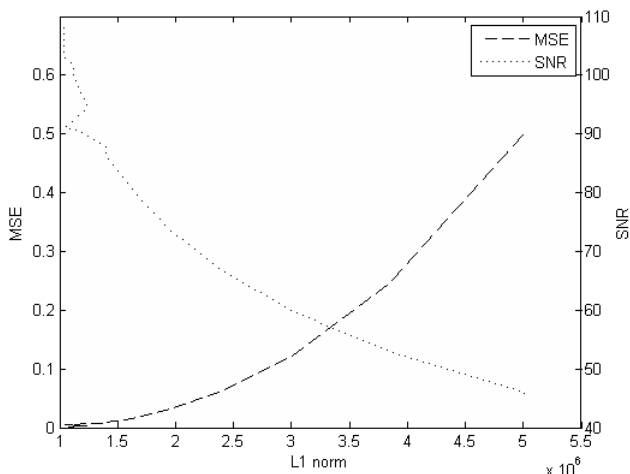


Fig. 8 Desempenho do método (MSE e SNR) em função da esparsidade (norma L1) para o sinal *doppler*. O gráfico possui 220 pontos para cada função.

Para o sinal *heavysine* o gráfico não possui mínimos locais, o que mostra que para qualquer conjunto de parâmetros, um novo conjunto que diminua a norma L1 resultará em um desempenho melhor. Para os demais sinais, há mínimos locais, mas o ponto de melhor desempenho ainda foi, como para *heavysine*, o mínimo global em relação a norma L1. Com SNR inicial de 40dB, o pior ponto de operação em todos os sinais gerou SNRs em torno de 45dB, e em nenhum caso houve mínimos locais com SNR abaixo de 90dB.

Foi possível observar, portanto, que em todos os casos analisados o conjunto de parâmetros que gerou a representação mais esparsa gerou, também, o melhor desempenho. Em todos os casos o ponto de melhor desempenho gerou SNRs acima de 100dB, aproximadamente 55dB superiores aos pontos de pior desempenho. Nenhum conjunto de parâmetros gerou sinais com MSE e SNR piores que o sinal contaminado.

Para os testes realizados, portanto, duas afirmações foram confirmadas. Em primeiro lugar, o desempenho do método depende grandemente dos parâmetros da transformação (a serem definidos pelo usuário), como dito anteriormente. Em segundo lugar, a esparsidade da representação do sinal (que

pode ser aferida em situações reais) possui forte relação com o desempenho do método. Portanto, a minimização da norma L1 da representação do sinal poderá tornar o método de *wavelet shrinkage* adaptativo.

A Fig. 9 mostra, para cada sinal, o sinal recuperado no ponto de operação com menor norma L1.

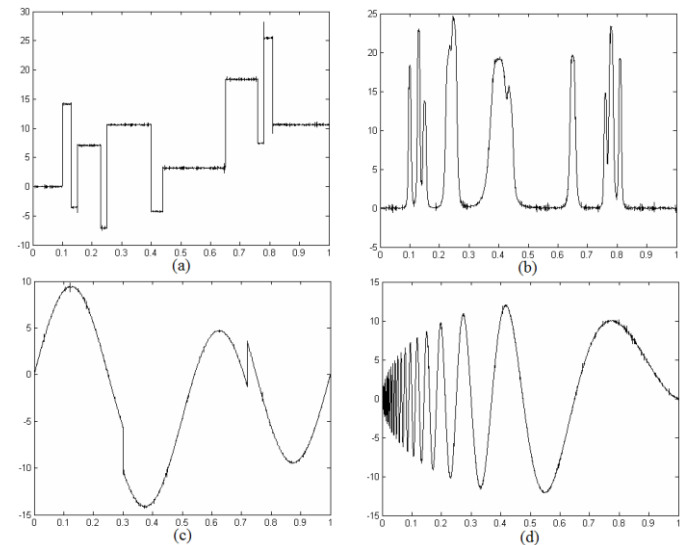


Fig. 9 Sinais após a aplicação do wavelet shrinkage. (a) *Blocks*, (b) *bumps*, (c) *heavy sine*, (d) *doppler*.

V. CONCLUSÕES

O método de *wavelet shrinkage* para *denoising* é simples e eficiente, porém não totalmente adaptativo. Foi possível observar que a adaptabilidade do método pode ser obtida pela norma L1 dos coeficientes *wavelet*, característica que influencia o desempenho final e é de fácil aferição. Um método, portanto, que busque a minimização na norma L1 no domínio *wavelet* poderá tornar o método de *shrinkage* mais eficiente e robusto para diversos sinais.

AGRADECIMENTOS

O autor Vinícius Ávila da Silva agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento da pesquisa.

REFERÊNCIAS

- [1] C. A. Metzler; A. Maleki; R. G. Baraniuk, *From denoising to compressed sensing*. IEEE Transaction on Information Theory, v. 62, n. 9, pp. 5117-5144, 2016.
- [2] D. L. Donoho e I. M. Johnstone, *Adapting to unknown smoothness via wavelet shrinkage*. Journal of the American statistical association, v. 90, n. 432, pp. 1200-1224, 1995.
- [3] C. Stein, *Estimation of the Mean of a Multivariate Normal Distribution*. The Annals of Statistics, v. 9, pp. 1135-1151, 1981.
- [4] I. Daubechies, *Ten lectures on wavelets*. Society for industrial and applied mathematics, 1992.
- [5] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.
- [6] G. Yu, S. Mallat, E. Bacry, *Audio denoising by time-frequency block thresholding*. IEEE Transactions on Signal processing, v. 56, n. 5, pp. 1830-1839, 2008.
- [7] N. Hurley, S. Rickard, *Comparing measures of sparsity*. IEEE Transactions on Information Theory, v. 55, n. 10, pp. 4723-4741, 2009.

APÊNDICE I

Transformações wavelet ortogonais utilizadas: haar, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, sym7, sym8, coif1, coif2, coif3, coif4 e coif5.