

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Escolha da Base *Wavelet* na Recuperação de Dados Sísmicos via *Compressive Sensing*

Francisco Iranildo Ferreira do Nascimento Gomes¹

Programa de Pós-Graduação em Ciências e Engenharia do Petróleo, UFRN, Natal, Brasil

Éberton da Silva Marinho²

Programa de Pós-Graduação em Ciências e Engenharia do Petróleo, UFRN, Natal, Brasil

Tiago Cavalcanti da Rocha³

Programa de Pós-Graduação em Ciências e Engenharia do Petróleo, UFRN, Natal, Brasil

Gilberto Corso⁴

Departamento de Biofísica e Farmacologia, Centro de Biociências, UFRN, Natal, Brasil

Liacir dos Santos Lucena⁵

Programa de Pós-Graduação em Ciências e Engenharia do Petróleo, UFRN, Natal, Brasil

Resumo. Estudamos a técnica *Bayesian Compressive Sensing* com várias bases *Wavelets* aplicadas à recuperação de traços sísmicos. Em uma análise comparativa, a *Symlet 9* apresentou a melhor performance, seguida pela *Daubechies 6*, *Coiflet 3* e, por último, *Biorthogonal 3.9*.

Palavras-chave. *Bayesian Compressive Sensing*, *Wavelets*, Traço Sísmico, Recuperação de sinais, *Symlet*.

1 Introdução

A exploração de petróleo enfrenta crescentes desafios. Novos problemas como o imageamento de regiões profundas, por exemplo, o ambiente do pré-sal, mostram que os métodos tradicionais de sísmica atingiram os seus limites de eficácia. Na busca de melhoras na resolução das imagens

¹ ironrde@gmail.com

² ebertonsm@gmail.com, eberton.marinho@ifrn.edu.br

³ tiago.cavalcanti@unp.br

⁴ gfcorso@gmail.com

⁵ liacir.lucena@gmail.com

da subsuperfície são necessários novos enfoques e novas técnicas. Desta forma, faz-se necessário um estudo e desenvolvimento de métodos inovadores na exploração sísmica [5,16].

Como na exploração de hidrocarbonetos o volume de dados sísmicos envolvido é imenso, exige-se uma grande capacidade de processamento e armazenamento de dados. Uma vez que muitos desses dados apresentam redundância, trabalhar com representações esparsas torna-se bastante viável neste contexto. Ademais, representá-los esparsamente significa reduzir a sua dimensão, para tanto é usada a Transformada *Wavelet* [7, 16] ao novo paradigma *Compressive Sensing* (CS) [7] no processo de aquisição comprimida de sinais sísmicos tendo sempre em vista uma posterior recuperação com boa performance.

Neste trabalho, consideramos o problema da recuperação de um sismograma real obtido da literatura. Sobre estes dados aplicamos o método *Bayesian Compressive Sensing* (BCS) [9] para diferentes taxas de amostragem, tendo como bases as *Wavelets: Biorthogonal 3.9, Coiflet 3, Daubechies 6 e Symlet 9*. Por fim, computamos a performance de cada *Wavelet* comparando o erro relativo entre o sinal original e o recuperado por CS.

2 Metodologia

2.1 Aquisição e processamento de dados sísmicos

O processamento de sinal sísmico é uma tarefa importante em sondagem geofísica e representa um desafio permanente na exploração de petróleo [1]. A figura 1 abaixo, à esquerda, ilustra uma sondagem sísmica, figura adaptada de <http://www.ektinteractive.com/exploration/exploration-fundamentals/>. Nesta representação se pode ver uma fonte sísmica emitindo uma onda que se propaga pela subsuperfície e um conjunto de geofones que registram as ondas mecânicas. Após a detonação do tiro, diversas reflexões desta frente de onda são captadas pelos receptores. O sinal captado por cada geofone é chamado de traço sísmico [14], à direita da figura são representados traços sísmicos. Em (a) temos a série de dados originais do traço sísmico (traço de número 16) com 1501 pontos; e em (b) temos a superposição do traço sísmico recuperado (em vermelho) e do original (em azul). Neste exemplo ilustrativo foi usada uma taxa de amostragem de 30% e a base *Wavelet Coiflet 3*.

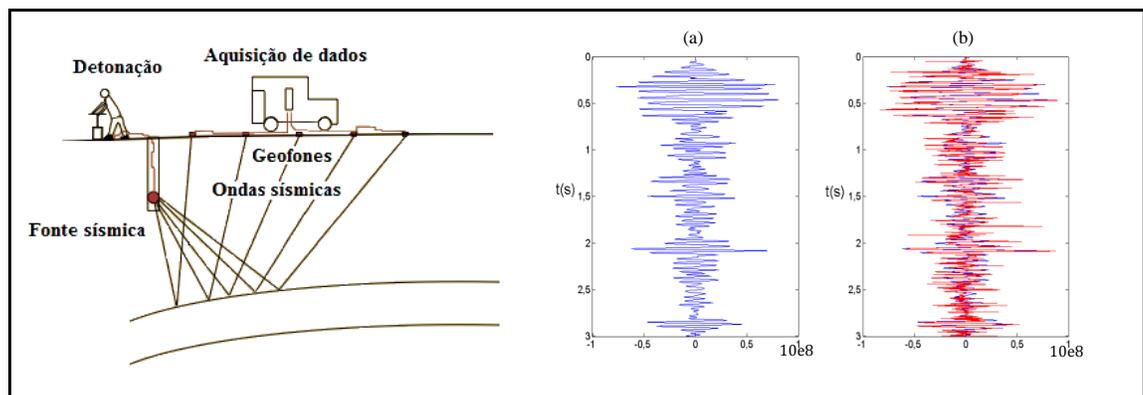


Figura 1: À esquerda, representação esquemática de uma aquisição sísmica. À direita, ilustração do processo de recuperação do sinal com, (a), a série de dados do traço sísmico original, (b), a superposição do traço sísmico recuperado (em vermelho) e do original (em azul).

2.2 Funções de Base *Wavelets* Discreta

2.2.1 Visão geral sobre *Wavelets*

A transformada *Wavelet* é uma ferramenta matemática para decompor sinais em suas partes constituintes. Com esta técnica é possível analisar as frequências presentes no sinal distintamente com uma precisão adequada [3]. Os algoritmos de *Wavelets* processam dados em diferentes escalas ou resoluções e independentemente do objeto de interesse ser uma série temporal ou uma imagem bidimensional este método oferece uma representação elegante e detalhada [10, 11]. Se pudermos escolher as *Wavelets* que melhor se adaptam a série temporal, ou truncarmos os coeficientes de *Wavelets* menores do que um valor previamente estabelecido, o dado original se torna esparsamente representado [3].

2.2.2 *Wavelets* discretas

Uma *Wavelet* é uma função no domínio real e compacto, tal que a família de funções $\psi_{j,k}(i) = 2^{-j/2}\psi(2^{-j}i - k)$ onde j e k são inteiros arbitrários e ψ seja uma base ortonormal em $L^2(\mathbb{R})$ [3]. Nesse trabalho, utilizamos a Transformada *Wavelet* Discreta (*Discrete Wavelet Transform* - DWT) visto que a informação contida nos traços sísmicos aqui trabalhados está discretizada em vetores coluna de 1501 componentes. Há muitas razões pelas quais as *Wavelets* trabalham bem nas mais diversas aplicações e provavelmente a mais importante seja o fato da transformada *Wavelet* discreta ser uma transformada local [4]. Em nossa implementação usamos a função *wavedec* do MATLAB, disponível em http://www.mathworks.com/help/wavelet/ref/wavedec.html?s_tid=srchtitle, que realiza uma análise *Wavelet* unidimensional multinível com cada uma das *Wavelets* específicas ou com um filtro de decomposição *Wavelet* específico. Com relação às características das quatro *Wavelets* que usamos, a *Biorthogonal 3.9* não apresenta a propriedade de ortogonalidade, embora seja a única simétrica dessas quatro. Por outro lado, a *Daubechies 6* é assimétrica, enquanto a *Symlet 9* e a *Coiflet 3* apresentam a propriedade de quase simetria [13].

2.3 *Bayesian Compressive Sensing*

Antes de introduzirmos o BCS, precisamos apresentar, embora brevemente, o funcionamento da técnica CS. Se $x \in \mathbb{R}^N$ é um sinal e uma matriz $\Psi := [\psi_1, \psi_2, \dots, \psi_N]$ for uma base de \mathbb{R}^N , tem-se uma transformada

$$x = \Psi s \tag{1}$$

em que s representa os coeficientes da transformada. Para a maioria dos sinais naturais x , a maioria das componentes do vetor s tem amplitude desprezível [17]. Dizemos que x é K -esparso se, e somente, se apenas $K \ll N$ coeficientes de s são significativos.

A condição de incoerência é muito importante para a CS. Para as medições incoerentes, reescrevemos $s = s_k + s_e$, em que s_k é a componente significativa K e s_e são os $N - K$ menores coeficientes definidos como zero [17]. Portanto, temos

$$x = \Psi s = \Psi s_k + \Psi s_e = \Psi s_k + n_e \quad (2)$$

onde cada elemento de n_e pode ser modelado por uma gaussiana de média zero com variância apropriada. O próximo passo consiste em usar uma matriz de sensoriamento Φ , que é $N \times N$, para medir s . Geralmente, assume-se que Φ satisfaz a conhecida propriedade da isometria restrita (*Restricted Isometry Property* - RIP) [17]. Uma matriz Φ satisfaz a RIP de ordem K se existe um $\delta_K \in (0, 1)$ tal que

$$(1 - \delta_K) \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_K) \|x\|_2^2, \quad (3)$$

vale para todo $x \in \Sigma_K = \{x : \|x\|_0 \leq K\}$.

Se uma matriz Φ satisfaz a RIP de ordem $2K$, então podemos interpretar (3) dizendo que Φ preserva aproximadamente a distância entre qualquer par de vetores K -esparso [2]. Seja Φ contendo entradas gaussianas independentes e identicamente distribuídas. Então a matriz Φ é incoerente com qualquer Ψ fixo. As medições serão calculadas por

$$y = \Phi x \quad (4)$$

em que $y \in \mathbb{R}^M$ com $M \ll N$ [17].

Podemos aplicar o enfoque Bayesiano [9, 15] na recuperação de traços sísmicos, associando o vetor s desconhecido a uma distribuição de probabilidade a priori $p(s|\gamma)$, que modela nosso conhecimento prévio a cerca dos coeficientes deste vetor. A distribuição condicional $p(y|s, \beta)$, onde $\beta = 1/\sigma^2$ é a precisão do ruído, modela o vetor y como um conjunto de observações de um processo aleatório. Para ambas as distribuições, a presença do *relevance vector machine* (RVM) é indicada pelos hiperparâmetros γ e β que são atualizados a cada iteração até que o critério de convergência seja alcançado. Realizamos uma pré-seleção das doze *Wavelets* seguintes: Daubechies (db3, db6 e db9), Coiflets (coif1, coif3 e coif5), Biortogonais (bior2.2, bior2.4 e bior3.9) e Symlets (sym3, sym5 e sym9).

3 Análise Comparativa entre as *Wavelets*

A partir do sismograma escolhido [12] testamos o método de reconstrução *Bayesian Compressive Sensing* para taxas de amostragem de 10, 20, 30, 40 e 50% de cada um dos 33 traços sísmicos do sismograma. Assumindo que o sinal original é um vetor representado por x e o sinal recuperado por x' , o erro relativo é dado por $\|x - x'\|/\|x\|$. Utilizamos quatro *Wavelets* (*Biorthogonal 3.9*, *Daubechies 6*, *Coiflet 3* e *Symlet 9*) como bases de transformação do sinal original em um sinal esparso. Para cada taxa de amostragem realizamos a recuperação do sinal

100 vezes, estimamos o erro relativo e computamos a média aritmética dos erros. Na execução dos testes pelo método BCS foi usada a função *BCS_fast_rvm* da biblioteca SparseLab versão 2.1 disponível em <https://sparselab.stanford.edu/>.

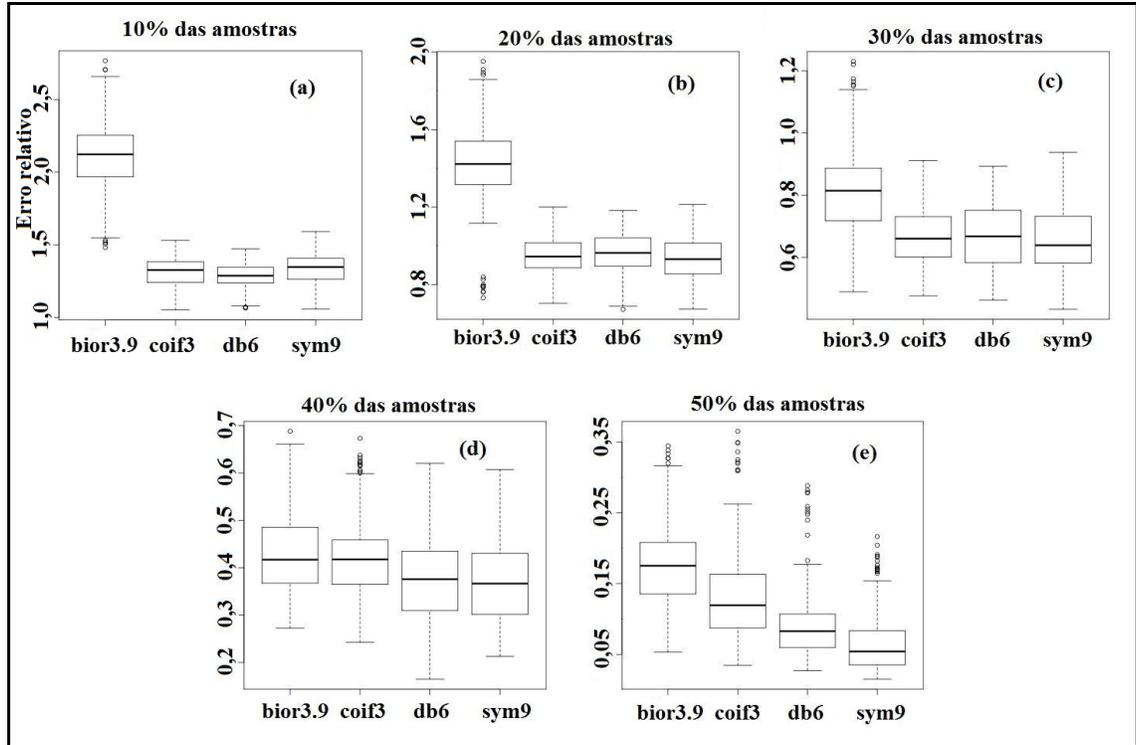


Figura 2: Gráficos do tipo *box plot* dos erros relativos da reconstrução via CS para as bases *Wavelets* com taxas de amostragem de 10, 20, 30, 40 e 50%.

A Tabela 1 apresenta resultados comparativos dois a dois para todas as bases *Wavelets* estudadas. Observa-se que a *bior3.9* apresenta $p < 0,05$ em relação a todas as demais *Wavelets* (exceto para 40%), e, pela Figura 2 e Tabela 1, concluímos que esta é a pior base para recuperação, dentre as estudadas. Para 50% todas as wavelets são diferentes e a *sym9* apresenta a melhor performance conforme indica a Figura 2. Em resumo, é clara a melhor performance da *sym9* para taxas de amostragem de 30%, 40% e 50%.

Tabela 1: Comparação dois a dois entre bases *Wavelets* usando *post-test* ANOVA.

Wavelets	valor de p				
	10%	20%	30%	40%	50%
coif3-bior3.9	< e-05	< e-05	< e-05	0,313	< e-05
db6-bior3.9	< e-05	< e-05	< e-05	< e-05	< e-05
sym9-bior3.9	< e-05	< e-05	< e-05	< e-05	< e-05
db6-coif3	0,067	0,91	0,98	< e-05	< e-05

sym9-coif3	0,092	0,64	0,12	< e-05	< e-05
sym9-db6	< e-05	0,25	0,26	0,413	< e-05

4 Conclusões e trabalhos futuros

Utilizamos um conjunto real de dados sísmicos sobre o qual aplicamos o *Bayesian Compressive Sensing*. Para isso, adotamos o critério da taxa do erro, $\|x - \hat{x}\|_2 / \|x\|_2$, da recuperação do sinal original com 100 iterações usando o método BCS no qual a *Daubechies 6*, a *Coiflet 3*, a *Symlet 9* e a *Biortogonal 3.9* tiveram melhor desempenho. Dessas quatro, em nossa análise a *Wavelet Biortogonal 3.9* se mostrou a pior base de todas para as taxas de amostragens estudadas. Para taxas de amostragens maiores (30, 40 e 50%), a *Symlet 9* e a *Daubechies 6* se mostraram as bases mais promissoras, sendo que, dentre as duas, a *Symlet 9* apresentou a melhor performance para a taxa de amostragem de 50%.

Para sermos inteiramente honestos, a técnica do CS e sua importância em geofísica vai além da aquisição compacta de sinais. O CS é usado também na compactação de imagens [17] e na regularização de dados. Esta última aplicação se refere ao problema de dados faltantes no sismograma que precisam ser de alguma forma interpolados. O CS tem sido usado para recuperar dados de geofones que não funcionam durante a aquisição do sinal sísmico [8, 6].

A escolha da base *Wavelet* mais apropriada à recuperação de sinais utilizando o CS pode reduzir consideravelmente o volume de dados adquiridos durante a exploração sísmica. O potencial impacto econômico desta técnica nos faz prosseguir neste caminho e na exploração de um leque ainda maior de famílias *Wavelet*.

Referências

- [1] G. Corso, P.S. Kuhn, L.S. Lucena, Z.D. Thomé, Seismic ground roll time-frequency filtering using the gaussian wavelet transform, (2001), DOI: 10.1016/S0378-4371(02)01379-1.
- [2] M. Davenport, The restricted isometry property, OpenStax-CNX, (2011), to appear: <https://legacy.cnx.org/content/m37171/1.6/>.
- [3] P. C. de Lima, Wavelets: uma introdução, Matemática Universitária, vol. 33, 13-44, (2002).
- [4] P. J. V. Fleet, Discrete Wavelets Transformations – An Elementary Approach with Applications, John Wiley & Sons, (2008).
- [5] F. J. Herrmann, M. P. Friedlander and Ö. Yilmaz, Fighting the curse of dimensionality: compressive sensing in exploration seismology, IEEE Signal Processing Magazine, vol. 29, 88-100, (2011), DOI: 10.1109/MSP.2012.2185859.

- [6] F. J. Herrmann and G. Hennenfent, Non-parametric seismic data recovery with curvelet frames, *Geophysical Journal International*, vol. 173, 233-248, (2008), DOI: 10.1111/j.1365-246X.2007.03698.x.
- [7] F. J. Hermann, H. Wason and T. T. Y. Lin, Compressive sensing in seismic exploration: an outlook on a new paradigm, *CSEG Recorder*, vol. 36, 19-33, (2011), to appear in: *CSEG Recorder*.
- [8] B. Jafarpour, V. K. Goyal, D. B. McLaughlin and W. T. Freeman, Transform-domain sparsity regularization for inverse problems in Geosciences, *Geophysics*, R69-R83, (2009), DOI: 10.1190/1.3157250.
- [9] S. Ji, Y. Xua and L. Carin, Bayesian Compressive Sensing, *IEEE Transactions on Signal Processing*, vol. 56, (2008), DOI: 10.1109/TSP.2007.914345.
- [10] S. Mallat, *A wavelet tour of signal processing the Sparse way*, Academic Press, (2008).
- [11] S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet represent, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, 674-693, (1989), DOI: 10.1109/34.192463.
- [12] W. A. Mousa, Abdullatif A. Al-Shuhail, *Processing of Seismic Reflection Data Using MATLAB*, Morgan & Claypool, Cap. 4, (2011), DOI: 10.2200/S00384ED1V01Y201109SPR010.
- [13] D. P. Radunović, *Wavelets - From Math to Practice*, Springer-Verlag, (2009).
- [14] J. M. T. Romano, R. R. Lopes e M. Tygel, *Processamento Digital de Sinais Geofísicos, XXIX Simpósio Brasileiro de Telecomunicações*, (2011), to appear: http://www.decom.fee.unicamp.br/~rlopes/conf/sbirt11_Proc_Sismico.pdf.
- [15] M. W. Seeger and H. Nickisch, Compressed sensing and Bayesian experimental design, in *ICML*, 912–919, (2008).
- [16] A. S. Spanias, S. B. Jonsson and S. D. Stearns, Transform Methods for Seismic Data Compression, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, 407 - 416, (1991), DOI: 10.1109/36.79431.
- [17] S. Yang, Z. Zhang, H. Du, Z. Xia and H. Qin, The Compressive Sensing Based on Biorthogonal Wavelet Basis, *2010 International Symposium on Intelligence Information Processing and Trusted Computing*, 479-482, (2010), DOI: 10.1109/IPTC.2010.146.