Sparse reconstruction of correlated multichannel activity

Sem Peelman[1], Joachim Van der Herten[1],
Maarten De Vos[2,3], Wen-shin Lee[1], Sabine Van Huffel[2,4] and Annie Cuyt[1]

Abstract—**Parametric methods for modeling sinusoidal signals with line spectra have been studied for decades. In general, these methods start by representing each sinusoidal component by means of two complex exponential functions, thereby doubling the number of unknown parameters. Recently, a Hankel-plus-Toeplitz matrix pencil method was proposed which directly models sinusoidal signals with discrete spectral content. Compared to its counterpart, which uses a Hankel matrix pencil, it halves the required number of time-domain samples and reduces the size of the involved linear systems.**

The aim of this paper is twofold. Firstly, to show that this Hankel-plus-Toeplitz matrix pencil also applies to continuous spectra. Secondly, to explore its use in the reconstruction of real-life signals. Promising preliminary results in the reconstruction of correlated multichannel electroencephalographic (EEG) activity are presented. A principal component analysis preprocessing step is carried out to exploit the redundancy in the channel domain. Then the reduced signal representation is successfully reconstructed from fewer samples using the Hankel-plus-Toeplitz matrix pencil. The obtained results encourage the future development of this matrix pencil method along the lines of well-established spectral analysis methods.

I. INTRODUCTION

Nowadays, there is the possibility of recording an almost unlimited amount of biomedical signals (MRI, EEG, ECG). Additionally, there is the desire to record this data in a mobile setting [5], which is often associated with wireless settings. However, the limited battery life is a bottleneck in continuous wireless transmission and therefore interest is raised in compression techniques and sparse methods. The goal is to efficiently reduce the number of data samples without reducing the information content. The general framework of Compressed Sensing (CS) exploits the sparse nature of the signals for this purpose, and has gained a lot of attention in signal processing [6]. Several attempts have been made to explore the possibility of using CS techniques in EEG [2], [1]. As EEG is not sparse in the original time nor the frequency domain, it is still a topic of ongoing debate if EEG can be sparsely represented in any domain [14].

Many spectral analysis tools can be used to efficiently capture the information content of a signal, e.g., [13], [12],

[1]Dept. Mathematics & Computer Science, Universiteit Antwerpen, Antwerpen, Belgium. Email: {sem.peelman, joachim.vanderherten, wen-shin.lee, annie.cuyt}@ua.ac.be.

[2]SCD/SISTA, Dept. of Electrical Engineering, KU Leuven, Leuven, Belgium. Email: sabine.vanhuffel@esat.kuleuven.be

[3]Dept. of Psychology, University of Oldenburg, Oldenburg, Germany. Email: maarten.de.vos@uni-oldenburg.de

[4]iMinds Future Health Department, Belgium.

[8]. But these methods normally represent each sinusoidal component by a sum of two complex exponentials, thereby doubling the number of unknown parameters and the required number of samples.

This paper presents a novel sampling technique based on a modified matrix pencil method [7], which directly reflects the damped sinusoidal components and thus halves the number of samples, as well as the size of the corresponding linear systems. We combine a Principal Component Analysis (PCA)-based compression step with the new sampling technique and illustrate its performance with an epileptic EEG example.

II. MATHEMATICAL BACKGROUND

We outline the two mathematical tools that facilitate the experiments in Section III. Section II-A recapitulates the established statistical technique of PCA. Section II-B applies an existing matrix pencil method [8], [11], [7] for sparse polynomial interpolation to trigonometric interpolation. The resulting matrix structure becomes a more complicated Hankel-plus-Toeplitz structure.

A. Principal component analysis

Let the matrix $X \in \mathbb{C}^{d \times n}$ denote a dataset, consisting of n samples each having d dimensions. We assume that the centre of mass of our dataset lies in the origin, i.e. the mean of each row of X is zero. Principal component analysis is widely used to reduce the dimensionality d of a dataset to its intrinsic dimension $\tilde{d}$, i.e. the number of independent variables that allows for a satisfactory representation of the dataset. It constructs $\tilde{d}$ orthonormal vectors $u_1, \ldots, u_{\tilde{d}} \in \mathbb{C}^d$, the principal components, such that the data have maximal spread in the subspace spanned by these components. Formally the i-th principal component is the solution to the constrained optimization problem

$$\arg\max_{u \in \mathbb{C}^d, \|u\|_2 = 1} \left\| u^T X \right\|_2^2 \quad \text{s.t.} \quad u^T u_j = 0 \text{ for all } j < i.$$

Equivalently, u_i is the i-th left singular vector of X with corresponding singular value σ_i. Define the matrix $U \in \mathbb{C}^{d \times \tilde{d}}$ whose columns are the principal components. A dimension reduced representation of the original dataset in terms of projections on the principal components is given by $Y = U^T X \in \mathbb{C}^{\tilde{d} \times n}$. An approximation of the original dataset by a linear combination of the principal components is $\tilde{X} = UY \in \mathbb{C}^{d \times n}$. Finally, in order to decide on the intrinsic dimension $\tilde{d}$ of X one can observe the so-called retained variance $\sum_{i=1}^{\tilde{d}} \sigma_i^2 / \sum_{i=1}^{d} \sigma_i^2$ and require this to exceed a certain threshold (see, e.g. [9]).

B. Sparse trigonometric interpolation

In this section we apply a matrix pencil method to reconstruct a K-sparse trigonometric sum from $2K$ samples in the time domain. Let

$$y(t) = \sum_{i=1}^{K} c_i \phi(\alpha_i; t) \tag{1}$$

where $\phi(\alpha; t) = \cos(2\pi\alpha t)$. The problem is to determine $c_i \in \mathbb{C}\backslash\{0\}$ and the pairwise distinct $\alpha_i \in \mathbb{C}$. We restrict ourselves to the case where $\mathrm{Re}(\alpha_i) \in [0, B]$ for some known $B \in \mathbb{R}^+$. The basic method was first proposed in [7] for the reconstruction of K-sparse trigonometric sums with integer frequencies ($\alpha_i \in \mathbb{N} \cap [0, B]$) and derived as a special case of sparse interpolation in the Chebyshev basis [11], [10]. We show that the frequencies may take values in a real interval ($\alpha_i \in [0, B]$) or even complex values ($\mathrm{Re}(\alpha_i) \in [0, B]$). The latter implies that the above matrix pencil method applies to an even broader class of functions, since $\cos(\mathrm{i}t) = \cosh(t)$ where $\mathrm{i} = \sqrt{-1}$.

The presented algorithm involves two stages. At first the frequencies $\alpha_1, \ldots, \alpha_K$ are obtained by solving a structured generalized eigenvalue problem. Once the frequencies are known, the coefficients $c_1, \ldots, c_K$ are found by solving a classic linear interpolation problem. These $2K$ unknowns are determined from $2K$ samples $\{y_s = y(t_s)\}_{s=0}^{2K-1}$, where $t_s = s\Delta$ for some $\Delta \in \,]0, 1/(2B)[$.

When downsampling the signal even further, below the Nyquist rate, an additional structured linear system needs to be solved to identify the most appropriate frequencies α_i in (1). The total computational complexity remains $\mathcal{O}(K^2)$ which compares favorably to other methods [3]. The full mathematical details of this new sparse interpolation technique are described in a forthcoming paper [4].

(i) Determining $\alpha_1, \ldots, \alpha_K$. In what follows we denote $\phi_{i,s} = \phi(\alpha_i; t_s)$. The following Lemmata will provide a way to determine $\phi_{1,1}, \ldots, \phi_{K,1}$. Since $0 \leq 2\pi \,\mathrm{Re}(\alpha_i)\Delta < \pi$ we can subsequently determine α_i uniquely from $\phi_{i,1}$ as $\arccos(\phi_{i,1})/(2\pi\Delta)$.

Lemma 1: The matrix $\Phi = [\phi_{k,l-1}]_{k,l=1}^{K} \in \mathbb{C}^{K \times K}$ is nonsingular.

Proof: Note that the values $\phi_{1,1}, \ldots, \phi_{K,1}$ are mutually distinct due to the choice of Δ. Now suppose there exists a vector $a = [a_1, \ldots, a_K]^T \in \mathbb{C}^K$ such that $\Phi a = 0$. Then for all $i = 1, \ldots, K$

$$\sum_{s=1}^{K} a_s T_{s-1}(\phi_{i,1}) = \sum_{s=1}^{K} a_s \phi_{i,s-1} = 0$$

where T_i denotes the i-th Chebyshev polynomial of the first kind. Since a polynomial of degree $K-1$ cannot have K distinct roots, this proves our Lemma. ∎

Lemma 2: The Hankel-plus-Toeplitz matrices $Y_1, Y_2 \in \mathbb{C}^{K \times K}$ given by

$$Y_1 = [y_{k+l-2} + y_{k-l}]_{k,l=1}^{K}$$

$$Y_2 = \frac{1}{2}\left([y_{k+l-1} + y_{k-l+1}]_{k,l=1}^{K} \right.$$
$$\left. + [y_{k+l-3} + y_{k-l-1}]_{k,l=1}^{K} \right)$$

satisfy

$$Y_1 = \Phi^T C \Phi \qquad\qquad Y_2 = \Phi^T C E \Phi$$

where $C = \mathrm{diag}(2c_1, \ldots, 2c_K)$ and $E = \mathrm{diag}(\phi_{1,1}, \ldots, \phi_{K,1})$. The matrix Y_1 is nonsingular.

Proof: The above matrix decompositions follow from the relation $2\phi_{i,k}\phi_{i,l} = \phi_{i,k+l} + \phi_{i,k-l}$. Looking at the (k,l)-th entry of the first matrix equality, we have

$$\left(\Phi^T C \Phi\right)_{(k,l)} = \sum_{i=1}^{K} 2c_i \phi_{i,k-1}\phi_{i,l-1}$$

$$= \sum_{i=1}^{K} c_i \left(\phi_{i,k+l-2} + \phi_{i,k-l}\right) = (Y_1)_{(k,l)}.$$

Analogously this can be done for Y_2. The nonsingularity of Y_1 follows from the fact that both matrices Φ and C are nonsingular. ∎

The values $\phi_{1,1}, \ldots, \phi_{K,1}$ are the eigenvalues of the structured generalized eigenvalue problem

$$Y_2 v = \lambda Y_1 v. \tag{2}$$

This follows from the Lemmata 1 and 2, since

$$Y_2 \Phi^{-1} = Y_1 \Phi^{-1} E$$

implies that the i-th column of Φ^{-1} is an eigenvector of (2) with corresponding eigenvalue $\phi_{i,1}$.

(ii) Determining $c_1, \ldots, c_K$. Once the frequencies $\alpha_1, \ldots, \alpha_K$ are known, the matrix Φ can be constructed explicitly. Solving the Vandermonde-like system of linear equations

$$\Phi^T c = y \tag{3}$$

with $y = [y_0, \ldots, y_{K-1}]^T$, leads to the coefficients $c = [c_1, \ldots, c_K]^T$. In general, any subset of K interpolation conditions from $\{y_s = y(t_s)\}_{s=0}^{2K-1}$ suffices to obtain the coefficient vector c.

Remark 1: The problems (2) and (3) only involve the samples $\{y_s = y(t_s)\}_{s=0}^{2K-1}$, because $y_s = y_{-s}$.

III. EXPERIMENT

Section III-A demonstrates the immediate application of the outlined matrix pencil method in the approximation of one channel of an EEG signal. In Section III-B we first exploit the redundancy in the channel domain of a multichannel EEG by a PCA preprocessing step, the matrix pencil method is afterwards used to approximate the dimension reduced signal representation.

A. Modeling singlechannel activity

Given $2K$ equidistant data points $\{y_k\}_{k=0}^{2K-1}$ with sampling period Δ, the presented matrix pencil method constructs a complex-valued K-sparse trigonometric interpolant of the form (1). Such an interpolant exists when Y_1 is nonsingular and all eigenvalues of (2) have multiplicity one. Note that in (3) we only impose interpolation in the first K data points, the remaining interpolation conditions are then automatically fulfilled.

To compare an approximation $\{\tilde{x}_s\}_{s=0}^{N-1}$ to a discrete-time signal $\{x_s\}_{s=0}^{N-1}$ we observe the cross-correlation defined as

$$\text{CC} = \frac{\sum_{s=0}^{N-1}(x_s - M_x)(\tilde{x}_s - M_{\tilde{x}})}{\sqrt{\sum_{s=0}^{N-1}(x_s - M_x)^2}\sqrt{\sum_{s=0}^{N-1}(\tilde{x}_s - M_{\tilde{x}})^2}}$$

where $M_x = \frac{1}{N}\sum_{s=0}^{N-1} x_s$ and $M_{\tilde{x}} = \frac{1}{N}\sum_{s=0}^{N-1} \tilde{x}_s$.

The presented method can be used to obtain a continuous approximation of a discrete-time signal or spectral analysis. As an illustration we consider a 30-second neonatal EEG fragment x_s with original sampling period $T = 1/256$ given in Figure 1a, which is band limited between 1 and $B = 20$ Hz. We construct an approximation based on the downsampled signal $y_s = x_{5s}$ with sampling period $\Delta = 5T = 5/256$. Each set of 16 consecutive samples is interpolated with a sparse model of the form (1) with $K = 8$. In total 106 interpolants are computed, these are connected in a continuous manner as the samples x_{75s} are used both as ending and starting point of different interpolants. Figure 1b shows a discrete version of the computed approximation with sampling period T. Compared to the original signal the discrete approximation has CC = 0.9988.
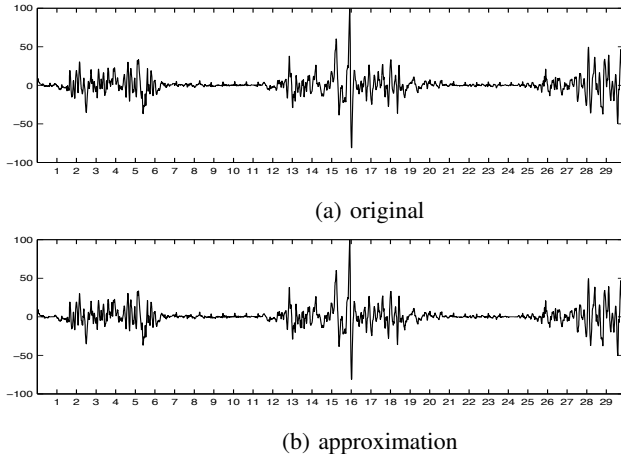

(a) original

(b) approximation

Fig. 1: Approximation of a 30-second neonatal EEG fragment by use of sparse trigonometric interpolation.

B. Modeling correlated multichannel activity

Video-EEG was recorded on 21-channel OSG EEG recorders (Rumst, Belgium) during a presurgical evaluation. Sampling frequency was 250 Hz and an average reference montage was used. The EEG was digitally filtered with a low-pass filter (0-20 Hz). A notch filter was applied to suppress the 50 Hz power-line interference. A 70-second long period containing seizure activity was selected for studying the performance of the new sparse representation method.

(i) PCA preprocessing step. To reduce the redundancy in the channel domain PCA was performed on submatrices containing consecutive one second fragments of the original recording. A fixed intrinsic dimension of $\tilde{d} = 11$ was decided on, which guaranteed an overall retained variance above 0.99. Figure 2a (top) shows a signal approximation in terms of the principal components for the second half of the recording. Figure 2a (bottom) visualizes the cross-correlation of this approximation with respect to the original signal, computed separately for each second and each channel. This cross-correlation never falls below 0.965 demonstrating that the morphology of the original signal was preserved in the PCA step.

(ii) Sparse trigonometric interpolation step. After performing PCA we are left with 770 discrete one-dimensional signals of one second with sampling period $T = 1/250$. Similar to the treatment in Section III-A we ran four different reconstructions for each of these signals with fixed $\Delta = 5T$, but different K. Choosing a fixed Δ but different K implies the use of the same downsampled signal for all four approximations. The following heuristic was adopted to guarantee a satisfactory approximation. For each discrete one-dimensional signal we compute the cross-correlation of the high-term approximations ($K_1 = 23$, $K_2 = 24$ and $K_3 = 25$) with respect to the low-term approximation ($K_0 = 2$). When two or more of these cross-correlations were below 0.85 we kept the low-term approximation. In the other cases we kept the high-term approximation with the highest cross-correlation. Table I shows how many times K_i ($i = 0, 1, 2, 3$) was used to come up with the final approximation.

K_0	K_1	K_2	K_3
55	151	168	396

TABLE I

After performing this heuristic the cross-correlations with respect to the 770 discrete signals were partitioned as given by Table II. We notice that 712 out of the 770 one-dimensional signals of one second were reconstructed with a cross-correlation of at least 0.99.

CC	> 0.99	$0.99 - 0.95$	$0.95 - 0.90$	< 0.90
	712	54	2	2

TABLE II

The trigonometric approximations were then substituted in the reverse PCA step to obtain an approximation to the

original recording. Figure 2b (top) shows the final result for the second half of the recording. Figure 2b (bottom) visualizes the cross-correlation of this approximation with respect to the original signal, computed separately for each second and channel and shows that it never falls below 0.95. It is observed that the trigonometric approximation step introduces an additional comparable loss in the accuracy of the approximation.

IV. CONCLUSION AND CURRENT RESEARCH

We present the application of a new matrix pencil method, that models activity directly as a sum of damped cosines, to the sparse reconstruction of EEG. The method reconstructs activity from $2K$ equidistant samples in a computationally efficient manner: the solution is obtained by solving $K \times K$ linear systems. This method is not restricted to EEG, but further validation is required on a broader range of biomedical signals, and in comparison with other methods. The Hankel-plus-Toeplitz matrix pencil can be further developed along the line of well-established spectral analysis methods.

REFERENCES

[1] A.M. Abdulghani, A.J. Casson, E. Rodriguez-Villega. Compressive sensing scalp EEG signals: implementations and practical performance. Med & Biol. Eng. & Comp. 1-9, 2011.

[2] S. Aviyente. Compressed sensing framework for EEG Compression. 14th Workshop on statistical signal processing 181-184, 2007.

[3] E. J. Candes, M. B. Wakin. An introduction to compressive sampling. IEEE Signal Proc. Magazine 25(2):21-30, 2008.

[4] A. Cuyt, W.-s. Lee, S. Peelman. Sparse trigonometric interpolation. In preparation, 2013.

[5] S. Debener, F. Minow, R. Emkes, K. Gandras, M. de Vos. How about taking a low-cost, small, and wireless EEG for a walk? Psychophysiology 49:1617-1621, 2012.

[6] D.L. Donoho, Compressed sensing. IEEE Trans. Inf. Theory 52(4):1289-1306, 2006.

[7] M. Giesbrecht, G. Labahn, and W.-s. Lee. Symbolic-numeric sparse polynomial interpolation in Chebyshev basis and trigonometric interpolation. In Proc. Workshop on Computer Algebra in Scientific Computation (CASC) 195-204, 2004.

[8] Y. Hua, T.K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. IEEE Trans. Acous, Speech and SP. 38:814-824, 1990.

[9] I. T. Jolliffe. Principal component analysis (second edition). Springer, 2002.

[10] E. Kaltofen and W.-s. Lee. Early termination in sparse interpolation algorithms. Journal of Symbolic Computation, 36(3-4):365-400, 2003.

[11] Y.N. Lakshman and B.David Saunders. Sparse polynomial interpolation in nonstandard bases. SIAM J. Comput., 24(2):387-397, 1995.

[12] R. Roy and T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans. Accoustics, Speech and Signal Proc., 37(7):984-995, 1989.

[13] R.O. Schmidt. Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas and Propagation, AP-34(3):276-280, 1986.

[14] Z. Zhang, T-P. Jung, S. Makeig, B.D. Rao. Compressed sensing of EEG for wireless telemonitoring with low energy consumption and inexpensive hardware, IEEE Trans. Biomed. Eng., 60(1):221-224, 2013.
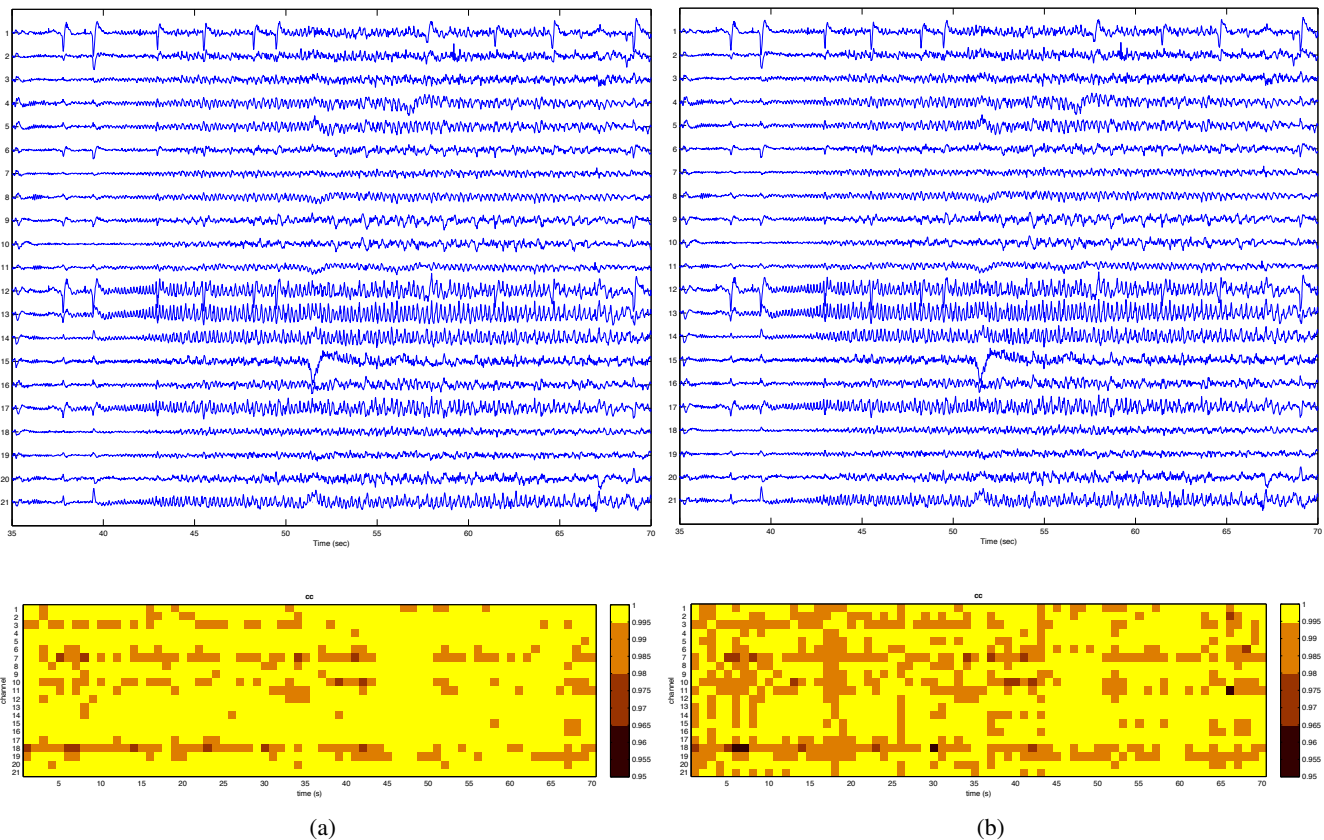
Fig. 2: (a) Approximation of EEG measurements (35s-70s) after the PCA step [top], CC with original measurements per second per channel [bottom]. (b) Approximation of EEG measurements (35s-70s) after the PCA step combined with sparse interpolation [top], CC with original measurements per second per channel [bottom].