Low-power EEG monitor based on Compressed Sensing with Compressed Domain Noise Rejection

Nicola Bertoni*, Bathiya Senevirathna§, Fabio Pareschi*,‡, Mauro Mangia†,‡,
Riccardo Rovatti†,‡, Pamela Abshire§,¶, Jonathan Z. Simon§,¶,‖, Gianluca Setti*,‡
* ENDIF – University of Ferrara, Italy. email:{nicola.bertoni, fabio.pareschi, gianluca.setti}@unife.it
† DEI – University of Bologna, Italy email:{mauro.mangia2, riccardo.rovatti}@unibo.it
§ Department of Electrical and Computer Engineering, University of Maryland, MD, USA. email: {bsenevir, jzsimon, pabshire}@umd.edu
¶ Institute for Systems Research, University of Maryland, MD, USA
‖ Department of Biology, University of Maryland, MD, USA

***Abstract*—Wireless sensor nodes capable of acquiring and transmitting biosignals are increasingly important to address future needs in healthcare monitoring. One of the main issues in designing these systems is the unavoidable energy constraint due to the limited battery lifetime, which strictly limits the amount of data that may be transmitted. Compressed Sensing (CS) is an emerging technique for introducing low-power, real-time compression of the acquired signals before transmission. The recently developed *rakeness* approach is capable of further increasing CS performance. In this paper we apply the rakeness-CS technique to enhance compression capabilities for electroencephalographic (EEG) signals, and particularly for Evoked Potentials (EP), which are recordings of the neural activity evoked by the presentation of a stimulus. Simulation results demonstrate that EPs are correctly reconstructed using rakeness-CS with a compression factor of 16. Additionally, some interesting denoising capabilities are identified: the high-frequency noise components are rejected and the 60 Hz power line noise is decreased by more than 20 dB with respect to the state-of-the-art filtering when rakeness-CS techniques are applied to the EEG data stream.**

I. INTRODUCTION

Wearable or implantable computing and communication systems allow convenient, real-time monitoring of vital signals, such as heartbeat and brain waves, and other physiological parameters. In particular, electroencephalography (EEG), the technique of measuring electrical signals generated within the brain by placing electrodes on the scalp, is an essential tool in identifying and studying several neurological disorders including epilepsy and schizophreny.

In this work we focus on Evoked Potentials (EPs), recordings of brain electrical activity evoked by an external stimulus that is repeatedly delivered to the subject. EP acquisition is usually carried out inside a laboratory using a wired cap, featuring a number of electrodes located at specific positions, cable-connected to one or more amplifiers followed by an Analog-to-Digital Converter. Offline signal processing is regularly performed after the signal has been acquired and stored as a mandatory analysis step in order to get rid of several noise components affecting the measurements, especially: *environmental noise* (external noise due to electromagnetic interference and 50-60 Hz power line hum), *sensor noise* (arising from problems like momentary variation of skin contact) and *physiological noise* (unwanted components of biological origin). Even after this filtering, the spontaneous EEG amplitude (alpha waves, beta waves, etc.) is usually much higher than the evoked activity, ranging from less than one to several microvolts, and significant signal averaging is commonly required to reveal the stimulus response. This averaging is carried out over many epochs (time intervals between two subsequent stimuli) so that signal contributions that are not time-locked to the stimulus are averaged out, as shown in Fig. 1.

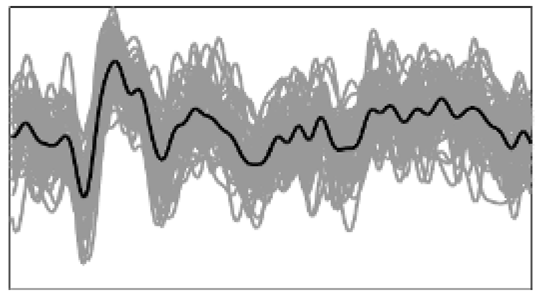

Fig. 1. Evoked potential recordings aligned with the stimulus trigger (grey lines), and their average (black line).

The clear advantage of a wireless health monitoring system is that the subject can be monitored during normal life activities without being forced to stay inside clinical facilities connected to the electronic apparatus. In portable systems the entire recording unit is battery powered, and the physical size of the batteries determines the overall device size and operational lifetime. For mobile recording applications, size and lifetime represent the most important design constraints and tradeoffs. The system must be tiny enough that it does not represent an obstacle for the patient's movements; unfortunately, this is often in contrast with the constraint on the minimum desired operational lifetime before recharging is needed. As an example, the aforementioned filtering algorithms typically demand too much computational power to be implemented in such a system.

The power consumed by the wireless device can be divided into two main contributions: power employed for data acquisition/processing and power used for data transmission. The latter contribution can be reduced by considering a low-power, real-time compression of the raw physiological data in the wearable device itself. However it is essential that the computational complexity of the data compression algorithm is low.

Accordingly, the Compressed Sensing (CS) paradigm has been recently introduced to provide data compression in wearable computing systems thanks to the low computational complexity [1], [2]. The basic idea behind CS is to move the burden of computational effort from the battery-powered wireless node to a base station where energy constraints are less stringent since the energy is taken directly from the power lines. Hence, the compression/encoding algorithm is designed to be simple (i.e. made of a few elementary operations) while its decompression/decoding counterpart is inherently non-linear and usually more time- and energy-consuming.

The purpose of this work is to move towards development

978-1-4799-5341-7/16/$31.00 ©2016 IEEE

of ad hoc wireless sensor nodes for biosignals by means of the CS framework, in particular the recently introduced *rakeness* approach [3]. The advantage is twofold. First, by embedding CS in the sensor node we can exploit its intrinsic compression capability to reduce the amount of data to be transmitted (and also the required energy). Then, the rakeness approach, by tuning the statistical properties of the CS acquisition algorithm with that of the input signal, can boost the performance of CS in terms of compression ratio.

An additional consequence of the statistical tuning of the CS acquisition algorithm is not only to increase the CS performance in terms of compression ratio, but also that the matched features are enhanced, while others are removed or strongly attenuated. In this way we can enhance features we are interested in, and reject unwanted components directly during acquisition. In other words, we are using CS to perform *implicit* signal filtering. We can use this to achieve an effect similar to that of existing denoising techniques, which are generally too expensive in terms of power consumption to be implemented directly in portable devices. The most interesting aspect of the proposed approach is that the filtering effect is performed directly as the signals are projected into the compressed domain, and do not require any additional, time-consuming, denoising steps.

The content of the paper is structured as follows. In Sec. II we provide an overview of CS theory and of a recently developed rakeness extension. Then, in Sec. III, we show some interesting results obtained employing rakeness-based CS to compress EEG, with particular emphasis on EP signals. Finally, we summarize the contributions.

II. Rakeness-based CS Framework

Let us focus on the discrete-time CS approach, where any instance of the input signal $x \in \mathbb{R}^N$ is represented by its Nyquist-rate samples. The CS framework is based on a *sparsity* assumption, i.e. that given a proper basis $\Psi \in \mathbb{R}^{N\times N}$, any signal instance can be expressed as the linear combination of only a few vectors of Ψ, i.e., $x = \Psi\alpha$ where the coefficient vector α has only $K \ll N$ non-null coefficients. The aim of CS is then to represent each realization x with a measurement vector $y \in \mathbb{R}^M$, with $M < N$. The ratio N/M is commonly referred to as *Compression Ratio (CR)*. The compressed measurements are achieved as a simple matrix multiplication (i.e. multiply and accumulate operations) between a set of sensing vectors $\phi_j, j = 1, ..., M$, arranged as the rows of a sensing matrix $\Phi \in \mathbb{R}^{M\times N}$:

$$y = \Phi x + \nu = \Phi\Psi\alpha + \nu = A\alpha + \nu \tag{1}$$

where $A = \Phi\Psi$ is a $M \times N$ matrix that links the sparse representation α to y, and ν is additive noise modeling the system non-idealities.

The reconstruction $\hat{x} = \Psi\hat{\alpha}, \hat{\alpha} \in \mathbb{R}^N$ of the original signal from y is an underconstrained problem, since $M < N$, and hence there are an infinite number of inverse solutions for (1). Under the assumption that the matrix A satisfies the *Restricted Isometry Property* (RIP) [4] and that M is $\mathcal{O}(K\log(N/K))$ [5], CS theory ensures that the correct solution is given by the following sparsity promoting optimization problem:

$$\begin{aligned} \hat{\alpha} &= \min \|\alpha\|_{\ell 1} \\ s.t. \quad & \|\Phi\Psi\alpha - y\|_{\ell 2} < \varepsilon \end{aligned} \tag{2}$$

where $\|\cdot\|_{\ell 1}$ and $\|\cdot\|_{\ell 2}$ represent the standard $\ell 1$ and $\ell 2$ norms respectively and ε takes into account the effect of ν.

The most convenient way to ensure RIP is to generate the ϕ_j as instances of independent and identically distributed (i.i.d.) Gaussian (or Sub-Gaussian) random variables. A common hardware friendly choice is to adopt an i.i.d. antipodal random process, where the probability to have +1 or -1 is the same [2]. In this way, the hardware requirements for implementing the multiply and accumulate operations to compute (1) are greatly simplified. Recently it has been shown [3] that, if the class of signals to be acquired exhibit a non-flat energy distribution, the CS performance can be improved taking advantage of an innovative concept named *rakeness* [1]. In this new scenario the ϕ_j are no longer constructed from simple i.i.d. random entries but rather their statistics are tuned in order to match that of the input signal, thus increasing the average energy of the y and ensuring better results in terms of reconstruction accuracy or, equivalently, to increase the CR given the desired reconstruction accuracy. To formalize this, let us define the rakeness ρ between two stochastic processes $\underline{\phi}$ and $\underline{x}$, generating the sensing sequences ϕ_j and the signal instances x respectively, as

$$\rho(\underline{\phi}, \underline{x}) = \mathbf{E}_{\underline{\phi},\underline{x}}\left[|\langle\phi_j, x\rangle|^2\right] \tag{3}$$

where $\mathbf{E}_{\underline{\phi},\underline{x}}[\,\cdot\,]$ denotes the statistic expectation over $\underline{\phi}$ and $\underline{x}$ and $\langle\cdot,\cdot\rangle$ the standard inner product. Under the assumption that the correlation matrices R^ϕ and R^x of the processes $\underline{\phi}$ and $\underline{x}$ are known, respectively, we can express the rakeness value as $\rho(\underline{\phi}, \underline{x}) = \sum_{j=1}^{N}\sum_{i=1}^{N} R^\phi{}_{i,j} R^x{}_{i,j}$ [2], with $R^\phi{}_{i,j}$ the (i, j)-th element of R^ϕ. The key concept is trying to maximize the "raked" energy preserving RIP. This is obtained by the following optimization problem

$$\begin{aligned} &\max_{\underline{\phi}} \rho(\underline{\phi}, \underline{x}) \\ &s.t. \begin{cases} \langle\phi_j, \phi_j\rangle = e \\ \rho(\underline{\phi}, \underline{\phi}) \le r\, e^2 \end{cases} \end{aligned} \tag{4}$$

where e represents the energy of each ϕ_j, and r is a non-critical parameter ensuring that the ϕ_j are random enough to preserve RIP [6]. The outcome of (4) is the correlation matrix R^ϕ of the stochastic process $\underline{\phi}$ to be used to generate the ϕ_j. Interestingly, by using a properly designed Linear Probability Feedback Process (LPFP) [7], it is possible to generate binary antipodal ϕ_j with a prescribed R^ϕ. This allows us to use the rakeness approach with the hardware advantages offered by binary antipodal sensing vectors.

In the following we exploit the rakeness approach in the acquisition of EEG signals. The advantage is twofold, enhancing both compression and noise suppression. First, we are able to achieve a very high CR. With the rakeness approach, we achieve $CR = 16$, while we only achieve $CR = 4$ with the standard CS approach. At the same time, statistically designing the ϕ_j actually acts to enhance some components of x and attenuating others, i.e., it results in a *filtering* effect on the input signal. We exploit this to introduce, along with the compression capability, a denoising feature in the rakeness-based CS framework.

In the next section the effectiveness of the proposed approach is demonstrated with application to EEG signals, where we are able to simultaneously achieve very high CR and online denoising that is comparable with that obtained from the state-of-the-art offline solutions.

III. Results

In this section the state-of-the-art denoising techniques previously introduced and the rakeness-based CS approach are tested on actual EP recordings taken from a normal-hearing subject who performed a simple auditory task, consisting of listening to one second spaced speech syllables.

[1]Online at http://cs.signalprocessing.it

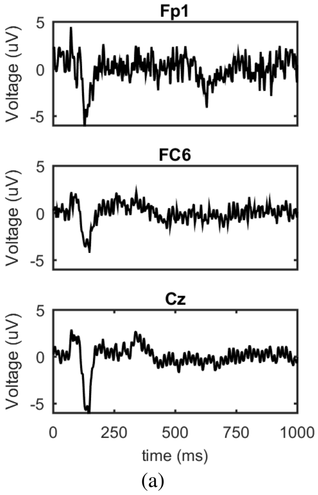

Fig. 2. Comparison between averaged auditory EP of channels Fp1, FC6 and Cz before (a) and after (b) denoising.

The analog waveforms were collected with 32 brain channels, labeled as in the International 10-20 system, plus two additional differential ocular channels, named *HEOG* and *VEOG*, used as noise references to reject artifacts arising from eye blinks. All channels are 1 -200 Hz band-pass filtered and organized into 700 epochs, each of one-second duration. Successively, the entire dataset is divided into two parts: the first 350 epochs are employed as a Training Set (TS) as explained in the following, while the second part, named Data Set (DS), is used for CS performance testing.

A. State-of-the-Art Offline Denoising

Both brain and reference channels are sampled at 512 Hz and filtered using offline denoising algorithms in the following order. First, Time-Shift PCA (TSPCA) [8] is applied for environmental noise removal. This filtering algorithm is obtained by delaying the signals collected by sensors used as noise reference channels, orthogonalizing them, projecting the brain sensors onto the noise derived basis, and removing the projections to obtain clean data. Subsequently, the resulting signals are filtered using Sensor Noise Suppression (SNS) [9]. This method is based on the assumption that every source of interest is picked up by more than one sensor. To reduce noise, each sensor signal is projected on the subspace spanned by its neighboring channels and replaced by its projection. In this process, wide-band noise and glitches, that are not present in the neighbors are eliminated, while shared features are retained. The last denoising step aims to remove unwanted physiological sources. A spatial filter is designed, using a blind source separation method known as Denoising Source Separation (DSS) [10] to partition recorded activity into stimulus-related and stimulus-unrelated components, based on a criterion of stimulus-evoked reproducibility.

Fig. 2 shows the EP average response of the TS brain channels *Fp1*, *FC6* and *Cz* before (Fig. 2(a)) and after (Fig. 2(b)) the denoising process. In the first two cases the noise components appear considerably reduced in the denoised signals, while for *Cz*, which is supposed to give the strongest auditory response [11], the effects of this method are mostly masked by averaging. Note that this denoising technique is not able to remove the power line noise, as clearly observable from the residual 60 Hz oscillations in Fig. 2(b).

B. Rakeness-CS Online Denoising

We first compare the system performance of the standard CS approach with the rakeness-based CS in terms of the maximum CR achievable that preserves the features of the average EP waveform. Next, we present results showing that the rakeness-based CS acquisition introduces desirable denoising properties, after signal reconstruction.

The CS framework is tested on the DS epochs with $N = 512$ (1 s epochs sampled at 512 Hz), using the orthogonal Daubechies 8 wavelet as sparsity basis and the SPGL1 toolbox[2] to solve (2) and retrieve the signal x from the compressed samples y. The input signal autocorrelation matrix, required to solve (4) in the rakeness-based approach, is estimated from the TS *offline denoised data* and particularly from the midline centered channel Cz. Furthermore, in order to increase the robustness of the estimation, each Cz epoch is replicated 100 times and randomly shifted in the epoch. In such a way, we obtain L instances $x_{\mathrm{Cz}}(\cdot, l)$, $l = 1, \dots, L$ of a "clean" auditory response, which are no more aligned to the stimulus trigger at the beginning of the time window. Then, the associated correlation matrix R^{Cz} is computed as follows:

$$R^{Cz} = \begin{bmatrix} r_1(0) & r_1(1) & r_1(2) & ... & r_1(N-1) \\ r_1(1) & r_2(0) & r_2(1) & ... & r_2(N-2) \\ r_1(2) & r_2(1) & r_3(0) & ... & r_3(N-3) \\ ... & ... & ... & ... & ... \\ r_1(N-1) & r_2(N-2) & r_3(N-3) & ... & r_N(0) \end{bmatrix} \quad (5)$$

with $r_n(i) = \frac{1}{L}\sum_{l=0}^{L-1} x_{Cz}(n,l)x_{Cz}(n+i,l)$, $n = 1,...,N, i = 0,...,N-1$. Then, using R^{Cz} and solving the rakeness optimization problem (4), we get the correlation matrix R^{ϕ} which, by means of the LPFP in [7], is used to generate the sampling vectors ϕ_j.

The simulation results are presented in Fig. 3. In case of CR=4 (Fig. 3(b)), the averaged reconstructed EPs using standard CS show a quite good match with the filtered EPs average (Fig. 3(a)) and with rakeness-based CS this matching is greatly improved. When CR=16 (Fig. 3(c)), the discrepancy between the standard-CS and the filtered EP dramatically increases while the rakeness approach still ensures acceptable performance.

As expected, employing the rakeness approach considerably improves the achievable CR with negligible performance degradation. Using this approach, a smaller number of compressed samples is needed to achieve a target reconstruction accuracy and hence the power consumption in the portable device can be greatly reduced.

Next we highlight the extremely interesting denoising properties offered by the rakeness approach, as exemplified by Fig. 4. Here, we focus on one single epoch looking at its multichannel representation (scalp map) shown in Fig. 4(a) where one can notice a strong component located in the right-frontal hemisphere, identifiable with the dark area localized around the electrode FC6. As confirmed by the representation of the scalp potential of FC6 over time (Fig. 4(b)), there is a sharp noise peak around 700 ms that is completely removed by both the offline filtering and the rakeness-based CS. This demonstrates that some desirable properties of the TS have been implicitly imparted into the acquired signal during the rakeness optimized compression process.

Interestingly enough, by looking at the power spectral density plot in Fig. 4(c), one can note an additional advantage of the rakeness-based approach with respect to the state-of-the-art one. The spectral density of the offline denoising shows a strong component around 60 Hz that is clearly an interference

[2]Available online at https://www.math.ucdavis.edu/~mpf/spgl1/

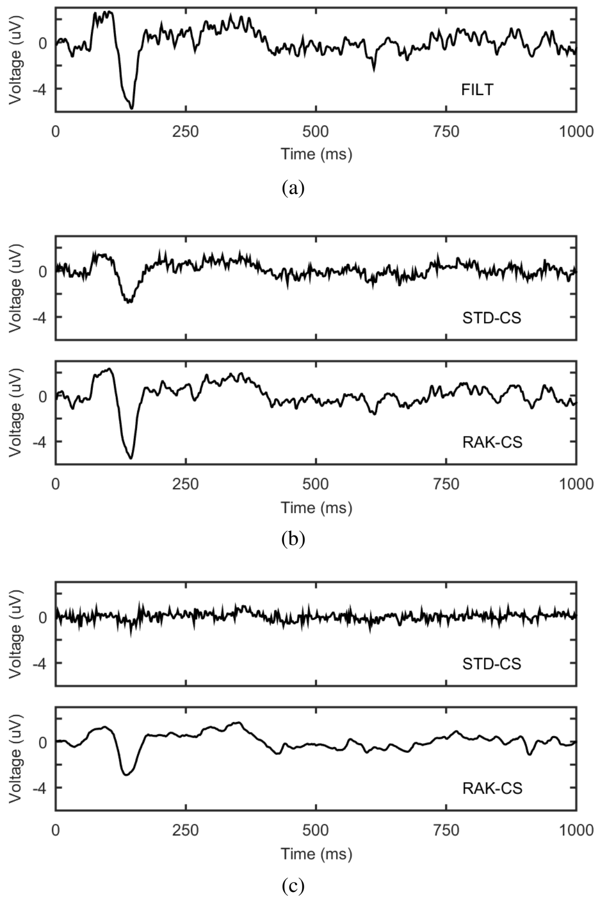

Fig. 3. Comparison between the standard CS and the rakeness-based CS performance on the channel Cz: offline filtered and averaged (a); standard-CS vs rakeness-CS acquired, reconstructed and averaged with CR=4 (b) and CR=16 (c).

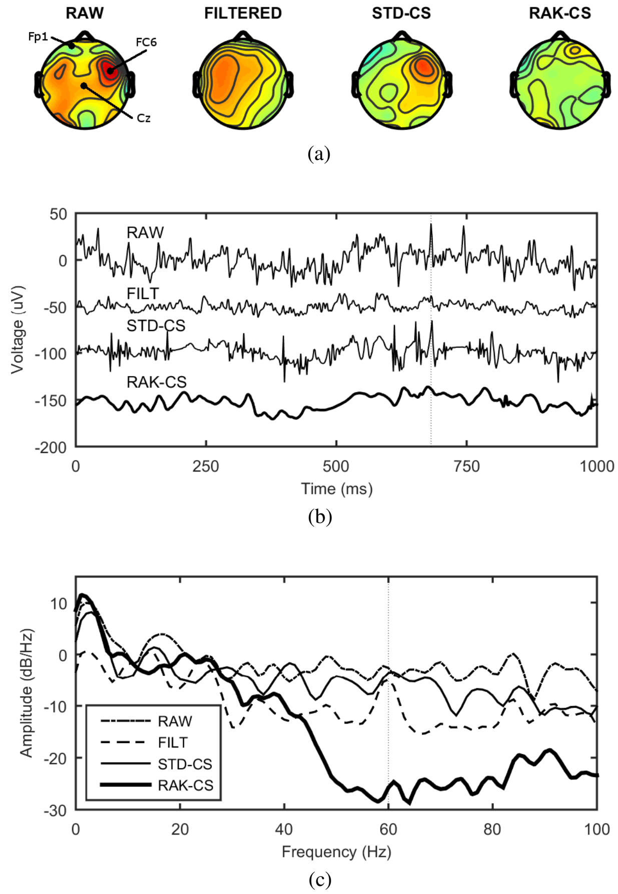

Fig. 4. Rakeness-based CS denoising features. Comparison between the raw and denoised signals with the standard-CS and rakeness-CS reconstructions: (a) scalp maps, (b) channel Fp6 single epoch with vertical dotted line indicating the point in time where the scalp maps are calculated, (c) power spectral density with vertical dotted line indicating 60 Hz line noise.

from the power supply. The rakeness-CS decreased this environmental noise power by more than 20 dB in comparison with the state-of-the-art denoising technique.

IV. Conclusion

In this work, a novel approach for reducing the power consumption in portable health monitoring systems has been presented along with simulation results on a real dataset of EEG Evoked Potentials.

The proposed method, based on a recent extension of Compressed Sensing theory, called rakeness, outperforms standard CS in terms of quality of reconstruction for a given Compression Ratio, with very good performance up to CR=16. Moreover, some interesting filtering capabilities are introduced, such as suppression of high-frequency noise peaks and 60 Hz power line noise, proving the effectiveness of the proposed methodology also for denoising purposes.

Acknowledgements

Thanks to Alessandro Presacco for his time and expertise in making the EEG recordings. This work was supported by the University of Maryland "Engineering Systems for Medicine" Seed Grant and by the University of Ferrara 2015 "Giovani Ricercatori" (Young Researchers) programs.

References

[1] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[2] F. Pareschi, P. Albertini, G. Frattini, M. Mangia, R. Rovatti, and G. Setti, "Hardware-algorithms co-design and implementation of an analog-to-information converter for biosignals based on compressed sensing," *IEEE Trans. Biomed. Circuits Syst*, 2016, dOI: 10.1109/TBCAS.2015.2444276.

[3] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 59, no. 5, pp. 1001–1014, May 2012.

[4] R. G. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, pp. 253–263, 2008.

[5] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[6] N. Bertoni, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Correlation tuning in compressive sensing based on rakeness: A case study," in *Electronics, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on*, Dec 2013, pp. 257–260.

[7] R. Rovatti, G. Mazzini, and G. Setti, "Memory- m antipodal processes: Spectral analysis and synthesis," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 1, pp. 156–167, Jan 2009.

[8] A. de Cheveigné and J. Z. Simon, "Denoising based on time-shift pca," *Journal of Neuroscience Methods*, vol. 165, no. 2, pp. 297 – 305, 2007.

[9] A. de Cheveigné and J. Z. Simon, "Sensor noise suppression," *Journal of Neuroscience Methods*, vol. 168, no. 1, pp. 195 – 202, 2008.

[10] A. de Cheveigné and J. Z. Simon, "Denoising based on spatial filtering," *Journal of Neuroscience Methods*, vol. 171, no. 2, pp. 331 – 339, 2008.

[11] B. Lütkenhöner and J. Mosher, "Source analysis of auditory evoked potentials and fields," in *Auditory Evoked Potentials - Basic Principles and Clinical Applications*, R. F. Burkard, J. J. Eggermont, and M. Don, Eds. Lippincott Williams & Wilkins, 2007, pp. 546–569.