Near-lossless multi-channel EEG compression based on matrix and tensor decompositions

Justin Dauwels†, *Senior Member, IEEE,* K Srinivasan, *Member, IEEE,* M Ramasubba Reddy, Andrzej Cichocki, *Senior Member, IEEE*

***Abstract*—A novel near-lossless compression algorithm for multi-channel electroencephalogram (MC-EEG) is proposed based on matrix/tensor decomposition models. Multi-channel EEG is represented in suitable multi-way (multi dimensional) forms to efficiently exploit temporal and spatial correlations simultaneously. Several matrix/tensor decomposition models are analyzed in view of efficient decorrelation of the multi-way forms of MC-EEG. A compression algorithm is built based on the principle of "lossy plus residual coding," consisting of a matrix/tensor decomposition based coder in the lossy layer followed by arithmetic coding in the residual layer. This approach guarantees a specifiable maximum absolute error between original and reconstructed signals. The compression algorithm is applied to three different scalp EEG datasets and an intracranial EEG dataset, each with different sampling rate and resolution. The proposed algorithm achieves attractive compression ratios compared to compressing individual channels separately. For similar compression ratios, the proposed algorithm achieves nearly five-fold lower average error compared to a similar wavelet-based volumetric MC-EEG compression algorithm.**

***Index Terms*—Arithmetic coding, compression, electroencephalogram (EEG), multi-channel EEG, singular value decomposition (SVD), Parallel factor decomposition (PARAFAC)**

I. Introduction

A tensor is a multidimensional array, where the elements are addressed normally by more than two indices. Tensors are natural generalization of vectors (one-way tensor) and matrices (two-way tensor). Tensors are often huge and hence appropriate tensor decomposition is used to extract (or analyze) its useful properties. Tensor decomposition algorithms find its application in a variety of fields including chemometrics, psychometrics, data mining, graph analysis, neuroscience, and image/signal processing [1, 2]. Tensor decomposition algorithms can efficiently capture dependencies in the tensor with very few elements, and hence finds a natural extension towards data compression. In this paper, we use tensor decompositions to compress multi-channel electroencephalograms (MC-EEG).

This work has been presented in part at the NIPS 2010 Workshop "Tensors, Kernels, and Machine learning" and at ICASSP 2011.

†Corresponding author. Justin Dauwels and K. Srinivasan are with School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore-639798, Singapore. email: jdauwels@ntu.edu.sg, srinivasan_k@ntu.edu.sg

M Ramasubba Reddy is with Biomedical Engineering Group, Department of Applied Mechanics, IIT Madras, Chennai-600036, India. email: rsreddy@iitm.ac.in

A. Cichocki is with RIKEN BSI, Laboratory for Advanced Brain Signal Processing, Wako-Shi, Saitama-351-0198, Japan, and Systems Research Institute in Polish Academy of Science, Warsaw, Poland. email: cia@brain.riken.jp

Electroencephalogram (EEG) is the record of the electrical activity of the human brain, which is usually recorded at several locations on the scalp. EEG reflects the underlying state of the human brain and hence sophisticated analysis of EEG can accurately identify sleep stages, seizures, depth of anesthesia, and other cognitive processes. To evaluate epilepsy surgery, high spatiotemporal resolution brain recordings are often performed where the recorded data may contain extracellular potentials and high frequency oscillations. For example, a 320 micro- and macro-electrode array recording with 32 kHz sampling rate at 18-bit resolution generates nearly 3 terabytes of data per day [3]. Generally, EEG datasets can be quite large, ranging from gigabytes of data per day for scalp recordings [4] to terabytes per day for intracranial recordings [3]. Wearable EEG systems are being developed for long-term recordings where EEG is stored for analysis [5, 6]. Such portable systems need to operate under a low-power budget and hence cannot transmit large volumes of data. By compressing the signals before transmission, power consumption may be reduced significantly. As the application areas of EEG continue to expand, the need for EEG compression will become even more pressing.

EEG compression can be categorized into lossless and lossy schemes. The former exploits only the redundancies present in the signal, achieves a modest amount of compression, and give an exact reconstruction of the original signal. On the other hand, lossy compression achieves a higher amount of compression compared to the former by discarding the least significant information while retaining all relevant diagnostic information. Lack of legislation and approved standards in the field of signal compression made clinicians to ask for exact reconstruction [7]. Lossless compression may perhaps seem an ideal solution, however, it yields negligible compression rates which makes it practically useless for many applications. As EEG data sets continue to grow, especially for long-term EEG recordings, high compression rates are required, and lossless compression is then not a viable option. As an alternative, near-lossless compression provides much higher compression rates at small and negligible reconstruction error (distortion), so that the reconstructed EEG signals have sufficient accuracy for most purposes.

In near-lossless compression, the original signal $\mathbf{x} \in \mathbb{Z}^N$ and the reconstructed signal $\tilde{\mathbf{x}}$ only differ by a small positive value δ:

$$||\mathbf{x} - \tilde{\mathbf{x}}||_\infty = \max_{0 \le i < N} |x(i) - \tilde{x}(i)| \le \delta. \qquad (1)$$

In this paper, we consider a near-lossless coding scheme based on tensor decomposition with guarantee (1). A precise control of error is achieved in each sample of the reconstructed signal, providing the desired accuracy for both automated analysis and visual inspection by clinicians.

Single-channel EEG compression is widely studied, and can be categorized under lossless, near-lossless, and lossy methods (see [8] and references therein). Predictive-based coders are competitive in lossless [9] and near-lossless [10, 11] scenarios, but they do not support progressive transmission and hence they are of little use in practical scenarios. Combining progressive transmission and guaranteed maximum distortion (in L^∞ sense) will be crucial in real-time transmission and clinical settings.

In multi-channel EEG, spatially adjacent channels are significantly correlated, known as inter-channel correlation. Each individual channel has temporal correlations also, referred as intra-channel correlation. Single-channel compression algorithms, when applied to MC-EEG compression, exploit only the temporal correlations and the inter-channel correlation remains unutilized. In literature, one can find few multi-channel compression algorithms that exploit both inter-channel and intra-channel correlations, where these correlations are often considered independent and removed using separate techniques. The existing MC-EEG compression algorithms can be categorized into lossless [7, 12] and lossy [13, 14] schemes. In [15], we proposed the idea of arranging the multichannel EEG in the form of a matrix or tensor, exploiting the inter- and intra-channel correlations by matrix/tensor decompositions, and their potential application towards compression. Further, in [16], we developed compression algorithms using image/volumetric wavelet coders that supports progressive quality, progressive resolution, and guarantee a maximum distortion bound in L^∞ sense (cf. (1)). Here we explore the use of matrix/tensor decompositions, and extend our earlier study [15] to design near-lossless compression algorithms for MC-EEG. First, as in [16], we represent MC-EEG in various multi-way forms, and apply matrix or tensor decompositions to exploit both spatial and temporal correlations simultaneously. We construct a simple and efficient coding procedure to compress the matrix/tensor decompositions. Similarly to the coders in [16], they support progressive quality and guarantee the maximum distortion bound (eq. (1)). The compression schemes presented in this paper provide a substantially more favorable compression vs. distortion trade-off compared to the wavelet-based schemes proposed in [16].

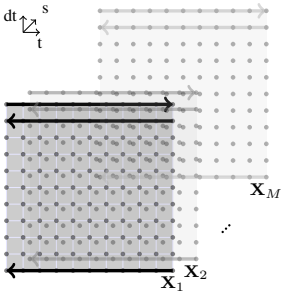


Fig. 1. Formation of t/dt/s three-way tensor from multi-channel EEG. The matrices $(\mathbf{X}_1, \ldots, \mathbf{X}_M)$ formed from single channel EEG are stacked to form volumetric data. Matrix formation from single-channel EEG is shown in bold arrows for the first slice (matrix $\mathbf{X}_1$). The signal $x_1(\cdot)$ is divided in segments of length N. The first segment, $x_1(1)$ to $x_1(N)$, is arranged from left to right as the first row, the second segment, $x_1(N+1)$ to $x_1(2N)$, is arranged from right to left as the second row, etc.

This paper is organized as follows. In Section II, we explain how MC-EEG is arranged in multi-way forms and we describe the various multi-way decomposition schemes. We explain the general idea behind two-stage coding in Section III, and describe our matrix/tensor decomposition based lossy coder. In Section IV, we elaborate on the performance measures used to evaluate our MC-EEG compression schemes. We present our results in Section V, and provide concluding remarks in Section VI.

II. Formation of matrix/tensor from multi-channel EEG

In MC-EEG, spatially adjacent channels have substantial correlation among them along with temporal correlations in each individual channel. Previous studies [7, 12]–[14] address spatial and temporal correlations independently and use different techniques to remove them. Here, as in our previous study [16], the MC-EEG is presented in various multi-way forms which include two-, three-, and four-way forms to exploit the spatial and temporal correlations *simultaneously*.

A. Two-way tensor formation from multi-channel EEG

Each individual channel of MC-EEG is arranged as a row to generate a two-way tensor $\mathbf{X}$ (matrix). To exploit spatial correlation, adjacent EEG channels are arranged as adjacent rows, because they are substantially correlated. Therefore, the EEG represented as matrix $\mathbf{X}$ is locally smooth: each entry of $\mathbf{X}$ is similar compared to its immediate row and column neighbors. However, the correlation decreases as one move farther along either rows or columns.

B. Three-way tensor formation from multi-channel EEG

We consider two ways to form a three-way tensor from MC-EEG [16]. In our previous studies [8, 19], we obtained improved compression performance by arranging single-channel EEG in matrix form. Here we extend this idea by stacking the matrices formed from each individual EEG channel in a three-way tensor, as depicted in Fig. 1. Adjacent slices in the tensor are formed from adjacent EEG channel by scanning the electrodes in spiral fashion [12]. We refer to this three-way tensor as "t/dt/s", where x, y, and z directions reflect temporal (t), delayed temporal (dt), and spatial (s) variations, respectively. The k-th slice $\mathbf{X}_k$ of the $\mathcal{X}$, extracted from channel k, can be written as:

$$\mathcal{X}^{(k)}_{t/dt/s} = \{\mathbf{X}_k | k = 1, \ldots, M\}, \tag{8}$$
$$= \begin{bmatrix} x_k(1) & x_k(2) & \cdots & x_k(N) \\ x_k(2N) & x_k(2N-1) & \cdots & x_k(N+1) \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdots & x_k(N^2) \end{bmatrix}_{(N \times N)}.$$

TABLE I
DESCRIPTION OF THE MATRIX/TENSOR DECOMPOSITION MODELS

Form of EEG data	Model	Description
Matrix (two-way tensor) $\mathbf{X} \in \mathbb{R}^{M \times N}$	Singular value Decomposition (SVD)	$\mathbf{X} \approx \sum_{i=1}^{r} \mathbf{u}_i \sigma_i \mathbf{v}_i^T$, (2) $\mathbf{u}_i$, $\mathbf{v}_i$ - left and right singular vectors, σ_i - singular value.
	Column Row Decomposition [17] (CUR)	$\mathbf{X} \approx \mathbf{CUR}$, (3) $\mathbf{C}$ and $\mathbf{R}$ - selected columns and rows of $\mathbf{X}$. $\mathbf{U} = \mathbf{W}^{-1}$, where $\mathbf{W} = \mathbf{C} \cap \mathbf{R}$.
Matrix (three-way tensor) $\mathcal{X} \in \mathbb{R}^{N \times N \times M}$	Parallel factor Decomposition (PARAFAC)	$\mathcal{X} \approx \sum_{i=1}^{R} \lambda_i (\mathbf{a}_i \circ \mathbf{b}_i \circ \mathbf{c}_i)$, (4) $\mathbf{a}$, $\mathbf{b}$, & $\mathbf{c}$ - normalized factors along the three modes, λ_i - PARAFAC weight.
	TUCKER Decomposition	$\mathcal{X} \approx \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}$. (5) $\mathcal{G} \in \mathbb{R}^{P \times Q \times R}$ - core tensor, $\mathbf{A} \in \mathbb{R}^{N \times P}$, $\mathbf{B} \in \mathbb{R}^{N \times Q}$, & $\mathbf{C} \in \mathbb{R}^{M \times R}$ - basis matrices
	Fiber-Sampling Tensor Decomposition [18] (FSTD)	$\mathcal{X} \approx \mathcal{U} \times_1 \mathbf{C}_1 \times_2 \mathbf{C}_2 \times_3 \mathbf{C}_3$ (6) $\mathbf{C}_1$, $\mathbf{C}_2$, and $\mathbf{C}_3$ - matrices formed by fibers along the mode, $\mathcal{U} = \mathcal{W} \times_1 \mathbf{W}_{(1)}^{\dagger} \times_2 \mathbf{W}_{(2)}^{\dagger} \times_3 \mathbf{W}_{(3)}^{\dagger}$, where $\mathbf{W}_{(i)}^{\dagger}$ - pseudo-inverse of the intersection sub-tensor.
	Compact tensor Decomposition (TT)	$\mathcal{X} \approx \mathcal{U} \times_1 \mathbf{G}_1 \times_3 \mathbf{G}_3$ (7) where $\mathcal{U} \in \mathbb{R}^{R_1 \times N \times R_3}$, $\mathbf{G}_1 \in \mathbb{R}^{N \times R_1}$, $\mathbf{G}_3 \in \mathbb{R}^{M \times R_3}$.

Description is provided only for three-way EEG tensor, as extension to four-way tensor is straightforward. Please refer [15] for detailed description on the decomposition models.

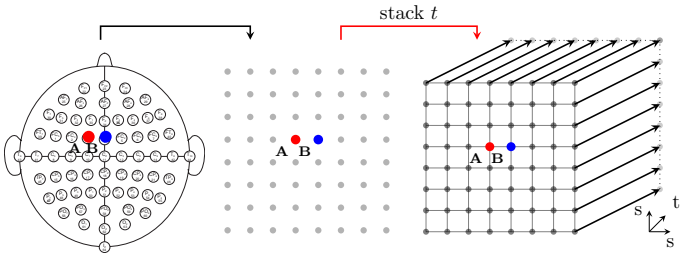


Fig. 2. Formation of s/s/t three-way tensor from MC-EEG. 64-channel EEG montage is shown on the left. At a particular time instance t, samples from all the channels are arranged in the form of matrix, as shown in the middle. The adjacent entries in the matrix stem mostly from spatially adjacent channels. For e.g., the spatially adjacent channels (A and B) correspond to adjacent entries in the matrix. At subsequent time instances, similar matrices are formed which are stacked to along z-direction to construct the s/s/t volume.

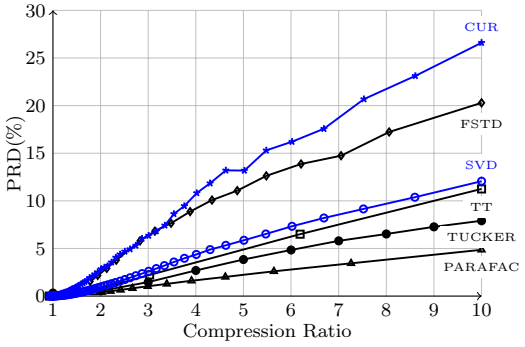


Fig. 3. Compression ratio vs. distortion performance of the Matrix/Tensor decomposition schemes for t/dt/s tensor.

We also consider an alternative and simple method to form a tensor from MC-EEG; the procedure is illustrated in Fig. 2. We call this three-way tensor as "s/s/t", because $x-y$ plane reflects the spatial correlations and the temporal correlations can be found along the z direction. The k-th slice of the three-way tensor may be written as:

$$\mathcal{X}_{s/s/t}^{(k)} = \{x_{(i,j)}(k) | i = 1, \ldots, N_1, j = 1, \ldots, N_2\}, \quad (9)$$
$$= \begin{bmatrix} x_{(1,1)}(k) & x_{(1,2)}(k) & \cdots & x_{(1,N_2)}(k) \\ x_{(2,1)}(k) & x_{(2,2)}(k) & \cdots & x_{(2,N_2)}(k) \\ \vdots & \vdots & \ddots & \vdots \\ x_{(N_1,1)}(k) & x_{(N_1,2)}(k) & \cdots & x_{(N_1,N_2)}(k) \end{bmatrix}_{(N_1 \times N_2)},$$

where i and j refer to the position in the $x-y$ plane, whereas the slice number k refers to the time index. The dimension of the $x-y$ plane is limited by the number of channels and hence the slice in the $x-y$ plane may be square or rectangular.

C. Four-way tensor formation from multi-channel EEG

The t/dt/s and s/s/t tensors capture the spatial and temporal variations respectively along the z direction *only*. The compression performance may enhance if those variations are arranged along two dimensions. To test this idea, we construct a four-way tensor from t/dt/s three-way tensor, where the spatial variations are captured along two dimensions. We split the t/dt/s tensor along the z direction with steps of size K, resulting in L smaller t/dt/s tensors. These smaller t/dt/s tensors are used to form a four-way tensor, named "t/dt/s/s" tensor, as it captures temporal, delayed temporal, and spatial variations along the first, second, and last two dimensions, respectively. More specifically, t/dt/s tensor is split into subtensors of size K along z (spatial) direction, and the t/dt/s/s tensor is formed from those subtensors as follows:

$$\mathcal{X}_{t/dt/s/s}(:,:,:,i) = \mathcal{X}_{t/dt/s}(:,:,iK:(i+1)K-1), \quad (10)$$

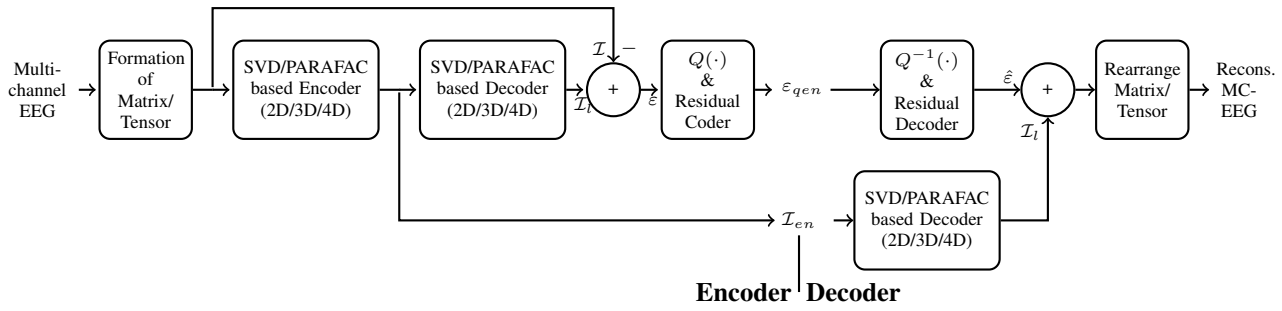


Fig. 4. Near-lossless compression of multi-channel EEG based on SVD/PARAFAC-based coding. $Q(\cdot)$ refers to uniform quantizer (eq. (13)), $Q^{-1}(\cdot)$ refers to dequantizer (eq. (14)).

where $i = 0, 1, \ldots, L-1$. The value of K must be chosen to yield subtensors of equal size in order to form the 4-dimensional t/dt/s/s tensor. Moreover, small values for K often result in numerical problems with the decomposition schemes. Hence we choose intermediate values of K such that subtensors are of similar in size in order to stack them to form a four-way tensor.

D. Rate-Distortion performance of matrix/tensor decompositions on EEG Matrix/tensor

Matrix and tensor decomposition models are widely used as compact representations of large datasets [1]. In our previous study [15], we analyzed the low-rank approximation capability of the above matrix/tensor decomposition models mentioned in the Table I. The distortion (PRD; explained in Section IV-B) is obtained by varying the rank in the matrix/tensor decomposition; the amount of compression achieved is also calculated for that particular rank. The rate-distortion (or compression vs. distortion) performance is shown in Fig. 3. For brevity, the results of tensor decomposition for t/dt/s tensor are presented, as other tensor types also yield similar results. The results were obtained by varying the dimensions in the matrix/tensor decomposition, including number of singular values (eq. (2)) and rank-one tensors (eq. (4)), columns/rows (eq. (3)) or fibers (eq. (6)). We computed Tucker and TT decomposition for all possible sizes of the core tensor $\mathcal{G}$ and $\mathcal{U}$ respectively. As can be seen from Fig. 3, the tensor-based compression schemes clearly outperform matrix-based compression schemes, especially at large compression ratios. CUR and FSTD, which approximates the whole matrix/tensor with few entries/fibers, performs very poor compared to all other decompositions. From Fig. 3, SVD and PARAFAC decompositions provide the best compression results for the matrix and tensor form of MC-EEG respectively; hence, SVD and PARAFAC decomposition is used to build the compression algorithm.

III. Two-stage coder based on matrix/tensor decomposition

Figure 4 shows the block diagram of the proposed matrix/tensor decomposition based two-stage near-lossless coder for MC-EEG. We denote the multi-way form of EEG by $\mathcal{I}$. At the encoder side, $\mathcal{I}$ is compressed by SVD/PARAFAC-based coder resulting in the compressed data $\mathcal{I}_{en}$, yielding the lossy approximation $\mathcal{I}_l$ of the original data $\mathcal{I}$. The residue $\varepsilon = \mathcal{I} - \mathcal{I}_l$ is quantized, leading to $\hat{\varepsilon}$, and next compressed by residual coding, leading to ε_{qen}. Both $\mathcal{I}_l$ and ε_{qen} are stored, and comprise the near-losslessly compressed EEG data. At the decoder side, as illustrated in Fig. 4(b), $\mathcal{I}_{nl}$ is obtained by combining the lossy reconstruction $\mathcal{I}_l$ and the decompressed residual $\hat{\varepsilon}$, i.e., $\mathcal{I}_{nl} = \mathcal{I}_l + \hat{\varepsilon}$. Finally, $\mathcal{I}_{nl}$ is rearranged to yield the near-losslessly reconstructed EEG signal(s). The following relations can be readily confirmed from Fig. 4:

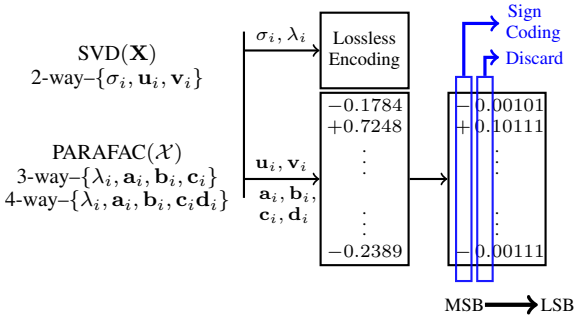


Fig. 5. Illustration of SVD/PARAFAC encoder used in Fig. 4.

$$\mathcal{I} = \mathcal{I}_l + \varepsilon, \tag{11}$$

$$\mathcal{I}_{nl} = \mathcal{I}_l + \hat{\varepsilon}. \tag{12}$$

Therefore, it follows that $||\mathcal{I} - \mathcal{I}_{nl}||_\infty = ||\varepsilon - \hat{\varepsilon}||_\infty$, and hence $||\varepsilon - \hat{\varepsilon}||_\infty \leq \delta$ is equivalent to $||\mathcal{I} - \mathcal{I}_{nl}||_\infty \leq \delta$. The residual ε is uniformly quantized to generate quantization indices ε_q:

$$\varepsilon_q = \begin{cases} \lfloor \frac{\varepsilon+\delta}{2\delta+1} \rfloor & \varepsilon > 0, \\ \lfloor \frac{\varepsilon-\delta}{2\delta+1} \rfloor & \varepsilon < 0, \end{cases} \tag{13}$$

where $\lfloor \cdot \rfloor$ denotes the integer part of the argument. The quantized residual ε_q is then losslessly encoded by residual coding procedure. Hence, $\mathcal{I}_{en}$ and ε_{qen} is transmitted as output of the lossy coding layer and residual coding layer respectively.

At the decoder end, the residual bitstream ε_{qen} is decoded to yield ε_q, followed by a dequantizer defined to guarantee $||\varepsilon - \hat{\varepsilon}|| \leq \delta$:

$$\hat{\varepsilon} = (2\delta + 1)\varepsilon_q. \tag{14}$$

By adding the lossy reconstruction $\mathcal{I}_l$ and the dequantized residual $\hat{\varepsilon}$, we obtain the final near-lossless reconstruction $\mathcal{I}_{nl}$ with guarantee $||\mathcal{I} - \mathcal{I}_{nl}|| \leq \delta$.

The SVD/PARAFAC-based encoder consists of two steps,

whose illustration is shown in Fig. 5. In first step, the multi-way form of MC-EEG is decomposed by SVD or PARAFAC. On decomposition, SVD generates singular values (σ_i) and vectors ($\mathbf{u}_i$ & $\mathbf{v}_i$), whereas PARAFAC yields weights (λ_i) and loadings (three-way–$\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i$; four-way–$\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i, \mathbf{d}_i$). In the second step, σ_i and λ_i are converted to binary with full precision (lossless coding). Singular vectors ($\mathbf{u}_i$, $\mathbf{v}_i$) and PARAFAC loadings ($\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$, $\mathbf{d}_i$) are encoded by binary arithmetic coding; the coding procedure for one such vector is shown in Fig. 5. Consider a vector $\mathbf{u}_i$ whose entries are fraction having value less than unity. All the elements of the vector are converted to binary with fixed precision. The sign bits are encoded by binary arithmetic coding, the value "0." is common to all elements and hence discarded. Then, each bit-plane from MSB to LSB is separately compressed by binary arithmetic coding, generating individual bitstreams one from each bit plane. Other vectors in the i^{th} decomposition are compressed in the same way. These bit streams are then concatenated to produce final bitstream corresponding to i^{th} decomposition. This procedure is repeated for all singular values (or weights) and vectors (or loadings) in the order of decreasing significance, thus generating a progressive quality bitstream.

In the residual coding stage, if the symbol size M of the residual is small ($M \leq 128$), then residual is arithmetic coded directly, else the residual stream is split into sub streams of smaller symbol size and each stream is arithmetic coded separately [16, 20].

IV. Performance Measures

Here we describe the measures used to assess different MC-EEG compression algorithms.

A. Compression Ratio

We compare the performance of the different algorithms by means of compression ratio, the factor of reduction in file size, given by:

$$CR = \frac{L_{orig}}{L_{comp}}, \tag{15}$$

where L_{orig} and L_{comp} refer to bitstream length of the original and compressed sources, respectively.

B. Distortion Measures

The distortion between original signal $\mathbf{x}$ and reconstructed signal $\tilde{\mathbf{x}}$ is given by the error signal $\mathbf{e} = \mathbf{x} - \tilde{\mathbf{x}}$. We use two distortion measures to assess the performance of the algorithms: an average distortion measure based on mean-square-error, known as PRD, and worst-case distortion measure based on maximum-absolute-error, known as PSNR.

The percent root-mean-square distortion (PRD) is given by the following equation:

$$\text{PRD}(\%) \quad = \quad \sqrt{\frac{\sum_i e(i)^2}{\sum_i x(i)^2}} \times 100. \tag{16}$$

PRD gives the amount of energy present in the error signal as a percentage of the energy of original signal. PRD reflects large variations and often insensitive to small variations in $\mathbf{e}$.

TABLE II
EEG DATASETS USED FOR TESTING THE ALGORITHMS

Dataset Name	Channels	f_s	Resolution	Duration
EEG-MMI [21]	64	80 Hz	12	20 min
BCI3-MI [22]	118	100 Hz	16	10 min
BCI4-MI [23]	64	1000 Hz	16	10 min
Intracranial [24]	36	5000 Hz	11	25 min

The second distortion measure, peak signal-to-noise ratio (PSNR), is defined as follows:

$$\text{PSNR}(\mathbf{x}, \tilde{\mathbf{x}}) = 10 \log_{10} \left(\frac{2^Q - 1}{\text{MAE}(\mathbf{x}, \tilde{\mathbf{x}})} \right) \quad \in [0, \infty], \tag{17}$$

where Q refers to the resolution (in bits) of the original signal $\mathbf{x}$. The numerator is the range of the signal and denominator refers to the maximum absolute error (MAE) (1) between original and reconstructed signal. The MAE (1) is the L^∞ norm between $\mathbf{x}$ and $\tilde{\mathbf{x}}$, which depends on the sample with the largest error and does not reflect the amount of error present in the other samples. Two entirely different error signals can have the same value of MAE (and PSNR). Therefore, PRD is computed along with PSNR for assessing the performance of the compression algorithms.

V. Results and Discussions

We apply the proposed tensor-based MC-EEG compression algorithm (Fig. 4) to four datasets, whose properties are given in Table II. A block size of 1024 samples is used from each channel. For example, from 64-channel EEG data (EEGMMI, BCI3 & BCI4), we generate tensors of the following sizes: $32 \times 32 \times 64$ (t/dt/s), $8 \times 8 \times 1024$ (s/s/t), and $32 \times 32 \times 8 \times 8$ (t/dt/s/s). We form similar tensors for EEG data with different number of channels. We compare the results of the proposed algorithm (Fig. 4) with wavelet-based single-channel [8] and image/volumetric coding algorithms [16]. Bi-orthogonal wavelet 4.4 is chosen as the mother wavelet in those algorithms [8]. For unbiased comparison, the dimensions of the matrix/tensor size are kept same in all the compression algorithms. In SVD/PARAFAC coder, the first ten components are chosen from the decomposition model. They typically preserve nearly 75-80% of the energy in the original signal. In single-channel compression, EEGs from each channel are arranged in the form of matrix of size 32×32 before compression, as using large matrix sizes resulted in negligible improvement in compression performance [8].

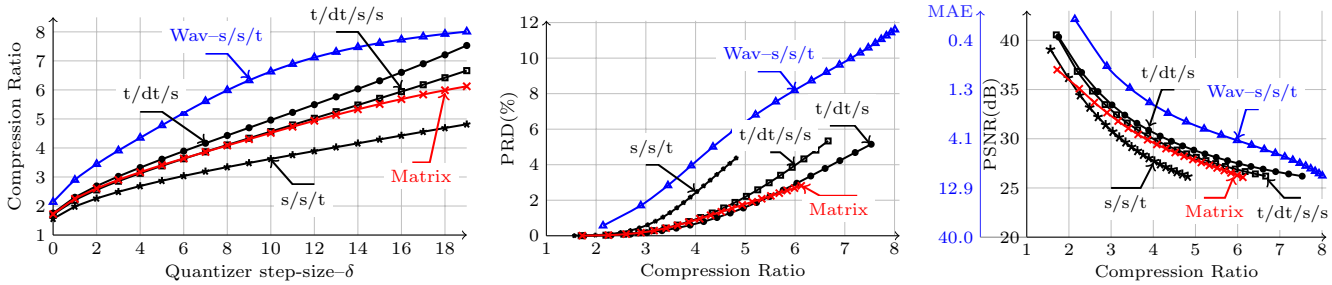
We study the performance by varying the step-size δ of the quantizer; for each step-size we calculate the compression ratio and two distortion measures (PRD and PSNR). In Fig. 6(a), we show how the compression ratio increases with the quantizer step-size δ for the EEG-MMI dataset. We also provide PRD and PSNR distortion with compression ratio in Fig. 6(b) & (c) respectively. Detailed numerical results for all three EEG datasets are presented in Table III for SVD/PARAFAC-based two-stage coding, single-channel, and wavelet-based image/volumetric coding.

From Fig. 6(a), we can ascertain that compression ratio increases with the step-size of the quantizer for all the compression algorithms. Volumetric wavelet coding yields higher

TABLE III

LOSSLESS/NEAR-LOSSLESS COMPRESSION PERFORMANCE OF MATRIX/TENSOR-BASED MC-EEG COMPRESSION ALGORITHMS. MULTI-CHANNEL COMPRESSION INCLUDE SVD-BASED (SVD) AND PARAFAC-BASED (PARAFAC–T/DT/S,–S/S/T, & –T/DT/S/S) COMPRESSION ALGORITHMS.

LOSSY LAYER CODER	**TOL.**	**PARAMETERS**											
	δ	CR	PRD	PSNR	CR	PRD	PSNR	CR	PRD	PSNR	CR	PRD	PSNR
Datasets →		**EEG-MMI**			**BCI3-MI**			**BCI4-MI**			**Intracranial**		
SVD-Matrix		1.72	0.01	37.01	1.75	0.00	48.00	1.39	0.00	47.89	1.56	0.04	21.14
PARAFAC–t/dt/s		**1.77**	**0.00**	**40.34**	**1.82**	**0.00**	**44.10**	1.63	0.00	44.84	**1.96**	**0.00**	**31.10**
PARAFAC–s/s/t		1.56	0.00	39.06	1.56	0.00	40.45	1.43	0.00	41.37	1.11	0.01	29.01
PARAFAC–t/dt/s/s	0	1.70	0.00	40.56	1.71	0.00	44.18	**1.63**	**0.00**	**45.48**	1.55	0.00	31.14
Single-Channel		1.73	0.66	42.17	1.68	0.10	51.23	1.51	0.10	51.30	2.31	0.63	36.15
Wavelet-Image		1.99	0.57	42.15	1.75	0.07	51.22	**1.51**	**0.08**	**51.29**	2.09	0.46	36.27
Wavelet-s/s/t		**2.14**	**0.57**	**42.14**	**1.91**	**0.07**	**51.18**	1.49	0.08	51.18	**2.35**	**0.47**	**36.15**
SVD-Matrix		3.41	0.43	31.06	2.83	0.01	40.44	2.00	0.01	40.42	3.14	0.27	19.29
PARAFAC–t/dt/s		**3.61**	**0.43**	**31.57**	**3.04**	**0.01**	**39.54**	2.55	0.01	39.75	**5.56**	**0.25**	**22.51**
PARAFAC–s/s/t		2.86	0.45	31.73	2.37	0.01	38.01	2.10	0.02	38.44	1.81	0.29	22.19
PARAFAC–t/dt/s/s	5	3.38	0.43	31.59	2.73	0.01	39.56	**2.55**	**0.01**	**39.89**	3.42	0.29	22.52
Single-Channel		3.28	7.30	31.74	2.66	1.31	40.77	2.27	1.16	40.77	7.62	5.70	25.72
Wavelet-Image		4.12	6.22	31.73	2.84	0.81	40.77	**2.27**	**0.90**	**40.77**	5.67	4.89	25.72
Wavelet-s/s/t		**4.78**	**5.96**	**31.73**	**3.23**	**0.81**	**40.76**	2.21	0.90	40.76	**7.74**	**4.34**	**25.71**
SVD-Matrix		4.52	1.35	28.59	3.83	0.02	37.81	2.29	0.05	37.77	4.08	0.59	18.05
PARAFAC–t/dt/s		**4.96**	**1.56**	**28.84**	**3.73**	**0.02**	**37.29**	**3.00**	**0.05**	**37.42**	**9.90**	**0.70**	**19.80**
PARAFAC–s/s/t		3.62	1.57	28.75	2.76	0.02	36.36	2.40	0.05	36.59	2.16	1.05	19.63
PARAFAC–t/dt/s/s	10	4.56	1.56	28.85	3.26	0.02	37.32	2.99	0.05	37.50	4.90	1.04	19.80
Single-Channel		4.21	13.77	28.93	3.17	1.98	37.96	**2.64**	**2.22**	**37.96**	11.25	7.38	23.26
Wavelet-Image		5.65	10.99	28.92	3.40	1.55	34.96	2.63	1.72	37.96	9.44	7.45	22.90
Wavelet-s/s/t		**6.63**	**9.21**	**28.92**	**4.03**	**1.55**	**37.95**	2.56	1.72	37.95	**12.13**	**5.40**	**22.90**



(a) Improvement in CR with quantizer step-size (δ) (b) CR vs. PRD (%) (c) CR vs. PSNR (dB)

Fig. 6. Compression performance of the SVD/PARAFAC-based two-stage near-lossless compression for EEG-MMI dataset. SVD-based coder is labelled as "Matrix"; PARAFAC-based coder is labelled as t/dt/s, s/s/t, and t/dt/s/s depending on the underlying multi-way representation of the MC-EEG. The best performing algorithm from our previous study [16], wavelet-based volumetric coding on s/s/t tensor labeled "Wav-s/s/t", is reproduced for comparison.

CR compared to SVD/PARAFAC-based coders for a particular value of step-size δ. The PSNR of volumetric wavelet coder is slightly higher (Fig. 6(b)), however, the PRD is clearly vastly larger (Fig. 6(c)) compared to SVD/PARAFAC based coders. For a particular value of CR, PARAFAC-t/dt/s and SVD (Matrix) yields smaller PRD, followed by PARAFAC-t/dt/s/s and PARAFAC-s/s/t; PRD of wavelet volumetric coding is at least twice as large compared to PARAFAC-t/dt/s. On the other hand, the PSNR of wavelet based coder (Wav-s/s/t) is slightly higher than SVD/PARAFAC based coders. However, this difference is small and practically negligible in linear scale (shown on the left of PSNR), and usually does not introduce a noticeable change in the appearance of signal.

The compression ratio for a given step size δ is largest for the EEG dataset (EEG-MMI) with lowest sample frequency f_s and lowest amplitude resolution, compared to the datasets with higher f_s and resolution (BCI3-MI & BCI4-MI). SVD/PARAFAC based near-lossless compression achieve a higher compression compared to single channel scheme, but slighlty lesser than its wavelet-coder counterpart. On the other hand, the average error of tensor based schemes are very less compared to the wavelet based schemes, which is an important fact to be taken into consideration. The PSNR values are higher (and hence the MAE values are lower) for the datasets with higher amplitude resolution, particularly, the PSNR values are higher for BCI3-MI and BCI4-MI (both 16 bit resolution) than for EEG-MMI (12 bit). Likewise, the PRD is lower for BCI3-MI and BCI4-MI than for EEG-MMI. For datasets with higher amplitude resolution, a given step size δ corresponds to smaller error.

The original signal and near-losslessly reconstructed signals are shown in Fig. 7 for one channel of 64 channel EEG (Table III, $\delta = 10$, EEG-MMI). The reconstructed signal closely resembles the original signal, and it is difficult to adjudge the performance of the algorithms by visual inspection of reconstructed and error signals. The error measures, however, are reported in Table III. The best performance in terms of CR, PRD, and PSNR for each data set and δ-value is indicated in bold face for SVD/PARAFAC-based schemes; the best performance among single-channel, image/volumetric wavelet coding scheme is also indicated in bold-face [16]. Note that the PRD and PSNR is comparable for each method,

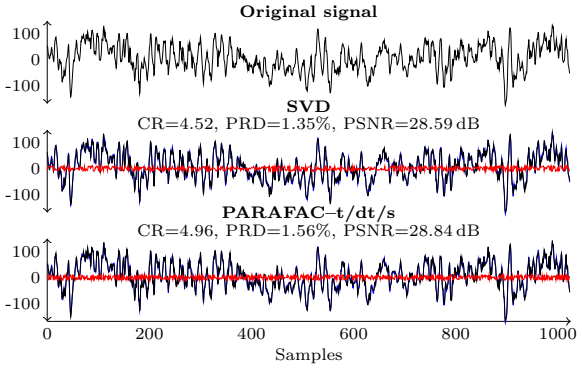


Fig. 7. Original and reconstructed signals of EEG-MMI dataset for SVD, PARAFAC-t/dt/s compression schemes; both the algorithms operate with a quantizer step-size $\delta = 10$. Other PARAFAC-based compression schemes yield similar results. The reconstructed signals are superimposed with the original signal in blue. Error between the original and reconstructed signal is superimposed and shown in red for easier comparison. Y-axis refers to quantization values.

for a given dataset and δ value.

Clearly, MC-EEG compression algorithms achieve larger compression ratios (CR), for a particular value of error δ, compared to the single-channel method. This observation confirms that inter-channel correlations may be exploited to achieve better compression.

For a fixed CR, SVD/PARAFAC-based algorithms yield a slightly lower PSNR and substantially lower average distortion (PRD) compared to wavelet-based volumetric coder. This is a key advantage of the SVD/PARAFAC-based two-stage coding. Moreover, when the sampling rate or resolution of signal is high (BCI3-MI & BCI4-MI), PARAFAC-based algorithm performs better compared to single-channel and image/volumetric wavelet coder.

VI. Conclusion

In this paper, we proposed novel near-lossless compression schemes for MC-EEG based on matrix/tensor decompositions. Multi-way representations allow us to efficiently exploit temporal and spatial correlations in MC-EEG. Several matrix/tensor decomposition models are analyzed and compared for compressing multi-channel EEG. Both SVD and PARAFAC seem to yield the most favorable results for compressing two-way and three- & four-way representations respectively. We designed a two-stage compression scheme: first we compress the SVD & PARAFAC components in a lossy fashion, and next the residuals are compressed by quantization and modified arithmetic coding.

We compared the proposed compression scheme with a single-channel and a wavelet-based MC-EEG compression scheme developed in our earlier work [16]. For given compression ratios, the proposed near-lossless compression scheme yields two-fold to five-fold smaller average distortion compared to wavelet-based volumetric coding. In conclusion, tensor decomposition is an attractive approach to near-lossless compression of MC-EEG.

References

[1] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, September 2009.

[2] A. Cichocki, R. Zdunek, A. H. Phan, and S. ichi Amari N, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Wiley, 2009.

[3] B. H. Brinkmann, M. R. Bower, K. A. Stengel, G. A. Worrell, and M. Stead, "Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data," *J. Neuroscience Meth.*, vol. 180, no. 1, pp. 185 – 192, 2009.

[4] R. Agarwal and J. Gotman, "Long-term EEG compression for intensive-care settings," *IEEE EMBS Mag.*, vol. 20, no. 5, pp. 23–29, Sep./Oct. 2001.

[5] S. Baker and D. Hoglund, "Medical-grade, mission-critical wireless networks," *IEEE EMBS Mag.*, vol. 27, no. 2, pp. 86 –95, march-april 2008.

[6] A. Casson, D. Yates, S. Smith, J. Duncan, and E. Rodriguez-Villegas, "Wearable electroencephalography," *IEEE EMBS Mag.*, vol. 29, no. 3, pp. 44 –56, may-june 2010.

[7] G. Antoniol and P. Tonella, "EEG data compression techniques," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 2, pp. 105–114, Feb. 1997.

[8] K. Srinivasan, J. Dauwels, and M. R. Reddy, "A two-dimensional approach to lossless EEG compression," *Biomed. Signal Proc. Control*, vol. 6, pp. 387–394, 2011.

[9] N. Sriraam and C. Eswaran, "An adaptive error modeling scheme for the lossless compression of EEG signals," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 5, pp. 587–594, Sep. 2008.

[10] N. Memon, X. Kong, and J. Cinkler, "Context-based lossless and near-lossless compression of EEG signals," *IEEE Trans. Biomed. Eng.*, vol. 3, no. 3, pp. 231–238, Mar. 1999.

[11] N. Sriraam and C. Eswaran, "Performance evaluation of neural network and linear predictors for near-lossless compression of EEG signals," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 87–93, Jan. 2008.

[12] Y. Wongsawat, S. Oraintara, T. Tanaka, and K. Rao, "Lossless multi-channel EEG compression," in *Proc. IEEE Int. Symp. Circuits & Syst.*, sep. 2006, pp. 1611–1614.

[13] D. Gopikrishna and A. Makur, "A high performance scheme for EEG compression using a multichannel model," in *High Performance Computing - HiPC 2002*, ser. Lecture Notes in Computer Science, vol. 2552. Springer Berlin / Heidelberg, 2002, pp. 443–451.

[14] Q. Liu, M. Sun, and R. Sclabassi, "Decorrelation of multichannel EEG based on hjorth filter and graph theory," in *6th Int. Conf. on Signal Proc.*, vol. 2, Aug. 2002, pp. 1516–1519.

[15] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy, and A. Cichocki, "Multi-channel EEG compression based on matrix and tensor decompositions," in *IEEE Int. Conf. Acoustics, Speech & Signal Proc.*, May 2011, pp. 629–632.

[16] K. Srinivasan, J. Dauwels, and M. Ramasubba Reddy, "Multichannel EEG compression: Wavelet-based image and volumetric coding approach," *IEEE Trans. on Inf. Technol. Biomed.*, 2012, (in press).

[17] M. W. Mahoney and P. Drineas, "CUR matrix decompositions for improved data analysis," *Proc. National Acad. of Sciences*, vol. 106, no. 3, pp. 697–702, 2009.

[18] C. F. Caiafa and A. Cichocki, "Generalizing the column-row matrix decomposition to multi-way array," *Linear Algebra and its Applications*, vol. 433, pp. 557–573, 2010.

[19] K. Srinivasan and M. R. Reddy, "Efficient pre-processing technique for lossless real-time EEG compression," *Electronics Letters*, vol. 46, no. 1, pp. 26–27, Jan. 2010.

[20] A. Said, "On the reduction of entropy coding complexity via symbol grouping: I redundancy analysis and optimal alphabet partition," HP Laboratories Palo Alto, Tech. Rep. HPL-2004-145, 2004.

[21] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–220, 2000.

[22] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Müller, "Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 993 – 1002, June 2004.

[23] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539 – 550, 2007.

[24] [Online]. Available: http://braintrust.seas.upenn.edu/