Enabling Advanced Inference on Sensor Nodes Through Direct Use of Compressively-sensed Signals

Mohammed Shoaib, Niraj K. Jha, and Naveen Verma
Department of Electrical Engineering, Princeton University, NJ 08544
Email: {mshoaib,jha,nverma}@princeton.edu

***Abstract*—Nowadays, sensor networks are being used to monitor increasingly complex physical systems, necessitating advanced signal analysis capabilities as well as the ability to handle large amounts of network data. For the first time, we present a methodology to enable advanced decision support on a low-power sensor node through the direct use of compressively-sensed signals in a supervised-learning framework; such signals provide a highly efficient means of representing data in the network, and their direct use overcomes the need for energy-intensive signal reconstruction. Sensor networks for advanced patient monitoring are representative of the complexities involved. We demonstrate our technique on a patient-specific seizure detection algorithm based on electroencephalograph (EEG) sensing. Using data from 21 patients in the CHB-MIT database, our approach demonstrates an overall detection sensitivity, latency, and false alarm rate of 94.70%, 5.83 seconds, and 0.199 per hour, respectively, while achieving data compression by a factor of 10×. This compares well with the state-of-the-art baseline detector with corresponding results being 96.02%, 4.59 seconds, and 0.145 per hour, respectively.**

I. Introduction

In next-generation sensor networks, the ability to provide advanced assessment over a large number of signal channels will be of critical value. In many applications, however, the signals involved are too complex to model adequately at the physical level. Data-driven techniques are emerging as a powerful approach for overcoming this challenge [1]. Out-patient monitoring networks are representative of such applications, since they require specific, clinically-relevant states to be modeled in physiological signals that are accessible through low-power sensors. As rich data become increasingly available, both through low-power sensing technologies and in the form of electronic records, supervised-learning methods provide efficient techniques for constructing and applying accurate signal models derived from the data. The problem, however, is that in networks formed from small-scale, energy-constrained nodes, communication and storage of data pose primary limitations.

Fig. 1 illustrates the bottlenecks. First, the signals of interest may be distributed and only invasively accessible [action potential, local field potential (LFP), electrocorticogram (ECoG), and EEG signals are shown as examples in Fig. 1]. In order to transmit these signals from an energy-constrained implant

Acknowledgments: This work was supported in part by the Qualcomm Innovation Ph.D. Fellowship and in part by NSF, and by the Gigascale Systems Research Center, one of six research centers funded under the Focus Center Research Program (FCRP), a Semiconductor Research Corporation entity.

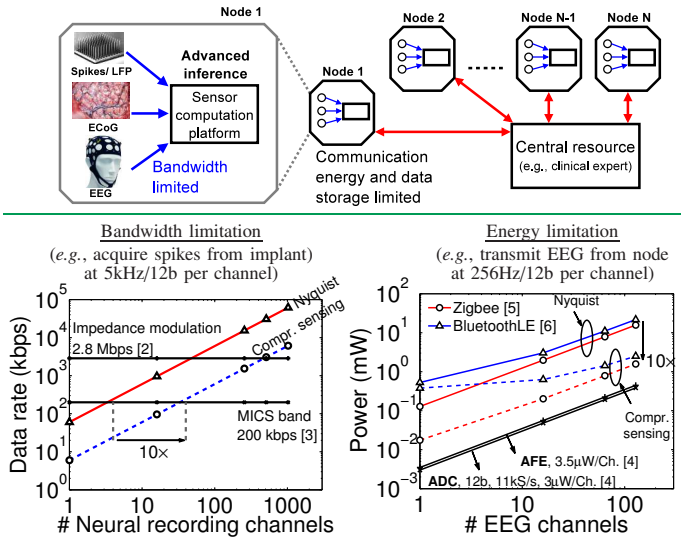

Fig. 1. Bandwidth and energy limitation of systems for physiological signal acquisition. Compressive sensing (compression by a factor of 10 is illustrated) can help alleviate the strict limitations imposed by Nyquist sampling.

to an external computation platform, transcutaneous or MICS-band links may be utilized [2], [3]. However, as shown for the example of neural spike signals (bottom left of Fig. 1), such links can pose strict bandwidth limitations for even moderate channel counts. Second, data storage on a sensor platform and communication to centralized resources or gateway devices face capacity and energy limitations in a network, particularly as the number of such nodes increases. For instance, the communication energy (bottom right of Fig. 1) imposes limits on wireless EEG acquisition systems [4]–[6].

Efficient signal representation techniques could thus play a critical role in networks composed of energy-constrained nodes. In this paper, we unite such a technique with advanced on-sensor signal analysis. The signal representation we focus on is compressive sensing [7], which has gained popularity in small-scale sensors because of its potential to enable very low-energy compression as well as its applicability to a broad range of signals. An important challenge, however, is that it requires highly intensive computations for signal reconstruction [7], [8]. We thus present a methodology that enables the direct use of compressively-sensed signals in a supervised-learning framework to construct low-energy biomedical detectors for decision support on sensor nodes within a network. At the same time, this approach allows the signals to be reconstructed for offline analysis by clinical experts. Our specific contributions are as follows.

978-3-9810801-8-6/DATE12/© 2012 EDAA

- We propose a new methodology for transforming linear and nonlinear filtering operations into the compressed domain, enabling the computation of features from physiological signals. Features correspond to specific biomarkers, which substantially improve the performance of medical detectors. Previous work in signal classification using compressively-sensed data [9]–[11] has only investigated theoretical bounds and has not considered methods for extracting specific features.
- We demonstrate our methodology on a seizure detector based on spectral-energy features extracted from compressively-sensed EEG. In fact, spectral features are generic biomarkers for neural field potentials and hold relevance for a broad range of neurological applications (*e.g.*, brain-machine interfaces [12], sleep disorders [13], *etc.*). Using patient data from the CHB-MIT [14] database, we validate the computed features with mutual information and Kullback-Leibler (KL) divergence metrics, and compare the end-to-end detector performance with a baseline algorithm that uses Nyquist samples. Comparable performance is obtained while achieving high data compression ratios (of more than 10×).
- We analyze the hardware implications of implementing a detector using the proposed approach. The hardware operations required are compared to optimized Nyquist-domain methods, where optimizations, such as folded FIR and polyphase filter implementations, are possible.

II. Background

Given a signal $\mathbf{u} \in \mathbb{R}^N$, which is k-sparse in a dictionary $\mathbf{\Psi}$, the theory of compressive sensing states that we can collect M $[M \ll N]$ samples from $\mathbf{u}$, using simple measurement vectors ϕ_i, to extract sufficient information for accurate reconstruction with a high probability. Several choices of the measurement matrix $\mathbf{\Phi} = \{\phi_i\}_{1 \le i \le M}$ are possible. A random matrix, whose elements are ± 1 according to a Gaussian probability distribution, is often sufficient [7], [8], and leads to low-energy compression via simple additions.

A. Related Work

Techniques for dimensionality reduction and random projections in machine learning have recently received interest. In [9], a theoretical error bound is presented for the use of compressively-sensed signals with a discriminative classifier [*e.g.*, support-vector machine (SVM)]. Similar explorations on regression methods are described in [10], which show an estimation-error bound of $O(logM/\sqrt{M})$ (M is the number of compressed samples). The use of generative classifiers with random projections applied to data has also been explored [11].

The approach we take in this paper expands on the theoretical concepts behind the application of such machine-learning frameworks. An important distinction is that we also focus on methods for extracting specific signal features. In medical detectors, these features correspond to biomarkers in physiological signals that are determined to have some correlation with the clinical states we are interested in detecting. By isolating physically-meaningful biomarkers through the direct use of compressively-sensed signals, our methodology thus enables high-performance detectors that are also compatible with efficient methods for representing data within the sensing system and network.

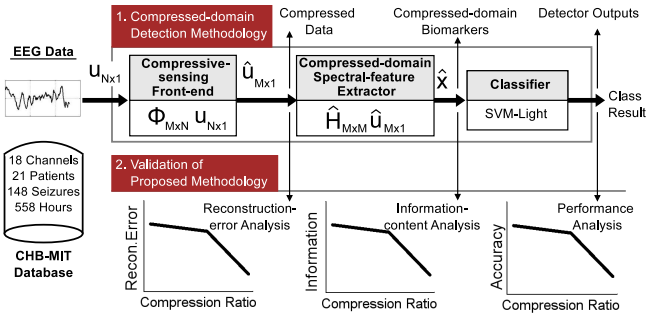

Fig. 2. The proposed detector comprises a compressive-sensing front-end, a compressed-domain feature extractor, and a classifier. We quantitatively validate the approach by analyzing the reconstruction error, the information content of the computed features, and the final detector performance.

III. Overview Of The Proposed Technique

An overview of our system architecture, which is implemented in MATLAB, is shown in Fig. 2. The detector consists of a sensing front-end, a compressed-domain spectral-feature extractor, and a classifier (for which we use an SVM). The figure also shows our methodology of quantitative validation at various stages for analyzing the reconstruction error, the information content of the computed features, and the final detector performance. An overview of the proposed methodology, which is demonstrated on an epileptic seizure detector, is presented in the rest of this section.

A. Compressed-domain Seizure Detector

In this section, we describe the sub-systems used in the proposed compressed-domain seizure detector.

Compressive-sensing front-end. 18 channels of raw EEG from several patients are available from the CHB-MIT database [14] as digitized samples recorded at 256 Hz. N Nyquist samples (corresponding to an epoch of two seconds) are applied as input to the sensing front-end shown in Fig. 2, where they are projected onto an $M \times N$ random measurement matrix $\mathbf{\Phi}$, resulting in a compressively-sensed signal representation of M ($\ll N$) samples. The number of Nyquist and compressively-sensed samples are therefore related by the compression ratio (R_{CS}), which is defined as

$$R_{CS} = N/M. \tag{1}$$

Compressed-domain feature extractor. The compressively-sensed samples from each EEG channel are processed to compute features using the functions $\hat{\mathbf{H}}_\mathbf{1}, \ldots, \hat{\mathbf{H}}_\mathbf{8}$ (only one instance of these functions is shown in Fig. 2). Each of these functions corresponds to the compressed-domain equivalent of applying a bandpass filter to the Nyquist signal. The functions thus effectively isolate the frequency content in eight bands. The energies within these bands are derived, which correspond to the biomarkers required for seizure detection [15].

Classifier. The outputs from the compressed-domain feature extraction block are fed to a classifier. To implement the SVM classifier, the open-source SVM-Light [16] package is used.

B. Validation of Methodology

The proposed approach is validated via the three types of analyses described below. We show that both the information content in the computed features and the final performance of the detector correspond to the reconstruction error of the compressively-sensed signal.

Information analysis. The first step in our validation is to quantify the change in the information contained in the compressed-domain features relative to that contained in the Nyquist-domain features. We use the metric of mutual information, which measures how the computed features alter the uncertainty of classifying a data instance. We observe that the information content initially degrades minimally with R_{CS}, but then, beyond a threshold, which is around ten, it degrades significantly.

Performance analysis. The performance metrics we use in our evaluation are sensitivity, detection latency, and the number of false alarms per hour. We observe that there is a direct correlation between the decay in information content and the performance of the compressed-domain detector. Initially, the performance degrades modestly with increasing R_{CS}. However, with larger compression ratios, the latency and false alarms increase, and sensitivity drops significantly.

Reconstruction error analysis. Reconstruction error is used to understand the information-content and detector-performance trends observed. We find that these trends correspond to the degradation in reconstruction error and thus accurate compressed-domain detection can be achieved in a manner that follows the reconstruction accuracy. This is promising, since compressive sensing can often tolerate high compression ratios while maintaining reasonable reconstruction accuracy.

IV. Detection With Compressively-sensed Data

In this section, we formulate the computations of the baseline seizure detection algorithm to permit transformation to the compressed domain. We then present the proposed compressed-domain approach.

A. Epileptic Seizure Detection Using the EEG

Fig. 3 illustrates the baseline algorithm, which relies on patient-specific SVM training [15]. The algorithm processes a two-second epoch of each EEG channel using eight bandpass filters (with passbands of 0-3 Hz, 3-6 Hz, ..., 21-24 Hz). Clinical studies have shown that EEG spectral energies can serve as biomarkers that indicate the onset of a seizure [17]. As shown in the figure, the energy from each filter is then represented by summing the squared value of the output samples to form a feature vector, which is then used for classification by an SVM. The detector is applied to 558 hours of data from 21 patients (corresponding to 148 seizures) and achieves an average sensitivity, latency, and specificity of 96.02%, 4.59 sec., and 0.145 false alarms per hour.

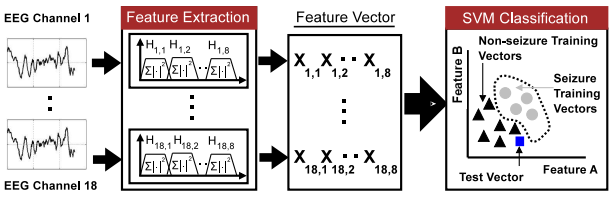

Fig. 3. The baseline epileptic seizure detection algorithm employing spectral-analysis feature extraction and SVM classification.

Feature-extraction formulation. Let each epoch of Nyquist-sampled EEG per channel be denoted by an N-dimensional vector $\mathbf{u} \in \mathbb{R}^N$. The eight bandpass filters can be represented by eight matrices $\mathbf{H_j} \in \mathbb{R}^{N\times N}$, $j = 1, 2, \ldots, 8$. Thus, a filtering operation can be formulated as a linear transformation given by

$$\mathbf{f_j} = \mathbf{H_j u}, \tag{2}$$

where $\mathbf{f_j} \in \mathbb{R}^N$ represents the filtered signal of N samples. To form the feature vector, we compute the energy in each of the frequency bands defined by $\mathbf{H_j}$. The following accumulation operation is thus used to derive the energy:

$$x_j = \mathbf{f_j}^T \mathbf{f_j}, \tag{3}$$

where x_j is the j^{th} dimension of the feature vector.

B. Compressed-domain Feature Computation Methodology

Let the inputs to the proposed compressed-domain detector be M ($\ll N$) compressively-sensed EEG samples per channel. These samples can be represented as $\hat{\mathbf{u}} \in \mathbb{R}^M$, and are related to the Nyquist samples $\mathbf{u}$ by

$$\hat{\mathbf{u}} = \mathbf{\Phi u}, \tag{4}$$

where $\mathbf{\Phi}$ is the measurement matrix whose elements are ± 1 *i.i.d* samples from a Gaussian normal distribution $\mathbb{N}(0, 1)$.

Now, in place of the original bandpass filters $\mathbf{H_j}$, we propose to use new matrix transformations $\hat{\mathbf{H}}_\mathbf{j} \in \mathbb{R}^{M\times M}$, $j = 1, 2, \ldots, 8$, which derive compressed-domain representation of the spectral components by directly using the compressively-sensed EEG samples.

In particular, in order to compute the desired features, $\hat{\mathbf{H}}_\mathbf{j}$ must be such that it yields an output signal from which the energy of the original band-isolated signal $\mathbf{f_j}$ can be obtained. Suppose that we are able to find a suitable matrix $\hat{\mathbf{H}}_\mathbf{j}$ that yields the output signal $\hat{\mathbf{f}}_\mathbf{j}$ such that

$$\hat{\mathbf{f}}_\mathbf{j} = \hat{\mathbf{H}}_\mathbf{j} \hat{\mathbf{u}}, \tag{5}$$

where $\hat{\mathbf{f}}_\mathbf{j} \in \mathbb{R}^M$ is a vector of M samples. We will show below that $\hat{\mathbf{f}}_\mathbf{j}$ can represent the energy of $\mathbf{f_j}$ if it is chosen to be the random projection of $\mathbf{f_j}$ given by

$$\hat{\mathbf{f}}_\mathbf{j} = \mathbf{\Phi f_j}. \tag{6}$$

This leads to the following relationship [from Eqs. (2), (4), and (5)]:

$$\hat{\mathbf{H}}_\mathbf{j} \mathbf{\Phi u} = \mathbf{\Phi H_j u}, \tag{7}$$

which allows us to derive $\hat{\mathbf{H}}_\mathbf{j}$. However, since $\mathbf{\Phi}$ is not a square matrix, the above relationships represent an underdetermined

set of equations. They can be solved using the Moore-Penrose pseudo-inverse $\mathbf{\Phi}^{\dagger} \in \mathbb{R}^{N \times M}$ of $\mathbf{\Phi}$ for a closed-form representation given by

$$\hat{\mathbf{H}}_{\mathbf{j}} = \mathbf{\Phi} \mathbf{H}_{\mathbf{j}} \mathbf{\Phi}^{\dagger}. \tag{8}$$

Proof that the energy of $\mathbf{f_j}$ can be represented by $\hat{\mathbf{f}}_{\mathbf{j}}$: A corollary from the Johnson-Lindenstrauss lemma states that the inner products are preserved under random projections [18]. We exploit this by applying it to the inner products of vectors $\mathbf{f_j}$ and $\mathbf{f_j^T}$ and their projections $\hat{\mathbf{f}}_{\mathbf{j}}$ and $\hat{\mathbf{f}}_{\mathbf{j}}^{\mathbf{T}}$. Thus, we have the following condition:

$$Pr\left(\left|\mathbf{f_j}^T \mathbf{f_j} - \hat{\mathbf{f}}_{\mathbf{j}}^T \hat{\mathbf{f}}_{\mathbf{j}}\right| \geq \delta\right) \ll 1, \forall j \in [1, 8] \tag{9}$$

where δ is a probabilistic error bound for the extracted features that justifies the assumption in Eq. (6). □

The required spectral energy (to compute the feature vector), in the compressed domain, can thus be obtained from $\hat{\mathbf{f}}_{\mathbf{j}}$ as follows:

$$\hat{x}_j = \hat{\mathbf{f}}_{\mathbf{j}}^T \hat{\mathbf{f}}_{\mathbf{j}}. \tag{10}$$

The feature vector from the above equation can now be used with an SVM classifier for seizure detection.

V. Analysis and Experimental Results

We quantify the performance of the proposed compressed-domain detection approach by analyzing how the derived features affect the conditional entropy of the seizure or non-seizure class values associated with the data. We then provide an experimental evaluation of the detection algorithm using compressively-sensed EEG.

A. Information Analysis of the Proposed Methodology

Let $C \in \{0, 1\}$ be the set of class values (*i.e.*, seizure or non-seizure) with cardinality $\mathbf{E}$ corresponding to all two-second epochs of EEG data in the CHB-MIT database. Further, let $P(C)$ represent the probability distribution of these values. The initial uncertainty in the feature vector set $\mathbf{X}$ (for all dimensions $\mathbf{D}$) prior to classification is measured by the Shannon entropy given by

$$S(X) = -\sum_{k=1}^{\mathbf{D}} P(\mathbf{x_k})\ logP(\mathbf{x_k}). \tag{11}$$

The average uncertainty after determining the $\mathbf{E}$ class labels is given by the conditional entropy:

$$S(X|C) = -\sum_{C=0,1} P(C) \sum_{k=1}^{\mathbf{D}} P(\mathbf{x_k}|C)\ logP(\mathbf{x_k}|C), \tag{12}$$

where $P(\mathbf{x_k}|C)$ is the conditional probability of the feature vector dimension derived from the Nyquist-sampled EEG, given the set C. The amount by which the uncertainty is decreased is, by definition, the mutual information $I(X;C)$ between the feature vector set X and the class values. A common measure for evaluating the ability of a classifier to discriminate between data instances thus exploits the following property [19]:

$$I(X;C) = S(X) - S(X|C). \tag{13}$$

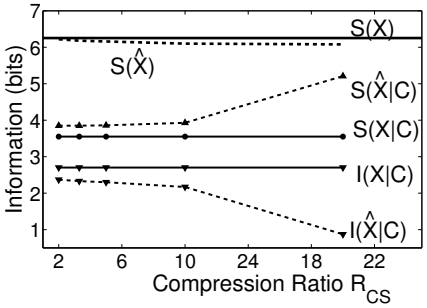

Fig. 4. KL divergence between the feature vectors in the original and compressed-domain detectors is retained up to a large compression ratio.

Intuitively, mutual information is the amount by which the knowledge provided by the set of class values C decreases the uncertainty about the feature vector set X [20]. It can therefore be used as a metric to validate the information content in the features derived by the compressed-domain detection methodology. This can be done as follows:

$$I(\hat{X};C) = S(\hat{X}) - S(\hat{X}|C). \tag{14}$$

After undergoing the feature extraction process, consider two separate instances of the SVM classifier operating on feature vector sets X and $\hat{X}$ corresponding to the baseline and the compressed-domain detectors, respectively. A measure of closeness in the result of the two classifiers is the KL divergence [19], which is defined as

$$D_{KL}(X||\hat{X}) = I(X;C) - I(\hat{X};C). \tag{15}$$

A KL divergence close to zero is desired to maintain correspondence between the two approaches. Fig. 4 shows numerical results for the analysis of information content averaged over 1 million epochs of EEG data from 21 patients in the CHB-MIT database. $S(X) = 6.25$ bits is the initial entropy in class values. After the extraction of spectral-energy features from Nyquist-sampled EEG data, the conditional entropy $S(X|C)$ drops to 3.55 bits. Using compressively-sensed EEG data and the proposed transformation of Eq. (8) for feature extraction, the conditional entropy in class values $S(\hat{X}|C)$ increases to 3.85 bits for $R_{CS} = 2$ and continues to increase to 3.85, 3.86, and 3.93 bits for $R_{CS} = 3.3$, 5, and 10, respectively. Compressing the EEG signals beyond this point increases $S(\hat{X}|C)$ to 5.21 bits. Consequently, compressing EEG signals, up to a ratio of approximately 10, increases $D_{KL}(X||\hat{X})$ minimally (*i.e.*, by 0.20 bits) and rapidly beyond this. Thus, the information content of the compressed-domain features is maintained up to a large compression ratio. To enable a calculation of the entropies, we quantized the probability distributions using 256 bins. Note that the *y-axis* in Fig. 4 is in bits, which means that the computation of individual entropies involves logarithm to base 2.

B. Performance Analysis of the Proposed Methodology

The compressed-domain detection methodology is further validated by simulating the end-to-end seizure detection performance using an SVM classifier.

Compressively-sensed EEG over 18 channels in the CHB-MIT database is processed one epoch at a time, using the

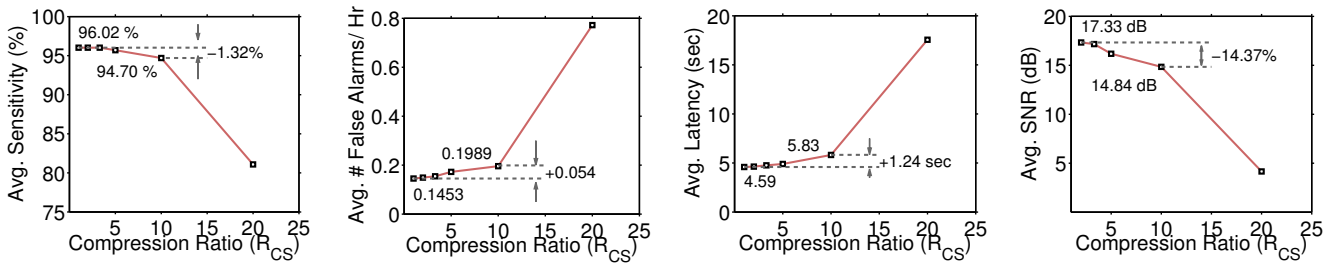

Fig. 5. Average performance of the compressed-domain seizure detection algorithm with direct use of compressively-sensed signals. The results were calculated using EEG data from 21 patients in the CHB-MIT database. The figure on the right shows the reconstruction accuracy versus R_{CS}.

transformation functions $\hat{\mathbf{H}}_\mathbf{j}$, $j = 1, 2, \ldots, 8$, from Eq. (8). Thus, feature vectors of 144 dimensions corresponding to the biomarkers are generated at a rate of 0.5 Hz. These vectors are used to train and test the SVM classifier in a patient-specific manner [15]. A leave-one-out cross-validation scheme is employed for measuring the performance of the detection algorithm. Fig. 5 shows the scaling in average performance (over 21 patients). The degradation in sensitivity is less than 1.32% up to a compression ratio of 10×, beyond which it begins to drop more significantly. The scaling in the number of false alarms per hour and the latency also follow a similar trend. The mean latency of detection increases by 1.24 seconds while the specificity of the algorithm degrades by 0.054 false alarms per hour at $R_{CS} = 10$. The degradation in performance exhibits the expected correlation with information loss in the feature vectors, as discussed in the previous section.

C. Reconstruction Error Analysis

We evaluate the accuracy of signal reconstruction from compressively-sensed EEG. Our aim is to observe whether any correlation exists between the detector performance and the reconstruction error.

In the compressive-sensing front-end (described in Fig. 2), a random measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ is used, along with a Gabor basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$, for sparse representation. The use of a Gabor basis enables an efficient representation and accurate reconstruction of EEG signals [21]. Gradient projection is used for signal reconstruction in the sparse reconstruction algorithm [22]. The signal-to-noise ratio (SNR) in dB is calculated based on the original ($\mathbf{u_i}$) and reconstructed epochs of the EEG. Fig. 5 shows the degradation in SNR as a function of the compression ratio averaged over 21 patients. With increasing compression ratio, the SNR of the reconstructed EEG drops minimally up to R_{CS} of approximately 10, and rapidly beyond that. This behavior correlates with the deterioration in performance of the compressed-domain detector. We thus observe a rationale for why decision-support computations can be supported directly by compressively-sensed EEG even in the case of substantial compression ratios.

VI. Hardware Implementation Analysis

In this section, we present an analysis of the hardware implementation cost for the compressed-domain detection algorithm. Since the proposed approach requires formulating feature computation as a matrix multiplication, it precludes the use of filter optimizations, such as multiply-accumulate (MAC)-stage folding and polyphase decomposition. On the other hand, the compressed-domain approach has the benefit of having to process fewer input samples. Thus, a trade-off related to the compression ratio is presented.

Filter optimizations. Fig. 6 shows the structure of the compressed- and Nyquist-domain detectors. MAC operations are the dominant computation in the system and are used for feature extraction in two stages: (i) the application of the spectral-analysis filters of Eq. (2) (identified as MAC0 in the figure), and (ii) the energy accumulation process of Eq. (3) (identified as MAC1). The number of operations per epoch are shown in dark boxes below the MAC units ($N = 512$ is the number of Nyquist samples, $k = 64$ is the filter order used, and $F = 8$ is the number of spectral-analysis filters). For the band-pass filtering (BPF) stage, kN operations are performed per epoch per channel in the cascade-form FIR implementation. They can be reduced to $kN/2$ by exploiting the symmetry of the filter coefficients $\mathbf{H_j}$ (necessary for a detector implementation with linear phase). Further computational reductions can be achieved through down-sampling prior to processing by a polyphase filter. Thus, F also corresponds to the polyphase down-sampling possible as a result of BPF.

The optimizations for Nyquist-domain processing are possible because filtering actually corresponds to convolution, allowing the matrix $\mathbf{H_j}$ to have a regular structure (with several zeroed entries), as shown in Fig. 6. With compressed-domain processing, however, the entries of the $\hat{\mathbf{H}}_\mathbf{j}$ matrix are determined by the random projection matrix $\mathbf{\Phi}$, disrupting

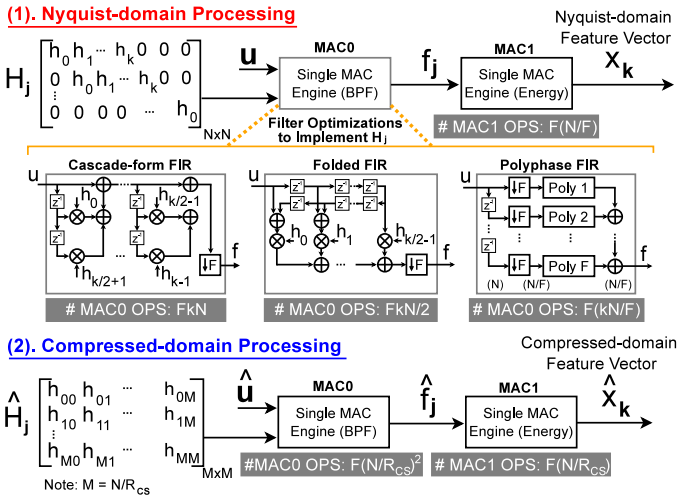

Fig. 6. The structure of the convolution matrix $\mathbf{H_j}$ in the Nyquist domain enables filter optimizations to reduce the number of MAC operations.

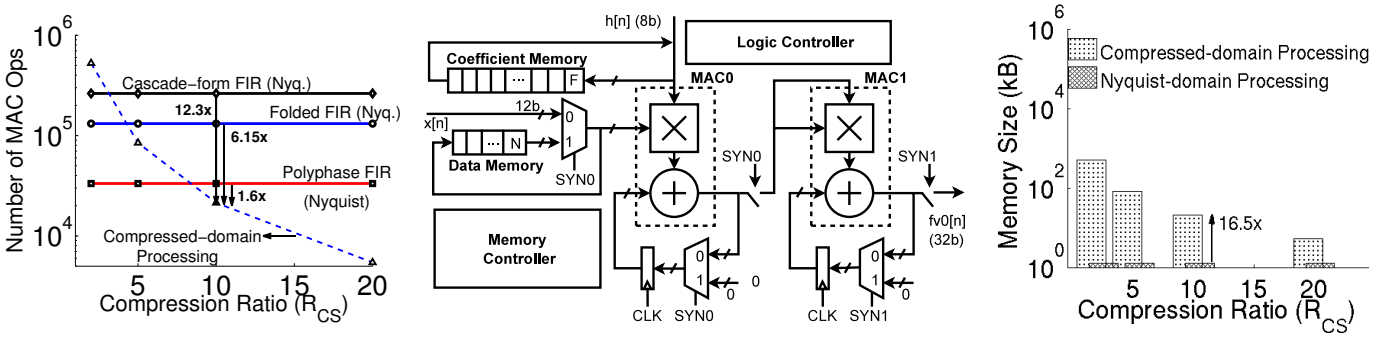

Fig. 7. Hardware analysis for detection in the compressed domain shows a subtle trade-off between an efficient implementation and data storage.

the regularity, and precluding the optimizations above. $\hat{\mathbf{H}}_\mathbf{j}$ thus potentially consists of $(N/R_{CS})^2$ non-zero coefficients. Therefore, in the most generic case, the compressed-domain implementation involves a matrix multiplication resulting in $(N/R_{CS})^2$ operations in MAC0. Fig. 7 shows a practical implementation of the detector using a single MAC engine for each of the operations, MAC0 and MAC1, described earlier. On the left, the figure shows the scaling in MAC operations with compression ratio R_{CS} (the number of MAC operations required for an optimized Nyquist implementation is also shown). We observe that at sufficient compression ratios (around $R_{CS} > 8$), compressed-domain processing can actually enable fewer hardware operations despite the optimizations possible in the Nyquist implementation. At $R_{CS} = 10$, the number of MAC operations in the compressed domain are 1.6×, 6.15×, and 12.3× fewer as compared to the polyphase, folded, and cascade-form FIR architectures, respectively. Note that random projections in the compressed domain can be performed without any MAC operations and are excluded from this analysis.

Another key consideration in the implementation is the amount of memory required by the detectors (shown at the right in Fig. 7). We need to store kF coefficients for the Nyquist implementation versus $k(N/R_{CS})^2$ for the compressed-domain detector. There is thus a subtle computation-memory trade-off in the implementation of the compressed-domain detector. The system-level gains from reduced communication and computation costs, however, are significant, typically far exceeding the memory overheads in a low-power sensor platform.

VII. Conclusions

An ensemble of sensor nodes monitoring complex physical systems presents high volumes of data both for storage and communication. We presented a methodology to enable accurate on-sensor decision-support computations based on direct use of compressively-sensed signals; such signals provide an efficient method for representing the data. We presented an efficient compressed-domain detector for epileptic seizures. We observed that the performance of the compressed-domain algorithm is preserved up to the accuracy with which the signals can be reconstructed. We also observed that the gains from reduced computation and communication potentially far exceed the memory overheads of hardware implementation on a low-power sensor platform.

References

[1] J. B. Predd, S. R. Kulkarni, and H. V. Poor, "Distributed learning in wireless sensor networks," *IEEE Signal Proc. Magazine*, vol. 23, no. 4, pp. 56–69, 2006.

[2] S. Mandal and R. Sarpeshkar, "Power-efficient impedance-modulation wireless data links for biomedical implants," *IEEE Trans. Biomedical Circuits and Systems*, vol. 2, no. 4, pp. 301–315, Dec. 2008.

[3] J. Pandey and B. P. Otis, "A sub-100 μW MICS/ISM band transmitter based on injection-locking and frequency multiplication," *IEEE J. Solid-State Circuits*, vol. 46, pp. 1049–1058, May 2011.

[4] N. Verma *et al.*, "A micro-power EEG acquisition SoC with integrated feature extraction processor for a chronic seizure detection system," *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 804–816, Apr. 2010.

[5] Texas Instruments Inc., Oct. 2007, "Low-cost low-power 2.4 GHz RF transmitter," http://focus.ti.com/docs/prod/folders/print/cc2550.html.

[6] Texas Instruments Inc., Jul. 2011, "2.4 GHz bluetooth low-energy SoC," http://ti.com/lit/ds/symlink/cc2540.pdf.

[7] D. Donoho, "Compressed sensing," *IEEE Trans. Information Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[8] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Information Theory*, vol. 52, pp. 5406–5425, Dec. 2006.

[9] R. Calderbank *et al.*, "Compressed learning: Universal sparse dimensionality reduction and learning in the measurement domain," *Technical Report, Dept. of Computer Science, Princeton University*, Dec. 2009.

[10] O. A. Maillard and R. Munos, "Compressed least-squares regression," *Neural Information Proc. Systems*, pp. 1–9, 2009.

[11] R. J. Durrant and A. Kaban, "Compressed Fisher linear discriminant analysis: Classification of randomly projected data," in *Proc. ACM Int. Conf. Knowledge Discovery and Data Mining*, Jul. 2010, pp. 1119–1128.

[12] A.-T. Avestruz *et al.*, "A 5 μW/channel spectral analysis IC for chronic bidirectional brain-machine interfaces," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 3006–3024, Dec. 2008.

[13] W. B. Mendelson *et al.*, "Frequency analysis of the sleep EEG in depression," *Psychiatry Research*, vol. 21, no. 2, pp. 89–94, 2007.

[14] Physionet, "CHB-MIT Physionet database," http://www.physionet.org/physiobank/database.

[15] A. Shoeb and J. Guttag, "Application of machine learning to seizure detection," in *Proc. Int. Conf. Machine Learning*, Jun. 2010.

[16] T. Joachims, "SVM-Light, support vector machine," http://svmlight/joachims.org.

[17] J. Gotman *et al.*, "Frequency content of EEG and EMG at seizure onset: Possibility of removal of EMG artefact by digital filtering," *EEG and Clinical Neurophysiology*, vol. 52, no. 6, pp. 626–639, Aug. 1981.

[18] S. Dasgupta and A. Gupta, "An elementary proof of the Johnson-Lindenstrauss lemma," *Random Structures and Algorithms*, vol. 22, no. 1, pp. 60–65, 2002.

[19] J. W. Fisher, M. Siracusa, and K. Tieu, "Estimate of signal information content for classification," in *Proc. IEEE DSP Wkshp. and IEEE Signal Proc. Education Wkshp.*, 2009, pp. 353–358.

[20] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Trans. Neural Networks*, vol. 5, no. 4, pp. 537–550, 1994.

[21] S. Aviyente, "Compressed sensing framework for EEG compression," in *Proc. IEEE Int. Wkshp. Statistical Signal Proc.*, Aug. 2007, pp. 181–184.

[22] M. A. T. Figueiredo *et al.*, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics in Signal Proc.*, vol. 1, no. 4, pp. 586–597, 2007.