A Compressed-domain Processor for Seizure Detection to Simultaneously Reduce Computation and Communication Energy

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

Abstract— **In low-power sensing systems, communication constraints play a critical role; *e.g.*, biomedical devices often acquire physiological signals from distributed sources and/or wireless implants. Compressive sensing enables sub-Nyquist sampling for low-energy data reduction on such nodes. The reconstruction cost, however, is severe, typically pushing signal analysis to a base station. We present a seizure-detection processor that directly analyzes compressively-sensed electroencephalograms (EEGs) on the sensor node. In addition to alleviating communication costs while also circumventing reconstruction costs, it leads to computational energy savings, due to the reduced number of input samples. This provides an effective knob for system power management and enables scaling of energy and application-level performance. For compression factors of 2-24×, the energy to extract signal features (over 18 channels) is 7.13-0.11μJ, and the detector's performance for sensitivity, latency, and specificity is 96-80%, 4.7-17.8 sec., and 0.15-0.79 false-alarms/hr., respectively (compared to baseline performance of 96%, 4.6 sec., and 0.15 false-alarms/hr.).**

I. Introduction

In compressive sensing, an N-sample signal is multiplied by an $M \times N$ projection matrix $\mathbf{\Phi}$ to create an M-sample signal (with $M \ll N$) [1]. This approach for data compression is possible under the following conditions: (1) the N-sample signal is sparse in a secondary basis $\mathbf{\Psi}$, and (2) $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are incoherent with each other and satisfy the restricted isometry property [2]. For signals that satisfy the first condition, a $\mathbf{\Phi}$ whose elements are set to ± 1 randomly with a uniform probability satisfies the second condition with high probability [2]. Such a choice for $\mathbf{\Phi}$ enables low-energy compression, applicable to a broad range of signals; this has recently been exploited in biomedical sensors [3]. However, the signal reconstruction cost for compressive sensing is severe [4], limiting such devices to primarily serve as nodes for transmitting raw data to a base station. In advanced medical devices, however, there is a need to also analyze signals on the node [5]. Such devices typically work by extracting signal features based on physical biomarkers and then feeding these features to high-performance classifiers to detect targeted physiological states [5], [6]. After compressive sensing, however, the critical challenge for signal analysis is that the physical biomarkers are obscured due to the random projection. In this paper, we present an IC for EEG-based seizure detection that performs analysis directly on the compressively-sensed signals.

Fig. 1 illustrates the concept. Nyquist-domain EEG is known to be sparse in the Gabor basis $\mathbf{\Psi_G}$ [7]. This enables low-energy compression using a random projection matrix $\mathbf{\Phi_G}$. The resulting EEG signal is compressed, but also substantially altered. Nonetheless, a representation of the desired signal features can be obtained by *transforming* the feature-extraction computations based on $\mathbf{\Phi_G}$. This results in a *compressed-domain seizure detector*. The resulting transformation not only overcomes the limitations imposed by signal reconstruction, but also leads to computational energy savings due to a reduction in the number of input samples that need to be processed (by an amount equal to the compression factor, ξ). The IC thus provides a previously unexplored approach to processor power management.

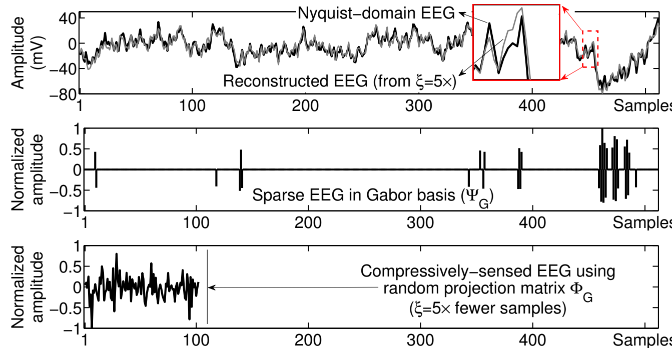

Fig. 1. Nyquist-domain EEG is sparse in the Gabor basis (center panel), enabling substantial compression (bottom panel); although accurate reconstruction is possible (as shown in top panel), reconstruction is computationally intensive, motivating signal analysis directly using the compressed signal.

II. System Approach

Fig. 2(a) shows a block diagram of the baseline Nyquist-domain seizure detector [8]. The biomarkers correspond to the spectral energy distribution of each EEG channel, computed over eight frequency bins for a 2-second epoch. FIR filtering is an extremely common DSP operation in sensor applications. In this case, a feature vector (FV) is extracted through a bank of eight FIR band-pass filters (BPF0,. . . , BPF7) followed by energy accumulators, applied to each EEG channel. The resulting FVs are then used for training and classification by a support vector machine (SVM) classifier. Next, the algorithmic formulation for compressed-domain processing as well as the

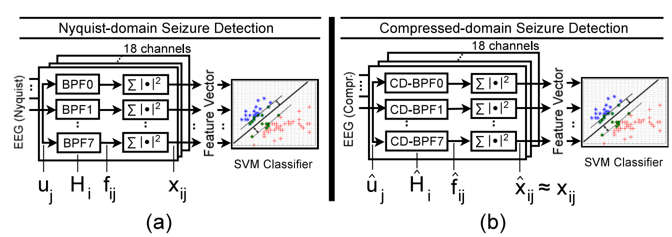

Fig. 2. The seizure detection algorithm involves feature extraction and classification using an SVM. In the compressed domain, we transform the Nyquist-domain BPFs $\mathbf{H_i}$ to the compressed-domain BPFs $\mathbf{\hat{H}_i}$.

opportunity this offers for processor power management are described.

A. Algorithmic Formulation

For transformation to the compressed domain, an FIR filter (i) for an EEG channel (j) can be formulated as matrix multiplication, namely of an input signal $\mathbf{u_j}$ by a matrix $\mathbf{H_i}$ to compute the filtered signal $\mathbf{f_{ij}}$ [Fig. 2(a)]. As described in [9], this makes possible the compressed-domain system shown in Fig. 2(b). In the proposed system, a corresponding matrix $\hat{\mathbf{H}}_\mathbf{i}$ is constructed such that the desired filtered signal can be represented by $\hat{\mathbf{f}}_\mathbf{ij}$ directly using the compressed input signal $\hat{\mathbf{u}}_\mathbf{j}$. The energy ($x_{ij}$) of $\mathbf{f_{ij}}$ can then be derived from $\hat{\mathbf{f}}_\mathbf{ij}$ by exploiting the properties of random projections. Each $\hat{\mathbf{H}}_\mathbf{i}$ thus effectively forms a *compressed-domain band-pass filter* (CD-BPF).

First, to construct $\hat{\mathbf{H}}_\mathbf{i}$, we specifically aim to find a matrix that, using $\hat{\mathbf{u}}_\mathbf{j}$ ($= \mathbf{\Phi u_j}$), leads to the corresponding random projection of the filtered signal $\hat{\mathbf{f}}_\mathbf{ij}$ ($= \mathbf{\Phi f_{ij}}$). This gives the following relationship:

$$\hat{\mathbf{f}}_\mathbf{ij} = \hat{\mathbf{H}}_\mathbf{i}\hat{\mathbf{u}}_\mathbf{j} \Leftrightarrow \mathbf{\Phi f_{ij}} = \hat{\mathbf{H}}_\mathbf{i}\mathbf{\Phi u_j} \Leftrightarrow \mathbf{\Phi H_i u_j} = \hat{\mathbf{H}}_\mathbf{i}\mathbf{\Phi u_j}. \tag{1}$$

Since $\mathbf{\Phi}$ is not a square matrix (but rather an $M \times N$ matrix with $M \ll N$, to enable substantial compression), Eq. (1) represents an overdetermined set of equations and cannot be directly solved for $\hat{\mathbf{H}}_\mathbf{i}$. It can, however, be solved in the least-square sense as follows:

$$\begin{aligned} \hat{\mathbf{H}}_\mathbf{i}\mathbf{\Phi} &= \mathbf{\Phi H_i} \\ \hat{\mathbf{H}}_\mathbf{i}\mathbf{\Phi\Phi^T} &= \mathbf{\Phi H_i \Phi^T} \\ \Rightarrow \hat{\mathbf{H}}_\mathbf{i} &= \mathbf{\Phi H_i} \underbrace{\mathbf{\Phi^T}\left(\mathbf{\Phi\Phi^T}\right)^{-1}} = \mathbf{\Phi H_i \Phi_R^\dagger}. \end{aligned} \tag{2}$$

where $\mathbf{\Phi_R^\dagger}$ is called the right pseudo-inverse of $\mathbf{\Phi}$ (since $\mathbf{\Phi}$ is not a square matrix, it does not have a true inverse).

To validate the end-to-end performance of this approach and to compute the signal features, we must next derive the energy of the band-limited signal. In order to do this, we take advantage of the *Johnson-Lindenstrauss theorem* and a resulting corollary, which states that for a subset of vectors, the inner products are preserved under random projections [10]. In the seizure-detection application, the filtered epoch data can be represented by a small set of basis vectors. This implies that the inner product of $\mathbf{f_{ij}}$ with itself can be represented by the inner product of $\hat{\mathbf{f}}_\mathbf{ij}$ with itself. Since the inner product of $\mathbf{f_{ij}}$ with itself represents the spectral energy, the desired signal features can be computed as follows:

$$x_{ij} \approx \hat{x}_{ij} = \hat{\mathbf{f}}_\mathbf{ij}^\mathbf{T}\,\hat{\mathbf{f}}_\mathbf{ij}. \tag{3}$$

Fig. 3 shows the performance of this approach using 558 hrs. of patient data from [11]. As shown, performance very close to the Nyquist-domain seizure detector [Fig. 2(a)] is retained up to large compression factors ξ ($= N/M$).

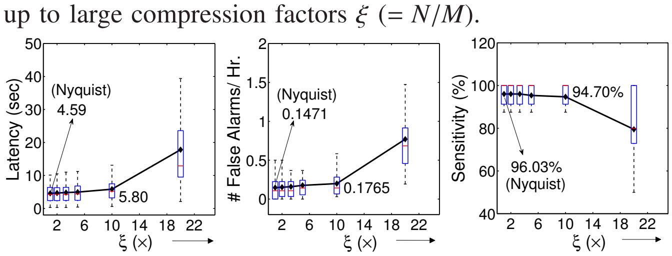

Fig. 3. Detector performance (shown over 21 patients) is maintained up to large ξ ($\sim 10\times$).

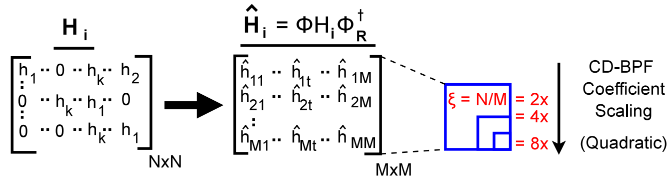

Fig. 4. $\hat{\mathbf{H}}_\mathbf{i}$, derived using $\mathbf{H_i}$ and $\mathbf{\Phi}$, disrupts the regularity and zeros in $\mathbf{H_i}$. The complexity of the CD-BPFs thus scales quadratically with ξ.

B. Processor Architecture with Power Management

Given the performance *vs.* compression-factor profile of the compressed-domain seizure detector, scalability in ξ can be exploited as a knob for system power management. An important consequence of the construction method described above, however, is that the CD-BPF matrices $\hat{\mathbf{H}}_\mathbf{i}$ (which are of dimensionality $M \times M$) do not retain the regularity of $\mathbf{H_i}$ (which are of dimensionality $N \times N$). As shown in Fig. 4, the rows of $\mathbf{H_i}$ are simply shifted to realize convolution with the BPF impulse response, and several of the entries are zero, depending on the filter order k. In $\hat{\mathbf{H}}_\mathbf{i}$, however, the shifted impulse responses and the zero entries are disrupted due to multiplication by random matrices $\mathbf{\Phi}$ and $\mathbf{\Phi_R^\dagger}$. As shown in Fig. 4, the number of multiplications required thus no longer depends on the filter order, but rather *quadratically* on the number of input samples M used to represent each epoch of EEG data. This scaling can potentially reduce the number of multiplications required, leading to the energy-scalable processor architecture whose block diagram is shown in Fig. 5. The processor consists of a compressed-domain feature extractor (CD-FE), which includes a CD-BPF and energy-accumulator block. The coefficients for the CD-BPF are pulled from a scalable SRAM bank. By disrupting the regularity, the $\hat{\mathbf{H}}_\mathbf{i}$ matrices necessitate that a larger number of distinct coefficients be stored, potentially increasing the memory requirements. Scalability in the SRAM bank is thus an important aspect of power management. This is achieved through the use of multiple subarrays, which enable fine-grained power-gating as well as reduced bit-line and word-line access energy. The total bank size in our implementation is 32kB. An SVM classifier is also integrated to perform real-time seizure detection using the derived FVs. It is assumed that compressed signals are input to the processor. However, for the case of Nyquist inputs, a compressive-projection frontend (CPF) is also included to explicitly multiply inputs by a random projection matrix $\mathbf{\Phi}$ to exploit a reduced number of samples for processor energy savings.

III. Low-energy Compressed-domain Processor

Fig. 6 shows the circuits used in the compressed-domain processor. Since FV computations are not throughput-limited (*i.e.,*

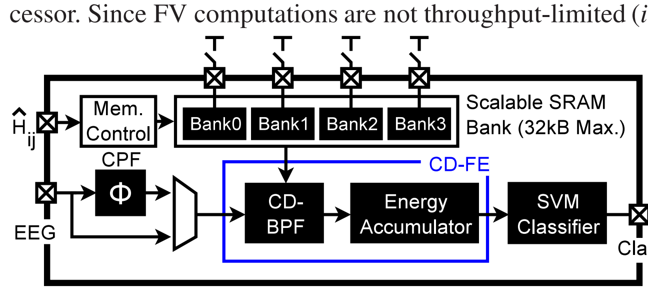

Fig. 5. Architecture block diagram of energy-scalable, compressed-domain seizure detector.

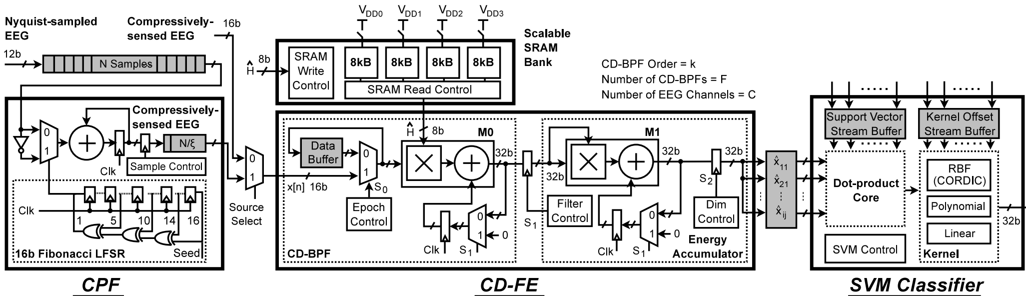

Fig. 6. Circuits used in the compressed-domain processor for seizure detection.

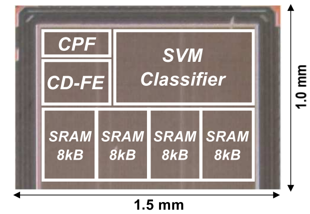

Fig. 7. Die photo of IC.

TABLE I

IC PERFORMANCE SUMMARY.

Energy-scalable, compressed-domain processor in 0.13μm LP CMOS				
Supply voltage: Logic (SRAM)	1.2-0.44 V (0.7/0.42 V)	Subblock energy/feature vector		
EEG sampling rate	256 Hz	CPF		85-7.3 pJ
Clock frequency	10.2-0.3 MHz	CD-FE		156-2 nJ
CPF compression factor	2-24×	SRAM		7.0-0.1 μJ
Feature computation rate	0.5 Hz	SVM	RBF	16.9-26.8 μJ
BPF order	64		Poly4	11.1-17.2 μJ
CD-BPF memory size	0.3-32 kB		Linear	55.6 nJ

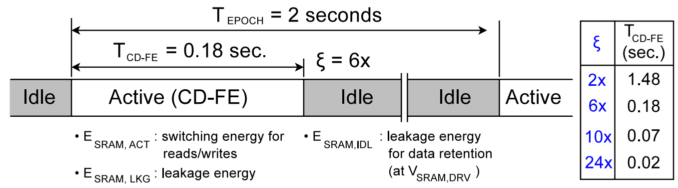

Fig. 8. Summary of energy components contributing to total SRAM energy (the $\xi = 6\times$ case is shown for illustration).

one FV is derived every 2 seconds), the CD-FE sequentially computes each feature. The CD-FE can be configured to compute up to eight spectral features ($i = 0, \ldots, 7$) for each EEG channel (j) over as many as 18 channels, yielding a maximum FV dimensionality of 144. The control pulse S_0 initiates CD-BPF computations. A multiply-accumulate (MAC) unit (M0) is used to perform the matrix multiplications required for $\hat{\mathbf{H}}_\mathbf{i}$, with each output-vector element being registered by the control pulse S_1. Energy accumulation over the output vector is then performed by a second MAC unit (M1). After spectral-energy extraction [which thus requires $(N/\xi)(N/\xi + 1)$ MAC operations], each feature ($\hat{x}_{ij}$) is derived and stored in an FV buffer when triggered by control pulse S_2. Four subarrays are used in the scalable SRAM bank. As described in Sec. IV, this provides a substantial power-management range while also balancing the hardware overheads incurred by finer SRAM granularity. Each subarray can be independently power-gated (from off-chip).

The derived FVs are classified for seizure detection using the SVM block. The SVM can apply linear, polynomial, or radial-basis function (RBF) transformations (*via* an embedded CORDIC engine). The support vectors are derived from offline training of the classifier and are provided through a dedicated interface. The classification result is encoded in the MSB of the SVM output (MSB = 1 for seizure detected, MSB = 0 for no seizure detected).

The CPF is selectable for front-end signal compression. It uses a 16b linear feedback shift register (LFSR) to implement multiplication with a random projection matrix $\mathbf{\Phi}$, as shown in Fig. 6.

IV. Measurement Results

The IC was prototyped in a 0.13μm LP CMOS process from IBM. The die photograph and performance summary are shown in Fig. 7 and Table I, respectively. Nyquist EEG signals are sampled at a rate of 256 Hz, and the CD-BPFs are derived for Nyquist-domain BPFs of order $k = 64$ (based on the filter specifications required for seizure detection [6]). The CPF permits EEG compression by a factor of ξ = 2-24×, consuming 85-7.3pJ of energy. The results presented next consider the impact of compression-factor scaling on processor energy.

SRAM Energy Components. The SRAM consumes a substantial portion of the total CD-FE energy. Its optimization to exploit scalability with respect to ξ is thus a key focus. The detector processes an EEG epoch every T_{EPOCH} = 2 sec. However, the optimal operating frequency (and supply voltage) for CD-FE is determined by the minimum-energy point [12], which yields a throughput that allows the active computations to be completed in less than 2 seconds. The SRAM energy is thus the sum of the active-mode ($E_{SRAM,CD\text{-}FE}$) and idle-mode ($E_{SRAM,IDL}$) energies for each subarray that is enabled. During the active mode, the SRAM is operated at the minimum operational supply voltage of 0.7 V, where it can operate at 920

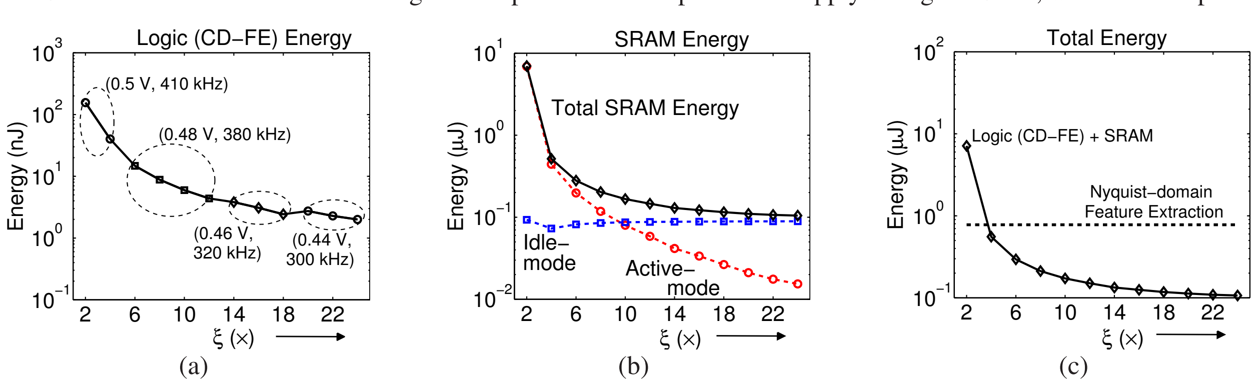

Fig. 9. Logic and SRAM energy for the CD-FE (with the optimal logic V_{DD} in the range 0.5-0.44 V).

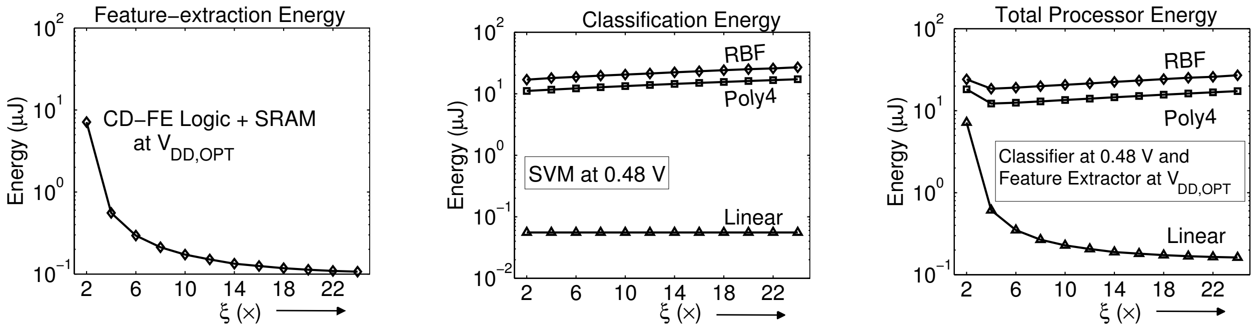

Fig. 10. ξ and V_{DD} scaling enable energy scalability in the processor.

kHz, providing sufficient performance for CD-FE operation at the minimum-energy point. During the idle mode, the SRAM is operated at its minimum data-retention voltage ($V_{SRAM,DRV}$) of 0.42 V. The duration of the active mode ($T_{CD\text{-}FE}$) depends on the chosen compression factor and varies in the range 1.48-0.02 sec. for ξ = 2-24×. Fig. 8 summarizes the SRAM operating modes and energies, where both active-access ($E_{SRAM,ACT}$) and leakage ($E_{SRAM,LKG}$) energies are considered during the active mode. Considering the subarray scalability, the total SRAM energy is thus given as follows (where the number of subarrays N_{SUB} varies in the range 1-4 depending on ξ):

$$E_{SRAM} = [E_{SRAM,CD\text{-}FE} + E_{SRAM,IDL}] \cdot N_{SUB}$$

Total Feature-extraction Energy vs. ξ. Figs. 9(a) and 9(b) show the logic and SRAM energies and Fig. 9(c) shows the total energy for compressed-domain feature extraction (results are for 18 EEG channels with eight CD-BPFs). At $\xi > 4\times$, the total energy of compressed-domain processing is less than that projected for Nyquist-domain processing (for the filter orders considered). For each compression factor, the energy is reported for the optimal supply voltage ($V_{DD,OPT}$), which minimizes CD-FE's active-switching and leakage energies as well as the SRAM energy (including SRAM energy causes the CD-FE $V_{DD,OPT}$ to change with ξ). Fig. 9(a) shows the CD-FE energy, with $V_{DD,OPT}$ annotated. Fig. 9(b) shows that the total SRAM energy, which scales significantly with ξ, eventually begins to saturate due to the granularity limit of the four subarrays; a finer granularity would enhance scaling at the cost of hardware overhead.

Processor Energy vs. ξ. Fig. 10 shows the effect of compression-factor scaling on the total processor energy, where the SVM operates at its minimum-energy point of 0.48 V, the CD-FE operates at the minimum-energy points specified in Fig. 9(a), and the SRAMs operate at 0.7/0.42 V during the active/idle modes. Nonlinear SVMs (RBF and Poly4) consume significant energy, while SVMs with a linear kernel incur minimal energy, causing the energy-scaling characteristics to be dominated by CD-FE. For the nonlinear cases, the SVM energy is considerable and actually leads to an optimal compression factor of ξ = 5×. This happens because the SVM models obtained from classifier training become somewhat more complex as the compression factor increases. The resulting increase in classifier energy opposes the reduction in CD-FE energy.

V. Conclusions

Random projections in compressive sensing obscure the sensed signals, thus preventing the use of Nyquist-domain algorithms for signal analysis. Signal reconstruction, however, is energy-intensive and is not desirable on low-power sensor nodes. We presented the design of a processor that enables on-node signal analysis directly using compressively-sensed data. The approach relies on the analytical construction of a compressed-domain BPF, and allows performance to be retained up to high compression factors ($\xi \sim 10\times$). In addition to communication energy savings through end-to-end data reduction in a system, this enables a mode of power management where the computational energy scales strongly with the compression factor due to a reduction in the number of input samples that need to be processed.

Acknowledgments

This work was supported in part by the Qualcomm Innovation Fellowship and in part by the Gigascale Systems Research Center, one of six research centers funded under the Focus Center Research Program (FCRP), a Semiconductor Research Corporation entity. IC fabrication was provided by MOSIS. We thank Kyong Ho Lee for help with the IC layout.

References

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

[2] 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.

[3] F. Chen, A. P. Chandrakasan, and V. Stojanoviċ, "A signal-agnostic compressed sensing acquisition system for wireless and implantable sensors," in *Proc. Conf. Custom Integrated Ccts.*, Sep. 2010, pp. 1–4.

[4] 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.

[5] A. Csavoy, G. Molnar, and T. Denison, "Creating support circuits for the nervous system: Considerations for brain-machine interfacing," in *Proc. Int. Symp. VLSI Circuits*, Jun. 2009, pp. 4–7.

[6] N. Verma *et al.*, "A micro-power EEG acquisition SoC with integrated seizure detection processor for continuous patient monitoring," in *Proc. Symp. VLSI Circuits*, Jun. 2009, pp. 62–63.

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

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

[9] M. Shoaib, N. K. Jha, and N. Verma, "Enabling advanced inference on sensor nodes through the direct use of compressively-sensed signals," in *Proc. IEEE DATE Conf.*, Mar. 2012, pp. 437–443.

[10] 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.

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

[12] A. Wang and A. P. Chandrakasan, "A 180-mV subthreshold FFT processor using a minimum energy design methodology," *J. Solid-State Circuits*, vol. 40, no. 1, pp. 310–319, Jan. 2005.