



Zhang, J., Duncan, K., Suo, Y., Xiong, T., Mitra, S., Tran, T. D., and Etienne-Cummings, R. (2016) Communication channel analysis and real time compressed sensing for high density neural recording devices. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 63(5), pp. 599-608.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/121420/>

Deposited on: 04 August 2016

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Communication Channel Analysis and Real Time Compressed Sensing for High Density Neural Recording Devices

Jie Zhang[†], *Student Member, IEEE*, Kerron Duncan[†], *Student Member, IEEE*, Yuanming Suo, *Student Member, IEEE*, Tao Xiong, *Student Member, IEEE*, Srinjoy Mitra, *Member, IEEE*, Trac Duy Tran, *Fellow, IEEE* and Ralph Etienne-Cummings, *Fellow, IEEE*,

Abstract—Next generation neural recording and Brain-Machine Interface (BMI) devices call for high density or distributed systems with more than 1000 recording sites. As the recording site density grows, the device generates data on the scale of several hundred megabits per second (Mbps). Transmitting such large amounts of data induces significant power consumption and heat dissipation for the implanted electronics. Facing these constraints, efficient on-chip compression techniques become essential to the reduction of implanted systems power consumption. This paper analyzes the communication channel constraints for high density neural recording devices. This paper then quantifies the improvement on communication channel using efficient on-chip compression methods. Finally, This paper describes a Compressed Sensing (CS) based system that can reduce the data rate by $> 10\times$ times while using power on the order of a few hundred nW per recording channel.

Index Terms—Neural Recording, Compressed Sensing, Dictionary learning, Ultra-Wide Band, Compression

I. INTRODUCTION

Moore's law has predicted the scalability of integrated circuit and sensors in the past few decades. Technology scaling has enabled high-density sensor arrays which find applications in both the scientific and consumer domain. But unfortunately, battery capacity scaling does not follow such an aggressive trend. Battery energy density has stayed largely constant over the decade while transistors have doubled every two years. As a result, many modern sensors are forced to sample, process and transmit more data within the power impoverished environment.

High density neural recording devices for neuroscience experiments and Brain Computer Interfaces (BCI) are among the sensors that are facing these challenges. These systems call for high performance in power impoverished implanted environments. The neural recording tools have significantly evolved from single electrode in the 1950s to multi-electrode array in the 1990s [1] [2]. Technology further allows high density recording electrodes to be fabricated onto the same substrate as the acquisition circuitry [3] [4]. The trend of

high density electrode integration is expected to continue as researchers are currently designing silicon probes that contains >1000 channels [5]. The researchers are also exploring the possibility of designing a distributed network of sub- $100\mu m$ recording devices that can float in the brain to record and stimulate the brain tissue [6] [7].

For neural recording devices, a system power consumption break down has shown that data readout/transmission accounts for the majority system power consumption [8]. Transmission power consumption is directly proportional to data rate: for a wired connection, dynamic power is consumed every time a transition occurs at the output of the circuit. For wireless channels, more power has to be consumed to transfer data at faster speed to maintain acceptable Bit Error Rate (BER) [9]. Aside from battery energy and weight limitations, these implanted devices must also be low power to prevent heating of the surrounding brain tissue [10]. For implanted devices in human studies, Federal Communications Commission (FCC) also sets strict limit on the tissue Specific Absorption Rate (SAR) allowed from an implanted device in human. The total power dissipation, which includes the system power dissipation, wireless output power and inductive power delivered to the implants, must be bounded by the safety SAR.

Previously, many authors have investigated efficient on-chip compression methods for transmission data rate reduction [11] [13]–[17]. In this paper, we demonstrate a power efficient Compressed Sensing technique to reduce communication data rate between the on-chip recording circuit and off-chip receiver. This approach guarantees high compression rate (>10) while using low power digital circuit that occupies small chip area ($< 0.11mm^2/channel$, in $180nm$ CMOS process) [16] [17]. Unlike many previous compressed sensing systems, our system's sensing and signal recovery steps operate in real-time. Because of this feature, we can also monitor signal recovery quality in real-time. Signal recovery quality is directly related to compression rate. In a neural recording experiment, our system's compression rate can be dynamically adjusted to improve reconstruction quality. In section II of the paper, we first discuss the need for an efficient on-chip data compression method by analyzing the transmission channel constraints. Then in Section III, we outline the alternative approaches that have been published in the past. In section IV we introduce our compressed sensing approach using examples of public

J. Zhang, K. Duncan, T. Xiong, T. Tran and R. Etienne-Cummings are with Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD, USA

S. Mitra is with School of Engineering, University of Glasgow, Glasgow, Scotland, UK

[†] These authors contributed equally to the manuscript

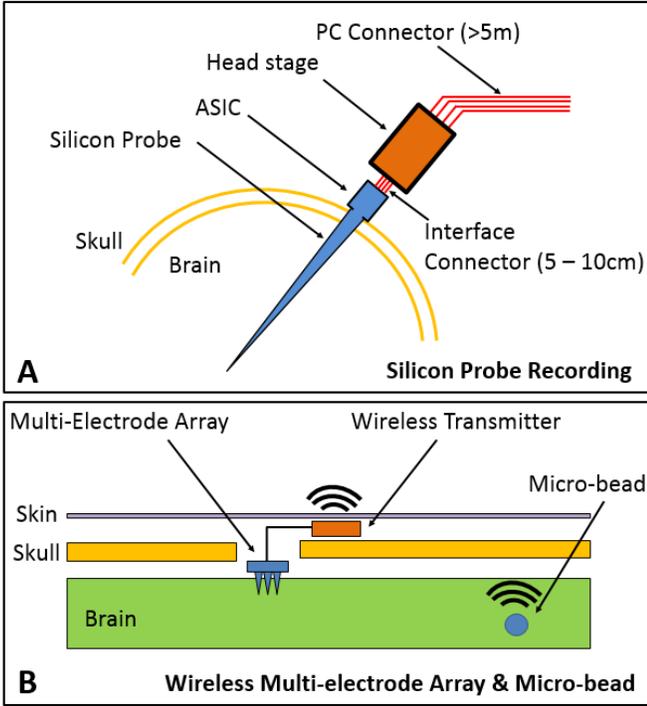


Fig. 1. (A) Wired silicon probe recording experiment setup. (B) Wireless Multi-electrode Array and miniature recording device setup.

available in-vivo recordings. Finally, we conclude the paper in section V.

II. THE NEED FOR ON-CHIP DATA COMPRESSION

As the high density integration continues, the amount of data generated by the device also grows. A 1000-electrode silicon probe generates data on the order of 300Mbps with each electrode sampling at 30KHz while providing at least 10-bit resolution per sample. Transmitting such high volume of data dissipates a significant of power in both wired and wireless communication channels. In the rest of this section we first analyze data transmission power consumption for both wired and wireless communication links. We then address how on-chip compression can be used to reduce power consumption and improve overall system performance.

A. Compression in Wired data transmission

Fig. 1 (A) shows a typical setup for recording experiment using high density silicon probe [3]. The neural probe consists of recording contacts and Application Specific Integrated Circuit (ASIC) fabricated on the same silicon substrate. The recording contacts are inserted into the brain while the ASIC is placed outside of the skull. The recording contacts, equipped with active electrodes, amplify the signals before sending them to the ASIC for additional filtering and digitization. The ASIC then transmits the digitized signal through a short PCB trace (5 to 10cm) to the head-stage, which then drives a long cable (>10m) to reach the computer.

In the wired recording setup, the power is typically delivered to the electronics through additional wires. Thus battery

capacity and weight are not the major constraint in this setup. However, the power dissipation of the head stage and the ASIC still need to be low to avoid heat generation which could increase the temperature of the brain to cause tissue damage. The heat dissipation at the ASIC has to be kept low especially due to its proximity to the brain tissue.

The dynamic power consumption at the output of the ASIC can be calculated as:

$$P_{dynamic} = \rho_t (C_L \cdot V_{dd}^2 \cdot f) \quad (1)$$

Where C_L is the load capacitance of the chip I/O, f is the data rate or clock frequency, and ρ_t is the probability where a signal transition occurs at the output [18]. To compute dynamic power, we take $\rho_t = 0.5$ and $C_L = 30pF$. The load capacitance includes wire bonding capacitance, connector interface and Printed Circuit Board (PCB) trace ($\approx 1pf/cm$) to the head stage. Each electrode has to sample the signal at 30 KHz with 10 bits of resolution. A 1000-electrode silicon probe generates around data at rate of 300 Mbps. Using standard digital output drivers with 3.3V power rail, the dynamic power consumption is calculated to be around 50 mW. To compare, we also estimated power consumption for the rest of the circuit. One electrode readout circuit typically consumes around 20uW [3]. A linear extrapolation indicates that 1000-recording channels would consume around 20mW. Therefore, the power consumed in wired transmission at high data rate accounts for 62.5% of the entire system power consumption.

50 mW of power consumption may also induce significant amounts of heat dissipation. Since silicon has high thermal conductivity, the heat may travel along the probe and increase the temperature of the brain tissue. The exact determination of the heat transfer is still an active area of study. Previous study has shown that 13 mW of heat dissipation by the ASIC in proximity of the electrode array induces around $1.3^\circ C$ [10] temperature increase in the tissue, which is higher than the maximum permissible temperature increase in the cortex of $1^\circ C$.

The power consumption in wired transmission is directly proportional to the data rate. Thus efficient on-chip compression method is important to reduce power consumption. With $10\times$ compression rate, the power consumption can be decrease to 5mW. As a result, overall system power consumption can decrease from 80mW to 25mW.

B. Compression in Wireless Data Transmission

Fig.1(B) shows recording setup using wireless multi-electrode array (MEA) and distributed miniature brain implants. Wireless communication method is commonly used for neuroscience experiment where the animal must perform free movement tasks [11] [19] [20]. They are also commonly used for human experiments such as prosthetics and Brain Computer Interface (BCI). The miniature brain implants are the new generation of devices under research [6] [7]. They are small enough to float within the deep brain to record neural activity and transmit the data wirelessly to the receiver outside of the brain. These devices are either powered through wireless power transfer or through a small battery carried

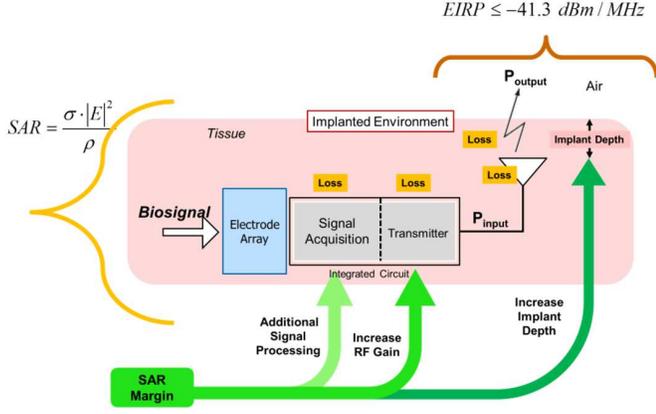


Fig. 2. System power dissipation has to be bounded within the safety SAR limit. The total transmitter output has to be bounded with safety EIRP limit.

by the animal. Hence, the wireless communication protocol must be power efficient to increase battery life, minimize RF radiation injected into the tissue and avoid heating the brain tissue.

1) *Constraints on Wireless Channel:* Ultra-wideband (UWB) has been proven to be an efficient band of communication for biotelemetry links [21] [22] [23]. There are two important constraints on the UWB biotelemetry channel of implant neural recording devices. They are the Equivalent Isotropic Radiated Power (EIRP) and the Specific Absorption Rate (SAR). These factors are extremely important for human related experiments. EIRP is the limitation placed on the output power transmitted into free space. The FCC spectral mask in this spectrum is limited to ≤ -41.3 dBm/MHz of maximum EIRP, which is a function of the operational bandwidth (BW) of the wireless system. EIRP for a UWB communication system using the maximum or minimum, 7500 MHz or 500 MHz bandwidth is equal to $556 \mu W$ and $37 \mu W$, respectively.

The SAR limits govern the energy radiated by the antenna absorbed and dissipated in the tissue and is described by:

$$SAR = \frac{d}{dt} \frac{dW}{dm} = \frac{d}{dt} \frac{dW}{\rho \cdot dV} = \frac{\sigma \cdot E^2}{\rho} \quad (2)$$

SAR is a function of the time derivative (d/dt) times the mass derivative (d/dm) of energy (W) in the tissue. Mass (m) can be calculated as the density of the tissue (ρ) by its volume (dV). Further derivations show that SAR can also be calculated as the product of the tissue conductivity (σ) and electric field (E) squared divided by density. The FCC SAR absorption limit is 1.6 W/kg in $1g$ of tissue which limits the performance of the wireless link since it constrains the amount of RF energy dissipated in the system which can affect transmission range [24]. Tissue losses due to the inefficiency of the antenna design and transmission loss through tissue directly impacts the SAR results. As shown in Fig. 2, the total loss from each block of the implanted system needs to be within the SAR safety limit.

Considering an implanted system with SAR calculated to be < 0.8 W/kg, with a limitation of 1.6 W/kg. This results in > 0.8 W/kg of safety margin that is now available to the other sub-systems in the transceiver.

2) *Analysis of the wireless channel using Link Budget:* There are numerous considerations involved in optimization of the wireless link of an implanted device [9]. We create a link budget to effectively evaluate the wireless channel. Link budget is a method of calculating the communication channel parameters of power, gain and losses in the system transmit and receive paths.

In the creation of the link budget the Friis equation for narrowband communication defines the received power as:

$$P_{rx} = P_{tx} \cdot G_{tx} \cdot G_{rx} \cdot \left(\frac{c_0}{4\pi \cdot d \cdot f} \right)^2 \quad (3)$$

Equation (3) is a function of transmit (G_{tx}) and receive (G_{rx}) antenna gains, operational frequency (f), speed of light (c_0), distance between antennas (d) and transmit (P_{tx}) and receive (P_{rx}) power [25]. The path loss, attenuation of RF energy as it propagates through the wireless channel environment is expressed as:

$$PL = \frac{P_{tx}}{P_{rx} \cdot G_{tx} \cdot G_{rx}} = \left(\frac{c_0}{4\pi \cdot d \cdot f} \right)^2 \quad (4)$$

This expression can be enhanced by changing the exponent value of 2 to n , which takes into consideration the variability of the wireless indoor environmental including factors such as whether the antennas (transmit and receive) are in line-of-sight (LOS) or non-line-of-sight (NLOS) and number of barriers (walls, floors and furniture) in the path [25]. We must also consider the statistical variation in the path loss. Thus, the refined path loss equation using the log-normal shadowing model is given by:

$$PL(dB) = -20 \cdot \log_{10} \left(\frac{4\pi \cdot d \cdot f}{c_0} \right) - 10 \cdot n \cdot \log \left(\frac{d}{d_0} \right) - X_\sigma \quad (5)$$

Where d_0 is the reference distance of the model, n is the path loss exponent. X_σ is the log-normal statistical zero-mean variation parameter with a standard deviation of σ , both of which are indoor environment dependent. For wide band applications f becomes the geometric mean (\bar{f}) of the upper frequency (f_U) and lower frequency (f_L) in the bandwidth:

$$\bar{f} = \sqrt{f_U \cdot f_L} \quad (6)$$

The received power equation for wide band applications (including the geometric mean of the frequency range), in logarithmic form, is now complete and given as:

$$P_{rx}(dB) = P_{tx}(dB) + G_{tx}(dB) + G_{rx}(dB) - 20 \cdot \log \left(\frac{4\pi \cdot d_0 \cdot \bar{f}}{c_0} \right) - 10 \cdot n \cdot \log \left(\frac{d}{d_0} \right) - X_\sigma \quad (7)$$

With path loss simplified into three parts, bandwidth $PL(f_U, f_L)$, distance/environment $PL(d, n)$ and statistical $PL(\sigma)$ dependent components given by:

$$PL = PL(f_U, f_L) - PL(d, n) - PL(\sigma) \quad (8)$$

3) *Effects of Compression on the Wireless Channel:* The impact of compression on the wireless channel can be evaluated when looking at the performance improvements to range, EIRP and SAR. For instance, with $10\times$ compression rate, a 100 Mbps data stream can be reduced to 10 Mbps. Reduction of data rate has a significant improvement to the overall system link.

The receiver sensitivity, the minimum power level required at the receiver in order to detect the signal, is a function of the thermal noise level of the receiver ($k \cdot T \cdot BW$), where k is the Boltzmann constant, T is the system temperature in Kelvin (assumed to be $290K$ for this calculation), bandwidth of the system (B), the noise figure (NF) of the receiver and required signal-to-noise ratio (SNR) for a particular bit error rate and modulation technique. SNR is a function of energy required per bit (E_b), thermal noise in 1 Hz of bandwidth (N_o), system bandwidth (B) and system data rate (R_b) and is given by:

$$SNR = \left(\frac{E_b}{N_o}\right) \cdot \left(\frac{R_b}{B}\right) \quad (9)$$

For an expected bit error rate (BER) of 10^{-6} using Phase Shift Keying (PSK) in a system results in a (E_b/N_o) ratio of 10.5 dB, as described in Fig.3(A). Fig.3(A) also shows that a reduction in data rate results in an improved sensitivity of the receiver, allowing it to receive lower power signals that have been attenuated due to path loss. This benefit reduces the transmitter output power requirements at a given distance.

To visualize the performance improvement in the communications link we use equation (7) to plot the received power versus distance for a given link (implanted transmit antenna to a receiver located several meters away). With assumptions for the transmit and receive antenna RF performance based on prior modeling and simulations we consider two data rates of interest, 10 Mbps and 100 Mbps. From the link budget perspective, the received power meets the sensitivity level requirements of the 100 Mbps system and results in a transmission distance $<2m$, as shown in Figure 3(B). With $10\times$ data compression at 10 Mbps, the same power at the transmitter can extend communication range to 5m. This model assumes a non-line of sight communication channel in the UWB bandwidth of 7500 MHz and bit error rate of 10^{-6} . Also shown in Fig.3(B), with compression, a 10 dB reduction in transmit power can still achieve the desired transmission distance of $<2m$. Furthermore, with reduced power, the SAR margin improves significantly.

In summary, this section establishes that implanted wireless electronics are bounded by FCC limitations of SAR. This limitation bounds the amount of transmitted power available in the implanted system. Also, when considering the UWB spectrum as the communication channel of choice, a variation on the standard Friis equation must be used to extend its use to wide bandwidth communication systems. With this refined version of the Friis equation and various aspects of the communication channel path loss, we have plotted received power for two data rates (10 Mbps and 100 Mbps) to show that transmit power can be reduced by around 10 dB when using compression techniques to reduce data rate.

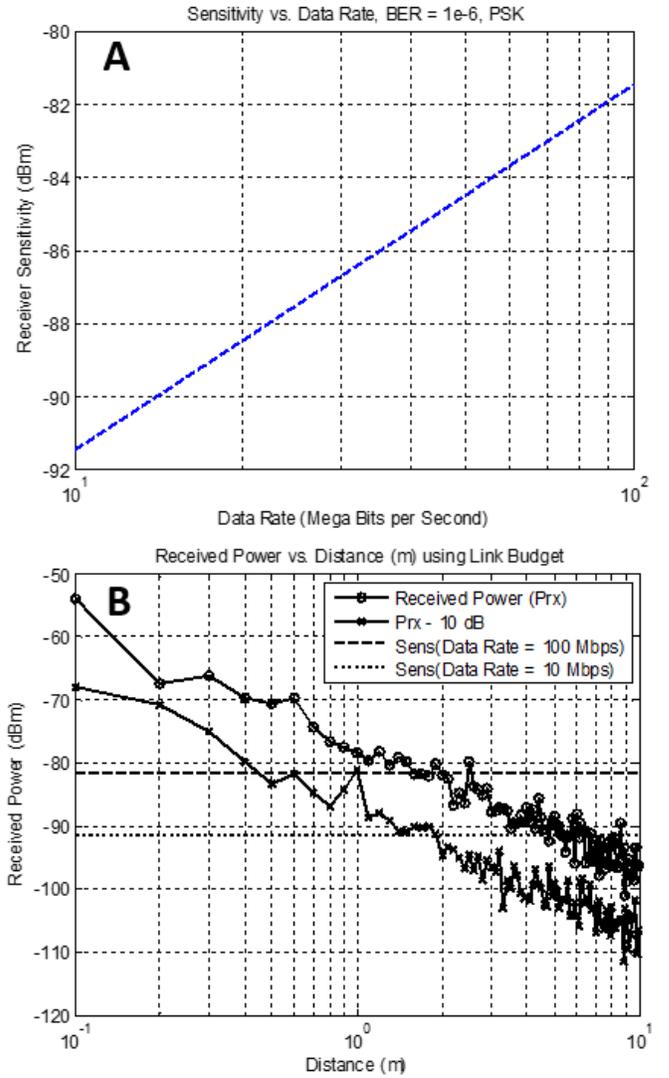


Fig. 3. (A) Receiver Sensitivity vs Data Rate (B) Received Power vs. Transmission Distance

III. PREVIOUS ON-CHIP COMPRESSION SYSTEMS

As demonstrated in previous section, efficient on-chip compression is essential in optimization of communication channel and reducing its power consumption. Previously, on-chip wavelet transform processor was used to first transform the neural signals into their wavelet representation [12] [26]. The processor then applies thresholding which only allows most significant wavelet coefficients to be transmitted off-chip. Although on-chip wavelet transform achieves high compression rate (CR), its implementation requires large hardware resources due to the implementation of DSP operations for the filter banks. Additional memory blocks also need to be inserted between different levels of the wavelet transform for data storage. These additional components increase the area of the chip. Hence, due to large area requirement, wavelet transform is not scalable for high density neural recording systems. A detail analysis on hardware resource of the wavelet transform can be found in [12] and [16].

Many prior multi-electrode array designs also relied on

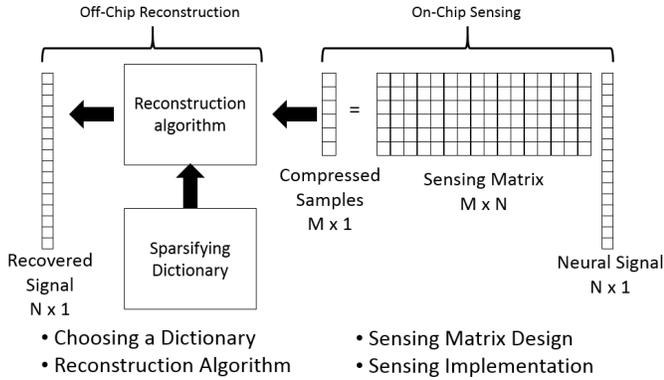


Fig. 4. Overview and considerations of a Compressed Sensing system

spike detection and windowing techniques to reduce transmission bandwidth [11] [13] [14] [27]. After a spike is identified through a threshold crossing detector, a small 1-2 ms long window around each spike is retained. This event based compression method achieves a decent CR for electrodes with sparse neuronal firing rates. When the aggregate firing rate of all detectable neurons is high (e.g. > 150 Hz), however, the CR is greatly reduced. Furthermore, since the original signal is lost forever, there is no way of correcting for any lost spike or a change in threshold. A detail comparison of the previous work can be found in [16].

A number of CS systems were also developed in the past [8] [28] [29] [30]. The advantage of CS system is that the sensing operation can be efficiently implemented on-chip. It also allows user to choose suitable dictionary to recovery signal. These sparsifying dictionaries can be optimized for different tasks such as spike recovery, clustering or classification.

Fig.4 outlines considerations to design a CS system. On-chip wise, the neural signal is compressed through multiplication with sensing matrix, \mathbf{A} , of size $M \times N$. A random Bernoulli sensing matrix consists of random distribution of -1 and 1. It is often the most optimal choice due to the simplicity of implementation. The sensing operation of a random Bernoulli matrix costs less hardware resources than other random matrices (e.g. Gaussian random matrix). It can be implemented efficiently without using any multipliers. Since the matrix can only have two values, each entry in the matrix can be represented with a single bit. Implementing one row of the matrix requires only one accumulator. When the corresponding matrix row entry is a 1, the accumulator adds the signal to the accumulated value. When the corresponding matrix row entry is a 0, the accumulator subtracts the signal from the accumulated value. Additionally, a number of sensing matrix optimization techniques could also be used to improve reconstruction quality, while constraining the on-chip sensing matrix to have only entries of -1 and 1 [31] [32].

Previous work of the on-chip sensing block can be classified into two major categories: analog implementation and digital implementation. Analog implementation performs multiplication and accumulation using analog switches and integrators [28] [29] [30]. Digital implementation moves the multiplication and accumulation step to the digital domain after A/D

conversion [8] [16] [17]. Analog implementation saves power of the ADC, while digital implementation occupies smaller area and less prone to mismatch. Several works have also analyzed both architectures' power efficiency [8] [33]. For high density neural recording probe, the digital implementation is more advantages due to smaller chip area.

During signal recovery, off-chip system has to receive the compressed value (y), and reconstruct the signal (x) through solving a convex optimization problem using Basis Pursuit (BP). BP algorithms are computationally intensive. Thus, for real-time applications, greedy algorithms such as the matching pursuits are preferred over BP [34] [35]. The reconstruction process also utilizes a sparsifying dictionary, \mathbf{D} , where the neural signal, \mathbf{x} , can be represented with only very few coefficients. Choice of sparsifying dictionary is an important step, as CS compression rate and recovery quality is closely related to sparsity. The number of compressed measurement needed for exact reconstruction, M , has to be greater than the theoretical bound [36]:

$$M \geq C \cdot S \cdot \log\left(\frac{L}{S}\right) \quad (10)$$

Where S is the sparsity of the signal in the sparsifying dictionary. L is the number of columns of the sparsifying dictionary and C is an arbitrary constant. Previously, discrete wavelet and Gabor dictionaries are commonly used as sparsifying dictionary. But the compression rate fails to increase above $2 \times$ while guaranteeing acceptable signal reconstruction quality [37].

IV. SIGNAL DEPENDENT COMPRESSED SENSING SYSTEM

In our previous work, we addressed the disadvantage of previous CS systems by introducing a signal dependent sparsifying dictionary [16] [17] [38] [39]. In neural recording experiment, a recording electrode detects neural action potentials (a.k.a spikes) from one or more neurons in its proximity. Every neuron's spike seen at the electrode is unique and is determined by the neuron's morphology and its relative location to the electrode. Based on this property, neuroscientists could separate and label the recorded spikes to corresponding neurons. From a compression point of view, if we know the shape of the action potential then we can train a dictionary to sparsely represent spikes and other signals seen at this particular electrode. Previously we have shown that signal dependent dictionary allows more compact representation of the neural action potential compared to commonly used time-frequency dictionary such as wavelet and Gabor transform. As a result, we can achieve $> 10 \times$ compression rate while guaranteeing signal reconstruction quality and spike clustering accuracy. We demonstrate our approach in this section.

A. Learning a Dictionary

Sparse dictionary Learning is a process where a dictionary is constructed to represent a set of training data using very few dictionary items. That is, given a training dataset

$\mathbf{W} = \{\mathbf{w}_{i=1}^P\}$, where $\mathbf{w}_i \in \mathbf{R}^N$, Dictionary learning finds a dictionary, \mathbf{D} , to minimize this objective function:

$$\begin{aligned} \operatorname{argmin}_{\mathbf{D}, \{\mathbf{v}_i\}_{i=1}^L} \sum_{i=1}^L \|\mathbf{w}_i - \mathbf{D}\mathbf{v}_i\|_2^2 \\ \text{s.t. } \|\mathbf{v}_i\|_0 \leq s_0, 1 \leq i \leq P \end{aligned} \quad (11)$$

where $\mathbf{D} \in \mathbf{R}^{N \times L}$ is the signal dependent dictionary. P denotes the size of the dictionary and $\mathbf{v}_i \in \mathbf{R}^L$ is the sparse vector representing the training data \mathbf{w}_i in \mathbf{D} . s_0 is the bound on the ℓ_0 -norm of the sparse vector.

Given that the neuron registers unique action potential shape at the recording electrode, a good collection of training data needs to include these characteristic shapes. In this example, we use first 8s of extracellular recording from a Hippocampus of a rat as our training data. The rest of the 5-minute data are then used for testing. This recording is carried out using a tetrode experiment setup. We only consider recording from one of the electrode in this example. We refer the reader to our previous work for compression techniques related to multi-electrode recording [17] [39].

From 8s training signal, shown in Fig.5, we acquired 200,000 training data each of size 64 using a moving window with length $N=64$. Each window corresponds to around $2.6ms$ with 25KHz sampling rate, a typical duration of a spike. Some examples of the training data are shown in Fig.5. Some of them are spike signals while others consist of band passed noise, local field potentials (LFP) and spikes from distant neurons.

The training data is used to train an over-complete dictionary, \mathbf{D} , of size 64×5000 . Examples of the dictionary items are also shown in Fig.5. There are many dictionary learning algorithms that can perform this task. The K-SVD algorithm is the most commonly used dictionary algorithm due to its simplicity and fast speed [41]. Therefore, in this example we used K-SVD to train our over-complete dictionary. In our previous work, we have also developed an unsupervised dictionary learning framework that embeds spike spectral clustering and group structures. This algorithm gives the best performance for neural signal reconstruction [39]. The code for this dictionary learning method is available at our lab's homepage [42]

B. On-Chip Sensing

Fig.6 demonstrates the common digital implementation of on-chip sensing operation. The digitized neural signal, \mathbf{x} , of length $N \times 1$ is multiplied by a sensing matrix \mathbf{A} of size $M \times N$, where $M \ll N$. Matrix \mathbf{A} consists of entries either 0 or 1. Each row of multiplication is implemented by a digital accumulator. M rows of accumulator are needed to compute the compressed samples, \mathbf{y} , of size $M \times 1$. In our example, N is 64. M is set by the compression rate (CR), where $M = N/CR$. Hence, with more the compression, less hardware resources are utilized. The digital accumulators can also be implemented efficiently using digital circuit operating at neural signal Nyquist rate of 25 KHz. Our previous chip shows that 25 CS Channels with SPI interface consumes around $0.83\mu W$ at $0.53V_{DD}$ [17].

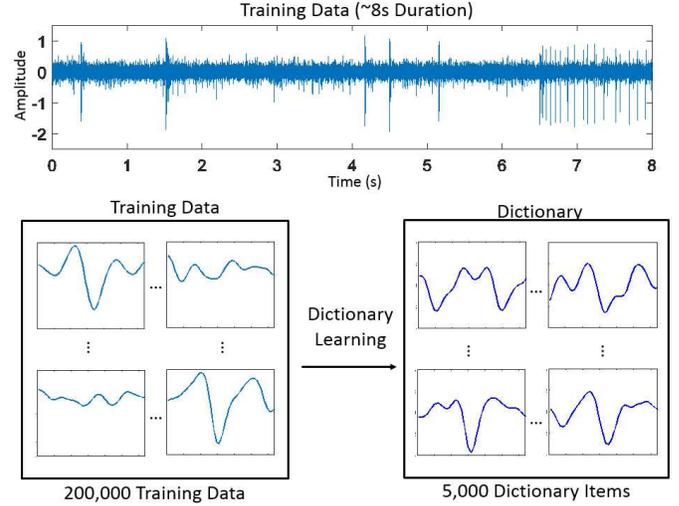


Fig. 5. Dictionary Learning using 8s of training data

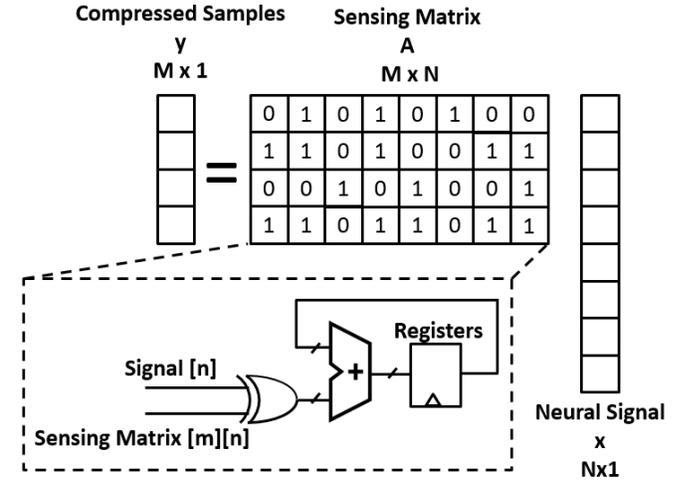


Fig. 6. On-Chip digital sensing implementation

Note that CR is also dependent on number of bits in the digitized samples (\mathbf{x}) and the compressed measurements (\mathbf{y}). Detail studies on reconstruction quality with respect to the number of bits in the compressed measurement can be found in [8], [33] and [43]. These results have shown that the resolution of the compressed measurement does not affect the overall reconstruction quality significantly at low M . In our implementation and analysis in this paper, the number of bits is the same (10 bits) for both the digitized samples and the compressed measurements. In term of implementation, a 14-bit accumulator is used to accumulate the 10 bits sampled values to avoid overflow. During data readout, only the 10 MSBs of the accumulator are transferred out of the chip for recovery.

C. Off-Chip Signal Reconstruction

After receiving compressed samples, \mathbf{y} , at the receiver, we can solve for sparse vector $\hat{\mathbf{v}}$, the approximation to the sparse representation of \mathbf{x} in the sparsifying dictionary, \mathbf{D} :

$$\hat{\mathbf{v}} = \underset{\mathbf{v}}{\operatorname{argmin}} \|\mathbf{v}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{D}\mathbf{v} \quad (12)$$

With $\hat{\mathbf{v}}$, we can then reconstruct the estimate for the neural signal:

$$\hat{\mathbf{x}} = \mathbf{D}\hat{\mathbf{v}} \quad (13)$$

where $\hat{\mathbf{x}}$ is the reconstructed neural signal. Fig.7 illustrates the recovered signal and the original signal. The recovered sparse vector, $\hat{\mathbf{v}}$, and the corresponding dictionary item are also shown in this figure. This is at a compression rate of 10.7x, where only 6 compressed samples are transmitted. Both the spikes and the inter-spike variation are well reconstructed using the trained dictionary.

A number of methods can be used to solve the optimization problem in equation (12) [34] [35] [36]. For real-time reconstruction, matching pursuit methods are preferred [34] [35]. Matching pursuit approximates the solution to equation (12) by solving the following objective function:

$$\hat{\mathbf{v}} = \underset{\mathbf{v}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{D}\mathbf{v}\|_2 \quad \text{s.t.} \quad \|\mathbf{v}\|_0 < s_0 \quad (14)$$

The reason we choose greedy pursuit method rather than convex optimization is to prioritize computational speed for real-time reconstruction. We implemented the Orthogonal Matching Pursuit using MATLAB on an Intel Core-i7 computer with 16G RAM. The reconstruction time for each of the spike window is less than 2ms, less than a typical duration of a spike.

We use the metric Signal to Noise and Distortion Ratio ($SNDR_x$) to measure reconstruction quality. Subscript x denotes that $SNDR$ is calculated with the knowledge of the original signal. It is defined as:

$$SNDR_x = \frac{1}{T} \sum_{i=1}^T 20 \log \frac{\|x\|_2}{\|x_i - \hat{x}_i\|_2} \quad (15)$$

where x_i is the i th window in the neural recording, belonging to a recording with total T windows, and \hat{x}_i is the reconstructed spikes from compressed measurements. The $SNDR_x$ for the signal shown in Fig.7 is 12.5 dB.

D. Reconstruction Quality Evaluation

One of the downsides of previous CS work is that there is no mechanism to measure the reconstruction quality in real-time. This is simply because they have no knowledge of the original signal when the chip only transmits the compressed samples. The end users therefore have no knowledge on how well the recovered signal resembles the uncompressed neural data.

The inability to measure reconstruction quality also affects performance of our signal dependent dictionary CS approach. During a long term recording, electrode shifting may cause the recorded spike shape to change. In other cases, a previously quiet neuron might start firing action potential when they are activated during a certain task in the experiment. A signal dependent dictionary in the CS framework needs to

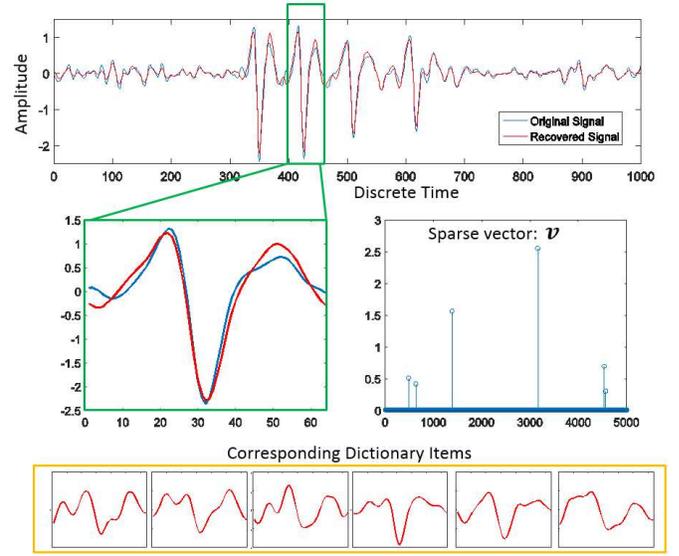


Fig. 7. Signal reconstruction results using over-complete dictionary at CR of 10.7x

be adaptable to accommodate changes in the neural signals that may occur during the recording. Without adaptation, the recovery quality would degrade over time because the learned dictionary can no longer represent spikes sparsely.

To address this challenge, we introduce a mechanism to measure reconstruction quality in real time. Due to the inability to calculate $SNDR_x$, we calculate Signal to Noise and Distortion Ratio measured in compressed domain ($SNDR_y$) as the metric for recovery performance. $SNDR_y$ is defined as:

$$SNDR_y = \frac{1}{T} \sum_{i=1}^T 20 \log \frac{\|y\|_2}{\|y_i - \hat{y}_i\|_2} \quad (16)$$

$$\hat{y}_i = \mathbf{A}\hat{x}_i \quad (17)$$

where y_i is the CS measurement of i th window within a neural recording containing total of T windows, and \hat{y}_i is the CS measurements estimated from the recovered spike \hat{x}_i . When a signal is not well reconstructed, the reconstruction error can also be reflected in the CS measurements after a linear mapping using the sensing matrix, \mathbf{A} .

Intuitively, a dictionary, \mathbf{D} , is trained to represent specific spike shapes. But when spike shapes vary, the signal can no longer be represented sparsely in this dictionary. Therefore, when we try to find the solution to equation (14) using the old dictionary, it is going to find the s_0 coefficients that best represent the signal. But since this new signal can no longer be represented sparsely, s_0 coefficients are no longer enough to represent it. Therefore, a larger error would exist between \mathbf{y} and $\hat{\mathbf{y}}$, leading to reduced $SNDR_y$ measurement. It is also important to know that $SNDR_y$ is an average measurement that computes over T spikes. The larger the T value, the more correlation it has with $SNDR_x$.

To verify that $SNDR_y$ can effectively be used as a metric to measure recovery quality, we must show that these two metrics

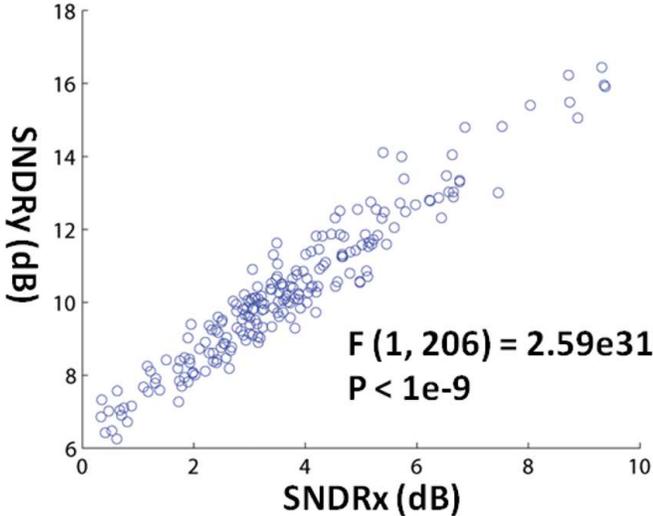


Fig. 8. $SNDR_x$ and $SNDR_y$ values for all the recording electrode over 15 weeks of recordings [17]

are highly correlated. In our previous work, we calculated the $SNDR_x$ and $SNDR_y$ correlation using a 15 week silicon probe chronic recording setup [17]. In this fifteen week long recording experiment, a high density recording array containing 32 electrodes on a single shank ($70\mu m$ wide) is implanted in the thalamus of a rat's brain when it is under anesthesia. We acquired one minute of raw data from the implanted electrodes every week for fifteen weeks. We then calculated $SNDR_x$ and $SNDR_y$ values for every electrode for each week with $CR=8$. Figure 8 is taken from our previous work to demonstrate the result. In Fig.8, a regression analysis between $SNDR_x$ and $SNDR_y$ results in $F(1, 206) = 2.59 \times 10^{31}$ and $P < 10^{-9}$, suggesting a strong linear relationship between two metrics, $SNDR_x$ and $SNDR_y$. Therefore, we could use $SNDR_y$ as a real-time metric to evaluate signal recovery quality.

E. Dictionary Update using Compressed Measurement

Instead of measuring $SNDR_y$ and update the dictionary using raw signal, we can also adapt the dictionary using the compressed measurements \mathbf{y} . The objective function of Dictionary Learning from the compressed measurements can be formulated as,

$$\min \sum_{i=1}^N \left(\frac{1}{2} \|\mathbf{y}_i - \mathbf{A}_i \mathbf{D} \mathbf{v}_i\|_2^2 + \lambda \|\mathbf{v}_i\|_1 \right), \quad (18)$$

where λ is the regularization parameter and \mathbf{v} is the sparse vector. The objective function of dictionary learning is to find the dictionary that can recover the signal well (by minimizing l_2 norm of representation error) and force the representation to be sparse (by regularizing the l_1 norm of the sparse coefficients). Note that different sensing matrices \mathbf{A}_i should be adopted for different training data to capture all the information when projecting the training signals onto one low dimensional subspace.

As pointed out in [45], the sparse coding step could be simply modified by changing the dictionary \mathbf{D} to $\mathbf{A}\mathbf{D}$ and we

just need to derive a dictionary update scheme for equation (18). By rewriting the objective function, we could get the dictionary update as

$$\hat{\mathbf{d}}_k = \left(\sum_{i=1}^N |a_{i,k}|^2 \mathbf{A}_i^\top \mathbf{A}_i \right)^\dagger \sum_{i=1}^N a_{i,k} \mathbf{A}_i \hat{\mathbf{y}}_i^k. \quad (19)$$

and

$$\mathbf{d}_k = \hat{\mathbf{d}}_k / \|\hat{\mathbf{d}}_k\|_2. \quad (20)$$

where \mathbf{d}_k is the k -th dictionary atom of \mathbf{D} , $\mathbf{v}_{i,k}$ is the k -th value of the sparse coefficient vector \mathbf{a}_i , and $\hat{\mathbf{y}}_i^k = \mathbf{y}_i - \mathbf{A}_i \sum_{j \neq k} \mathbf{d}_j \mathbf{v}_{i,j}$. For more details, readers could refer to [45].

We evaluate the performance of Dictionary Learning from Compressed Sensing Measurements (DL-CS) both qualitatively and quantitatively. Using Easy2 and Difficult2 datasets from the Leicester neural signal database [44], we compare the performance of dictionaries learned by Dictionary Learning from full length signal (regular DL), DL-CS and fixed DWT transformation. As seen in Fig. 9, we show the recovery performance of learned dictionaries at different Compression Ratio (CR). It can be seen that regular DL has superior recovery performance (e.g., above 10 dB) and improves dramatically when we acquire more measurements. For DL-CS, the recovery result looks very noisy when the CR is as high as 10 due to the loss of information during the compression process. As CR becomes higher, the recovered signal starts to present the shape of the desired signal and gradually capture the geometric features of the signal while still having some artifacts. Nevertheless, DL-CS consistently outperforms DWT dictionary for its recovery performance. Thus, DL-CS could be used as a viable alternative for training a dictionary with good recovery performance without the need of training the dictionary at full acquisition rate. This enables the potential of having a continuous online dictionary learning system for neural recording. In the future, we will explore the performance DL-CS for classification tasks.

V. CONCLUSION

In this paper, we provide an analysis on the communication constraints for high density neural recording devices. The analysis shows that an efficient on-chip compression is essential to minimize power and heat dissipation at the recording sites. We then show that using signal dependent compressed sensing approach, we can provide around $10\times$ compression to the data rate with very efficient digital implementation. The implementation only requires 6 digital accumulators operating at signal Nyquist frequency of 20-30KHz. Based on our previous implementation, we estimate the power consumption of this system to be $0.2\mu W$. With linear interpolation, the compression system for 1000-electrode system will only consume $0.2mW$.

With $10\times$ compression, for the same 1000-electrode system, the power consumption for a wired data transmission can be reduced from $50mW$ to $5mW$. For a wireless system, reduction from 100Mbps to 10Mbps data rate means we can extend the transmission distance from $2m$ to $5m$ with same power at the transmitter. We can also use 10 dB less transmitter

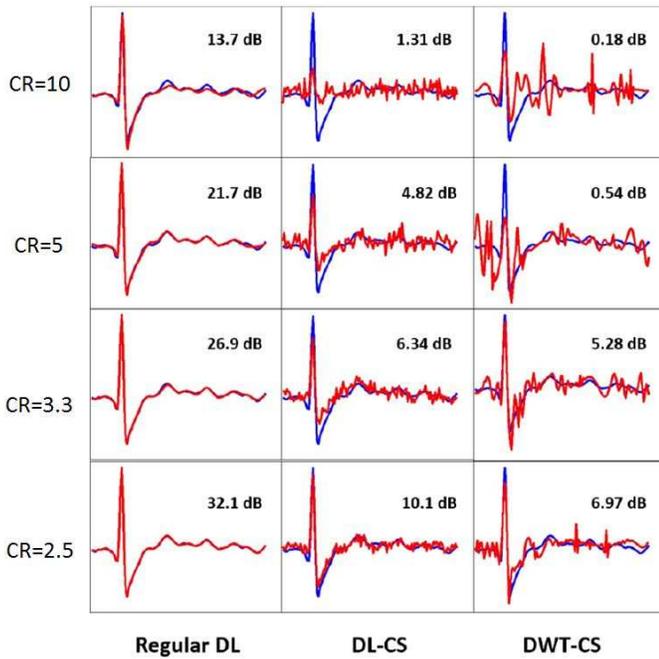


Fig. 9. Recovery results of a single spike using different dictionary choices at different CRs. The recovery results are measured using SNDR (dB). The groundtruth is plotted in blue and the recovered signal is plotted in red.

power to reach the same distance of $2m$. Thus we can limit the radiation injected into the tissue and increase the SAR margin.

Another key feature of our system is the ability to operate in real-time. CS reconstruction is lossy. The reconstruction quality is directly related to compression rate. In our system, real-time sensing, reconstruction and quality measurement give user instant feedback on the quality of the CS reconstruction. This feature allows the user to find the most optimal compression rate for the recording experiment.

To reinforce the real-time feature, we also show that a sparsifying dictionary can be built from the compress measurements directly. This allows the user to construct a dictionary without using raw training data. Therefore, this feature has the possibility to integrate the dictionary learning process with the reconstruction block.

REFERENCES

- [1] Hubel, David H., and Torsten N. Wiesel. Receptive fields of single neurons in the cat's striate cortex. *The Journal of physiology*, 148.3 (1959): 574-591. 2002.
- [2] Maynard, E. M., Nordhausen, C. T., and Normann, R. A. The Utah intracortical electrode array: a recording structure for potential brain-computer interfaces. *Electroencephalography and clinical neurophysiology*, 02.3 (1997): 228-239. 2008
- [3] Lopez, C. M., Andrei, A., Mitra, S., Welkenhuysen, M., Eberle, W., Bartic, C., Puers, P., Yazicioglu, R.F., and Gielen, G. An implantable 455-active-electrode 52-channel CMOS neural probe. *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International*, pp. 288-289
- [4] Shulyzki, R., Abdelhalim, K., Bagheri, A., Salam, M. T., Florez, C. M., Perez Velazquez, J. L., Carlen, P., Genov, R. An implantable 455-active-electrode 52-channel CMOS neural probe. *Biomedical Circuits and Systems, IEEE Transactions on*, 9(1), 2015 34-49.
- [5] C. M. Lopez, S. Mitra, J. Putzeys, B. Raducanu, M. Ballini, A. Andrei, S. Severi, M. Welkenhuysen, C. Van Hoof, S. Musa *et al.*, "A 966-electrode neural probe with 384 configurable channels in 0.13 μ m SOI

- CMOS," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 2016, pp. 392-393.
- [6] Seo, D., Carmena, J. M., Rabaey, J. M., Maharbiz, M. M., and Alon, E. Model validation of untethered, ultrasonic neural dust motes for cortical recording. *Journal of neuroscience methods*, 244(2015), 114-122.
- [7] Khalifa, A., Zhang, J., Leistner, M., and Etienne-Cummings, R. A Compact, Low-Power, Fully Analog Implantable Microstimulator *International Symposium on Circuits and Systems (ISCAS), 2016 IEEE*
- [8] F. Chen, A.P. Chandrakasan, and V.M. Stojanovic. Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors. *Solid-State Circuits, IEEE Journal of*, 47(3):744-756, Nov. 2006.
- [9] Duncan, K., and Etienne-Cummings, R. Selecting a safe power level for an indoor implanted UWB wireless biotelemetry link *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, (pp. 230-233). 2006.
- [10] Kim, S., Normann, R. A., Harrison, R., and Solzbacher, F. Preliminary study of the thermal impact of a microelectrode array implanted in the brain. In *Engineering in Medicine and Biology Society Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, (pp. 2986-2989).
- [11] Mitra, S., Putzeys, J., Lopez, C. M., Pennartz, C. A., and Yazicioglu, R. F. 24 Channel dual-band wireless neural recorder with activity-dependent power consumption *Analog Integrated Circuits and Signal Processing*, 83.3 (2014), 317-329
- [12] K.G. Oweiss, A. Mason, Y. Suhail, A.M. Kamboh, and K.E. Thomson. A scalable wavelet transform vlsi architecture for real-time signal processing in high-density intra-cortical implants. *Circuits and Systems I. IEEE Transactions on*, pages 1266-1278, June. 2007.
- [13] B. Gosselin and M. Sawan. An ultra low-power cmos automatic action potential detector *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, pages 346353, Aug. 2009
- [14] B. Gosselin, A. E. Ayoub, J. F. Roy, M. Sawan, F. Lepore, A. Chaudhuri, and D. Guitton. A mixed-signal multichip neural recording interface with bandwidth reduction *Biomedical Circuits and Systems, IEEE Transactions on*, pages 129141, June. 2009.
- [15] Zhang, J., Suo, Y., Mitra, S., Chin, S. P., Tran, T. D., Yazicioglu, F., and Etienne-Cummings, R. Reconstruction of neural action potentials using signal dependent sparse representations. *IEEE International Symposium on Circuits and Systems*, May 19-24. 2001.
- [16] Zhang, J., Suo, Y., Mitra, S., Chin, S. P., Hsiao, S., Yazicioglu, R. F., Tran, T. D., Etienne-Cummings, R., An efficient and compact compressed sensing microsystem for implantable neural recordings. *Biomedical Circuits and Systems, IEEE Transactions on*, 2014, 8(4), 485-496
- [17] Zhang, J., Mitra, S., Suo, Y., et al, A closed-loop compressive-sensing-based neural recording system *Journal of neural engineering*, 2015, 12(3), 036005
- [18] Chandrakasan, A. P., Sheng, S., and Brodersen, R. W. Low-power CMOS digital design *IEICE Transactions on Electronics*, 1992, 75(4), 371-382.
- [19] Borton, D. A., Yin, M., Aceros, J., and Nurmikko, A. An implantable wireless neural interface for recording cortical circuit dynamics in moving primates. *Journal of neural engineering*, 2013, 10(2), 026010.
- [20] Gao, H., Walker, R. M., Nuyujukian, P., Makinwa, K. A., Shenoy, K. V., Murmann, B., and Meng, T. H. HermesE: A 96-channel full data rate direct neural interface in 0.13um CMOS. *Solid-State Circuits, IEEE Journal of*, 2012, 47(4), 1043-1055.
- [21] Ebrazeh, A., and Mohseni, P. 30 pJ/b, 67 Mbps, Centimeter-to-Meter Range Data Telemetry With an IR-UWB Wireless Link. *Biomedical Circuits and Systems, IEEE Transactions on*, 2014, 9(3), 362-369
- [22] Elzeftawi, M., and Theogarajan, L. A 10pJ/bit 135Mbps IR-UWB transmitter using pulse position modulation and with on-chip LDO regulator in 0.13m CMOS for biomedical implants. *Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless), 2013 IEEE Topical Conference on*, 2013, 37-39
- [23] K. Chul and S. Nooshabadi, A 200Mbps, 0.66nJ/b DTR UWB receiver for high data rate wireless biotelemetry applications. *Biomedical Circuits and Systems Conference (BioCAS), 2010 IEEE*, 2013, 21-24
- [24] Fields, R. E. Evaluating Compliance with FCC Guidelines for Human Exposure to Radiofrequency Electromagnetic Fields. , 1997
- [25] Proakis, J., *Digital Communications*, 2nd edition , New York: McGraw-Hill, 1989.
- [26] Kamboh, A. M., Oweiss, K. G., and Mason, A. J. Resource constrained VLSI architecture for implantable neural data compression systems. *IEEE International Symposium on Circuits and Systems (ISCAS), 2009*, pp. 1481-1484

- [27] Chae, M., Liu, W., Yang, Z., Chen, T., Kim, J., Sivaprakasam, M., and Yuce, M. A 128-channel 6mW wireless neural recording ic with on-the-fly spike sorting and uwb transmitter. *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, pp. 146-603
- [28] Pareschi, F., Albertini, P., Frattini, G., Mangia, M., Rovatti, R., and Setti, G. Hardware-Algorithms Co-Design and Implementation of an Analog-to-Information Converter for Biosignals Based on Compressed Sensing. *Biomedical Circuits and Systems, IEEE Transactions on*, 2015
- [29] Shoaran, M., Kamal, M. H., Pollo, C., Vanderghenst, P., and Schmid, A. Compact low-power cortical recording architecture for compressive multichannel data acquisition. *Biomedical Circuits and Systems, IEEE Transactions on*, 2015 8(6), 857-870.
- [30] Gangopadhyay, D., Allstot, E. G., Dixon, A. M., Natarajan, K., Gupta, S., and Allstot, D. J. Compressed sensing analog front-end for bio-sensor applications. *Solid-State Circuits, IEEE Journal of*, 2014 49(2), 426-438.
- [31] Y. Suo, J. Zhang, T. Xiong, P. S. Chin, R. Etienne-Cummings, and T. D. Tran, "Energy-efficient multi-mode compressed sensing system for implantable neural recordings," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 8, no. 5, pp. 648-659, 2014.
- [32] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 59, no. 5, pp. 1001-1014, 2012.
- [33] D. E. Bellasi and L. Benini, "Energy-efficiency analysis of analog and digital compressive sensing in wireless sensors," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 62, no. 11, pp. 2718-2729, 2015.
- [34] Tropp, J., and Gilbert, A. C. Signal recovery from random measurements via orthogonal matching pursuit. *Information Theory, IEEE Transactions on*, 2007 53(12), 4655-4666.
- [35] Needell, D., and Tropp, J. A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis and Computational Harmonic Analysis*, 2009 26(3), 301-321.
- [36] E. Candes, Romberg J, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, pages 489-509, Feb. 2006
- [37] Bulach, C., Bihl, U., and Ortmanns, M. Evaluation study of compressed sensing for neural spike recordings. *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* pp. 3507-3510
- [38] Suo, Y., Zhang, J., Xiong, T., Chin, P. S., Etienne-Cummings, R., and Tran, T. D. Energy-Efficient Multi-Mode Compressed Sensing System for Implantable Neural Recordings. *Biomedical Circuits and Systems, IEEE Transactions on* 2014 8(5), 648-659.
- [39] Xiong, T., Zhang, J., Suo, Y., Tran, D. N., Etienne-Cummings, R., Chin, S., and Tran, T. D. An unsupervised dictionary learning algorithm for neural recordings. *Circuits and Systems (ISCAS), 2015 IEEE International Symposium on* pp. 1010-1013
- [40] Henze, D. A., Borhegyi, Z., Csicsvari, J., Mamiya, A., Harris, K. D., and Buzsáki, G. Intracellular features predicted by extracellular recordings in the hippocampus in vivo. *Journal of neurophysiology* 2000 pp. 84(1), 390-400.
- [41] M. Aharon, M. Elad, and A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. on Signal Processing*, 54:4311-4322, Nov. 2006.
- [42] <http://engineering.jhu.edu/csms/closed-loop-compressed-sensing-system-neural-recording/>
- [43] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A case study in low-complexity ecg signal encoding: How compressing is compressed sensing?" *Signal Processing Letters, IEEE*, vol. 22, no. 10, pp. 1743-1747, 2015.
- [44] Quiroga, R. Quian, Zoltan Nadasdy, and Yoram Ben-Shaul. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural computation*, 16.8 (2004): 1661-1687.
- [45] Studer, C., and Baraniuk, R. G. Dictionary learning from sparsely corrupted or compressed signals. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3341-3344.

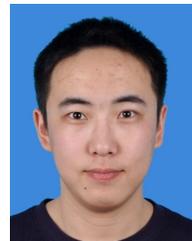


Jie Zhang received his B.Sc. degree in Electrical Engineering, in 2010, from The Johns Hopkins University, Baltimore, MD, where he is currently pursuing the Ph.D. degree in Electrical Engineering. He has been an International Scholar with the Ultra Low Power - Biomedical Circuit group at IMEC, Belgium, from October 2011 to July 2012. His research focuses on compressive sensing, computational imaging, image sensors and camera system, analog and mixed signal circuits design for low power biomedical applications



Kerron Duncan received his B.S. and M.S degree in Electrical Engineering from Morgan State University in 2000 and 2003, respectively. He joined Northrop Grumman ES in 2001 and is currently a Power Systems Technical Lead and supports many airborne Radar systems. He is also pursuing a PhD at Johns Hopkins University in the area of wireless biotelemetry trade space modeling and simulation and is interested in sub-system architecture models which include "push factors" for various subsystems such that their interdependences (cost, size, weight

and performance) with respect to performance can be used in model based engineering efforts and trades.



Yuanming Suo received the B.S. degree in Electrical Information Science and Technology from Sun Yat-sen University in 2004. From 2004 to 2007, he worked at China Netcom as a computer network engineer. He received the M.S.E in Electrical and Computer Engineering from University of Alabama in Huntsville in 2009, and Ph.D. in Electrical and Computer Engineering from Johns Hopkins University in 2015. His research focuses on developing algorithms for dictionary learning, matrix completion and representations for high-dimensional data.

His other research interests include image analysis in medical/remote sensing applications and compressed sensing.



Tao Xiong received the B.S. degrees in microelectronics and economics from Peking University in 2012 and the M.S.E degree in electrical and computer engineering from Johns Hopkins University in 2014. From 2009 to 2012, he was a research assistant at National Key Laboratory of Science and Technology on Micro and Nano Fabrication of Peking University. Since 2013, he has been a research assistant at Computational Sensory-Motor Systems Laboratory and Digital Signal Processing Laboratory at Johns Hopkins University. Currently,

he is working toward the M.S.E degree. in applied mathematics and statistics and Ph.D. degree in electrical and computer engineering, both at Johns Hopkins University. His research is focused on the deep learning, computer vision, compressive sensing, and VLSI design for biomedical applications.ns.



Srinjoy Mitra received his B.S. degree in physics and electronics from Calcutta, India and his M.Tech. degree from the Indian Institute of Technology, Bombay, India. After spending a short time in the electronics industry he received his Ph.D. from the Institute of Neuroinformatics, ETH Zurich in 2004. Between 2008 and 2010 he worked as a post-doctoral researcher at Johns Hopkins University, Baltimore, USA. He then joined the medical electronics team at IMEC, Belgium and worked there as a senior scientist until early January 2016. At IMEC

he had taken up lead roles in various industrial and public funded projects primarily related to bio-potential recording. For the last few years Dr. Mitra led multiple projects on neural implants for central and peripheral nervous system. He has recently joined the University of Glasgow, Glasgow, UK as a lecturer at the Biomedical Engineering Division. His primary research interest is in designing novel mixed-signal CMOS circuits for advancement in medical and neural electronics.



Ralph Etienne-Cummings received his B. Sc. in physics, 1988, from Lincoln University, Pennsylvania. He completed his M.S.E.E. and Ph.D. in electrical engineering at the University of Pennsylvania in December 1991 and 1994, respectively. Currently, Dr. Etienne-Cummings is a professor of electrical and computer engineering, and computer science at Johns Hopkins University (JHU). He was the founding Director of the Institute of Neuromorphic Engineering. He has served as Chairman of various IEEE Circuits and Systems (CAS) Technical Committees and was elected as a member of CAS Board of Governors. He also

serves on numerous editorial boards. He is the recipient of the NSF Career and Office of Naval Research Young Investigator Program Awards. He was a Visiting African Fellow at U. Cape Town, Fulbright Fellowship Grantee, Eminent Visiting Scholar at U. Western Sydney. He has won numerous IEEE and non-IEEE best paper awards at conferences and for Journal articles. He was recognized as a ScienceMaker, an African American history archive and was selected for the "Indispensable Blacks at Johns Hopkins" exhibit in 2016. He has published over 220 peer reviewed article and holds numerous patents on his work.



Trac D. Tran (S'94-M'98-SM08) received the B.S. and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, in 1993 and 1994, respectively, and the Ph.D. degree from the University of Wisconsin, Madison, in 1998, all in Electrical Engineering. In July of 1998, Dr. Tran joined the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD, where he currently holds the rank of Professor. His research interests are in the field of digital signal processing, particularly in sparse representation,

sparse recovery, sampling, multi-rate systems, filter banks, transforms, wavelets, and their applications in signal analysis, compression, processing, and communications. His pioneering research on integer-coefficient transforms and pre-/post-filtering operators has been adopted as critical components of Microsoft Windows Media Video 9 and JPEG XR the latest international still-image compression standard ISO/IEC 29199-2.

Dr. Tran was the co-director (with Prof. J. L. Prince) of the 33rd Annual Conference on Information Sciences and Systems (CISS'99), Baltimore, MD, in March 1999. In the summer of 2002, he was an ASEE/ONR Summer Faculty Research Fellow at the Naval Air Warfare Center Weapons Division (NAWCWD) at China Lake, California. He is currently a regular consultant for the U.S. Army Research Laboratory in Adelphi, Maryland. Dr. Tran has served as Associate Editor of the IEEE Transactions on Signal Processing as well as IEEE Transactions on Image Processing. He was a former member of the IEEE Technical Committee on Signal Processing Theory and Methods (SPTM TC) and is a current member of the IEEE Image Video and Multidimensional Signal Processing (IVMSP) Technical Committee.

Dr. Tran received the NSF CAREER award in 2001, the William H. Huggins Excellence in Teaching Award from The Johns Hopkins University in 2007, and the Capers and Marion McDonald Award for Excellence in Mentoring and Advising in 2009.