Proceedings of the 4th International
IEEE EMBS Conference on Neural Engineering
Antalya, Turkey, April 29 - May 2, 2009

A Highly Modular, Wireless, Implantable Interface to the Cortex

Faisal T. Abu-Nimeh[a], Awais Kamboh[a], Mehdi Aghagolzadeh[a], Uei-Ming Jow[c], Andrew Mason[a], Maysam Ghovanloo[c] and Karim Oweiss[a,b]

[a] Dept. of Electrical and Computer Engineering
[b] Neuroscience Program
Michigan State University, East Lansing, MI, USA

[c] School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA, USA

Email: {asd1815, kambohaw, aghagolz, mason, koweiss}@msu.edu , {jow0209, mghovan}@gatech.edu

Abstract— **The development of advanced neuroprosthetic systems and brain-machine interfaces for high-capacity, real-time, bi-directional communication with the nervous system is a major challenge to the emerging neural engineering discipline. In this paper, we summarize our first preliminary report on the design of a highly modular, wireless, adaptive, implantable large-scale interface to the cortex designed exclusively to permit faithful transmission of neural activity from high-density microelectrode array recordings to the outside world. The system is expected to augment the space of experimental design needed to improve our understanding of the nervous system functionality in freely behaving subjects interacting naturally with their surroundings. It will further accelerate the deployment of viable Brain Machine Interface technology in clinical applications.**

Keywords-component; Brain Machine Interface, neural signal processing, compressed sensing, low power implants, spike trains

I. Introduction

An essential step towards understanding how the brain orchestrates information processing at the cellular and population levels is to simultaneously observe the spiking activity of cortical neurons that mediate perception, learning, sensory and motor processing. Implantable high-density microelectrode arrays (HDMEAs) have enabled scrutinizing this activity at an unprecedented scale, and greatly accelerated our ability to monitor functional alterations of neural circuits in awake, behaving subjects. The concomitant neurobiological discoveries have paved the way for these devices to become an essential constituent in neuroprosthetic devices and Brain Machine Interfaces to provide assistive technology to people with severe disability [1].

Engineering cortically-controlled BMI systems is a daunting task. Aside from biocompatibility of the implanted probes, there are numerous challenges to overcome for these systems to be clinically viable. This paper primarily addresses the following question: how to engineer a wireless, adaptive and distributed interface to the cortex to simultaneously accommodate large data throughput from multiple areas in the cortex without compromising critical information in the neural activity? While the potential research and clinical applications of such system are countless, an immediate outcome is the ability to characterize functional connectivity among neuronal populations recorded from multiple areas in the cortex at a scale not possible with present functional brain imaging technologies such as fMRI.

In our extensive body of prior work, we have proposed multiple approaches designed to overcome the limitations of current neural data recording systems and accelerate their deployment from bench to bedside. In particular, we demonstrated that neural data denoising and compression [2], spike detection and sorting [3, 4], and firing rate estimation of individual neurons [5] can be all performed in a unified framework. The theoretical foundation of this framework builds heavily on the theory of *compressed sensing* that received significant attention in the signal processing community over the last few years [5-7]. It mainly exploits sparse representation of the recorded data prior to telemetry transmission, and as a result, significantly reduces the data bandwidth by tens of orders of magnitude. We have further reported on the suitability of computing this representation within the resource-constrained environment of an implantable system [8]. Herein, we briefly report on our ongoing work to fully design and build a complete system that integrates all the above functionality into a highly modular design that paves the way for deploying the next generation of BMI systems to clinical applications.

II. Theory

A. System Architecture

Figure 1 illustrates a conceptual representation of the Multiscale intra-Cortical Neural Interface System (MiCNIS®). The system design is highly modular and can be arbitrarily scaled depending on the number of electrode arrays, array size, number of brain areas of interest, the structure of each brain area, and the desired cortical response characteristics. Every implanted Neural Interface Node (NIN) is hardwired to either 1x32, 2x16, or 4x8 electrode arrays. It will process and wirelessly communicate (over a few millimeters) the 32-channel data to a Manager Interface Module (MIM) fixated extra-cranially. The MIM manages power and data transmission to and from NINs and features the complementary signal processing to what has already been achieved on each NIN [details below]. The MIM will also wirelessly communicate the information over a longer range to a central base station (CBS) equipped with a graphical user interface (GUI) to control the entire system operation and perform more advanced data analysis, if needed.

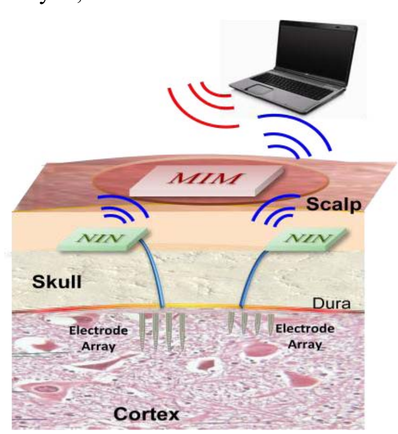

Figure 1. Conceptual representation of MiCNIS®.
NIN = Neural Interface Node; MIM = Manager Interface Module.

B. Neural Interface Node

The main purpose of each NIN is to acquire, filter, amplify, digitize, denoise and compress the neural data from 32 channels such that all the information in the neural signals is kept intact. It transmits the compressed data wirelessly to the extracranial MIM. In addition, it receives commands from MIMs to control the data acquisition modes and compression parameters as needed. The NIN is inductively powered from the external MIM, so no battery is required which further saves substantial space.

Figure 2 shows the block diagram of the NIN. It has three distinct hardware blocks, the analog front-end, the neuro-signal processor, and the power-data transceiver. The analog front end is a modified version of the amplifiers designed in [9] to include a frequency range relevant to the spike trains (300 Hz- 8 kHz) as well as Local Filed Potentials (LFPs) (1 Hz-100 Hz). We use a channel selection mechanism to signal the NIN to only keep the amplifiers of active channels running. With a gain of 59.5dB for 32 channels, current amplifiers were analyzed to consume 2.4 mW of power and 3.3 mm^2 of area in 0.5 μm process. For the ADC, a set of successive-approximation based ADCs with 10-bit resolution was designed. This multi-channel ADC converts the data in parallel from 32 channels and has been set up to output data in a time-multiplexed manner on a single data bus thereby eliminating the need for a separate multiplexer circuit. These ADCs are designed to interact with a global controller which has the capability to control the sampling rate of the converters based on the commands received from the MIM (e.g. spikes or LFPs). The resolution of the output channels can thus be increased to maximally utilize the available bandwidth and to provide higher resolution spikes to improve the subsequent signal processing. For 32 channels at sampling rate of 25 kHz, this block consumes ~0.63 mW of power and ~0.77 mm^2 of area in 0.5 μm CMOS process. The total NIN Analog Front end will consume ~3 mW of power, ~4 mm^2 of area.

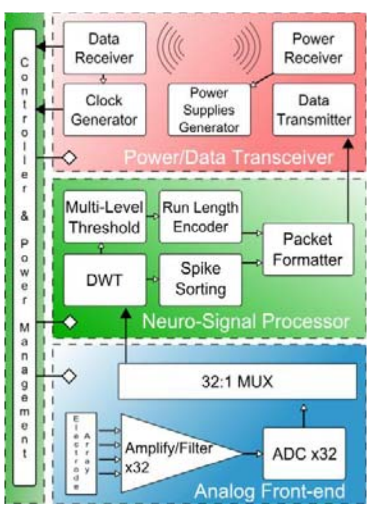

Figure 2: System diagram for the NIN

The Neurosignal processor computes 4 levels of Discrete Wavelet Transform (DWT) for 32-channel using the symmlet4 basis [8]. The module operates on the data instantaneously in a streaming fashion so there is no need for incoming data buffering. The resource requirements for the DWT block scale linearly with the number of channels and levels of decomposition [8]. These are dynamically controlled using command words from the global controller, which in turn can be fully programmed by the MIM.

Multi-level thresholding is required to fulfill multiple tasks at the DWT output. The data are compared against a scale-dependent set of thresholds programmable through the global controller. This feature allows adapting the NIN processing capability to SNR variability and recording conditions (e.g. electrode encapsulation, variable day-to-day neural yield). It further allows sorting spike trains in real time by sensing the most significant DWT coefficients in a given event. If a coefficient magnitude is sufficiently large for a given unit compared to those belonging to other units in a given DWT level, this threshold setting automatically becomes a sensing threshold for that unit and passes only time stamps for unit events surpassing the threshold [5].

The Controller has a Spatial Subspace Calculator block that exploits the correlation across channels to strip the redundancy in the data to improve denoising and compression. In addition, it acts as a channel selector to program the NIN to process data only from active channels where neural activity is present. The proposed block is much more efficient in extracting the spatial distribution of each event through the interleaved DWT computations across channels and incurs a very minimal extra hardware cost [8].

The data stream is further Run length encoded (RLE) at a bit-by-bit level. RLE is best suited for data with long repetitive strings of values. In addition, it is very conservative in the required resource as it only requires a parallel to serial converter and a counter. Because we expect long strings of zeros at the output of the threshold block due to the sparse nature of neural spike train data, RLE is a good lossless compression choice as it approaches the performance of near-optimal algorithms when given very long repetitive sequences. This block has a relatively negligible area and power consumption.

C. Wireless Operation

Wireless operation of the system is fundamental, particularly in clinical applications to reduce the risk of infection and patient discomfort, which may result from transcutaneous wires breaching the skin. This requirement, however, adds to the complexity of the system and constitutes a bottleneck in design of high performance implantable systems for applications such as auditory prostheses, retinal implants, and cortical brain-machine interfaces. Nevertheless, the "smart" and programmable processing capability of the NIN enables exceedingly large compression of the data throughout, which in turn shrinks to a large degree the design of the RF circuitry.

A short range wideband wireless link is required to transfer the compressed neural information from the implantable NINs to the external MIM across the scalp which is only a few millimeters thick. Due to conflicting requirements between power and data carriers, we have utilized a separate pair of coils for data transmission. We considered two types of data coils. First, a pair of vertical coils wound across the diameter of the power PSCs, which symmetry and orthogonal magnetic fields would minimize the interference with power carrier. Second, a pair of planar 8-figure coils, in which the electromotive force (EMF) induced from the power carrier in one loop opposes the same in the other loop. Therefore, the total EMF resulted from the power carrier interference, when the coils are perfectly aligned, can theoretically be very small [10]. To record the neural response and close any local or global control feedback loop, a wideband back telemetry link is required. We have chosen impulse radio ultra-wideband (IR-UWB) communication for this purpose. Hence, we have added a pair of miniature spiral UWB antennas to the wireless link for short range back telemetry across the skin.

We have considered the most energy efficient solution for both data and power transmission through near-field inductive links. The 3 pairs of coils and antennas are dedicated to power transmission, forward control data from the MIM to NIN, and back telemetry biological information from the NIN to MIM. For efficient power transmission from the MIM to one or more NIN modules, we will be using spiral planar coils with optimized geometries in order to maximize the coupling coefficient (k) between power coils, L1 and L2 as illustrated in Figure 3 [11]. We kept the power carrier frequency as low as possible (below 5 MHz) to minimize the skin effect in the coils and power absorption in the surrounding tissue.

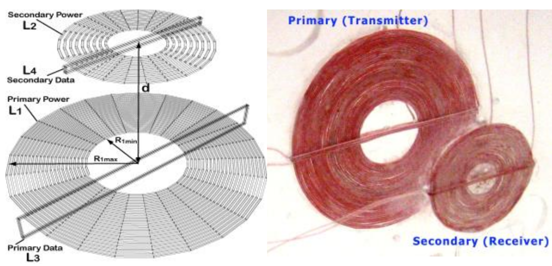

Figure 3: 3D rendering of the data and power transmission coils and fabricated prototypes using thin Litz wire (∅ = 20mm).

D. Manager Interface Module

All NINs are going to be wirelessly powered and controlled through an inductive link, established between the MIM and one or more NINs across the skin barrier. This will also result in a system architecture that is far more acceptable for chronic human BMI applications. The MIM system diagram is shown in Figure 4. The module's size constraints are not as restricted as the NINs. Therefore, the module will be battery powered. This battery, which will be the unique source of energy for the MIM and all the NINs that it communicates with, should last at least over the duration of the experiment and recharged afterwards. It should also be small and light enough to be easily carried by the animal. The MIM features an DWT block that is flexible to either perform an *extended* DWT (EDWT) for firing rate estimation [5] or an *expanded* DWT in case the MASSIT spike sorter is desired [4]. A longer range wireless link is required between the MIM and the external computing base station (CBS). This link will be bidirectional to permit the user to wirelessly adjust system parameters on the MIM and NIN modules on the fly.

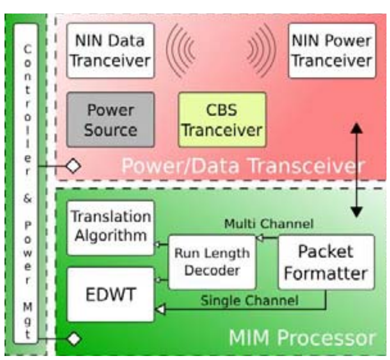

Figure 4: System diagram for the MIM

The MIM side transceiver is controlled by the MIM microcontroller while the CBS side is controlled through a high-speed USB2 interface. The MIM transceiver output is connected to a miniature omnidirectional patch antenna, which can be implemented on the MIM PCB. The planar structure of L1 and L2 power coils makes them an ideal choice for integration within the printed circuit (PCB) platforms of the NIN and MIM modules without requiring any off-chip components other than tuning capacitors

III. RESULTS

In this section, we report some preliminary results on the design and functionality of some of the system's components. We first demonstrate -from a signal processing perspective- that the sparse representation of neural signals with the DWT preserves all the information in the data to perform spike sorting later, if needed. Alternatively, we demonstrate that this sparse representation is even better suited to perform the spike sorting task. This is indicated by the superior spike class separability in figure 5 compared to that obtained through the traditional time domain/PCA spike class separability. Further details on how this performance was obtained are in [5]. This feature enables the system to directly perform firing rate estimation of individual neurons by extending the DWT formatted data as output from NIN to levels where the interspike interval becomes comparable to the time constant of the wavelet basis [5,7].

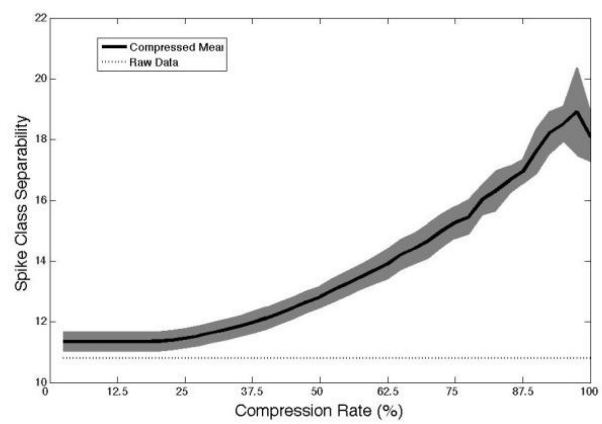

Figure 5: Spike class separability vs. compression rate [5].

Therefore, the system not only overcomes the severe bandwidth limitations of a wireless implantable system, but also enables adequate estimation of neuronal firing rates without the need to decompress, reconstruct, and sort the spikes *off-chip* in the traditional sense. The DWT block has been fabricated and the 3mm x 3mm chip is shown in Fig. 6.

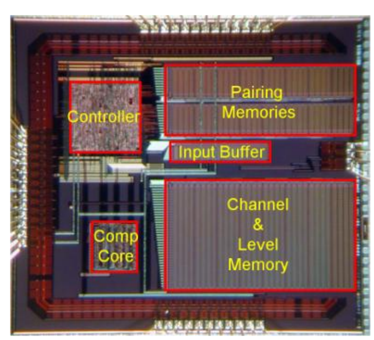

Figure 6: DWT chip fabricated in 0.5μm technology

The controller was synthesized using OSU's Standard Cells

Library, and all other blocks were custom designed for low power and low area. The active components of the prototype 32-channel, 4-level DWT implementation occupy roughly 3.84 mm^2. The layout for the threshold and RLE blocks requires about 0.95 mm^2 of area. Thus the combined compression system is expected to require approximately 5.75 mm^2 including global routing for a 0.5µm process, while consuming only 3mW of power. Figure 7 shows the resource allocation for the NIN hardware that conforms very closely to our measured size and power estimates. The Neurosignal processor that has been fully designed constitutes the largest portion of the module. We expect the entire module to fit within a 10 mm x 10 mm die size.

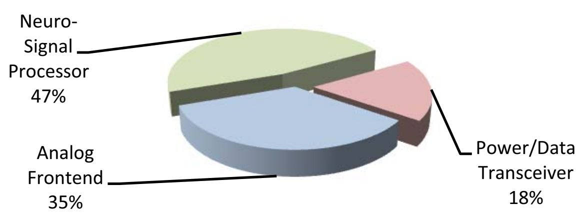

Figure 7: Area-Power product for the NIN module

For the wireless RF, Table I summarizes the set of design constraints that we considered to optimize the geometries of the prototype power transmission PSCs to be fabricated using 1-oz copper on FR4 substrates using a standard PCB fabrication process.

TABLE I
DESIGN CONSTRAINS IMPOSED BY APPLICATION AND FABRICATION PROCESS

Parameter	Symbol	Design Value
Implanted PSC outer diameter	d_{o2}	10 mm
PSCs relative distance	D	10 mm
Link operating frequency	f	5, 10, 13.56 MHz
Secondary nominal loading	R_L	500 Ω
Minimum conductor width	w_{min}	150 µm
Minimum conductor spacing	s_{min}	150 µm
Conductor thickness	t_c	38 µm
Substrate thickness	t_s	1.5 mm
Substrate dielectric constant	ε_{rs}	FR4, 4.4

Table II depicts the optimized geometries for the design requirements in Table I. It can be seen that parameters such as conductor width (w) and number of turns (n) are highly dependent on f, while the external PSC diameter (d_{o1}) is mostly affected by the PSCs relative distance (D) and the outer diameter of the implanted PSC (d_{o2}), which are kept constant in all three cases.

TABLE II
OPTIMIZED POWER PSC GEOMETRIES AND CHARACTERISTICS

Parameter	Set1		Set2		Set3	
Frequency	5 MHz		10 MHz		13.56 MHz	
Name	PSC11	PSC21	PSC21	PSC22	PSC31	PSC32
d_o (mm)	75	10	77	10	79	10
d_i (mm)	9.30	2.56	15.2	2.60	11.2	2.96
φ (fill factor)	0.779	0.592	0.670	0.588	0.751	0.543
n (turns)	9	12	6	10	6	8
w (µm)	3500	160	5000	220	5500	290
s (µm)	150	150	150	150	150	150
L (µH)	2.80	1.01	1.40	0.72	1.22	0.49
Q	76	22	99	34	102	36
Efficiency (%)	42.1		52.0		56.65	

IV. CONCLUSION

State-of-the-art neural recording systems simply negate the unique advantages and hinder the impending clinical application of HDMEA technology because of their inability to handle large data throughputs; their inability to provide continuous recording from freely-behaving, untethered subjects; and their inability to extract the critical biological information in real time with ultra-low power, small size electronics to answer persistent neuroscience questions.

In this paper, we have summarized our first report on a highly modular, wireless, multiscale neural interface system exclusively designed to overcome the aforementioned limitations. The system is highly programmable to allow the user to select an active number of channels, type of signal modality sought and type of information being extracted. It operates in real time and transmits all the critical biological information with highly efficient, reliable and miniaturized circuitry suitable to be fully implanted in freely behaving subjects. Taken together, it opens a whole new avenue of experimental investigations and accelerates the deployment of state of the art BMI systems to clinical applications.

V. ACKNOWLEDGEMENT

This work was supported by NIH grants # NS047516 and # NS062031.

REFERENCES

[1] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature,* vol. 442, pp. 164-171, 2006.

[2] K. G. Oweiss, "A systems approach for data compression and latency reduction in cortically controlled brain machine interfaces," *IEEE Transactions on Biomedical Engineering,* vol. 53, pp. 1364–1377, 2006.

[3] K. G. Oweiss, "Multiresolution Analysis of Multichannel Neural recordings in the Context of Signal Detection, Estimation, Classification and Noise Suppression." PhD dissertation, University of Michigan, 2002.

[4] K. G. Oweiss and D. J. Anderson, "Tracking Signal Subspace Invariance for Blind Separation and Classification of Nonorthogonal Sources in Correlated Noise," *EURASIP Journal on Advances in Signal Processing,* vol. 2007, 2007.

[5] M. Aghagolzadeh and K. Oweiss, "Compressed and Distributed Sensing of Neuronal Activity for Real Time Spike Train Decoding," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* 2009, in press.

[6] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory,* vol. 52, pp. 1289-1306, 2006.

[7] K. G. Oweiss, "Compressed and Distributed Sensing of Multivariate Neural Point Processes," *IEEE International Conference on Acoustics, Speech and Signal Processing,* vol. 2, 15-20 April 2007.

[8] 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," *IEEE Transactions on Circuits and Systems I,* vol. 54, pp. 1266-1278, 2007.

[9] R. R. Harrison, P. T. Watkins, R. J. Kier, and F. Solzbacher, "A Low-Power Integrated Circuit for a Wireless 100-Electrode Neural Recording System," *IEEE J. of Solid State Circuits,* vol. 42, p. 123-133, 2007.

[10] M. Ghovanloo and S. Atluri, "A Wide-Band Power-Efficient Inductive Wireless Link for Implantable Microelectronic Devices Using Multiple Carriers," *IEEE Transactions on Circuits and Systems I,* vol. 54, pp. 2211-2221, 2007.

[11] M. Ghovanloo and K. Najafi, "A wideband frequency-shift keying wireless link for inductively powered biomedical implants," *Circuits and Systems I: Regular Papers, IEEE Transactions on Circuits and Systems I,* vol. 51, pp. 2374-2383, 2004.