An Implantable Neuromorphic Sensing System Featuring Near-sensor Computation and Send-on-Delta Transmission for Wireless Neural Sensing of Peripheral Nerves

Yuming He, *Student Member, IEEE,* Federico Corradi, Chengyao Shi, *Student Member, IEEE,* Stan van der Ven, Martijn Timmermans, Jan Stuijt, Pieter Harpe, *Member, IEEE,* Lucas Lindeboom, Evelien Hermeling, Geert Langereis, Elisabetta Chicca, *Senior Member, IEEE,* Yao-Hong Liu, *Senior Member, IEEE*

***Abstract*— This paper presents a bio-inspired event-driven neuromorphic sensing system (NSS) capable of performing on-chip feature extraction and "send-on-delta" pulse-based transmission, targeting peripheral-nerve neural recording applications. The proposed NSS employs event-based sampling which, by leveraging the sparse nature of ENG signals, achieves a data compression ratio of >125×, while maintaining a low normalized RMS error of 4% after reconstruction. The proposed NSS consists of three sub-circuits. A clockless level-crossing (LC) ADC with background offset calibration has been employed to reduce the data rate, while maintaining a high signal to quantization noise ratio. A fully synthesized spiking neural network (SNN) extracts temporal features of compound action potential signals consumes only 13 µW. An event-driven pulse-based body channel communication (Pulse-BCC) with serialized address-event representation encoding (AER) schemes minimizes transmission energy and form factor. The prototype is fabricated in 40-nm CMOS occupying a 0.32-mm^2 area and consumes in total 28.2 µW and 50 µW power in feature extraction and full diagnosis mode, respectively. The presented NSS also demonstrates to extract temporal feature of compound action potential signals with 10-µs precision.**

***Index Terms*— Peripheral nerves, Neural recording, Body channel communication, Neuromorphic, Level-Crossing ADCs, Neural sensors, Electroneurogram (ENG), Action potentials, Spiking Neural Networks, Feature extraction.**

I. Introduction

THE peripheral nervous system (PNS) can be seen as a "highway" for propagating neuron firings, i.e., action potentials (AP), for the bidirectional communication between the central nervous system (CNS) and various organs. The electroneurogram (ENG) can be measured with a nerve cuff or a neural probe surrounding or penetrating the peripheral nerves, respectively. Nerve ENG provides rich clinical information for diagnosis and can be the source of modulating human health as electroceuticals [1][2]. Decoding of the firing pattern of afferent compound action potentials (CAPs), the result of summation of many APs from the individual axons in a nerve trunk, holds the promise for indirect sensing of clinically relevant information, e.g., inflammation status or glucose levels, which can be employed in future electroceutical closed-loop applications [3]. Next, the nerve conduction velocity (NCV) [4] is widely used as a diagnostic tool for various neuropathies. The requirement on temporal precision for such measurements is strict since the CAPs typically last for only a few milliseconds. This precision is especially critical for NCV studies [5], which measure the time difference between peaks of two CAPs recorded from two locations on the same nerve, as shown in Fig. 1. To achieve a high accuracy of NCV with a miniature nerve implant, temporal precision of the recording should be in the order of 10's of µs, since NCV of a myelinated nerve can be up to 120 meter/s. Better temporal precision of the recording allows the volume of the nerve implant (e.g., nerve cuff) to be further miniaturized. To achieve such temporal precision, the analog-to-digital converters (ADCs) in a conventional neural recording system need to have a sampling rate of 10's of kSample/s (kSps), which is 10-100× higher compared to the sampling of other electrocardiogram (ECG) signals. This increases the energy consumption not only in wireless transmission, but also in local processing, storage, and transportation of the data.

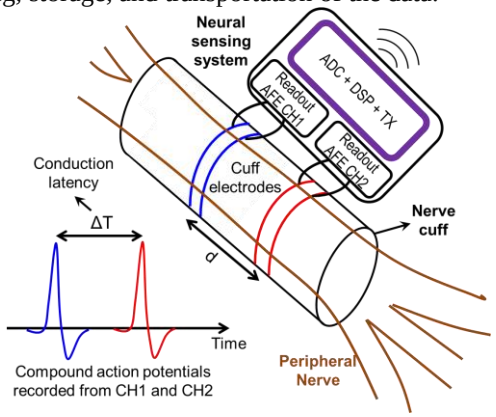


Fig. 1. Concept illustration of neural recording of peripheral nerves, and the conceptual illustration of nerve conduction velocity measurement.

In order to have high spatial selectivity, neural implants for peripheral nerves should be placed very close to the surface (or

inside) the nerve, as illustrated in Fig. 1. To avoid nerve tissue damages, such nerve implants should have strict volume and energy constrains. A nerve implant with a volume in the millimeter-scale is highly preferred. Since there is no sufficient volume for a battery, the electronic system should consume low energy well below 100 μW, to enable wireless power transfer.

Fig. 2 shows two conventional architectures of implantable sensing systems which typically consist of one or multiple channels of analog front-end (AFE), ADC, digital signal processing (DSP), memory, and a wireless transmitter (TX). These architectures are based on Nyquist sampling, but one replies on remote computation (Fig. 2(a)) and another one has embedded local computation (Fig. 2(b)). The first architecture is suitable for diagnosis purposes. The high-precision raw sensor data sampled by the Nyquist ADCs are sent wirelessly to a remote hub to perform further data processing. However, such "frame-based" sampling produces a large amount of data from neural recording, which will consume high wireless transmission energy [6][7]. In addition, performing the computation remotely may introduce potential privacy concerns.

Fig. 2(b) shows an alternative approach based on near-sensor local computation, e.g., feature extraction or classification, to reduce the burden on data transmission. However, this approach requires power- and volume-hungry computation and storage hardware, which is not affordable with a millimeter-scale nerve implant. In addition, such architecture may not be suitable in practice, if raw data are not available in case personalization or detailed diagnosis are needed. Reference [7] uses a local processor to extract ECG signal features, but it requires a relatively large memory (46 kByte) and high-power consumption (~60 μW) for detecting only the peak of the ECG signal, i.e., the R wave. Note that full ECG features (P, QRS and T waves) are still crucial for accurately detecting many cardiac abnormalities, e.g., arrhythmias.

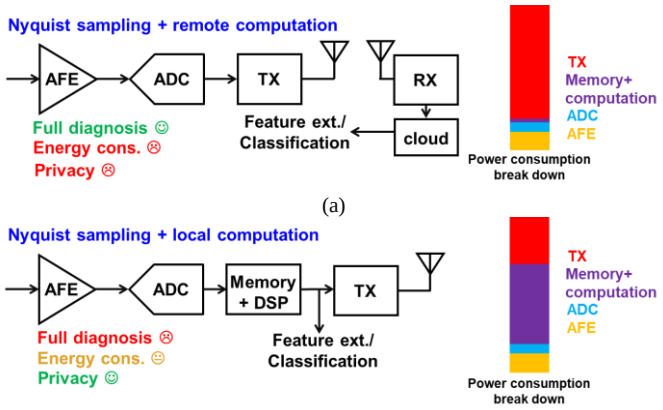


Fig. 2. Conventional architectures of sensing system. (a) Nyquist sampling with remote computation; (b) Nyquist sampling with local computation.

ENG signals have very sparse activity (typically <10 CAPs per second), but a high temporal resolution is still required. High redundancy will be generated, if these signals are sampled with convention high sampling rate Nyquist ADCs, thus leading to a poor system efficiency. Inspired from biology [8], the energy consumption of information processing and transportation can be significantly reduced if only the changes (i.e., delta) of the signal are processed, while information can still be recovered on the reception side with high resolution. One example of such a sensing system is our retina, whose neurons only fire when detecting temporal changes from photoreceptors. The action potentials fired from retina neurons are transmitted through an optic nerve with quite limited data capacity and energy budget, but our brain (the receiver of information) has no problem reconstructing high-quality images. This concept is also suitable for implantable neural sensing systems with very limited energy sources.

Fig. 3(a) illustrates one example waveform of a CAP. Instead of processing signals with a constant clock in every frame, the system is active only if there are CAPs. This significantly reduces the data rate, and thus the requirements of the hardware as well as the energy consumption. Furthermore, the temporal feature can be well preserved, without being limited by the sampling grid of Nyquist ADCs. To implement such bio-inspired sampling mentioned above, an analog to spike converter (ASC) based on level-crossing ADCs (LC ADC) can be employed to perform "delta encoding" [9], which reports (up or down) events when changes larger than a certain threshold are detected. As shown in Fig. 3(a), the ASC generates UP or DN pulses when the CAP signal crosses one quantization (or threshold) step with a positive or negative slope. This greatly reduces the temporal redundancy.

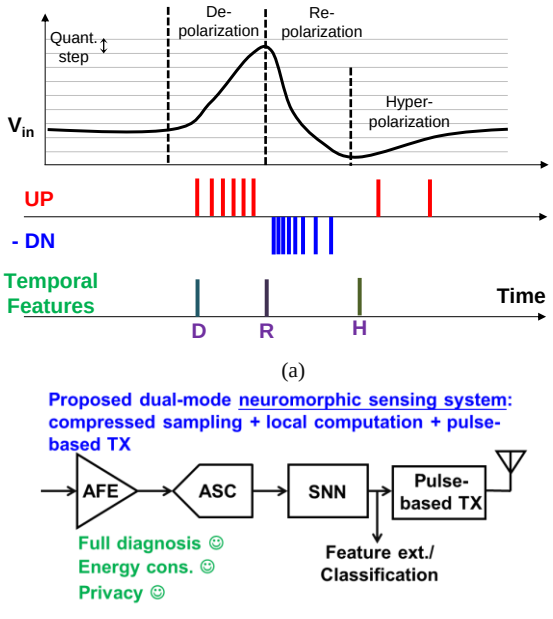


Fig. 3. (a) Delta encoding of ENG CAPs, and its temporal features D, R, and H (b) The conceptual block diagram of the proposed Neuromorphic Sensing Systems.

Fig. 3(b) shows this proposed neuromorphic sensing system (NSS) concept [10]. It includes an ASC for delta encoding, a spiking neural network (SNN) for local computation, and a pulse-based transmitter tailored for low-energy event-driven transmission. The event-based nature of the NSS not only improves the energy efficiency by exploiting the sparse nature of CAP signals, but also avoids a power-hungry system clock

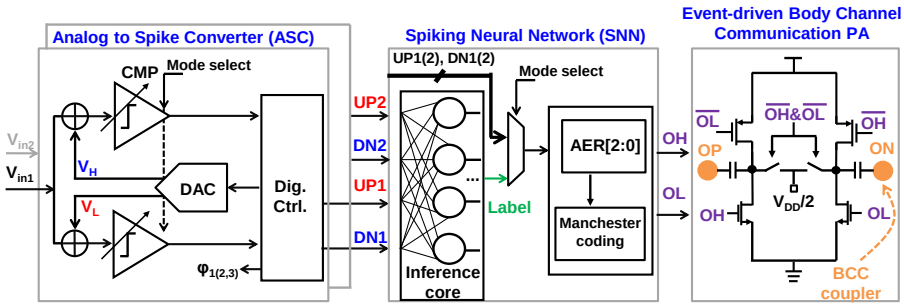


Fig. 4. The detailed block diagram of the proposed NSS.

generation and synchronization circuits. Most importantly, the temporal precision is no longer limited by the fixed sampling grid of the clock.

The proposed NSS is designed to support dual-mode operation: full diagnosis and feature extraction mode. The capability of dual-mode operation is crucial for implantable sensing devices since detailed diagnosis is required in case of urgent situations, e.g., implant failure. When the NSS is in the feature extraction mode, only the extracted temporal features (i.e., labels) are transmitted out for energy saving, or they can be decoded and fedback to the implant stimulator to perform closed-loop neuromodulations. Fig. 3(a) shows that three temporal features can be extracted from CAPs, i.e., depolarization (D), repolarization (R) and hyperpolarization (H). They can be detected from the polarity and density of the spike trains [11]. When full diagnosis is required, the NSS programs the ASC to have higher precision and the pulse-based TX to operate at a higher event transmission rate, so that raw CAP signals can be transmitted in full detail.

The rest of this article is organized as follows. Section II discusses the proposed architecture of NSS and its design trade-off. Section III describes the implementation of the circuits. The measurement results will be shown in Section IV. Finally, Section V presents the conclusions.

II. Proposed Architecture and Design Trade-Offs

This section discusses in detail the proposed NSS architecture. The design trade-off between quantization error and event transmission rate will also be provided.

A. Architecture Overview

Fig. 4 shows more detail of the proposed NSS. This work demonstrates two channels, which is the minimum channel number required for performing a conductive velocity study. The number of sensing channels can be easily scaled in the proposed NSS, based on the requirements and constrains from different clinical use cases.

Two ASCs are implemented as level-crossing ADCs (LC-ADCs). LC-ADCs perform delta encoding, so they also have better immunity to low-frequency noise than other event-driven ADCs [9].

Four outputs of two ASCs (UP1, DN1, UP2, and DN2) are connected to the SNN which is only active when there are CAPs. Thanks to the reduced temporal redundancy, the memory needed for the SNN to extract the signal features and generate corresponding labels is reduced by 2-10×. Note that this memory is distributed in SNN neurons, which also reduces the energy required for memory access.

In feature extraction mode, the SNN inference core can generate three temporal labels (D-R-H), which are then encoded with a serialized form of address-event representation (AER) [12]. The AER output is further encoded with Manchester code before the body channel communication (BCC) transmission, minimizing the residue charge in the tissue (to be detailed in Section III-C).

B. System Analysis and Design Trade-offs

To achieve the targeted temporal and amplitude resolution in neural recording, conventional Nyquist ADCs are typically designed with a high dynamic range up to 10-bit resolution and sampled with high frequency up to 30 kSps [13]. A large amount of data needs to be transferred to an external device wirelessly, resulting in a high data rate of up to 300 kbps each channel in full diagnosis mode. As reported in [14], the data transmission consumes more than 90% of the total system energy. The proposed NSS reduces the data rate by leveraging the sparse nature of CAP signals.

For a Nyquist sampling system, the ADC's clock defines the time stamps and thus the precise timing of the rest of the system (including wireless link) is not critical because the time stamps are already defined together with the data. However, in this clock-less event-based sampling system, the time stamps are set based on the timing of the received wireless data, i.e., time itself represent the time stamps, and thus timing variation in the entire chain affects temporal precision. The system timing resolution also determines the signal quality after reconstruction. The timing resolution can be limited by many parts of the NSS. One dominant limitation is the maximum event transmission rate the NSS can achieve, or equivalently how fast two events can be transmitted consecutively. If one event packet, i.e., an event with serial address-event representation, has a long length in time, the transmission of the following event must be delayed, which equivalently introduces a time-domain quantization error.

To minimize the length of event packets (or maximize the event transmission rate), the bit period must be reduced. However, this requires a higher speed of the Pulse-BCC PA, which also consumes more power. To understand the relation between the maximum required event transmission rate and the signal quality after reconstruction, analysis based on a numerical simulation are performed and validated. In this analysis, a synthetic action is converted to UP/DOWN events

by the ASC. Note the event rate is coupled to the number of bits (or the quantization steps) of the ASC, i.e., finer quantization steps produce more events.

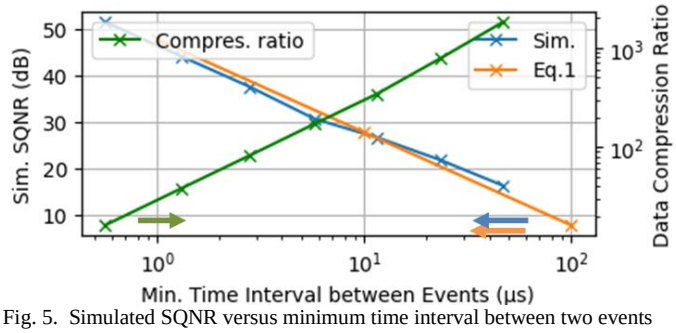


Fig. 5. Simulated SQNR versus minimum time interval between two events and the simulated data rate compression ratio.

Since the SNR of recorded APs in-vivo is typically limited to 20 dB [15] due to biological noise and wide signal bandwidth, the target of the maximum signal-to-quantization-noise-ratio SQNR is set to 25 dB in this work. Based on the analysis provided in [16],

$$SQNR = -20log10(\delta * f_{SIG}) - 14.2, \quad (1)$$

Where δ is the timing uncertainty of the event-based system; f_{SIG} is the signal bandwidth, which is ~1 kHz for the nerve's CAP signal. To achieve an SQNR of 25 dB, the system timing uncertainty should be less than 10 μs, which also sets the upper limit of the event packet length.

Fig. 5 shows the relation between the required timing resolution and the simulated SQNR, which matches well with the theoretical results from Eq. (1). Fig. 5 also shows a simulated data rate reduction compared to conventional Nyquist sampling and frame-based transmission. The firing rate of CAP is 10 Hz, and the AER overhead has been included for the event transmission. It shows that the proposed NSS can achieve higher than 200× of data reduction, while keeping SQNR above 25 dB.

III. Circuit Implementation

Three circuit innovations will be discussed in this section: (1) a background offset mitigation technique is proposed to enhance the offset tolerance of ASCs; (2) a fully synthesized low-power SNN is introduced, which is capable of on-chip temporal feature extraction; (3) a power amplifier (PA) with charge balancing and AER encoder for event-driven Pulse-based Body Channel Communication (Pulse-BCC) is introduced.

A. LC ADC with the Offset Calibration

The block diagram of the LC ADC is shown in Fig. . The ADC continuously monitors the input signal and generates an UP or DN event when the input change crosses a +1 LSB or -1 LSB threshold, respectively. To do so, the analog input signal is diminished by a voltage generated by a digital-to-analog converter (DAC), V_{DAC}. After the subtraction, a pre-amplifier amplifies the signal and two charge adders add and deduct 1 LSB from the signal, respectively. Two continuous-time comparators are employed to detect when the zero-crossings occur. If it happens, a digital control block (Dig. Ctrl.) will update V_{DAC} and wait for the next zero-crossing.

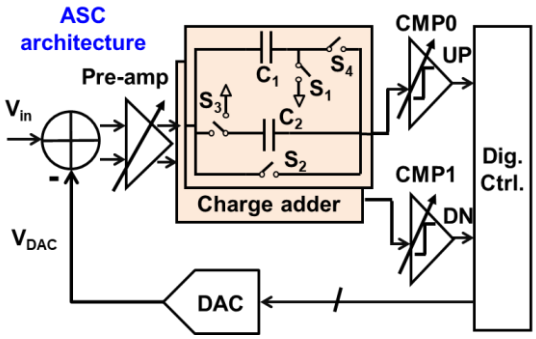


Fig. 6. Architecture of the LC ADC.

The LC-ADC employs two comparators to actively detect rising and falling zero-crossings. However, the offset difference of these two comparators, degrades SNDR of the LC ADC. Therefore, a background offset mitigation technique is proposed and shown in Fig. . Instead of using two separate comparators, the first stage (pre-amp) is shared between them so that the offset errors of the second stage are divided by the pre-amp gain *A*. Then, a double-sampling switched-capacitor circuit removes the offset error *e* of the pre-amp and stores the two threshold levels (V_H and V_L) in three steps.

First, the reference levels $\pm V_{LSB}$ and the offset *e* are amplified by *A* and stored on capacitor C1. Second, only the amplified offset error *e* is stored on another capacitor C2 whose capacitance is the same as C1. Third, as the normal operation, C1 and C2 are connected, the offset error *e* is cancelled, and the input-referred voltage is shifted by means of the capacitors is $\pm V_{LSB}$ for the threshold voltages V_H and V_L. These operations are controlled by three non-overlapping signals $\varphi_{1\text{-}3}$, which only toggle in the presence of UP or DN pulses. To support dual-mode operation, the pre-amp and the comparators can be programmed with different power and bandwidth.

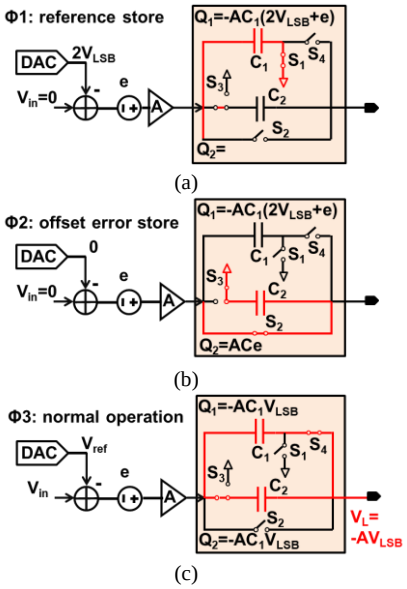


Fig. 7. Background ASC offset cancellation at different phases: (a) reference store; (b) offset error store; (c) normal operation.

B. *Fully Synthesized Low-power SNN*

Conventional Neural Network (NN) architectures either cannot support event-driven operation [17], or use analog-intensive neurons and synapses which are sensitive to PVT variations [18], and both still consume relatively high power (100's of µW). In this work, a low-power and fully synthesizable SNN is presented, as shown in Fig. 8. To exploit the sparse activity of the ENG signals, the network is made self-timed, resulting in near-zero dynamic power dissipation in the absence of any input activity. The SNN consists of two consecutive pools of fully recurrently connected spiking neurons and each pool contains 46 neurons.

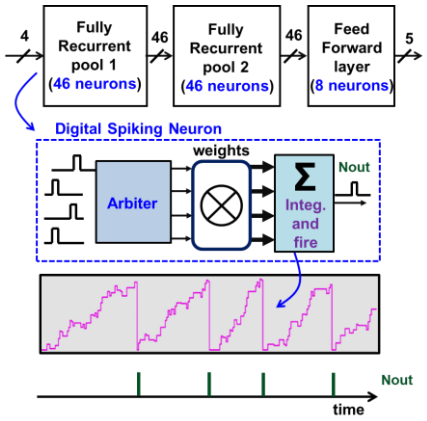


Fig. 8. Block diagram of the fully synthesizable SNN core.

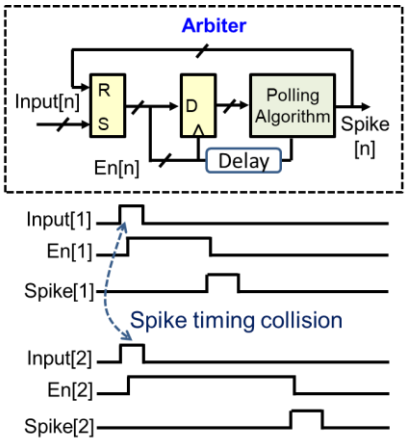


Fig. 9. Simplified block diagram of arbiter and its waveform.

These pools are followed by a fully-connected layer with 8 neurons. A stream of spikes (digital pulses) represents the network inputs and outputs, whereas the synaptic weights are stored with 8-b digital numbers. Note that digital spiking neurons provide more flexibility, lower power consumption and low sensitivity to PVT variations, compared to charge-based analog spiking neurons in [18].

Input spikes arriving at arbitrary times select a corresponding weight, which gets added to an accumulator. When the digital accumulator overflows, it produces an output spike similar to biological neurons' integrate-and-fire operation. To solve timing collisions, each neuron has an arbiter that adds a small time offset (100's of ns) to set the priority using a "Round-Robin" polling algorithm. The simplified block diagram of the proposed arbiter and its waveform are shown in Fig. 9. This SNN is fully synthesized by the standard digital design flow.

Since the temporal pattern of the CAP signals is important, the SNN needs to have a fast response time when extracting features. The spike arbiters determine the speed of our SNN. Every neuron layer contains such an arbiter. The largest one is at the first layer and has N=46 spike inputs. The delay T_{cycle} of the arbiter to process one spike is mainly determined by its embedded priority encoder. With T_{gate} being a typical logic gate delay in this technology, we found:

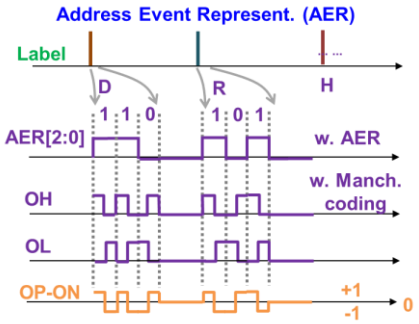
$$T_{cycle} = 3 * log2(N) * T_{gate}. \quad (3)$$

Then, the number of spikes per second that can be processed by the arbiter is $1/T_{cycle}$. For example, in 40nm CMOS, T_{gate}=40 ps. If N=46, then T_{cycle} = 3*8*40 ps = 1 ns. This delay has a negligible impact on temporal precision.

C. *Pulse-BCC PA and Address-Event Representation (AER)*

Instead of adopting high-frequency EM radiation [7][19][20], body-channel communication (BCC) is adopted in this work for the following reasons. First, it does not require an antenna, so the NSS volume can be miniaturized. Second, it has lower propagation loss inside the human body. And third, it also provides better privacy since the signals do not radiate [21].

The differential PA, as shown in Fig. 4, generates a positive output when OH is high and a negative signal when OL is high. When both OH and OL are low, the PA outputs are reset to half VDD, and it only dissipates 1.2µW.



(a)

AER & Manchester coding Table (@feature extraction mode)

	AER	OH
D	110	+- +- +-
R	101	+- -+ +-
H	100	+- -+ -+

AER & Manchester coding Table (@full diagnosis mode)

	AER	OH
UP1	011	-+ +- +-
DN1	010	-+ +- -+
UP2	001	-+ -+ +-
DN2	000	-+ -+ -+

(b)

Fig.10. (a) Output waveforms from AER and BCC-TX; (b) AER and Manchester coding tables from two operation modes

In feature extraction mode, three temporal labels generated by the SNN will be coded with a serialized 3-bit AER, as illustrated in Fig. 10. The corresponding "address" of each label will be attached to the polarity of the event (either up or down), using a serialized 2-bit Tenary code (with "+", "-", and "0"). The AER output is further coded with Manchester coding, to ensure the number of "+" and "-" are always equal, which is

important to ensure the charge balance. To transmit these AER coded labels, the differential Pulse-BCC PA either charges or discharges the tissue if the AER output is "+" or "-". When the AER output is "0", the reset circuit of the PA will short two ends of the BCC coupler and connect to half VDD, to further remove any residue charge in the tissue.

According to the discussion in Section II-B, the event packet length should be less than 10 µs to ensure that the timing resolution of the event-based transmission is fine enough not to degrade signal quality. The event packet length can be calculated by

$$T_{packet} = n * T_{bit} + T_{reset}, \quad (5)$$

where n represents the number of bits for the AER and Manchester code (in this case is 6), T_{bit} is the period of each bit, and T_{reset} is the time required for the reset circuit to remove the residue charge in the tissue. Note that T_{reset} heavily depends on the strength of the reset circuit and the BCC coupler geometry, and it is ~4 µs in this work. The Pulse-BCC PA has a programmable bit period in the range of 0.5-2µs.

IV. Measurement Results

As shown in Fig. , the proposed NSS is fabricated in 40-nm CMOS technology and it occupies only 0.32 mm^2, thanks to a reduced memory demand and the area-efficient Pulse-BCC.

To evaluate the performance of the LC-ADC, a sinusoidal wave is set as the input. In feature extraction mode, it achieves 30-dB SNDR with 7-µW. In full diagnosis mode, at a higher power of 17- µW, it achieves 72-dB SFDR and 59.5-dB SNDR which is shown in Fig. 5(a). Fig. 12(b) shows the simulated and measured SNDR with different induced comparator offsets and the improvement could be observed after employing a pre-amplifier.

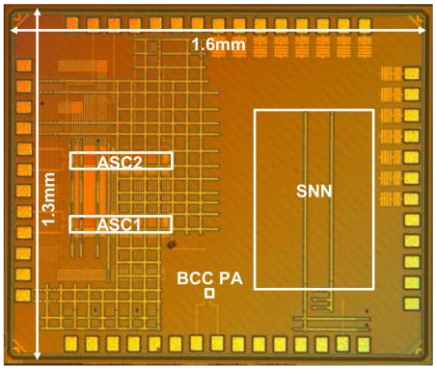
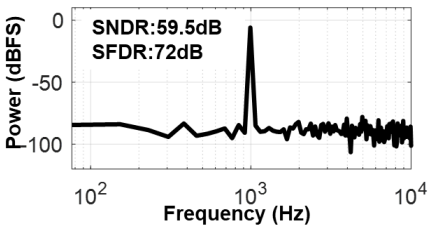
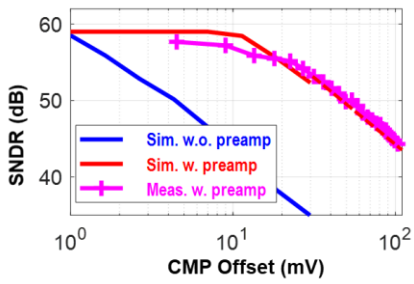


Fig. 11. Chip photo in 40nm CMOS.



(a)



(b)

Fig. 52. (a) Measured SNDR from ASC outputs. (b) Simulated and measured SNDR with different ASC comparator offset.

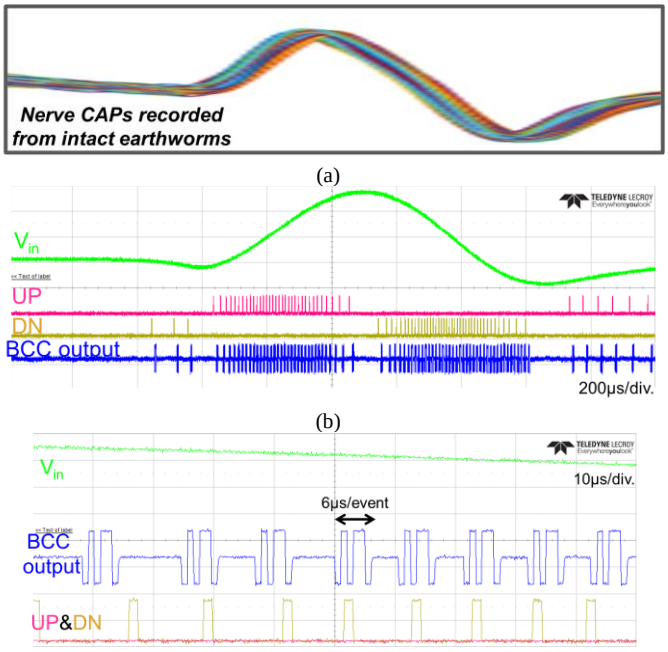


Fig. 13 (a) Nerve CAPs recorded from intact earthworms, and the measured NSS input, ASC and BCC-TX outputs in full diagnosis mode with (b) 2 ms span, (c) 0.1 ms span.

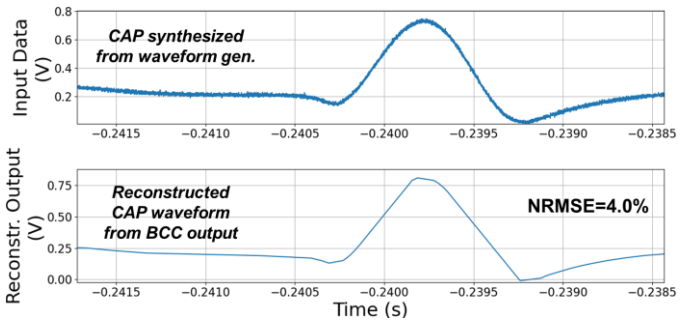


Fig. 14. Measured normalized RMS error between the reconstructed CAP waveform based on the TX output, and the ASC input signal.

The NSS is characterized with a synthetic CAP waveform according to our neural recordings from intact earthworm's medial giant nerve fiber, as shown in Fig. 13(a), which is widely used approach for studying neurophysiology [22]. It is provided as the input signal generated by a waveform generator, with an amplitude matched to the ASC's full scale and a frequency of ~10 Hz (i.e., 10 CAPs per second). Fig. 13(b) and 13(c) show the outputs of the ASC and the Pulse-BCC TX outputs in full diagnosis mode. Fig. 13(b) shows the time domain waveform

from the Pulse-BCC output with Manchester coded AER. The event packet length is measured to be ~6 µs. The reconstructed waveform from the Pulse-BCC TX output is shown in Fig. 14, where the reconstruction is implemented with a simple accumulation without any filtering. The measured Normalized RMS error (NRMSE) between the input and reconstructed signals is 4%, and the SQNR is estimated to be 27.9 dB, which is close to the analysis in Section II-B.

The measured power consumption of the SNN during the feature extraction is 2 µW and 11 µW for dynamic and leakage power, respectively. The SNN is only active when the input signals change substantially, leading to very low dynamic power. The leakage power is limited by the nanoscale transistors in the implemented process and can be further improved with advanced techniques or process [23].

The temporal resolution is typically limited by the sampling period (e.g., ~33 µs in [6][24]) in conventional frame-based sampling systems [25]. The temporal resolution in feature extraction mode is measured by overlaying multiple of extracted R labels of the CAP and measuring the timing uncertainty, as shown in Fig. 15. The SNN is pre-trained to recognize the R feature of the synthetic CAP. The measurement result shows that a temporal precision of 10 µs can be achieved with the proposed event-based NSS.

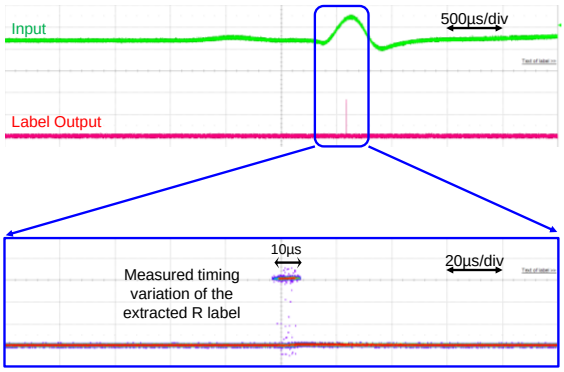


Fig. 15. Measured temporal precision of the extracted temporal feature R, i.e., peak of the CAP, by overlaying multiple measurements.

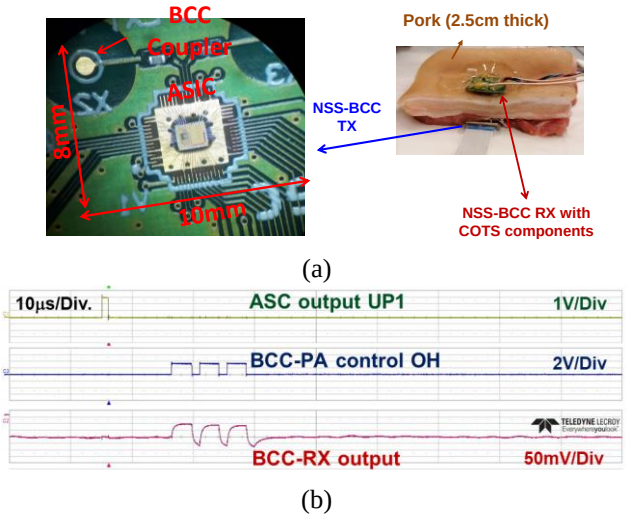


Fig. 16. (a) Measurement setup and the photo of the NSS module. The Pulse-BCC measurement is performed with a 2.5-cm porcine tissue. (b) Measured Pulse-BCC TX input and RX output signals.

Fig. 16(a) shows the setup for the evaluation of the Pulse-BCC link. The electronic module area of the NSS, including the coupler of the BCC TX, is only 72 mm^2. A PCB for the BCC receiver (RX) has been developed using commercially available components. A porcine tissue with 2.5-cm thickness has been used as the communication channel for the BCC link. Fig. 16(b) shows the TX input data and the received output from the RX.

The NSS consumes only 28.2 µW and 50.5 µW system power in the feature extraction and full diagnosis mode, respectively. Currently the system power consumption is limited by the leakage performance of the selected process. The comparison with state-of-the-art electrophysiology (ECG, ENG) sensing and data transmission systems is shown in Table I. Thanks to the bio-inspired event-based sampling, the data rate can be reduced by 125×, while achieving 10 µs temporal resolution. Note this compression ratio is with respect to a 30 kSps 10-b Nyquist ADC, and a CAP firing rate of 10 Hz. Although adaptive sampling techniques can also achieve a compression ratio of 7× [7], it requires complicated and power-hungry digital processing for mode control. Although finer temporal resolution can be achieved with a higher sampling frequency as in [20], this also leads to a higher system power consumption.

The presented NSS also has smallest system module area, thanks to the crystal-less event-based operation and the antenna-less body-channel communication. All these features make this NSS a very promising architecture for neural sensing of peripheral nerves.

Table I. Comparison with state-of-the-art implantable sensing system

	This work		[7] Kim TBIOCAS'14		[19] Shon Sensors'17	[20] Azin JSSC'11
Tech. (nm)	40		180		N.A.	350
Supply (V)	0.9/1/1.1		1.2		N.A.	1.5
System	ADC+SNN+TX		AFE+ADC+ DSP+TX		AFE+ADC +TX	AFE+ADC+ DSP+TX
Applications	Implantable ENG (PNS)		Wearable ECG		Implantable ENG (PNS)	Implantable ENG (Brain)
	Feat. Extr.	Full diag.	Feat. Extr.	Full diag.		
ADC architecture	LC (event-based)		SAR (Nyquist)		NA (Nyquist)	SAR(Nyquist)
Data rate/ch= ADC sample rate × nr. of bits (bps)	<100 event/s	~2.4k event/s	512/64 (Adaptive) ×12b =877		10k×10b =100k	35.7k×10b =357k
Data compression	**>125×[A]**		7×		1×	1×
Temporal precision (µs)	**10**		>2000[B]		-	>28[B]
ADC ENOB	5	9.5	10.3		N.A.	9.1
TX freq./mod.	0.5-µs Pulse based BCC		2.4GHz BLE		400MHz FSK	433MHz FSK
Nr. of chan.	2		3		2	8
Power cons. (µW)						
TX	1.2	1.5	1000	13300	15510	200
DSP/NN	13	15	42	0	N.A.	26
ADC/ASC	14	34	18	38	N.A.	47
Total[C]	**28.2**	**50.6**	1060	13340	N.A.	274
Total power cons./Ch (µW)[C]	14.1	25.3	330	4430	N.A.	34
Core die area (mm^2)[C]	0.32		8.46		N.A	2.43
System module area (mm^2)	**72**		~500		924	N.A.
On-chip labeling	**D/R/H**	-	Peak (R) only		No	Peak(R) only

[A] Compared to 300 kbps [B] Limited by ADC sampling rate. [C] Estimated by excluding AFE.

V. Conclusion

This work presents a bio-inspired neuromorphic sensing system, including compressed sampling with delta encoding, event-based spiking neural network for local feature extraction, and an event-driven pulse-based body channel communication for miniaturization. This work targets the application of neural recording in peripheral nerve implants, which requires fine

temporal resolution. The analysis in Section II shows that there is a trade-off between event transmission rate, data compression ratio and the signal quality after reconstruction. The proposed NSS demonstrates the capability of supporting two recording modes: the full diagnosis mode which transmits raw sensing data with low power consumption, and the feature extraction mode which transmits only the extracted temporal feature with fine precision. The presented NSS architecture features high energy efficiency, miniature form factor, and high temporal resolution, making it a promising architecture for neural recording of peripheral nerves implants.

ACKNOWLEDGMENT

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 101001448).

REFERENCES

[1] S. Breit, A. Kupferberg, *et al*., "Vagus Nerve as Modulator of the Brain–Gut Axis in Psychiatric and Inflammatory Disorders", *Frontiers in Psychiatry*, Mar. 2018.

[2] K. Famm, *et al.*, "A jump-start for electroceuticals", *Nature*, Vol. 496, pp. 2836-2840, Apr. 2013.

[3] T. P. Zanos, *et al.*, "Identification of cytokine-specific sensory neural signals by decoding murine vagus nerve activity," *PNAS*, May 2018.

[4] R. B. Szlavik, *et al.*, "A Novel Method for Characterization of Peripheral Nerve Fiber Size Distributions by Group Delay," *IEEE Transactions on Biomedical Engineering*, vol. 55, pp. 2836-2840, Dec. 2008.

[5] J. Hoffer, *et al.*, "Single unit conduction velocities from averaged nerve cuff electrode records in freely moving cats," *Journal of Neuroscience Methods*, vol. 4, no. 3, pp. 211-225, Dec. 1981.

[6] Y. Luo, *et al.*, "A 74-μW 11-Mb/s Wireless Vital Signs Monitoring SoC for Three-Lead ECG, Respiration Rate, and Body Temperature," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, pp. 907-917, Oct. 2019.

[7] H. Kim, *et al.*, "A Configurable and Low-Power Mixed Signal SoC for Portable ECG Monitoring Applications," *IEEE Trans.s on Biomedical Circuits and Systems*, vol. 8, pp. 257-267, Apr. 2014.

[8] P. Lichtsteiner, *et al.*, "A 128× 128 120 dB 15 ms Latency Asynchronous Temporal Contrast Vision Sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566-576, Feb. 2008.

[9] C. Weltin-Wu, *et al.*, "An Event-driven Clockless Level-Crossing ADC With Signal-Dependent Adaptive Resolution," *IEEE Journal of Solid-State Circuits*, vol. 48, pp. 2180-2190, Sept. 2013.

[10] Y. He, *et al.*, "A 28.2 μW Neuromorphic Sensing System Featuring SNN-based Near-sensor Computation and Event-Driven Body-Channel Communication for Insertable Cardiac Monitoring," *2021 IEEE Asian Solid-State Circuits Conference (A-SSCC)*, pp. 1-3, Nov. 2021.

[11] Y. Liu, *et al.*, "Event-driven processing for hardware-efficient neural spike sorting," *Journal of Neural Engineering*, Jan. 2018.

[12] K. Boahen, *et al.*, "Point-to-Point Connectivity Between Neuromorphic Chips Using Address Events," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47, pp. 416-434, May. 2000.

[13] M. Ballini, *et al.*, "Intraneural active probe for bidirectional peripheral nerve interface," *IEEE Symposium on VLSI Circuits*, June 2017.

[14] S. M. A. Zeinolabedin, *et al.*, "A 16-Channel Fully Configurable Neural SoC with 1.52W/Ch Signal Acquisition, 2.79W/Ch Real-time Spike Classifier, and 1.79 TOPS/W Deep Neural Network Accelerator in 22nm FDSOI," in *IEEE Transactions on Biomedical Circuits and Systems*, Early Access, Jan. 2022.

[15] Jun, J., *et al.*, "Fully integrated silicon probes for high-density recording of neural activity," *Nature* 551, 232–236, Nov. 2017.

[16] W. Kuang, *et al.*, "Correction and Comment on "An Adaptive Resolution Asynchronous ADC Architecture for Data Compression in Energy Constrained Sensing Applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 4, pp. 1097-1099, Apr. 2013.

[17] C. Frenkel, *et al.*, "A 0.086-mm^2 12.7-pJ/SOP 64k-Synapse 256-Neuron Online-Learning Digital Spiking Neuromorphic Processor in 28-nm CMOS," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, pp. 145-158, Feb. 2019.

[18] S. Moradi, *et al.*, "A Scalable Multicore Architecture With Heterogeneous Memory Structures for Dynamic Neuromorphic Asynchronous Processors (DYNAPs)," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, pp. 106-122, Feb 2018.

[19] Shon, Ahnsei, *et al.*, "An Implantable Wireless Neural Interface System for Simultaneous Recording and Stimulation of Peripheral Nerve with a Single Cuff Electrode." *Sensors*, vol. 18, no.1, Jan. 2018.

[20] M. Azin, *et al.*, "A Battery-Powered Activity-Dependent Intracortical Microstimulation IC for Brain-Machine-Brain Interface," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 4, pp. 731-745, April 2011.

[21] J. Jang, *et al.*, "A Four-Camera VGA-Resolution Capsule Endoscope System With 80-Mb/s Body Channel Communication Transceiver and Sub-Centimeter Range Capsule Localization," *IEEE Journal of Solid-State Circuits*, vol. 54, pp. 538-549, Feb. 2019.

[22] N. Steinmetz, *et al.*, "Teaching Basic Neurophysiology Using Intact Earthworms," *Journal of Undergraduate Neuroscience Education*, vol. 9, 2010.

[23] A. Kumar, *et al.*, "A 0.065-mm^2 19.8-mW Single-Channel Calibration-Free 12-b 600-MS/s ADC in 28-nm UTBB FD-SOI Using FBB," *IEEE Journal of Solid-State Circuits*, July 2017.

[24] T.-F. Wu, *et al.*, "A Flash-Based Non-Uniform Sampling ADC With Hybrid Quantization Enabling Digital Anti-Aliasing Filter," *IEEE Journal of Solid-State Circuits*, vol. 52, pp. 2335-2349, Sept. 2017.

[25] M. Petelczyc, *et al.*, "Impact of observational error on heart rate variability analysis," *Sports Medicine*, May. 2020.