

A 16-Channel Wireless Neural Recording System-on-Chip with CHT Feature Extraction Processor in 65nm CMOS

Arda Uran¹, Kerim Ture¹, Cosimo Aprile^{1,2}, Alix Trouillet¹, Florian Fallegger¹, Azita Emami³, Stéphanie P. Lacour¹, Catherine Dehollain¹, Yusuf Leblebici^{1,4}, Volkan Cevher¹

¹EPFL, Lausanne, Switzerland, ²Kandou Bus, Lausanne, Switzerland, ³California Institute of Technology, Pasadena, CA ⁴Sabancı University, Istanbul, Turkey

Wireless implantable neural recording chips enable multichannel data acquisition with high spatiotemporal resolution in situ. Recently, the use of machine learning approaches on neural data for diagnosis and prosthesis control have renewed the interest in this field, and increased even more the demand for multichannel data. However, simultaneous data acquisition from many channels is a grand challenge due to data rate and power limitations on wireless transmission for implants. As a result, recent studies have focused on on-chip classifiers (Fig. 1 top), despite the fact that only primitive classifiers can be placed on resource-constrained chips. Moreover, robustness of the chosen algorithm cannot be guaranteed pre-implantation due to the scarcity of patient-specific data; waveforms can change over time due to electrode micro migration or tissue reaction, highlighting the need for robust adaptive features.

To address these issues, this work presents a wireless neural recording system-on-chip with a Compressed Hadamard Transform (CHT) processor, which serves both as a feature extractor (FE) for classification and as a data compressor for waveform reconstruction. Fig. 1 (bottom) shows the block diagram of the system. A 16-channel 10-bit analog front-end (AFE) based on [1] amplifies and digitizes signals within the local field and action potential bands with adjustable gain and bandwidth. Since the activity of an electrode location cannot be guaranteed a priori, each channel can be configured individually to remain off, to output raw data, or to output features. The CHT processor realizes the matrix product of a 64x10-bit input window with 8 selected rows from the 64x64 Hadamard matrix. The selection can be configured per-channel and on-the-fly which eliminates the inefficiency caused by computing and transmitting nondescriptive features. The effective data compression ratio is 5 as the bit width increases from 10 to 16. The 8x16-bit features are allocated on the output data stream in 13 consecutive 10-bit packets, padded with 51 zero packets. Therefore, the average data rate reduces from 3.2 Mbps to 640 kbps at 20 kHz sampling rate. The resulting power savings in the impulse radio ultra-wideband transmitter (IR-UWB TX) compensates for the power overhead of the CHT processor.

Fig. 2 shows the hardware implementation of the matrix product based on an accumulator bank. Contrary to other spectral transforms, there is no need for multipliers or a coefficient memory as Hadamard coefficients (± 1) are generated on-the-fly without any significant overhead. For waveform reconstruction, transmitted features are brought back to the time domain by applying the inverse Hadamard transform. Group selection of indices operates as a band-pass filter. The lowest indices capture the most of the energy, hence the highest reconstruction fidelity. Fig. 3 (left) shows the preclinical validation of the system in-vivo. Cortical responses to optogenetic stimulation were recorded in rats using a 16-channel soft micro-electrocorticography array, microfabricated using thin-film and silicone processing technology inspired from the e-dura process [2]. Fig. 3 (right) displays the raw and the low-pass reconstructed waveforms and multi-unit spiking activity obtained in real time.

For classification, the most descriptive features vary between patients and channels. Therefore, the indices are tailored per patient and per channel using the feature importance obtained from training on raw data. Classification experiments were performed offline on the iEEG.org and the CHB-MIT seizure datasets using the XGBoost package with leave-one-out cross-validation. The classifier is an ensemble of 8 decision trees with depth 4. Fig. 4 presents the experiment setup and the results obtained on each patient. An initial training was performed on all available channels and all 64 indices to determine the most descriptive 16 channels and 8 indices per-channel from the constructed tree. The main training was performed using only those channels and indices, which simulates the chip

configuration. The chosen indices vary between patients, and span the entire spectrum. The average sensitivity for the CHB-MIT dataset is 92% and reaches 97.8% if only two of the outliers are excluded.

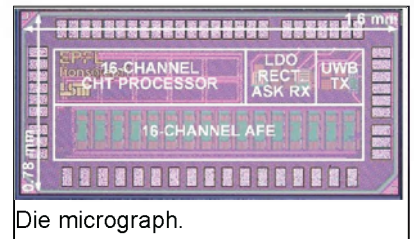


Fig. 5 depicts the wireless power and data transfer subsystem in detail. The magnetically induced AC voltage at 13.56 MHz frequency is converted to DC power by an active half-wave rectifier with delay compensation techniques. A regulated 1 V supply is provided by the low-dropout (LDO) regulator and maintained by an external capacitor. The chip clock is recovered from the power carrier. An amplitude-shift keying (ASK) demodulator was used for downlink transmission at 13.56 kbps. The uplink is established by the IR-UWB TX. The 6 GHz carrier frequency is generated by an active inductor-based LC oscillator. The oscillator and the buffer stage are modulated by 1.8 ns wide pulses to comply with the FCC mask and to save power. Measured at 3.2 Mbps and 640 kbps, the TX dissipates 33 μ W and 6.6 μ W, respectively, which correspond to 10.3 pJ/b energy efficiency.

The 16-channel prototype fabricated in 65nm CMOS has a footprint of 1.6 mm by 0.78 mm, 0.382 mm² of which is occupied by the core blocks. The total power consumption is 401.45 μ W including the power and command receiver (355 μ W). Fig. 6 (top) displays the power and area breakdowns. The power per channel including data transmission is 2.72 μ W for raw data and 2.9 μ W for features, meaning that 80% data rate reduction is achieved in return of only 6.7% increase in power. The area per channel including the AFE and the CHT processor is 0.021 mm², which makes the architecture feasible for up to a thousand channels.

Fig. 6 (bottom) compares the proposed design with the current state of the art [3-6]. The seizure detection performance with CHT features is similar to [5-6] on the same dataset. The CHT processor occupies less area than the conventional FEs based on fast Fourier transform (FFT), discrete wavelet transform (DWT), and finite impulse response (FIR) filters. At 20 kHz sampling rate, the transmitted features enable 312.5 class/s at the receiver side, which corresponds to 149 nJ/class or 9.3 nJ/class/channel energy efficiency for AFE, FE and TX. This is significantly less than the reported energy cost of on-chip classifiers, which makes sending multichannel data to off-chip, more advanced and flexible classifiers feasible. Moreover, new features can be computed on the receiver side as necessary thanks to the waveform recovery option.

Acknowledgement:

All animal experiments were approved by the Veterinary Office of the canton of Geneva in Switzerland and were in compliance with all relevant ethical regulations under animal license no GE 174_17. This work was supported by the ERC under the EU's H2020 Programme Grant 725594, Hasler Foundation under Project 16066, Bertarelli foundation and SNSF Grant CRSII5_183519.

References:

- [1] A. Uran *et al.*, "An AC-coupled wideband neural recording front-end with sub-1 mm²×fJ/conv-step efficiency and 0.97 NEF," SSC-L, Aug. 2020.
- [2] G. Schiavone *et al.*, "Soft, implantable bioelectronic interfaces for translational research," Adv. Mater., Mar. 2020.
- [3] H. Kassiri *et al.*, "Rail-to-rail-input dual-radio 64-channel closed-loop neurostimulator," JSSC, Nov. 2017.
- [4] G. O'Leary *et al.*, "A recursive-memory brain-state classifier with 32-channel track-and-zoom $\Delta\Sigma$ ADCs and charge-balanced programmable waveform neurostimulators," ISSCC, Feb. 2018.
- [5] S. Huang *et al.*, "A 1.9-mW SVM processor with on-chip active learning for epileptic seizure control," in JSSC, Feb. 2020
- [6] Y. Wang *et al.*, "26.3 A closed-loop neuromodulation chipset with 2-level classification achieving 1.5Vpp CM interference tolerance, 35dB stimulation artifact rejection in 0.5ms and 97.8% sensitivity seizure detection," ISSCC, Feb. 2020.

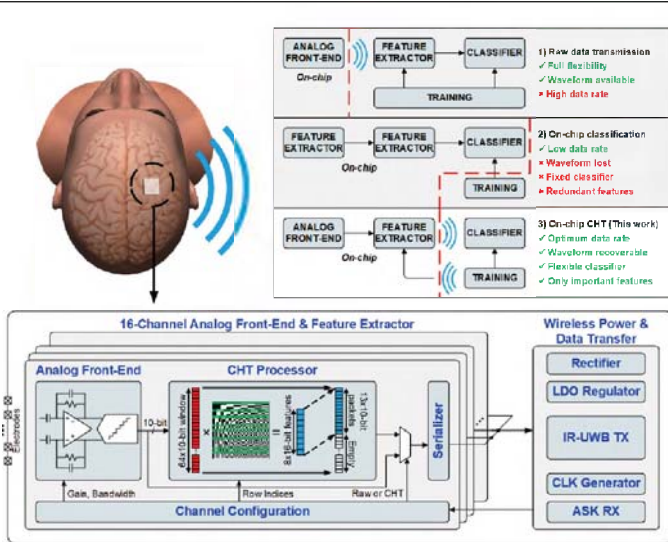


Fig. 1. Proposed neural recording approach (top) and the block diagram of the proposed system-on-chip (bottom).

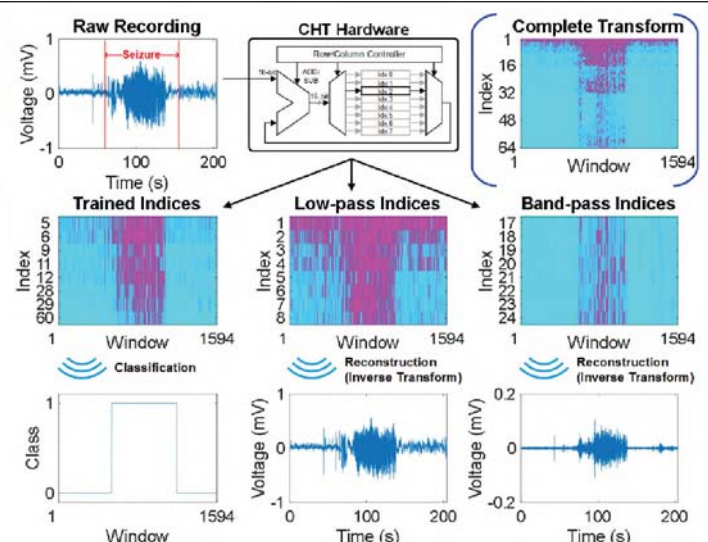


Fig. 2. Compressive Hadamard Transform hardware and its uses for seizure classification and filtered waveform reconstruction.

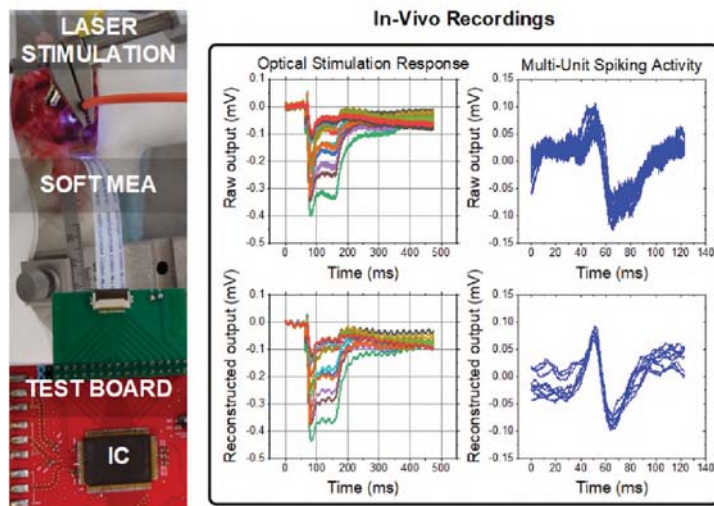


Fig. 3. In-vivo experiment setup (left) and comparison of raw and low-pass reconstructed waveforms (right).

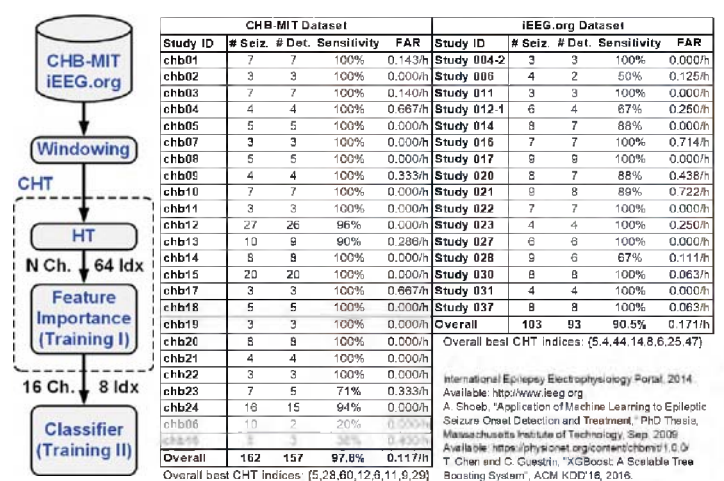


Fig. 4. Classifier experiment setup (left) and results (right) using CHT features on CHB-MIT and iEEG.org datasets.

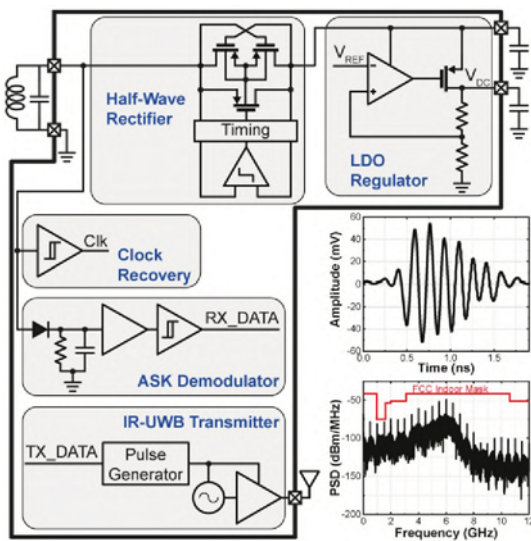
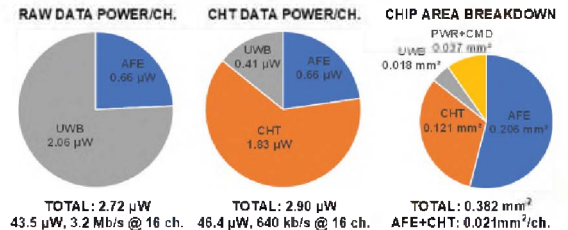


Fig. 5. Wireless power and data transfer block diagrams (left) and the measured IR-UWB pulse characteristics (right).



Technology [nm]	JSSC '17 [3]	ISCC '18 [4]	JSSC '20 [5]	ISCC '20 [6]	This work
V _{DD} [V]	1.2	1.2	0.58	1.5	1
Number of Channels	64	32	14	8	16
Blocks included	AFE+FE+TX	AFE+FE+CLF	FE+CLF	AFE+FE+CLF	AFE+FE+TX
Total Area [mm ²]	3.86	7.59	2.56	5.83	0.382
Total Power [mW]	0.47*	0.714**	1.9	0.133**	0.0464*
Feature Extractor	FIR+PLV	FIR+PLV/SE/CFC	FFT+SE	DWT+KDE	CHT
Dimensionality Reduction	-	Autoencoder	mRMR	ICA	XGBoost
Feature Dimension	n/r	125	18	48	128
Feature Computation	Fixed	Fixed	Fixed	Fixed	Adaptive
Waveform Reconstruction	No	No	No	No	Yes
Feature Extractor Area [mm ²]	1.285	0.475*	0.618*	0.817*	0.121
Classification Rate [class]	n/r	4	11.1	0.074	37.5
Energy/Classification [pJ]	n/r	168.6	170.9	14.2	0.149
Classifier	Threshold	SVM	SVM	SVM	XGBoost/DT
Dataset	Custom	EU	CHB-MIT	CHB-MIT	CHB-MIT
Sensitivity	75%	100%	96.6%	97.8%	97.8%
Specificity/False Alarm Rate	0.5/h	0.81/h	0.29/h	99.7%	0.12h

Fig. 6. Power and area breakdowns (top) and comparison with published work (bottom).