A 400-MHz Wireless Neural Signal Processing IC with 625× On-Chip Data Reduction and Reconfigurable BFSK/QPSK Transmitter Based on Sequential Injection Locking

Kok-Hin Teng, Tong Wu, Zhi Yang, and Chun-Huat Heng

Department of Electrical and Computer Engineering,
National University of Singapore, Singapore
Email: {a0033593, elewut, eleyangz, elehch}@nus.edu.sg

Xiayun Liu

Xilinx Asia Pacific,
Singapore
Email: liuxy0701@gmail.com

Abstract—An 8-channel wireless neural signal processing IC, which can perform real-time spike detection, alignment, feature extraction, and wireless data transmission is proposed. A reconfigurable BFSK/QPSK TX at MICS band is incorporated to support different data rate requirement. By using Exponential Component-Polynomial Component (EC-PC) spike processing unit with incremental principal component analysis (IPCA) engine, an overall 625× data reduction is achieved. The EC-PC unit is capable of detecting neural spikes of poor SNR. In TX, dual channels at 401 MHz and 403.8 MHz are supported by applying fixed injection and sequential injection locked techniques while attaining phase noise of -102 dBc/Hz@100 kHz offset. The measured EVM of 4.60%/9.55% with PA output power of -15dBm is achieved for QPSK@8 Mbps and BFSK@12.5 kbps. Fabricated in 65 nm CMOS with an area occupation of 1 mm^2, the design consumes a total current of 5∼5.6 mA with maximum energy efficiency of 0.7 nJ/b.

Keywords—EC-PC, neural signal processing, transmitter, injection locking, reconfigurable

I. INTRODUCTION

Recording action potentials, or spikes, fired by neurons in the brain is vital for neuroscience research and clinical treatment of neurological diseases. For example, in neural prostheses, spikes from a particular brain region are required to be identified and collected for information decoding, which are further used to control prosthetic devices. In order to ensure good signal quality and also impose less constraints to the subjects, neural implants are highly preferred to be lightweight, low-power, and wireless to avoid motion artifact and tissue infection caused by tethered connections.

In addition, neural recording and processing with increasing number of channels are desired to establish the causal connections between the neural activities and mental or physical behaviors, which pose extra challenges on 1) the transmission bandwidth due to increased data rate, and 2) the processing delay for information decoding to form closed-looped control. Recently reported wireless neural recording ICs [1], [2] are based on wired solution and performing offline signal processing, thus cannot meet these requirements. In this work, we proposes a wireless neural signal processing IC, which can perform real-time spike detection, alignment and feature extraction for up to 8 channels simultaneously. The system can be configured in different modes to transmit neural codes with different bandwidth requirements. A reconfigurable
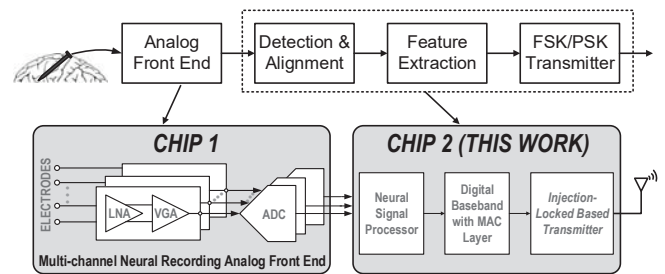

Fig. 1. System diagram of the envisioned wireless neural recording.

BFSK/QPSK TX is also incorporated to cater for different data rate requirement. The carrier frequency has been chosen to operate in MICS band, which has desired characteristic for biomedical and implant applications [3].

Fig. 1 shows a two-chip solution with multi-channel neural recording analog front-end (AFE) and the proposed wireless signal processing IC. The multi-channel neural recording AFE favors older CMOS technology (>0.35 μm) which exhibits higher voltage headroom and lower device flicker noise. On the other hand, advanced CMOS technology (<65 nm) enables area and power efficient solution for the digital signal processing and wireless TX. In this paper, we report the work related to the latter implementation.

Our recent work on neural spike detection based on an EC-PC algorithm [4] has enabled a reliable prediction of spikes in real-time even under poor SNR without any need of calibration on the algorithm parameters. In this work, we adopt the EC-PC decomposition mechanism with the spike alignment and feature extraction to further reduce the data rate for wireless transmission. We have also reported an energy efficient multi-channel reconfigurable TX architecture for this band [5]. In this work, we explore alternative technique to achieve frequency tuning and multi-channel support which is less hardware intensive. Sequential injection locking technique [6] has shown its frequency tuning capability in frequency synthesizer application. In this work, we attempt at incorporating such technique to our proposed TX.

This paper is organized as follows. Section II presents the proposed system architecture, followed by functional blocks in section III. Experimental results are shown in section IV. Finally, section V will conclude our findings.
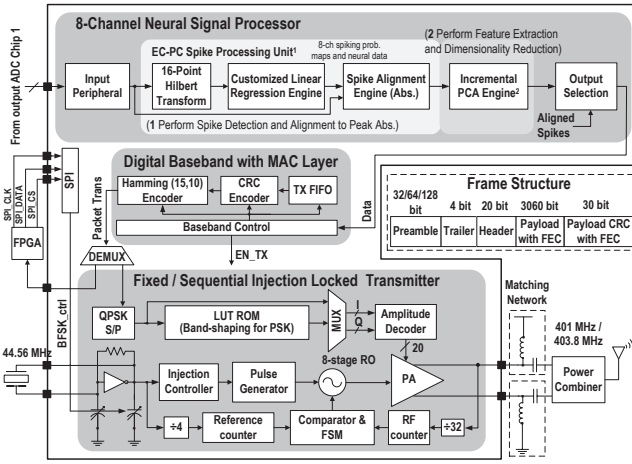
Fig. 2. Proposed wireless neural signal processing architecture.

Fig. 3. Diagrams of spike alignment and feature extraction and corresponding testing results including the 8-channel detected spike events.

II. SYSTEM ARCHITECTURE

As shown in Fig. 2, the digitized neural data from chip 1 will first be sent to the 8-channel neural signal processor. The processor provides two main functions. Firstly, the neural signal will go through an EC-PC spike processing unit which performs the spike detection and alignment. After which, an IPCA engine is employed for feature extraction. It transforms aligned spikes into a feature space with much lower dimensions where the separation of spikes into their originating neurons is much easier. Through the proposed neural signal processor, different levels of data rate reduction can be attained.

For the digital baseband, it has the option of transmitting raw or compressed data. The data will first go through MAC. The proposed data packet consists of preamble, trailer, header, payload with FEC, and payload CRC with FEC, as shown in Fig. 2. The payload length can vary from 10-bit to 2040-bit. The Hamming (15, 10) encoder is employed for FEC. Payload CRC is also incorporated to ensure the data integrity. The packet data is send to the injection-locked based transmitter and modulated using either BFSK or QPSK for transmission.

An 8-stage pseudo-differential Injection Locked Ring Oscillator (ILRO) based TX architecture with direct quadrature amplitude modulation at digital PA is employed. The injection controller is designed to select either fixed injection or sequential injection locked mechanism to apply for different channel transmission and power down unused blocks to improve power efficiency. A differential pulses with width narrower than half of the ILRO's period are generated in pulse generator block for injection mechanism. The four out of 16-phase output from ILRO coupled with 5-bit amplitude modulation from digital PA provides the desired QPSK modulation with band-shaping. For BFSK modulation, crystal frequency pulling is employed [7] to achieve frequency deviation of 88 kHz while selecting one fixed phase output from ILRO. The XO circuit, frequency calibration, digital band-shaping modulator, and digital power amplifier are similar to [5].
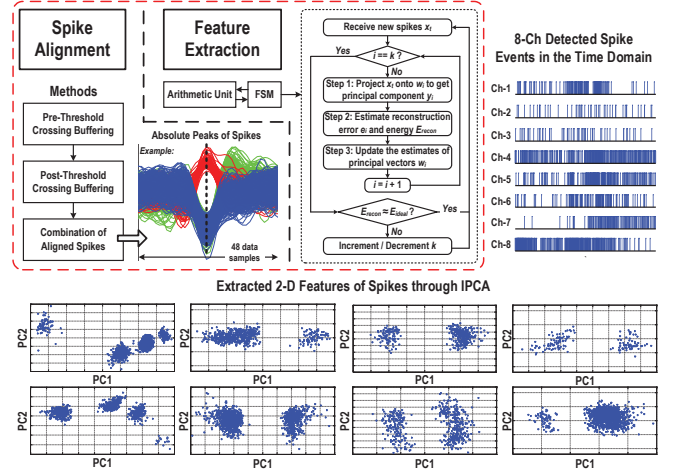
III. FUNCTIONAL BLOCKS AND VLSI IMPLEMENTATION

A. Neural signal processor

Within the EC-PC spike processing unit, the band-limited neural data from each channel are first Hilbert transformed. The output is then normalized with real-time estimated variance to approximate the probability density distributions of neural data. As the distributions consist of EC and PC portions, their parameters can be trained adaptively through a customized linear regression engine [4]. Once the EC and PC are extracted, spiking probability maps per channel can be derived which are used for spike detection.

For spikes detected in one channel, they are aligned to their absolute peaks as shown in Fig. 3. This allows truncation of spikes to only 48 samples centered around peaks, achieving significant data reduction. The channel ID and time stamp will then be incorporated to form spike data frames. Following the EC-PC spike processing unit, an IPCA engine is employed to map the time domain data into a two dimensional feature space to perform feature extraction. This leads to another 24× data reduction after the spike detection. IPCA avoids expensive computation of covariance matrix and can adaptively increment or decrement the number of principle components to achieve fast convergence with minimal reconstruction error. In the implementation, IPCA is designed to work in event driven mode for energy saving, i.e. the mapping is only done after each spike detected. The processing flow for IPCA and its testing results for 8-channel are shown in the Fig. 3.

As an estimation, the data rate of one channel is approximately 200 kbps (8-bit data with 25 kS/s). After EC-PC, only 48 samples are retained per detected spike. Assuming an averaged 20 spikes/s per channel, the data rate would be reduced by 26× to 7.68 kbps. With IPCA on, another 24× is achieved, resulting in an overall 625× data reduction.

B. Injection locked architecture

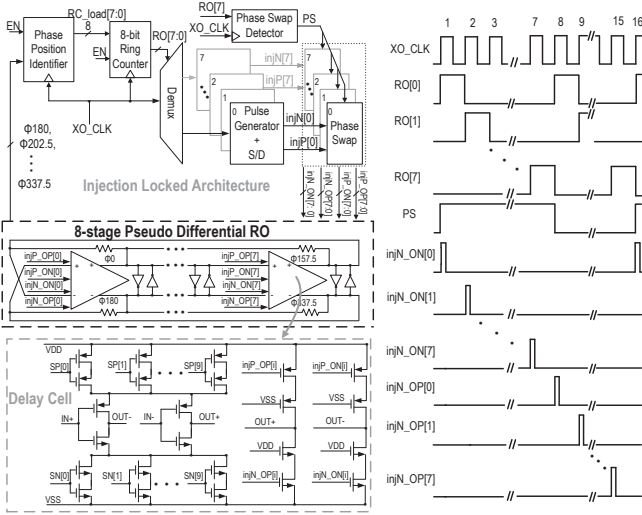From Fig. 4, sequential injection locking is introduced to achieve the desired frequency tuning. Initially, 8 phases (Φ_{180},

Fig. 4. Injection locked architecture with 8-stage pseudo differential ring oscillator.

Fig. 5. Measured phase noise under free running, fixed and sequential ILRO.

$\Phi_{202.5}$, Φ_{225}, $\Phi_{247.5}$, Φ_{270}, $\Phi_{292.5}$, Φ_{315}, and $\Phi_{332.5}$) from ILRO is sent to phase position identifier (PPI) to identify the phase edge closest to the XO. The large time difference between the injected reference and the output phase edge might cause the injection locking to fail [6]. The PPI then initializes the Ring Counter (RC) with the desired flip-flop (FF) setting. The RC output is used to de-multiplex the XO to one of the eight pulse generators (PG). As RC cycled through, the position of its "1" output will shift in sequence, and thus enable the PG in sequence. As each PG output is injected to different delay stage of the ILRO, sequential injection is thus achieved. In our design, 8-bit RC and 8 PG are sufficient for sequential-injection with 16-phase ILRO to reduce the chip area and power consumption. The injected phase will automatic swap the polarity when being injected into the Φ_0 to $\Phi_{157.5}$ delay cells by using the phase swap detector. The operating principle of this technique is illustrated in Fig. 4.

Through sequential injection, output frequency can be changed to $(N+1/16) \times f_{ref}$, where N is the Nth harmonics of the XO injection signal and f_{ref} is the XO frequency [6]. Different sequential injection scheme can be adopted to achieve different output frequency. In this implementation, N of 9 and f_{ref} of 44.56 MHz are chosen to achieve MICS band output. For fixed injection locking, PPI and RC are powered down. Only one fixed PG is selected without any phase swap. The resulting output frequency is $N \times f_{ref}$.

IV. Experimental Results

In this implementation, fixed injection and sequential injection are employed to generate output frequency of 401 MHz and 403.8 MHz respectively. As shown in Fig. 5, the injection locking technique improves the free running ILRO phase noise by at least 25 dB at 100 kHz offset. Both fixed and sequential injection exhibit similar phase noise performance. The QPSK output spectrum and EVM are shown in Fig. 6. A full MICS
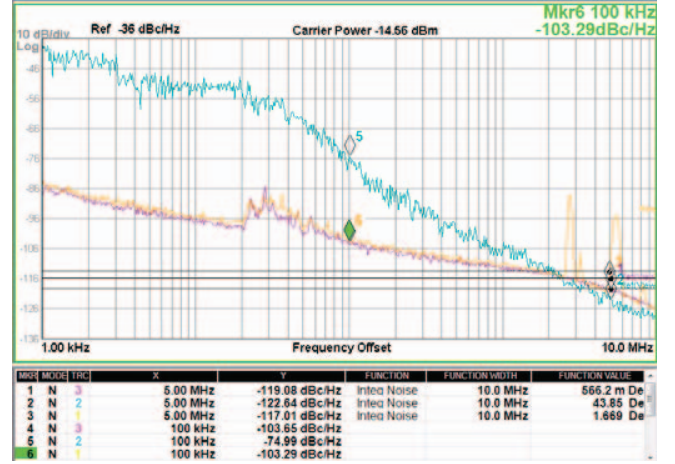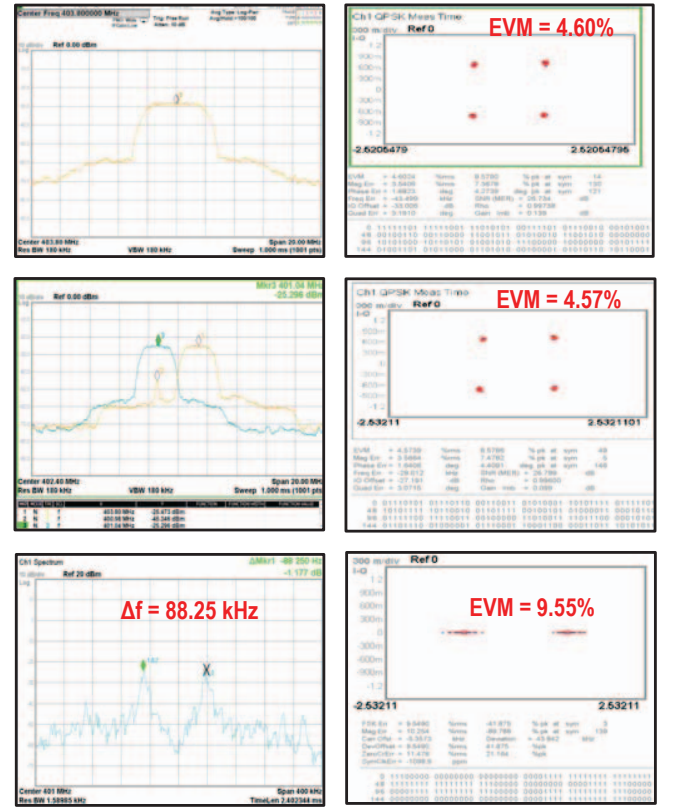

Fig. 6. Measured output spectrum and constellation diagram for band-shaped QPSK full MICS band at 8 Mbps, with two frequency channels at 4 Mbps and BFSK at 401 MHz.

band is fully occupied with 8 Mbps QPSK and achieves EVM of 4.60%. Alternatively, dual channel can be supported at 401 MHz and 403.8 MHz respectively with each channel having a data rate of 4 Mbps and EVM of 4.57%. As illustrated, due to band-shaping, QPSK output spectrum achieves ACPR closed to -30 dB. This is sufficient for raw neural data transmission without any compression. Sequential injection modulates the

TABLE I
PERFORMANCE COMPARISON

	[1]	[2]	[3]	[8]	**This Work**
Data Rate (Mbps)	0.708	0.1	90	1.5	8 (QPSK–1 Ch) 4 (QPSK–2 Ch) 0.0125 (BFSK)
Neural Signal Processor	No	No	Yes	Yes	Yes
No. of Channels	32	1	128	64	8*
TX Channel Selection	No	No	No	No	Yes
P_{DC} (mW)	3.3	0.4	6	3.7~6.6	5.6 (QPSK) 5.0 (BFSK)
VDD (V)	3	1	±1.65	1.2	1
Energy Eff. (nJ/b)	4.7	4	0.00178	2.5~4.4	0.7 (QPSK–1 Ch) 1.4 (QPSK–2 Ch) 400 (BFSK)
Frequency Carrier (MHz)	915/854.5 or 915/877	402/433	3000~5000	915	401~406
Modulation	FSK/OOK	FSK	UWB	FSK/OOK	QPSK/FSK
TX power (dBm)	-22	-16	-23**	-20~0	-15
Phase Noise (dBc/Hz)	NA	NA	NA	-96 dBc/Hz@100 kHz	-102 dBc/Hz@100 kHz
Integrated RMS jitter (10 MHz Bandwidth)	NA	NA	NA	1.06°	0.57° (fixed inj.) 1.67° (sequential inj.)
Process	0.5 μm	0.13 μm	0.35 μm	0.13 μm	65 nm
Active Area (mm^2)	16.4	2.5	63.36	12	1
FOM*** nJ/(bit·mW)	738.7	159.2	3.48	4.4	22.1

* AFE is not included in this design. ** Estimate from measured UWB pulse. *** $FOM = P_{DC}/(DataRate \times P_{out})$.
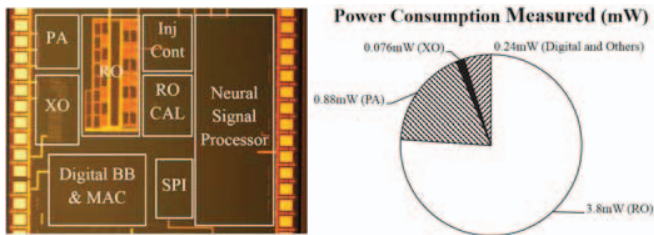

Fig. 7. Die micrograph and power consumption breakdown.

delay cell periodically which causes the spurs. For BFSK, 88.25 kHz deviation is measured based on frequency pulling. It achieves 9.55% EVM at 12.5 kbps, which is sufficient for the transmission of neural data with the highest data reduction.

The wireless neural signal processing IC, fabricated in 65 nm CMOS, occupied 1 mm^2 and consumes a total current of 5~5.6 mA. The benchmarking with the recent art is shown in Table I. The first 2 works do not consist of neural signal processor. [8] only performs FIR filtering on signals. [3] employs Non-linear Energy Operator algorithm for spike detection which does not work well with neural signal with poor SNR. In addition, PCA is used instead of IPCA for feature extraction, which requires the definition of covariance matrix and thus is more memory and hardware intensive. Our work performs real-time spike detection, alignment and feature extraction to achieve 625× data reduction. Through reconfigurable TX with dual channel support, different data rate can be catered using different modulations. Fig. 7 shows the die photo together with the power breakdown based on BFSK.

V. CONCLUSION

A 65 nm CMOS wireless neural signal processing IC with reconfigurable BFSK/QPSK transmitter for neural recording system is presented. The proposed signal processor architecture applies spike alignment and feature extraction together with the adopted EC-PC decomposition technique able to reduce the neural data by 625×. Sequential injection locked method is adopted to perform dual-channel transmission in MICS band with data rate up to 8 Mbps. The total power dissipation is 5~5.6 mW from a 1 V supply.

REFERENCES

[1] S. B. Lee and et al., "An inductively powered scalable 32-channel wireless neural recording system-on-a-chip for neuroscience applications," *ISSCC Dig. Tech. Papers*, pp. 120–122, 2010.

[2] S. Rai and et al., "A 500μw neural tag with 2μv rms afe and frequency-multiplying mics/ism fsk transmitter," *ISSCC Dig. Tech. Papers*, pp. 212–213, 2009.

[3] M. Chae and et al., "A 128-channel 6mw wireless neural recording ic with on-the-fly spike sorting and uwb tansmitter," *ISSCC Dig. Tech. Papers*, pp. 146–147, 2008.

[4] T. Wu and et al., "A 16-channel nonparametric spike detection asic based on ec-pc decomposition," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–15, Mar. 2015.

[5] X. Liu and et al., "A 103 pj/bit multi-channel reconfigurable gmsk/psk/16-qam transmitter with band-shaping," *Proc. ASSCC*, pp. 269–272, 2014.

[6] P. Park and et al., "An all-digital clock generator using a fractionally injection-locked oscillator in 65nm cmos," *ISSCC Dig. Tech. Papers*, pp. 336–337, 2012.

[7] J. Pandey and B. P. Otis, "A sub-100 w mics/ism band transmitter based on injection-locking and frequency multiplication," *IEEE J. Solid-State Circuits*, vol. 46, no. 5, pp. 1049–1058, May 2011.

[8] K. Abdelhalim and et al., "915-mhz fsk/ook wireless neural recording soc with 64 mixed-signal fir filters," *IEEE J. Solid-State Circuits*, vol. 48, no. 10, pp. 2478–2493, Oct. 2013.