The 5th Iranian International Conference on Microelectronics (IICM 2023)

October 25-26, 2023

2023 5th Iranian International Conference on Microelectronics (IICM) | 979-8-3503-6019-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/IICM60532.2023.10443192

An Asynchronous Strategy for Efficient Audio Processing for Better Perception in Cochlear Implants Based on Peak and Trough Detection

Amin Armin, Mohammad Yavari and Amir Kashi

Integrated Circuits Design Laboratory, Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), P.O. 15875-4413, Tehran 15914, Iran.

Emails: aminarmin77@aut.ac.ir, myavari@aut.ac.ir, kashi@aut.ac.ir

***Abstract*— In this paper, we propose a novel algorithm for processing acoustic signals in cochlear implants that is asynchronous and low power. Our algorithm stimulates the electrodes based on the frequency of the input sound signal and identifies its peaks and troughs. The algorithm behaves in such a way that until a peak or trough is identified, the subsequent blocks do not consume power. For example, the analog-to-digital converter is turned off until a peak or trough is detected. This ensures that the power consumption is proportional to the input sound signal and significantly lower compared to algorithms that continuously operate the entire system. Our approach has the potential to improve the perception of sound in cochlear implants and reduce power consumption, which are both important considerations in this field. We present simulation results to demonstrate the effectiveness of our algorithm in processing acoustic signals for cochlear implants.**

Keywords—Asynchronous Strategy, Cochlear Implant, Low Power, Temporal Fine Structure (TFS), Peak and Trough Detection, Sound signal processing, Amplitude and phase extraction, Sound perception.

I. INTRODUCTION

According to the latest announcement by the World Health Organization (WHO), approximately 7% of the world's population (500 million people) suffer from hearing loss. Also, according to the WHO, World Hearing Report, published on March 2, 2021, it warns that nearly 2.5 billion people worldwide - or one in four people - will live with some degree of hearing loss by 2050. At least 700 million of these people will need access to hearing care and other rehabilitation services unless action is taken [1].

Hearing loss affects both children and adults and requires various treatment approaches. While hearing aids are effective for many, individuals with severe hearing loss may still struggle even with aids. Cochlear implantation has been recommended for these individuals who cannot resolve their hearing problems with hearing aids [2]. Candidates for cochlear implants are those who do not experience significant improvement in hearing after using hearing aids for three to six months. Cochlear implants work by directly stimulating the auditory nerve to provide a sense of sound, offering an alternative pathway for damaged inner ears. Although not a cure for hearing loss, cochlear implants enable sound perception.

The signal processing block is one of the most important parts of the cochlear implant system, responsible for receiving information from the microphone, frequency division, processing, and ultimately sending it to the auditory nerve fibers [3]. Various ways and algorithms for audio signal processing for better perception have been proposed in recent years. Our study introduces a new algorithm for processing acoustic signals in cochlear implants. The algorithm is designed to be asynchronous and low-power, operating by detecting peaks and troughs in the input sound signal and stimulating the electrodes based on its frequency. By only consuming power when necessary, such as when a peak or trough is identified, our algorithm is able to significantly reduce power consumption compared to traditional algorithms that continuously operate the system.

The structure of this article is as follows. Section II offers a concise overview of the cochlear implant signal processor and outlines some of the previous strategies employed in this domain. Section III provides a detailed explanation of the proposed signal processor strategy. The results of simulations are shown in Section IV. Finally, section V presents the study's conclusion and offers concluding remarks.

II. REVIEW OF STRATEGIES FOR SIGNAL PROCESSING IN COCHLEAR IMPLANT

When broadband sounds, such as speech or music, enter the cochlea, they are filtered into a series of narrowband signals. Each of these signals has a slowly varying envelope (ENV), which represents the amplitude of the sound signal, and a rapidly oscillating carrier, known as the temporal fine structure (TFS). The timing and rate of action potentials in the auditory nerve convey the information about the ENV and TFS of the sound signal to the brain [4]. Extracting the ENV from sound is crucial for better hearing and understanding the intensity of sound. By extracting the envelope from the sound, an individual can understand more details of the sound and have a better understanding of speech and sound intensity.

Improving the extraction of TFS information in cochlear implants (CIs) can significantly impact the ability of individuals with hearing loss to perceive sounds in noisy environments and distinguish between competing voices [5][6]. Additionally, extracting TFS information is important for music perception and tonal languages. Individuals with CIs may have difficulty perceiving and appreciating music due to the limited number of electrodes used to stimulate the auditory nerve, resulting in a loss of pitch information. By improving the extraction of TFS information in CIs, researchers aim to provide a more accurate representation of music and improve music perception for individuals with hearing loss [7][8].

Several algorithms have been developed over the years, each with different advantages and limitations. One of the most well-known and widely used algorithms is the continuous interleaved sampling (CIS) strategy. The CIS method is a simple and efficient algorithm that uses a small number of electrodes to stimulate the auditory nerve. Instead of transmitting the original signal, this algorithm extracts the envelope of the signal and delivers it to the electrodes with a fixed pulse rate [9]. While the CIS method has been successful in providing improved speech perception for many individuals with hearing loss, it has limitations in representing the TFS of sound signals.

In order to enhance the ability of a bionic ear processor to transmit not only the speech signal envelope but also the phase and TFS to the auditory nerve fibers, several speech processing techniques have been suggested for use in the Bionic Ear (BE) processor. One of the signal processing strategies used in cochlear implants is the phase-locking zero-crossing detection (PL-ZCD) strategy. PL-ZCD uses a phase-locking algorithm to detect the phase of the incoming sound waveform at the zero-crossing points, which is then used to generate stimulation pulses synchronized with the sound waveform [10]. The advantages of PL-ZCD include improved temporal coding, reduced energy consumption, and a simple algorithm. However, PL-ZCD has limitations, as it is sensitive to noise, which can cause severe performance degradation [11], and it is not always clear what the zero-crossing point represents in terms of phase [12].

The other strategy that attempted to extract phase information is named frequency amplitude modulation encoding (FAME). The FAME strategy aims to enhance ENV and TFS separately and then combine them to provide a more natural and complete representation of sound for cochlear implant users [6]. Although phase extraction plays a crucial role in enhancing sound perception within the FAME strategy, its implementation through a separate pathway independent from domain extraction poses certain drawbacks. One notable limitation is the increased demand for processing power, making it less viable for devices with limited power capabilities. Therefore, while phase extraction enhances sound quality, alternative approaches should be considered for power-constrained devices to ensure optimal performance.

Another approach involves extracting phase information by identifying peak and trough points within the signal. This technique, known as the recording peak/trough technique, allows for the conveyance of more instantaneous frequency information beyond just the fundamental frequency to the stimulation electrodes [13].

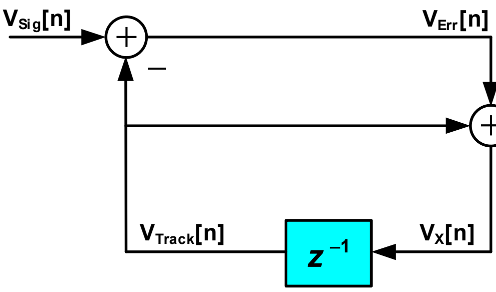

Fig. 1: Error Detection Block.

An example of this approach is phase-locked sampling (PL-PP) combined with peak instantaneous detection (PID) [14]. Additionally, Guo et al. presented an alternative implementation of this method where amplitude information is delivered to the electrodes precisely at the peaks and troughs of the signal [12]. One advantage of this method is that it enables simultaneous extraction of both phase and amplitude information by utilizing a peak-triggered sampling scheme. Since the amplitude information is embedded at the peak points, this approach offers the benefit of achieving phase and amplitude extractions simultaneously, enhancing the overall representation of the sound signal [12].

This article focuses on the peak/trough detection method that is considered the best strategy for providing superior TFS information, resulting in an enhanced perception experience. Compared to other strategies, this method has fewer drawbacks and is widely regarded as the optimal choice for peak picking. In the following section, we will delve into the details of the strategy and how it can be effectively implemented.

III. Proposed Signal Processing Strategy

As highlighted in the previous section, the algorithm that will be presented below places considerable importance on the peak picking strategy. This method boasts several benefits, including its ability to extract TFS in a superior manner, leading to an improved perception experience. Moreover, the strategy enables the simultaneous extraction of amplitude and phase, making it a highly efficient approach. Furthermore, the Zero Crossing block has been used as one of the simple and low-power consumption blocks in this algorithm for implementation purposes.

Before delving into the strategy and its functioning, it's important to take a closer look at Fig. 1, which is a crucial component of the proposed algorithm's circuit. Let's assume that the sound signal is sampled and recorded by a microphone with sampling frequency of 200 kHz and is then transmitted to the circuit's input as V_{sig}[n]. It is worth noting that the V_{Track}[n] signal is identical to the input V_{sig}[n] signal, but with a delay of one time cycle, as we will demonstrate below. This circuit is responsible for comparing the current signal with the previous cycle and generating the error value, V_{Err}[n], between these two signals, which is then sent to the output. Additionally, V_{Err}[n] is used to construct the previous cycle of the signal. In order to create the delayed signal, we can represent the mathematical equations for each node as follows. The equation at the V_{Err}[n] node is as

$$V_{Err}[n]=V_{Sig}[n]-V_{Track}[n] \tag{1}$$

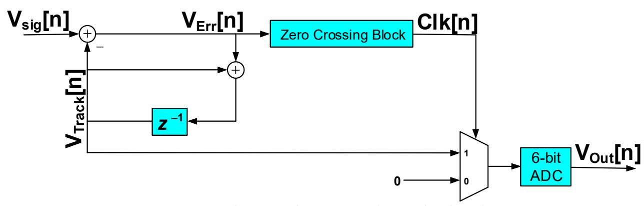

Fig. 2: Proposed Speech Processing Block Diagram.

Also, the equation at the V_X[n] node can be written as

$$V_X[n] = V_{Err}[n] + V_{Track}[n] \tag{2}$$

And the equation for the final node is as

$$V_{Track}[n] = V_X[n-1] \tag{3}$$

By inserting (2) in (3), the following equation is obtained by

$$V_{Track}[n] = V_{Err}[n-1] + V_{Track}[n-1] \tag{4}$$

By substituting the values from (1) into (4) and simplify it, it is denoted as

$$V_{Track}[n] = V_{Sig}[n-1] \tag{5}$$

Equation (5) demonstrates that the V_{Track}[n] is equivalent to the delayed V_{Sig}[n].

To summarize the circuit thus far, it subtracts the current signal from the signal of the previous stage, and then sends the resulting difference as V_{Err}[n] to the next stage. Furthermore, the same V_{Err}[n] is utilized to generate the delayed signal. Fig. 2 is the block diagram of the strategy based on the Peak Picking algorithm. As depicted, the zero-crossing block is employed to alter the sign of the V_{Err}[n] signal. To comprehend how the strategy functions, let us examine a sinusoidal signal as depicted in Fig. 3. As illustrated, during the ascending trend of the signal, each sample exceeds its previous sample, and conversely, during the descending trend, each sample is lower than its previous

Based on the definition of V_{Err}[n], which is the difference between the current signal and the previous stage, the error rate becomes positive during the ascending trend and negative during the descending trend. Consequently, if we observe a change in the sign of V_{Err}[n] from positive to negative, it indicates the point at which the input signal has reached its peak. Conversely, if the sign of V_{Err}[n] changes from negative to positive, it indicates the trough point.

Based on the explanations provided regarding Fig. 3, let us revisit the circuit diagram in Fig. 2. The audio signal, V_{Sig}[n], is fed into the input of the circuit, and then V_{Err}[n], which is the difference between the current signal and the previous stage, is obtained from the input signal. In the subsequent step, the zero crossing circuit is employed to detect whenever the sign of V_{Err}[n] changes. Thus, whenever the input reaches its peak which V_{Err}[n] sign changes from positive to negative, the output of the zero crossing circuit changes from Low to High for one cycle and then returns to Low. The same process is repeated for the trough points, where the sign of V_{Err}[n] changes from negative to positive. In summary, the output of the zero crossing block changes with each peak and trough detection.

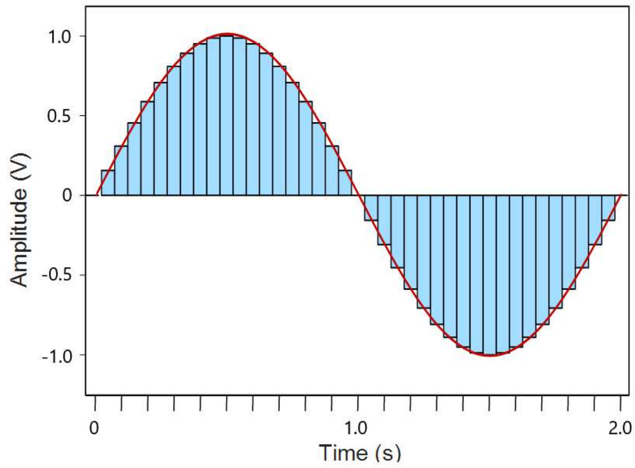

Fig. 3: Sine Function with Oversampling.

When output of the zero crossing block is High, the signal amplitude value of the previous stage, which is the maximum value in the peak and the minimum value in the trough, is transmitted to the next stage through a switch or multiplexer. Subsequently, the A/D digitizing task is initiated to process the input signal. The circuit operates in such a way that when the input signal is not at the peak or trough points, the A/D input is disabled to minimize power consumption.

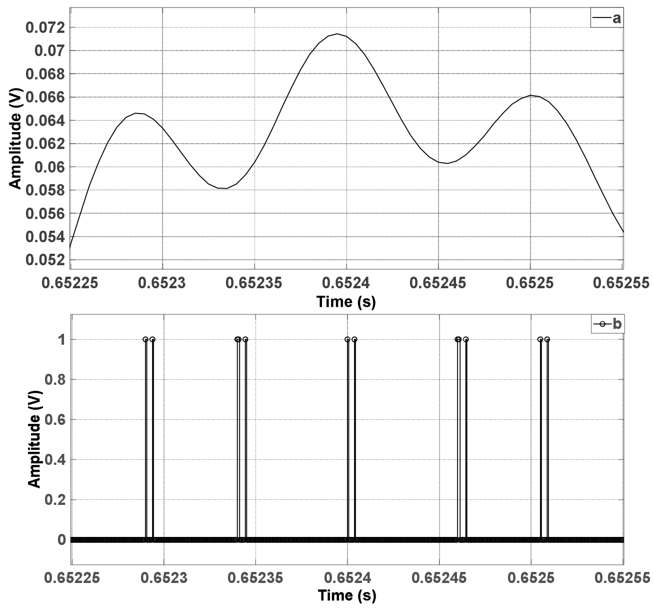

Fig. 4: (a) Part of Audio Input Signal and (b) Output of Proposed Strategy.

To summarize how this strategy works, it can be said that the circuit obtains the slope of the audio signal with a sampling frequency of 200 kHz by calculating the difference between the current signal and the previous stage. Whenever the slope of the signal changes, indicating that a peak or trough has been reached, the zero crossing block is utilized to detect the change. Subsequently, the amplitude value of the signal at the peak or trough is transmitted to the A/D input for one cycle. Until a peak or trough is reached, the circuit after the switch will not operate, which leads to a significant reduction in power consumption. As a result, the power consumption of the subsequent stages is only utilized when necessary, i.e., during the peak or trough areas. Thus, the power consumption varies with the input frequency, and the circuit consumes power only when required.

IV. System Simulation Resaults

To evaluate the effectiveness of the proposed approach described in Fig. 2, the Simulink MATLAB software was used for implementation. A sound signal containing a spoken sentence was selected as the input test signal, and it was sampled at a frequency of 200 kHz.

The results of the proposed strategy are presented in Fig. 4, where part (a) displays a section of the input sound waveform that contains multiple peaks and troughs. In part (b), the output waveform of the proposed approach is shown, which accurately identifies the peaks and troughs. To further analyze the results, Fig. 5 displays an enlarged view of the first trough in Fig. 4. The amplitude of the input signal at the moment of the trough is 0.06v, which corresponds to the digital value of 1 in 6 bit. Fig. 5, part (b) has been enlarged more in the time axis to better visualize the algorithm output. The value of the digitalized signal in this magnified section is serialized and encoded as 11000001, where the first bit indicates the start of sending the amplitude, and the second bit indicates whether the strategy detects a peak (if it is 0) or a trough (if it is 1). The next 6 bits also indicate the range of the detected signal, which is 1 as it should be. The encoded value, consisting of 8 bits, is sent to the next stage in a time cycle with a frequency of 200 kHz. This value is then utilized to stimulate the corresponding electrodes in the subsequent steps.

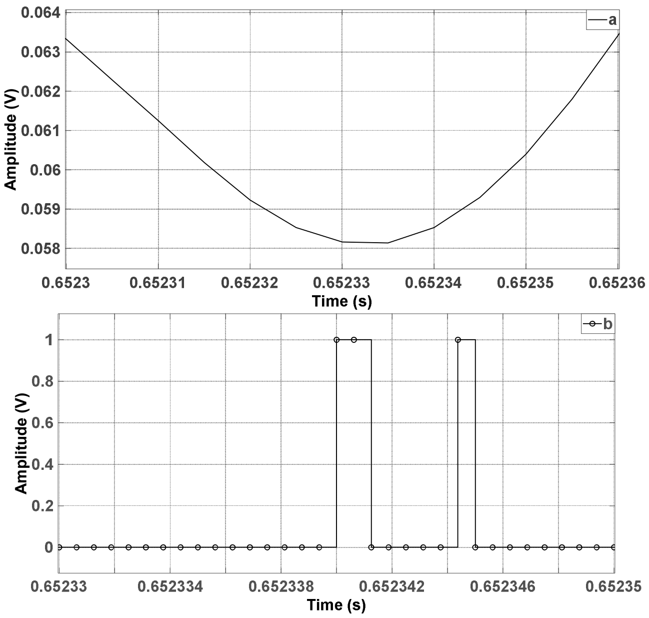

Fig. 5: (a) Magnification of the First Trough in Fig. 4 and (b) The Output of the Proposed Strategy in Selected Trough.

Guo et al. proposed an event-driven strategy based on the ADM algorithm in [12]. Guo et al. introduced an event-driven strategy based on the ADM (Amplitude Demodulation) algorithm in their publication [12]. This article represents a significant contribution to the field of signal processing in ear implant applications, as it aims to achieve optimal phase extraction with a focus on accuracy and speed. Recognized as one of the most up-to-date references in the field, the ADM strategy serves as a valuable benchmark for comparison with the strategy proposed in this conference paper.

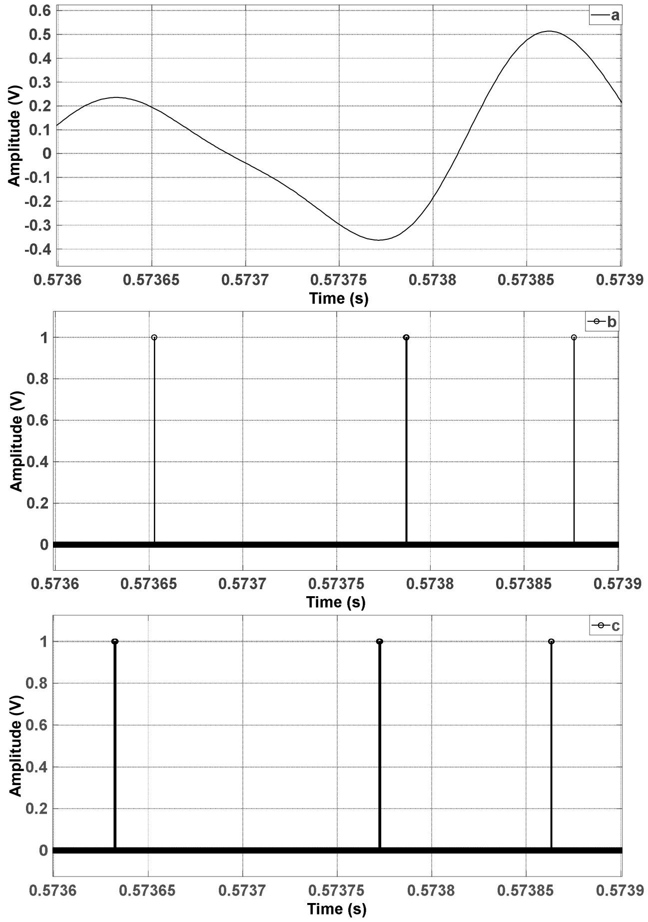

Fig. 6: (a) Part of Audio Input Signal and (b) Output of ADM Strategy and (c) Output of Proposed Strategy.

Fig. 6, part (a) displays an enlarged portion of an input sound signal. The outputs of the ADM strategy and the newly proposed strategy can be seen in parts (b) and (c), respectively. It is evident that the ADM strategy experiences significant delay compared to our strategy. Additionally, when the ADM strategy detects a peak or trough point, it sends a different amplitude value to the electrodes than the original input signal. In the ADM strategy, the phase error, known as $err_{ph}=(t_{polarity}-t_{peak})\times f_{input}$, is defined as the error in detecting the time of the peak or trough. However, in our proposed strategy, this error is highly optimized and nearly eliminated, resulting in more accurate peak and trough detection. As a result, our strategy outperforms the ADM strategy in terms of accuracy and precision.

Fig. 7 presents another section of the input audio signal, which is different from the previous analysis. In the ADM strategy, it is known that peaks and troughs that differ by less than 1 LSB (for a 6-bit counter) cannot be detected. In part (a) of Fig. 7, the amplitude of the input audio signal changes by less than 1 LSB. Part (b) shows the output of the ADM strategy, which only detects one of the peaks and fails to identify one peak and two intermediate troughs. As a result, the ADM strategy fails to accurately perceive this section of the sound. However, our proposed strategy successfully detects all the peaks and troughs with high accuracy, and transmits them to the electrodes.

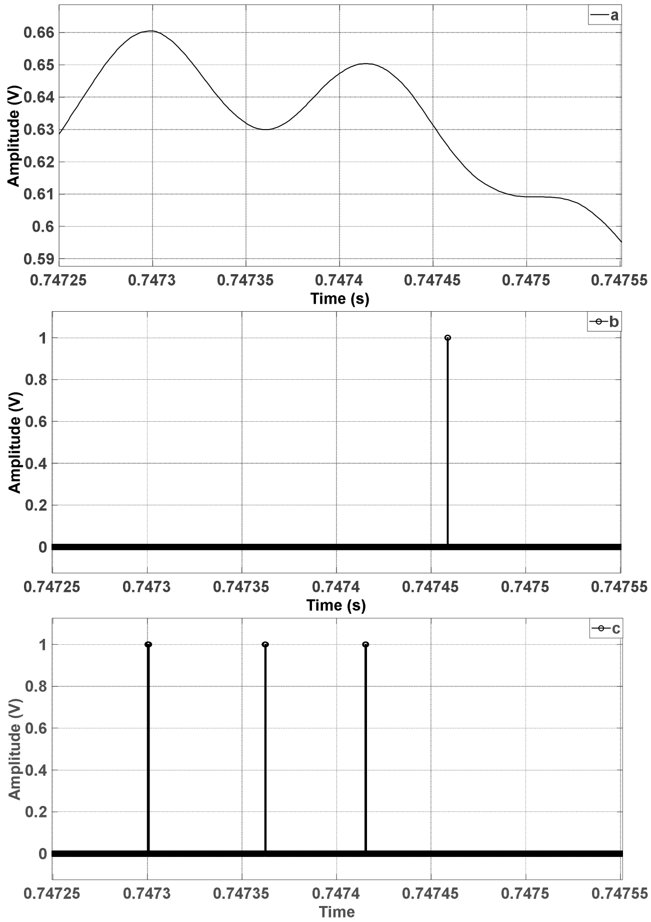

Fig. 7: (a) Shows a Section of the Audio Input Signal and (b) Output of ADM Strategy and (c) Output of Proposed Strategy.

Table I presents a qualitative comparison of various strategies. The strategies were evaluated based on their performance in phase extraction within the phase extraction line. It is worth noting that some strategies employ the amplitude envelope of the signal instead of the amplitude of the main signal. To assess the accuracy of the range, the difference between this range and the range of the original audio signal was measured and reported in the last line of the table. Our evaluation indicates that strategies employing signal push mechanisms exhibit low amplitude accuracy. Conversely, [12], which encounters a phase error, demonstrates medium accuracy. Our approach, which excels in peak and trough mining, is considered a high-accuracy strategy based on this parameter. Overall, this analysis provides valuable insights into the performance and characteristics of various signal processor strategies, emphasizing the significance of simplicity, power consumption, and accuracy in amplitude and phase extraction for optimal results.

TABLE I: COMPARISON BETWEEN SIGNAL PROCESSING STRATEGIES.

Reference	[9]	[15]	[10]	[14]	[12]	This work
complexity	Med*	High	Low	Low	Med*	Low
Sampling	Sync	Async	Sync	Sync	Async	Async
Phase Extraction	Very Low	High	Low	Med*	High	High
Power	Med*	High	Very Low	Low	Med*	Low
Amplitude Accuracy	Low	High	Low	High	Med*	High

*Med: abbreviation of Medium

V. CONCLUSION

This study proposes a novel strategy based on Peak Picking. The approach offers several advantages, including the simultaneous extraction of amplitude and phase information to better perception by recognizing the frequency of the input signal from the peaks and troughs. Additionally, the strategy utilizes simple circuits which reduces power consumption. Asynchronous timing of electrode stimulation is also implemented, which makes use of time cues of phase information. The stand-by capability of the strategy allows certain blocks, such as Analog-to-Digital Converter (ADC), to be deactivated when electrodes do not require stimulation. This effectively reduces power consumption by eliminating unnecessary operations and conserving energy. Overall, the proposed strategy offers several improvements over existing techniques and has the potential to advance the field of auditory prostheses.

REFERENCES

[1] https://www.who.int/ (Access Date: Mar. 3, 2023.)

[2] W. A. Yost, and D. W. Nielsen, *Fundamentals of hearing*: Academic Press New York, 2000.

[3] B. S. Wilson, C. C. Finley, D. T. Lawson, R. D. Wolford, D. K. Eddington, and W. M. Rabinowitz, “Better speech recognition with cochlear implants,” *Nature,* vol. 352, no. 6332, pp. 236-238, Jul. 1991.

[4] B. C. J. Moore, "The roles of temporal envelope and fine structure information in auditory perception," Acoustical Science and Technology, vol. 40, no. 2, pp. 61-83, Mar. 2019

[5] F.-G. Zeng, S. Rebscher, W. Harrison, X. Sun, and H. Feng, “Cochlear implants: system design, integration, and evaluation,” *IEEE reviews in biomedical engineering,* vol. 1, pp. 115-142, Nov. 2008.

[6] K. Nie, G. Stickney, and F.-G. Zeng, “Encoding frequency modulation to improve cochlear implant performance in noise,” *IEEE transactions on biomedical engineering,* vol. 52, no. 1, pp. 64-73, Dec. 2004.

[7] F.-G. Zeng, “Cochlear implants in China,” *Audiology,* vol. 34, no. 2, pp. 61-75, Jan. 1995.

[8] L. Bruns, D. Mürbe, and A. Hahne, “Understanding music with cochlear implants,” *Scientific reports,* vol. 6, no. 1, pp. 32026, Aug. 2016.

[9] R. Sarpeshkar, C. Salthouse, J.-J. Sit, M. W. Baker, S. M. Zhak, T.-T. Lu, L. Turicchia, and S. Balster, “An ultra-low-power programmable analog bionic ear processor,” *IEEE Transactions on Biomedical Engineering,* vol. 52, no. 4, pp. 711-727, Mar. 2005.

[10] R. Sarpeshkar, M. Baker, C. Salthouse, J.-J. Sit, L. Turicchia, and S. Zhak, "An analog bionic ear processor with zero-crossing detection." *Proc. IEEE Int. Solid-State Circuits Conf.,* pp. 78-79, Feb. 2005.

[11] J.-J. Sit, A. M. Simonson, A. J. Oxenham, M. A. Faltys, and R. Sarpeshkar, “A low-power asynchronous interleaved sampling algorithm for cochlear implants that encodes envelope and phase information,” *IEEE Transactions on Biomedical Engineering,* vol. 54, no. 1, pp. 138-149, Dec. 2006.

[12] N. Guo, S. Wang, R. Genov, L. Wang, and D. Ho, “Asynchronous Event-driven Encoder With Simultaneous Temporal Envelope and Phase Extraction for Cochlear Implants,” *IEEE Transactions on Biomedical Circuits and Systems,* vol. 14, no. 3, pp. 620-630, Apr. 2020.

[13] C. Sawigun, W. Ngamkham, and W. A. Serdijn, "Comparison of speech processing strategies for the design of an ultra low-power analog bionic ear." *Proc.IEEE EMBC*, pp. 1374-1377, Aug. 2010.

[14] C. Sawigun, W. Ngamkham, and W. A. Serdijn, "An ultra low-power peak-instant detector for a peak picking cochlear implant processor." *IEEE Conf. Biomed. Circuits Syst.*, pp. 222–225, Nov. 2010.

[15] J.-J. Sit, A. M. Simonson, A. J. Oxenham, M. A. Faltys, and R. Sarpeshkar, "A low-power asynchronous interleaved sampling algorithm for cochlear implants that encodes envelope and phase information," *IEEE Transactions on Biomedical Engineering,* vol. 54, no. 1, pp. 138-149, Dec. 2006.