Interspeech 2024
1-5 September 2024, Kos, Greece

DBD-CI: Doubling the Band Density for Bilateral Cochlear Implants

Mingyue Shi[1], *Huali Zhou*[1], *Qinglin Meng*[2], *Nengheng Zheng*[1]

[1]Guangdong Provincial Key Laboratory of Intelligent Information Processing, College of Electronics and Information Engineering, Shenzhen University, Shenzhen, 518052, China
[2]Acoustics Lab., School of Physics and Optoelectronics, South China University of Technology, Guangzhou, 510641, China

`mengqinglin@scut.edu.cn, nhzheng@szu.edu.cn`

Abstract

Cochlear implants (CIs) have limited spectral resolution due to a limited number of electrodes and inter-electrode current interaction, causing difficulties in speech-in-noise perception even for bilateral CI (BiCI) users. Previous studies have suggested alternately stimulating odd and even electrodes between the two sides to reduce current interaction in several ways, showing promising results if dichotic stimuli could be effectively integrated. To utilize the total electrode number of a pair of BiCIs, we propose Doubling the signal analysis Band Density (DBD) and encoding odd and even bands alternately for each side. Two preliminary vocoder-simulation experiments in spectral-temporally modulated ripple discrimination and speech-in-noise perception were carried out to compare DBD with the default setting of the advanced combination encoder (ACE) strategy. The proposed method showed promising benefits as well as limitations to be further resolved in theory and evaluated in BiCI users.

Index Terms: Cochlear implant, coding strategy, spectral resolution

1. Introduction

Cochlear implant (CI) can partially restore hearing perception in people with severe to profound sensorineural hearing loss [1]. It generates electrical signals that stimulate the auditory nerve fibers through 12-24 electrodes placed along the cochlea, replacing the damaged inner hair cell function [2][3]. Although current technology is sufficient to understand speech in quiet [4], speech perception in noisy environments is still a great challenge even for bilateral CIs (BiCIs) [5, 6, 7, 8]. The primary reason is the limited spectral resolution, which is attributed to the limited number of electrodes and the broad frequency range covered by each electrode. Significant current interaction among electrodes further exacerbates this limitation [9].

In most, if not all current BiCIs, signal processing operates independently on each side, without any exchange or integration of information between the two ears. In the literature, certain studies have endeavored to establish connections between the two sides, with the aim of further optimizing binaural cues or signal-to-noise ratio to enhance speech perception in noise. Among them, a series of methods have been proposed to stimulate the odd and even electrodes alternately and/or alternate between the left and right implants [10][11] [12]. Their shared hypothesis is that inter-electrode interaction could potentially be mitigated by increasing the distance or gap between electric pulses generated in the cochlea.

Specifically, in [10] and [11] the input spectrum was allocated to bands in the same way as a unilateral CI and alternately distributed between both ears (odd-indexed electrodes fired at one ear, even-indexed electrodes fired at the other) to increase the spacing between stimulation sites (less channel interaction) for each ear. Experiments showed that only a minority of BiCI users experienced improvements in consonant recognition and sentence in noise, but most did not experience consistent benefits [10] or even faced negative effects [11]. Aronoff et al. [12] suggested that one reason for this disadvantage was the uncontrolled place pitch matching between the two sides. Their study indicated that perceptually aligning the two ears may be necessary to optimize the performance of the alternate stimulation strategies left-right. Wohlbauer et al. [13] further extended the alternate strategy to an InterlACE strategy. InterlACE can activate all available electrodes for one ear by selecting even and odd electrodes in consecutive stimulation frames. This means that, for one ear, stimulation occurs with odd-indexed electrodes at one moment and even-indexed electrodes at the next moment. Experiments showed no significant improvement for InterlACE in the tasks of speech intelligibility and spectral ripple discrimination compared to the clinically used Advanced Combination Encoder (ACE) strategy in Cochlear® CI processors [14], although improvement for individual subjects in speech intelligibility was argued to be promising. In summary, previous efforts have primarily centered on alternately stimulating the left and right ears in an attempt to minimize inter-electrode current interactions.

In this study, we propose a novel alternately stimulating strategy, where the input spectrum is allocated to $2M$ bands and then the bands are alternately distributed to left and right implants each with M electrodes. In this way, this strategy Doubles the signal analysis Band Density (DBD) compared to the default density in a unilateral CI, so the strategy name is abbreviated as DBD. This idea stems from the fact that BiCIs have twice the number of electrodes compared to unilateral CIs. Current clinical BiCI strategies, such as the bilateral ACE, analyze input signal into M bands—equivalent to the number of electrodes in a unilateral implant. In theory, if dichotic information is appropriately integrated, DBD has the potential to improve spectral resolution for BiCI users. The DBD and ACE strategies for bilateral CIs are described in Sec. 2, and their performance was compared in two spectral resolution-related experiments with vocoder-simulated CIs as reported in Sec. 3. Promising results, DBD's limitations, and potential future work are discussed in Sec. 4.

2. DBD Strategy

Block digarms of the proposed DBD and the ACE strategy are presented in Figure 1 (a). In the ACE strategy, signals are processed in parallel by BiCIs, with the frequency bins for each ear grouped into M frequency bands (M is the number of elec-

10.21437/Interspeech.2024-1505

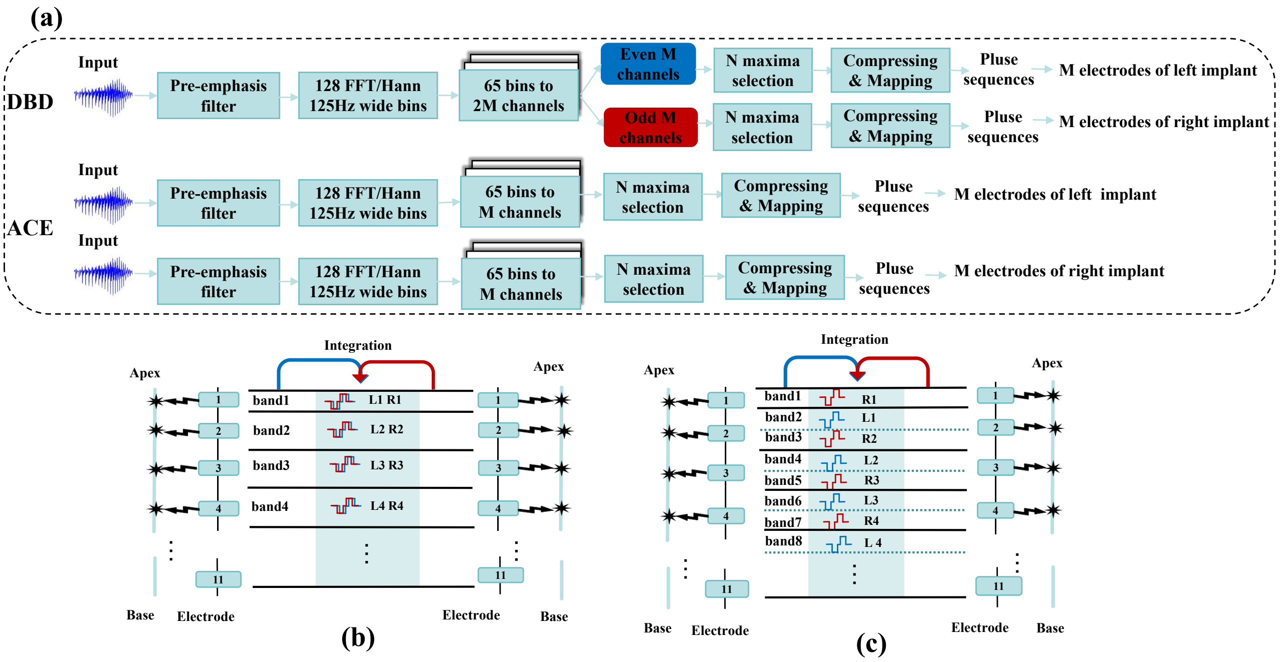

Figure 1: *Block diagram and relevant implementations. (a): Block diagrams of DBD and ACE showing the signal path from the input audio signal to pulse sequences. (b) & (c): Electrical stimulation patterns using ACE and DBD, respectively.*

trodes in an implant). DBD is based on the processing chain of the ACE strategy. It groups frequency bins into $2M$ bands, effectively doubling the density of frequency bands compared to ACE. Meanwhile, it alternates the spectrum information between odd and even frequency bands for each ear. Thus, DBD theoretically improves spectral resolution but has a higher requirement on the binaural integration abilities of the BiCI listeners. Subsequent steps include selecting the N electrode channels with the highest energy out of M channels (N-of-M maximum selection), along with logarithmic compression, acoustic-to-electric mapping and pulse generation, similar to ACE.

In the specific implementations in our experiment, DBD and ACE were compared under two conditions of total electrode number, i.e., M = 22 and M = 11. The four strategies are denoted as DBD22, DBD11, ACE22, and ACE11. Among these, 22 electrodes represent the default configuration for the ACE strategy. To account for cases with fewer electrodes, we also tested 11, with the band density being half of the 22 bands in the default ACE strategy. CCi-Mobile platform [15] was used for parameter configuration, featuring a 16 kHz sampling rate, 128 FFT and a Hanning window. After FFT, 65 bins were grouped into distinct frequency bands. For each electrode channel in the Map, the pulse rate was set at 900 pps, and electric levels ranged from a threshold (125) to a comfortable level (175), simulating a clinically typical electric dynamic range of 50.

Figures 1 (b) and (c) compare the ACE strategy with DBD, highlighting their differences in signal processing. Taking M = 11, i.e., ACE11 and DBD11, as an example and assuming a symmetric electrode distribution in both ears, the ACE strategy groups frequency bins into 11 frequency bands, corresponding to 11 electrode channels (the other 11 electrodes being deactivated). The process for mapping frequency bins to 11 bands begins by first excluding the initial three bins to reduce direct current (DC) interference, and the remaining 62 bins are then allocated across the bands based on a predefined sequence [1, 2, 2, 2, 3, 4, 5, 7, 9, 12, 15] specifying the bin count per band in ACE. Both ears of the BiCIs receive identical frequency information, corresponding to 11 frequency bands. In contrast, the DBD strategy doubles the band density to 22 frequency bands, which are interleaved across the electrodes of the two implants. Even-indexed frequency bands (e.g., band No. 2, 4, 6,..., 22) are assigned to the 11 electrodes of the left implant, while odd-indexed frequency bands (e.g., band No. 1, 3, 5,..., 21) are assigned to the 11 electrodes of the right implant. L1 and R1 (see Figures 1 (b) and (c)) indicate the actual frequency band positions corresponding to signals received by the first electrode in the left and right implants, respectively. Following this pattern, the actual frequency ranges of the signals received by all electrodes in both ears are marked (i.e., R1 L1 R2 L2,...). The stimulation sites for both strategies are the same, determined by the electrode numbering, with low frequencies located at the cochlear apex and high frequencies at the base. After binaural integration, DBD is expected to enhance spectral resolution by doubling the band density compared to ACE.

3. Experiments

3.1. Methods

Ten and eight normal hearing (NH) listeners participated in 1) spectral-temporally modulated ripple test (SMRT) and 2) speech perception tests, respectively. They were tested in a quiet room and used headphones to deliver sound at a comfortable and loud level. The original sounds were processed by a vocoder to simulate CI hearing. Before formal tests, a training session was conducted to familiarize listeners with the research interface and the testing procedure.

3.1.1. *Exp. 1: Spectral-temporally modulated ripple test*

Experiment 1 conducted spectral ripple tests, which are commonly used to measure CI users' spectral resolution ability [16][17]. In our study, we used a modified version of the spectral ripple test known as the spectral-temporally modulated ripple test (SMRT, [17]), i.e., SMRT 1.1.3 (see `https://www.ear-lab.org/smrt.html`). SMRT evaluates the participant's spectral resolution ability by requiring them to distinguish a spectrally rippled stimulus [17]. Theoretically, if the spectral resolution in CIs is improved, users' performance in the SMRT test should improve. The stimuli used for SMRT were generated using the MATLAB script developed by the Electrical Auditory Research Lab (EAR Lab, https://www.ear-lab.org/smrt.html).

A three-alternative forced-choice (3AFC) paradigm was used. Each subject completed eight adaptive tracks, i.e., twice for each of the four strategies (DBD22, DBD11, ACE22 and ACE11). In each trial, three stimulus sounds were presented. Among them, two stimuli were reference signals with a ripple intensity of 20 ripples/octave (rpo), one stimulus was a target signal with a lower rpo intensity. The task of the listener was to report which of the three stimuli was the oddball, that is, differed from the other two. The initial intensity of the target signal was set at 0.5 rpo, with a step size of 0.2 rpo. The density of the target stimulus signals varied using a 1-up/1-down adaptive procedure. The adaptive run terminated after 10 reversals, and the final result was the average of the last 6 reversals. A higher threshold in rpo value indicates that participants have better spectral resolution ability.

3.1.2. *Exp. 2: Speech Intelligibility*

Experiment 2 measured speech reception threshold (SRT), which refers to participants' ability to identify a certain percentage (e.g., 50%) of words in sentences at a signal-to-noise ratio (SNR). By measuring SRT, we aimed to determine whether the theoretically improved frequency resolution proposed in DBD could translate into perceptual benefits in speech perception. Lower SRTs indicate better speech perception ability. The speech materials were from the Mandarin Hearing in Noise Test (MHINT) [18] corpus, spoken by an adult male talker. It consists of 14 lists, each containing 20 sentences. The noise was a speech shape noise (SSN), which was generated using the method described in [19].

Each subject was tested in two sessions. In each session, the four strategies (i.e., DBD22, DBD11, ACE22, and ACE11) were tested in random order. These yield eight adaptive tracks in total. Each track were tested using one 20-sentence MHINT list. In total, each listener completed 160 sentences (8 tracks × 20 sentences). The listeners were instructed to repeat the target sentence as much as possible. An adaptive track with one list generated an SRT. For more details on the SRT calculation, please refer to the adaptive psychophysical procedure described in [20]. The geometric mean result of the two runs was used as the final SRT for each strategy.

3.1.3. *Vocoder Simulation*

Materials were processed using the Gaussian-enveloped tones (GET) vocoder [21] to simulate CIs. Vocoders play a crucial role as widely-used simulation tools, enabling researchers to evaluate how the auditory system processes degraded sounds under controlled conditions. The GET system was employed to convert the ACE output into a vocoded sound, establishing a sequential process where ACE and GET complementarily functioned together. Specifically, the electric pulse calculated using the ACE strategy is directly mapped into an envelope pulse by convolution with a Gaussian function, and finally modulated sinusoidally to generate simulated sound. For more details on the process and electric channel interactions, please refer to [21]. GET vocoder has exhibited comparable perceptual patterns in various tasks for NH and CI listeners (who used the ACE strategy in their clinical) in speech perception [22][23].

3.2. Results

The group and individual results for SMRT and speech perception, present a similar trend, as illustrated in Figure 2. To evaluate these effects, we conducted two separate two-way repeated measures analysis of variance (rm-ANOVA), each with strategy (ACE vs. DBD) and electrode number (M = 11 vs. M = 22) as independent variables, ripple discrimination thresholds and SRTs as dependent variables, respectively. Significant effects were observed, with both strategies and electrode number impacting ripple discrimination threshold and SRT. For ripple discrimination threshold, strategies showed an effect of $F(1, 9)$ = 27.35, p = 0.0005, and electrode number also showed an effect of $F(1, 9)$ = 27.35, p = 0.0005. For SRT, the effect of strategies was $F(1, 7)$ = 29.96, p = 0.0009, and for electrode number, it was $F(1, 7)$ = 35.39, p = 0.0006. Meanwhile, a significant interaction was observed for SRT ($F(1, 7)$ = 35.39, p = 0.0006), but not for ripple discrimination threshold ($F(1, 9)$ = 0.27, p = 0.61).

Specifically, DBD11 demonstrated a significant performance to ACE11, achieving a higher mean ripple discrimination threshold (DBD11: 1.55 rpo vs. ACE11: 1.02 rpo, $t(10)$ = 3.62, p = 0.03) and a better mean SRT (DBD11: 11.13 dB vs. ACE11: 14.78 dB, $t(8)$ = 10.08, p = 0.0001). However, differences in performance between the strategies DBD22 and ACE22 were not statistically significant both in ripple discrimination thresholds (DBD22: 3.72 rpo vs. ACE22: 2.94 rpo, $t(10)$ = 1.94, p = 0.51) and SRT (DBD22: 8.9 dB vs. ACE22: 8.8 dB, $t(8)$ = 0.18, $p >$0.99).

Finally, a detrimental effect was observed in DBD11, resulting in a decrease in performance compared to the ACE22, with ripple discrimination thresholds worsening by 1.39 rpo ($t(10)$ = 3.85, p = 0.02) and SRT deteriorating by 2.2 dB ($t(8)$ = 6.49, p = 0.002).

4. Discussions

4.1. Main findings and literature comparison

The findings in DBD11 and ACE11 suggest that double the band density does benefit auditory perception, similar to findings reported for normal hearing listeners [24]. It is indicated that spectral resolution influences spectral integration, which particularly affects sentences. However, the grouping rule negatively impacted the advantages of DBD22 compared to ACE22 (see Figure 3). For a typical CI, the default range of frequencies analyzed can be as wide as 100–8000 Hz. Due to its anatomical and design limitations, the electrode array is typically implanted to depths ranging from 8 to 21 mm [25], corresponding to cochlear place frequencies of no lower than 500–1500 Hz [26]. But the improvement in band density starts at the 10th channel under M = 22, and frequencies range from 1500 to 8000 Hz. Thus, focusing on improving frequency resolution within the mid to high-frequency range has led to DBD22 not consistently offering more advantages compared to the ACE22.

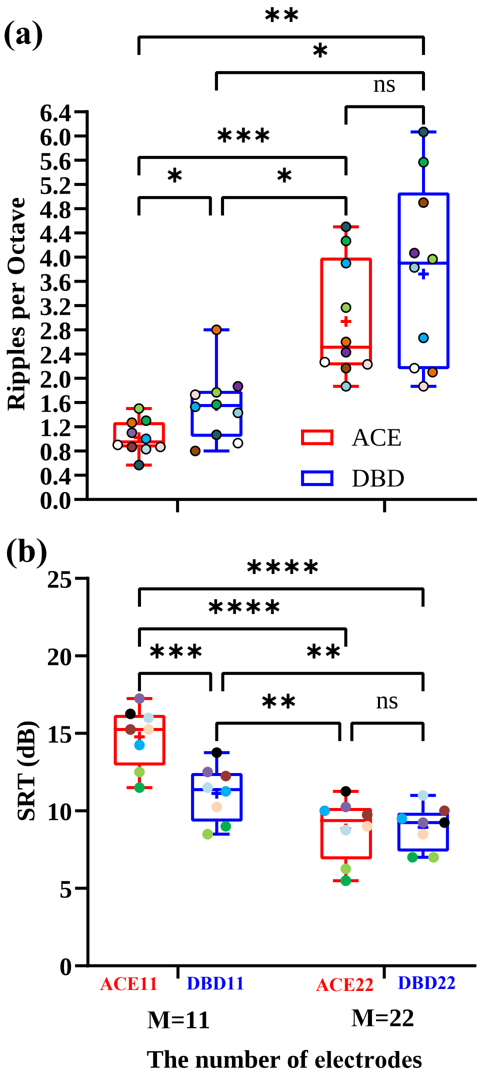

Figure 2: *(a): SMRT and (b): speech perception results with simulated CIs. Error bars indicate the standard deviations, and asterisks above denote the statistical significance. '+' indicates mean. Individuals are represented by discrete dots, and the same color belong to the same subject. ACE11 & DBD11: ACE and DBD strategies under M =11 condition. ACE22 & DBD22: ACE strategy and DBD strategies under M =22 condition.*

Meanwhile, the results of DBD11 and ACE22 indicated the odd - even (or interleaved) condition, not able to integrate the information as accurately as a condition where all channels were presented to BiCIs. This finding is consistent with previous research [11] [10], they found that interleaved spectrum with CI users does not have consistent benefits [10] or a notable detrimental effect [11].

4.2. Limitations

This study has several limitations. First, while the band density was improved, there may be an increase in the mismatch between the allocation of sound frequencies and the stimulation site, known as frequency-place mismatch. Previous research has shown that significant changes in frequency-place matching may lead to a decrease in speech intelligibility (refer to [27]). Thus the impact of frequency-place mismatch in DBD on actual BiCI users requires further verification in future work. Second, the DBD strategy utilizes electrodes from both sides to process signals for a single ear, neglecting binaural cues, which could potentially diminish the auditory experience.

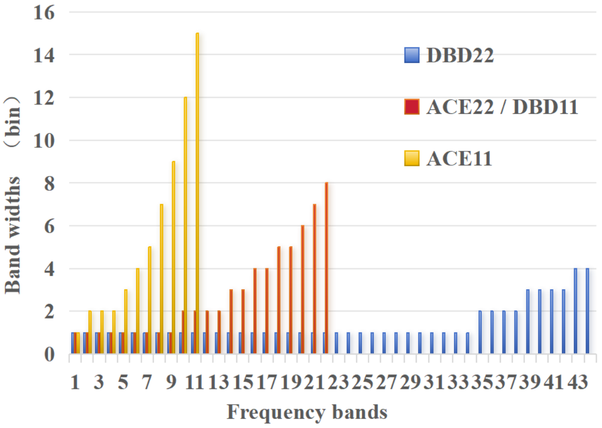

Figure 3: *Band Density in ACE and DBD for 11 and 22 Electrodes.*

5. Conclusions

This study introduces and evaluates the innovative Doubling the signal analysis Band Density (DBD) strategy, aimed at improving the auditory experience for BiCIs users by doubling the number of frequency bands and alternately encoding odd and even frequency bands. Preliminary vocoder-simulation experiments comparing DBD with the traditional ACE strategy have shown promising but limited benefits. One possible reason for these limitations could be that the 128-point frame length used not sufficiently reveal the DBD strategy's potential. Future studies are encouraged to investigate longer frame lengths and perform evaluations with actual BiCI users. Finally, this work opens up new avenues for enhancing the spectral resolution of BiCI users and lays a solid foundation for further research and practical applications.

6. Acknowledgements

We thank all the participants who generously donated their time to participate in this study. This work was supported by the Guangdong Basic and Applied Basic Research Foundation Grant (2022A1515011361), the Shenzhen Fundamental Research Program (20220809191805001), the China National Natural Science Foundation (12374448), and Guangdong Provincial Key Laboratory (2023B1212060076). Nengheng Zheng is the corresponding author.

7. References

[1] B. S. Wilson and M. F. Dorman, "The surprising performance of present-day cochlear implants," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 6, pp. 969–972, 2007.

[2] C. Jiang, S. Singhal, T. Landry, I. V. Roberts, S. R. de Rijk,

T. Brochier, T. Goehring, Y. C. Tam, R. P. Carlyon, G. G. Malliaras *et al.*, "An instrumented cochlea model for the evaluation of cochlear implant electrical stimulus spread," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 7, pp. 2281–2288, 2021.

[3] R. R. S. Sarreal, D. T. Blake, and P. T. Bhatti, "Development and characterization of a micromagnetic alternative to cochlear implant electrode arrays," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2116–2125, 2022.

[4] R. V. Shannon, F.-G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, no. 5234, pp. 303–304, 1995.

[5] F.-G. Zeng, "Challenges in improving cochlear implant performance and accessibility," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1662–1664, 2017.

[6] N. Zheng, Y. Shi, Y. Kang, and Q. Meng, "A noise-robust signal processing strategy for cochlear implants using neural networks," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 8343–8347.

[7] Q. Meng, N. Zheng, and X. Li, "A temporal limits encoder for cochlear implants," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 5863–5867.

[8] H. Zhou, A. Kan, G. Yu, Z. Guo, N. Zheng, and Q. Meng, "Pitch perception with the temporal limits encoder for cochlear implants," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2528–2539, 2022.

[9] S. Söderqvist, V. Sivonen, J. Koivisto, A. Aarnisalo, and S. T. Sinkkonen, "Spread of the intracochlear electrical field: Implications for assessing electrode array location in cochlear implantation," *Hearing Research*, vol. 434, p. 108790, 2023.

[10] R. S. Tyler, S. A. Witt, C. C. Dunn, A. Perreau, A. J. Parkinson, and B. S. Wilson, "An attempt to improve bilateral cochlear implants by increasing the distance between electrodes and providing complementary information to the two ears," *Journal of the American Academy of Audiology*, vol. 21, no. 01, pp. 052–065, 2010.

[11] A. Mani, P. C. Loizou, A. Shoup, P. Roland, and P. Kruger, "Dichotic speech recognition by bilateral cochlear implant users," in *International Congress Series*, vol. 1273. Elsevier, 2004, pp. 466–469.

[12] J. M. Aronoff, J. Stelmach, M. Padilla, and D. M. Landsberger, "Interleaved processors improve cochlear implant patients' spectral resolution," *Ear and hearing*, vol. 37, no. 2, p. e85, 2016.

[13] D. M. Wohlbauer, W. K. Lai, and N. Dillier, "Interlace sound coding for unilateral and bilateral cochlear implants," *IEEE Transactions on Biomedical Engineering*, 2023.

[14] J. Kiefer, S. Hohl, E. Stürzebecher, T. Pfennigdorff, and W. Gstöettner, "Comparison of speech recognition with different speech coding strategies (speak, cis, and ace) and their relationship to telemetric measures of compound action potentials in the nucleus ci 24m cochlear implant system," *Audiology*, vol. 40, no. 1, pp. 32–42, 2001.

[15] R. Ghosh, H. Ali, and J. H. Hansen, "Cci-mobile: A portable real time speech processing platform for cochlear implant and hearing research," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 3, pp. 1251–1263, 2021.

[16] B. A. Henry and C. W. Turner, "The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners," *The Journal of the Acoustical Society of America*, vol. 113, no. 5, pp. 2861–2873, 2003.

[17] J. M. Aronoff and D. M. Landsberger, "The development of a modified spectral ripple test," *The Journal of the Acoustical Society of America*, vol. 134, no. 2, pp. EL217–EL222, 2013.

[18] L. L. Wong, S. D. Soli, S. Liu, N. Han, and M.-W. Huang, "Development of the mandarin hearing in noise test (mhint)," *Ear and hearing*, vol. 28, no. 2, pp. 70S–74S, 2007.

[19] Q. Meng, X. Wang, Y. Cai, F. Kong, A. N. Buck, G. Yu, N. Zheng, and J. W. Schnupp, "Time-compression thresholds for mandarin sentences in normal-hearing and cochlear implant listeners," *Hearing Research*, vol. 374, pp. 58–68, 2019.

[20] Q. Meng, N. Zheng, and X. Li, "Mandarin speech-in-noise and tone recognition using vocoder simulations of the temporal limits encoder for cochlear implants," *The Journal of the Acoustical Society of America*, vol. 139, no. 1, pp. 301–310, 2016.

[21] Q. Meng, H. Zhou, T. Lu, and F.-G. Zeng, "Pulsatile gaussian-enveloped tones (get) for cochlear-implant simulation," *Applied Acoustics*, vol. 208, p. 109386, 2023.

[22] F. Kong, H. Zhou, Y. Mo, M. Shi, Q. Meng, and N. Zheng, "Comparable encoding, comparable perceptual pattern: Acoustic and electric hearing," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023.

[23] F. Kong, N. Zheng, X. Wang, H. He, J. W. Schnupp, and Q. Meng, "Cochlear-implant listeners listening to cochlear-implant simulated speech," in *INTERSPEECH 2023*, 2023, pp. 4988–4992.

[24] P. C. Loizou, A. Mani, and M. F. Dorman, "Dichotic speech recognition in noise using reduced spectral cues," *The Journal of the Acoustical Society of America*, vol. 114, no. 1, pp. 475–483, 2003.

[25] J. Lee, J. B. Nadol Jr, and D. K. Eddington, "Depth of electrode insertion and postoperative performance in humans with cochlear implants: a histopathologic study," *Audiology and Neurotology*, vol. 15, no. 5, pp. 323–331, 2010.

[26] D. D. Greenwood, "A cochlear frequency-position function for several species?29 years later," *The Journal of the Acoustical Society of America*, vol. 87, no. 6, pp. 2592–2605, 1990.

[27] Q.-J. Fu and R. V. Shannon, "Effects of electrode configuration and frequency allocation on vowel recognition with the nucleus-22 cochlear implant," *Ear and hearing*, vol. 20, no. 4, pp. 332–344, 1999.