Hearable Heart Beats: Recording the Heart Rate from the Ear

Bjarke Gaardbaek and Preben Kidmose

Abstract— **A single physiological measurement reflects the momentary physiology under the given circumstances. Longitudinal measurements of a key physiological health indicator, such as heart rate, provide a more comprehensive picture of the physiological condition compared to a single measurement. Recently, a cardiovascular phenomenon synchronized to the heartbeat were recorded from the ear using body-coupled microphones. This indicates that heart rate may be estimated using body-coupled microphones. The present study explores the feasibility of extracting time-varying heart rates using body-coupled microphones in the ears. Method: 10 subjects were recorded in a supine position for 5 minutes, with a body-coupled microphone (BCM) mounted in each ear. Heart rate (HR) was estimated from the BCM recordings using the PYIN pitch detection algorithm. Since PYIN is traditionally applied to speech data, a grid search was conducted to optimize hyperparameters for BCM-based HR estimation. The HR estimates were systematically compared to electrocardiography based HR estimations in terms of detection rate, error distribution, and root-mean-square error (RMSE). Results: In 9 out of 19 recorded ears, the RMSE was below 4 beats per minute (BPM). However, 4 out of 10 subjects exhibited both low detection rates and high RMSE, indicating substantial variability in performance across individuals. Discussion: The study revealed considerable inter-ear variability in both the detection rate and the HR error distribution. This variability likely reflects differences in how well cardiovascular activity is represented in the BCM signal, as well as the algorithm's ability to reliably extract the periodicity of the signal. The error distributions suggest that BCM-based HR estimation is unbiased. Conclusion: These findings demonstrate the feasibility of HR estimation using BCM in the ear.**

Clinical Relevance—**Ear-level sensing is an emerging technology with broad applications in health monitoring. This study demonstrates the feasibility of enhancing ear-centered sensing devices with BCM-based cardiovascular monitoring capabilities.**

I. Introduction

A single physiological measurement reflects the momentary physiology under the given circumstances. It does not capture temporal variations or responses under different conditions. Longitudinal measurements across diverse conditions provide a more comprehensive picture of the physiological condition, thus motivating monitoring outside controlled laboratory settings.

A key physiological health indicator of interest for longitudinal measurements is heart rate (HR). The primary modality used to measure HR is electrocardiography (ECG). This is measured using electrodes affixed to the subject, typically on the torso and extremities. To perform ECG measurements correctly, the electrodes must be affixed in the proper anatomical locations by a trained professional [1]. This limits traditional ECG to laboratory settings. A Holter monitor could be an alternative to bring ECG measurements outside the laboratory. This is still limited by electrodes needing to be mounted by trained professionals. For HR and heart rate variability (HRV) measurements, photoplethysmography have been widely used in wearable consumer electronics, such as smart and fitness watches. Such alternatives that do not require a trained professional allow HR measurements to be conducted outside laboratory settings.

Solid work has already been done on moving electroencephalography (EEG) out of the laboratory setting using ear-EEG [2]. This modality is measured using custom ear-pieces, which the subject can correctly mount without supervision [2], [3]. These ear-pieces provide a flexible platform that can be used by other sensing modalities.

Recently, blood pressure waves were recorded from the ear using a novel modality, the so-called body-coupled microphone (BCM) [4]. The blood pressure waves are synchronized with the heartbeat. This indicates that a BCM mounted in the ear could be a good alternative to ECG for determining HR. Therefore, we aim to extract time-varying heart rates using BCMs mounted in the ears.

II. Method

A. Experimental setup & paradigm

This manuscript is part of a larger study where 10 healthy subjects with no prior knowledge of cardiovascular symptoms were recruited. The subjects were affixed with ECG electrodes on the torso and a BCM mounted in each ear and asked to lie supine for 5 minutes. For further detail, see [4].

B. Heart rate detection

Extracting HR from a time-series can be considered a matter of finding periodicities in quasi-stationary signals. This is a challenge known from multiple domains. In speech processing this is known as pitch detection. The pitch, which constitutes the instantaneous fundamental frequency of a quasi-periodic time-series is detected. The HR can then be estimated as the pitch detected in the BCM signals.

A commonly employed algorithm for pitch detection in speech signals is the YIN algorithm [5], along with its advanced variant, the PYIN algorithm [6], which handles the thresholds within the algorithm in a probabilistic way. In

This work was supported by the Center for Ear-EEG, Aarhus University. (Corresponding author: Bjarke Gaardbaek.)

Bjarke Gaardbaek is with the Center for Ear-EEG, Department of Electrical and Computer Engineering, Aarhus University 8200 Aarhus N, Denmark blg@ece.au.dk

Preben Kidmose is with the Center for Ear-EEG, Department of Electrical and Computer Engineering, Aarhus University 8200 Aarhus N, Denmark pki@ece.au.dk

Manuscript 2184 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 9, 2025.

this work we used the librosa implementation of the PYIN algorithm [7].

A commonly employed algorithm for pitch detection in speech signals is the YIN algorithm [5]. An advanced extension to this is the PYIN algorithm [6], which handles thresholds within the algorithm in a probabilistic way. In this work we used the librosa implementation of the PYIN algorithm [7].

This implementation of the algorithm has certain hyperparameters which can be tuned. Since the algorithm is applied to pitch detection in a modality different from the one for which it was originally developed, we considered it essential to explore the hyperparameter space for this new modality. To this end, we conducted a grid search in the hyperparameter space, were, the mean of the beta distribution ($\mu_\beta = 0.05, 0.1, \ldots, 0.5$), the Boltzmann parameter ($\lambda = 0.5, 0.1, \ldots, 10$), and the switch probability ($P = 0.0002, 0.0004, \ldots, 0.002$). In summary, 2000 parameter sets were investigated.

The mean of the beta distribution is given by,

$$\mu_\beta = \frac{\alpha}{\alpha + \beta} \ , \tag{1}$$

where α and β are shaping parameters. The shaping parameters used for the grid search can be seen in Table I.

To allow a rigorous validation of the BCM-based estimated HR, the ground truth HR was derived from the ECG signals using the following procedure. First, the ECG R-peaks were identified from lead II with the algorithm used in [4]. This resulted in a list of indices, with ECG R-peaks, which was then converted to an impulse train, with the same sample distance as the original ECG signal. A Hann window were then slid with a fixed interval over the impulse train and HR was computed at the center of the window, as the sum of impulses within the window, normalized by the area of the window.

For both ECG and BCM signals, HR were extracted using 15 s windows which moved in steps of 100 ms. The BCM signals were bandpass filtered with a lower cut-off at 0.1 Hz and upper cut-off at 45 Hz. Additionally, a 50 Hz notch filter was used to attenuate power-line noise.

C. Evaluation

Mean square error (MSE) was chosen as the optimization criterion. The MSE is defined as,

$$\mathrm{MSE} = \frac{1}{N} \sum_{i=1}^{N} (e_i)^2 \ , \tag{2}$$

with the error, $e_i = \mathrm{HR}_{i,ECG} - \mathrm{HR}_{i,BCM}$, where $\mathrm{HR}_{i,ECG}$ and $\mathrm{HR}_{i,BCM}$ is the HR from the i^{th} interval estimated from the ECG and BCM signals respectively.

The hyperparameter search were executed as 10 runs in a leave-one-subject-out fashion. In each run, the MSE was computed for the 9 remaining subject, for each ear for each parameter set. The optimal parameter set of the run were found as the set with the lowest MSE across ears for a subset of the remaining subjects. The most frequent per run parameter set were identified as a single optimal parameter set. This single optimal parameter set were validated against each subject, where the error were computed as in (2) for each ear in addition to root means square error (RMSE).

III. Results

In the following the results from the hyperparameter grid search will be presented, as well as how the selected optimal parameter set performs. The selected optimal parameter set can be seen on Table II.

A comparison of performance on select subjects, using the optimal hyperparameter set with select segments of the underlying signal traces, can be seen on Fig. 1(a) and 1(b). Traces in the upper pane show BCM left ear, BCM right ear and ECG lead II signal respectively. These are in arbitrary units and have been normalized to 1 standard deviation, and biased for better visualization. Gray dots on the trace of ECG lead II denote the detected ECG R-peak used for the ECG HR estimates. Generally, the left upper pane is from an earlier time interval than the right.

The upper axis in the lower pane shows HR in beats per minute (BPM) for the BCM left ear, BCM right ear and ECG lead II respectively. The lower axis shows the error between BCM left ear, and BCM right ear and ECG lead II respectively, using (2). Additionally, a dashed line has been added to better visualize when the distribution of errors compared to an error of 0 BPM. Note that in the lower pane, BCM-based estimates are piecewise constant because the estimation window was moved in steps of 100 ms. The top panes have been colored, such that they match the colored regions of the lower pane. This indicates at which time intervals the segments in the upper pane coincide with the traces in the lower pane.

Results on validation can be seen on Fig. 2, where the distribution of errors for each ear for each subject can be seen. It should be noted that the left ear wire had snapped for subject 5 and was therefore been excluded.

Finally, summarized results from the validation can be seen on Table III. Here RMSE can be seen for each ear for each subject, as well as how many percent of the possible HR estimations were successful.

TABLE I

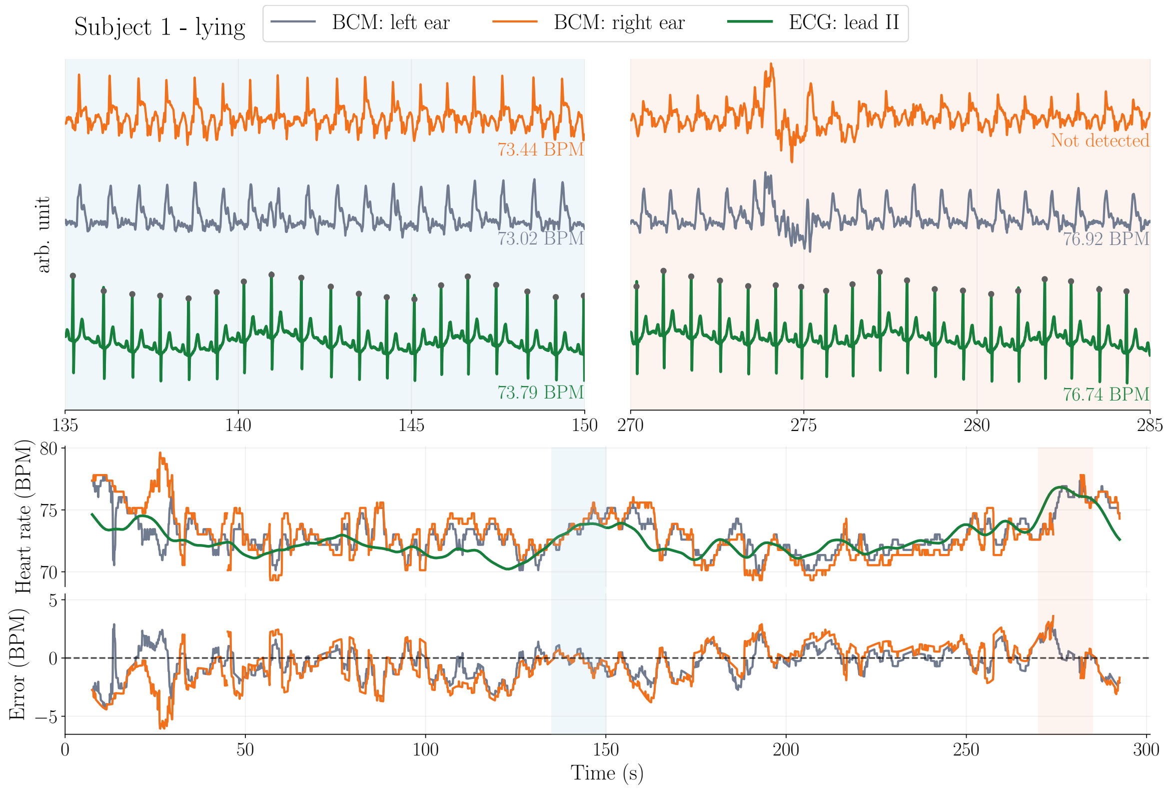
THE BETA DISTRIBUTION MEAN VALUE, AND THE CORRESPONDING SHAPING PARAMETERS, USED IN THE HYPERPARAMETER GRID SEARCH

	Values									
μ_β	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
α	2	2	3	2	2	3	2.1	2	2.7	3
β	38	18	17	8	6	7	3.9	3	3.3	3

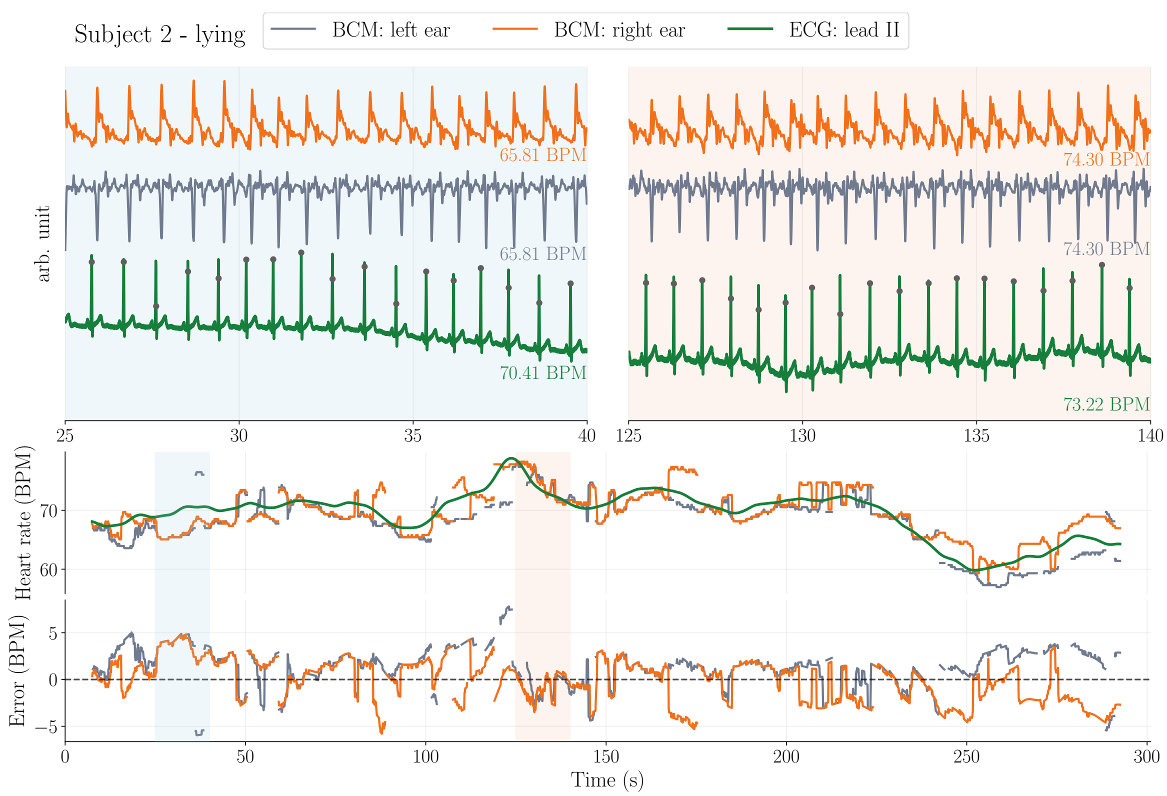
TABLE II

OPTIMAL HYPERPARAMETERS FOUND IN THE GRID SEARCH

	μ_β (α,β)	λ	P
Value	0.25 (2,6)	1.0	0.0008



(a) Performance on subject 1. (**Top, left**) HR detection is successful. (**Top, right**) HR detection has failed for BCM right ear.



(b) Performance on subject 2. (**Top, left**) HR detection is incorrect for both BCM left and right ear. (**Top, right**) HR detection is successful.

Fig. 1. Traces of BCM left ear, BCM right ear, and ECG signals respectively. Gray dots denote detected ECG R-peaks. (**Bottom**: **(a), (b)**) Upper axis shows estimated heart rate, while the lower axis shows the error for BCM left ear and BCM right ear. Colored rectangles denote the extent of segments shown in top panes.

TABLE III

SUMMARY OF RMSE AND PERCENT DETECTED FROM VALIDATION OF OPTIMAL PARAMETER SET FOR EACH EAR, FOR EACH SUBJECT.

	Subject	1	2	3	4	5	6	7	8	9	10
RMSE (BPM)	Left	1.21	2.42	4.21	7.07	-	11.03	7.24	3.28	7.55	3.85
	Right	1.70	2.21	3.59	14.22	5.17	15.42	10.57	3.03	24.01	7.77
Detected (%)	Left	99.5	93.1	92.4	91.3	-	40.9	66.6	95.0	80.7	93.4
	Right	97.2	99.7	52.9	27.7	72.5	29.0	86.4	99.4	70.8	84.0

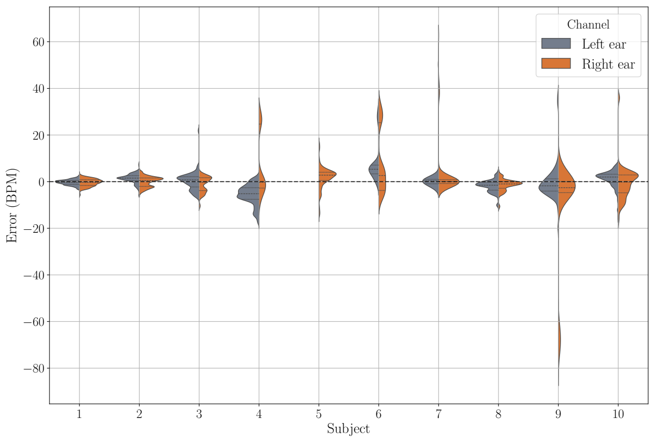


Fig. 2. Distribution of heart rate estimation error across all subjects for left and right ear, using optimal hyperparameter set. Dashed lines within distributions denote mean, and 1st and 3rd quantile.

IV. DISCUSSION

A. Selected hyperparameters

Based on a coarse grid search within the hyperparameter space, we selected the set of hyperparameters listed in Table II, corresponding to the lowest MSE. Notably, none of the selected hyperparameters are at the boundaries of the search space, indicating that the identified optimum is determined by the topography of the optimization criteria, rather than constrained by the grid search boundaries.

The mean of the beta distribution (μ_β) affects the probability of some periodicity being detected. As the probability of detecting some periodicity is increased, so may the probability of detecting a false positive increase. That is, the detection of some periodicity where there is none. In contrast, lowering the probability of detecting some periodicity may lead to the detection of a present periodicity being missed. That is, the probability of false negatives increase. As such, adjusting μ_β is a trade-off between false positive and false negative detections respectively. As the signals are generally highly periodic as indicated by inspecting the top panes of Fig. 1, favoring an increased μ_β may be ideal.

The Boltzmann shaping parameter (λ) shifts the mass of a Boltzmann distribution within the PYIN algorithm. This mass is either shifted towards the selection of smaller or larger cardiac cycle periods. A smaller λ shifts mass towards larger cardiac cycle period [7]. Assuming a normal resting heart rate of 60 BPM [8], a period is expected to be 1 s. This indicates that the periodicity of the cardiac cycle is in the scale of second. As such, a small λ is not surprising.

The switch probability (P) is the probability of changing from an HR detected state to a not detected state or vice versa. With the combination of subjects lying supine and minimizing artifacts, and a large μ_β, it should be reasonable to assume that some periodicity will be present for the duration of the recording. Because of this, it is not unreasonable to expect that the probability of shifting to a not detected state should be small. It should be noted that other physiological circumstances may not have these favorable conditions. As such, a conservative P might not be ideal if the presence of artifacts are increased.

B. Heart rate estimation performance

The results presented in Table III and Fig. 2 show a significant variability in the proportion of time intervals yielding an HR estimate, as well as in the accuracy and precision of the HR estimates. These results reflect a combination of how well cardiovascular activity is represented in the BCM signal and the detection algorithm's ability to reliably extract the periodicity of the available signal. In the assessment of the feasibility of extracting time-varying heart rates using ear-positioned body-coupled microphones, these two aspects have not been disentangled. Therefore, performance may be improved by optimizing the BCM and its placement within the ear, as well as by improving the detection algorithm.

In the assessment of these results we note that (1) the performance requirements will depend on the intended application of the HR estimate, and (2) the purpose of this study is to assess the feasibility of HR estimation using BCM, rather than to provide a quantitative validation of a system.

From Table III it is seen that three (1, 2, and 8) out of ten subjects had a high detection rate ($> 95\%$) on both ears, and additional three subjects (3, 4, and 10) had a high detection rate ($> 90\%$) on at least one ear. Seven out of these nine ears had a RMSE below 4 BPM, and nine out of the total nineteen ears had a RMSE below 4 BPM.

In general, we observe a strong correlation between detection rate and RMSE, with higher detection rates corresponding to lower RMSE. This relationship is likely driven by a reinforcing effect: (1) when the cardiovascular activity is well-represented in the BCM signal, HR estimation is feasible over a larger proportion of time, and (2) when HR can be reliably estimated due to the better representation of the cardiovascular activity, the resulting estimates achieve better precision and accuracy. This is a typical tradeoff between sensitivity and specificity and can be exemplified by subject 3, in which the RMSE of the right ear is possibly kept lower due to the low detection rate; if the threshold for detections were lowered, this would lead to a higher detection rate and consequently higher RMSE. We also observe a high correlation between high RMSE and the tendency for outliers in the HR-estimate as seen from the right ear of subject 4, 6, 7 and 9 in Fig. 2 Lastly, four (5, 6, 7, and 9) out of ten subjects had in general a low performance both in terms of low detection rate and high RMSE.

The error distributions presented in Fig. 2 show that the mean error varies across ears, with some ears having a positive mean error and others having a negative mean error, suggesting that the HR estimation is not in general biased. Additionally, ears with low variance and a low amount of outliers tend to have a mean error close to zero.

Segments where detection of HR is achieved with small error can be seen on Fig. 1(a) and 1(b) respectively. This is achieved even with the morphology of the BCM signals are noticeably different for the left and right ear on Fig. 1(b).

A segment where detection fails for the left ear can also be seen on the top right pane of Fig. 1(a). An artifact around 275 s may be the cause, it can however be hard to determine if this is the case, as a smaller but albeit similar artifact is visible at the same time for the right ear.

Additionally, a segment where HR detection is off can be seen on the top left pane of Fig. 1(b). No apparent reason indicates, why the HR estimation is off for this segment.

When subjects are lying supine, a difference in HR of above 15 BPM when measuring beat-to-beat HRV is normal [9, p. 383]. At a resting HR of 60 BPM, this amounts to a beat-to-beat change of more than 200 ms for a single heart rate or a local change in instantaneous frequency of more than 250 mHz.

The PYIN algorithm searches for periodicities utilizing autocorrelation. HRV reflects deviations in HR periodicity. Therefore, the larger the HRV, the less distinct the HR will appear as a periodicity in the autocorrelation function. Consequently, HR estimation under conditions with high HRV will likely be harder than conditions with low HRV.

Moreover, conventional ECG-based HRV estimation relies on R-to-R peak detection, utilizing information from approximately one heart cycle [10]. In contrast, methods based on autocorrelation, such as PYIN, typically estimate the autocorrelation over several periods of the underlying periodicity. Consequently, HRV metrics quantifying interbeat variability are likely to be smoothed out by the autocorrelation estimation, whereas metrics quantifying changes over a longer time horizon will remain unaffected.

In interpreting the study's results, it is important to note that measurements were conducted on participants in a supine position. Consequently, the data represents only a limited subset of cardiovascular activity variations that occur under other physiological conditions. Furthermore, other physiological conditions, such as physical exercise, may introduce various types of interference and artifacts that were not accounted for in the current study. Therefore, the results presented here may not be representative of other conditions.

V. Conclusion

The results demonstrate the feasibility of heart rate estimation from BCMs in the ear, paving the way for integrating BCM-based cardiovascular monitoring into ear-centered sensing devices.

References

[1] A. Hadjiantoni, K. Oak, S. Mengi, J. Konya, and T. Ungvari, "Is the Correct Anatomical Placement of the Electrocardiogram (ECG) Electrodes Essential to Diagnosis in the Clinical Setting: A Systematic Review," vol. 5, no. 2, pp. 182–200.

[2] K. B. Mikkelsen, Y. R. Tabar, S. L. Kappel, C. B. Christensen, H. O. Toft, M. C. Hemmsen, M. L. Rank, M. Otto, and P. Kidmose, "Accurate whole-night sleep monitoring with dry-contact ear-EEG," vol. 9, no. 1, p. 16824.

[3] Y. R. Tabar, K. B. Mikkelsen, N. Shenton, S. L. Kappel, A. R. Bertelsen, R. Nikbakht, H. O. Toft, C. H. Henriksen, M. C. Hemmsen, M. L. Rank, M. Otto, and P. Kidmose, "At-home sleep monitoring using generic ear-EEG," vol. 17.

[4] B. Gaardbaek and P. Kidmose, "On the Origin of Cardiovascular Sounds Recorded From the Ear," vol. 72, no. 1, pp. 210–216.

[5] Alain de Cheveigné and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," vol. 111, no. 4, pp. 1917–1930.

[6] M. Mauch and S. Dixon, "PYIN: A fundamental frequency estimator using probabilistic threshold distributions," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 659–663.

[7] B. McFee, M. McVicar, D. Faronbi, I. Roman, M. Gover, S. Balke, S. Seyfarth, A. Malek, C. Raffel, V. Lostanlen, Benjamin van Niekirk, D. Lee, F. Cwitkowitz, F. Zalkow, O. Nieto, D. Ellis, J. Mason, K. Lee, B. Steers, E. Halvachs, C. Thomé, F. Robert-Stöter, R. Bittner, Z. Wei, A. Weiss, E. Battenberg, K. Choi, R. Yamamoto, C. J. Carr, A. Metsai, S. Sullivan, P. Friesch, A. Krishnakumar, S. Hidaka, S. Kowalik, F. Keller, D. Mazur, A. Chabot-Leclerc, C. Hawthorne, C. Ramaprasad, M. Keum, J. Gomez, W. Monroe, V. A. Morozov, K. Eliasi, nullmightybofo, P. Biberstein, N. D. Sergin, R. Hennequin, R. Naktinis, beantowel, T. Kim, J. P. Åsen, J. Lim, A. Malins, D. Hereñú, Stef van der Struijk, L. Nickel, J. Wu, Z. Wang, T. Gates, M. Vollrath, A. Sarroff, Xiao-Ming, A. Porter, S. Kranzler, Voodoohop, M. D. Gangi, H. Jinoz, C. Guerrero, A. Mazhar, toddrme2178, Z. Baratz, A. Kostin, X. Zhuang, C. T. Lo, P. Campr, E. Semeniuc, M. Biswal, S. Moura, P. Brossier, H. Lee, and W. Pimenta. Librosa/librosa: 0.10.2.post1. [Online]. Available: https://zenodo.org/records/11192913

[8] D. Nanchen, "Resting heart rate: What is normal?" vol. 104, no. 13, pp. 1048–1049.

[9] I. Cygankiewicz and W. Zareba, "Chapter 31 - Heart rate variability," in *Handbook of Clinical Neurology*, ser. Autonomic Nervous System, R. M. Buijs and D. F. Swaab, Eds. Elsevier, vol. 117, pp. 379–393.

[10] M. Malik, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," vol. 93, pp. 1043–1065.

Manuscript 2184 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 9, 2025.