Ear-ECG Denoising Using Heart Sounds and the Extended Kalman Filter

Adam Pullin*, Jake Stuchbury-Wass*, Mathias Ciliberto*, Kayla-Jade Butkow*,
Philipp Lepold†, Tobias Röddiger† and Cecilia Mascolo*
*University of Cambridge, United Kingdom †Karlsruhe Institute of Technology, Germany
{alp78, js2372, mc2514, kjb85}@cam.ac.uk, philipp.lepold@kit.edu, tobias.roeddiger@kit.edu, cm542@cam.ac.uk

***Abstract*—Electrocardiogram (ECG) recording systems are increasingly being integrated into consumer wearable systems such as smartwatches, providing users with access to clinically-relevant information about their heart activity anytime, anywhere. The increasing adoption of in-ear wearables, known as earables, as well as their stable position on the body, makes them an attractive prospect for ECG integration. However, this comes with several challenges. Other biosignals, including those from the brain and surrounding muscles, are detectable at the ear in the same frequency bands with much higher amplitudes. This means that the ECG signal-to-noise ratio (SNR) can be extremely low at this location. The few existing denoising approaches mostly rely on autoencoders. In some cases they fail to recover the ECG morphology, and their black-box nature does not allow for explainability or understanding of limitations.**

To address these issues, we introduce a novel system to record and denoise ear-ECG signals, leveraging open-source hardware and the Extended Kalman Filter. In-ear audio recording of heart sounds is used to accurately determine timings of cardiac cycles. From these timings, a short-term ensemble average ECG signal is calculated, which is used to fit the parameters of a dynamical ECG model to an individual user. The Kalman filter is then applied to the full time series ECG for denoising, using the dynamical model for its state prediction steps, and heart sounds as phase measurements. We have evaluated the system with data collected from 18 participants. The results report a mean SNR of 6.4 dB, mean absolute QT interval error of 54 ms, and heart rate error of 3 BPM, demonstrating the system's potential for continuous, non-invasive, user-friendly ECG monitoring.

***Index Terms*—Earable, Electrocardiogram (ECG), Denoising.**

I. Introduction

The electrocardiogram (ECG) is a recording of the electrical activity of the heart, and is widely used for monitoring or diagnostic purposes. Cardiac potentials propagate from the heart in all directions, so ECGs can be recorded by measuring voltages between electrodes placed on the skin [1]. To record high quality data, some distance between the electrodes is required since differential signals are recorded. Electrode placements for clinical practice have been standardised to produce leads, each of which captures heart activity in a known direction [2]. Lead I, for example, measures the horizontal cross-body cardiac vector, normally between electrodes on the left and right arms. Normal, healthy ECGs are cyclical in nature, with five major peaks and troughs labelled as P, Q, R, S and T waves, as shown in Fig. 1.

We acknowledge the support of EPSRC through grant HearFit EP/Z53447X/1 and Arm.

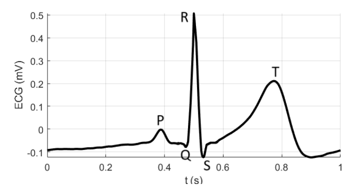

Fig. 1. Healthy single ECG cycle with labelled P, Q, R, S and T waves

In recent years ECG recording technology has started to be incorporated into wearable devices, allowing the scaling of cardiac health observation out of clinical settings. The most common of these are heart monitor chest straps such as the Polar H10, specifically designed for this single purpose, and smartwatches such as the Apple Watch. Both, however, have their drawbacks. Chest straps are uncomfortable, and since their only usage is measurement of cardiac activity they can be inconvenient, requiring users to alter their behaviour to wear additional devices. While smartwatches are more agreeable for day-to-day wearing, their method of recording ECGs requires users to touch the watch with the opposite hand. This blocks the use of both hands and vastly limits the scenarios in which data can be recorded, making passive monitoring impossible. Ear-ECG systems have the potential to solve these problems: earphones exist in comfortable form factors which users are used to wearing daily, and their symmetrical placement across the head provides the necessary inter-electrode distance in an orientation suitable for Lead I ECG measurement.

Implementation of ear-ECG systems is not without challenges. Although cardiac potentials propagate throughout the entire body, the relative narrowness of the neck means that by the time they reach the ears, ECG signals have been greatly attenuated [3]. Other potentials also exist in the body, including those originating from the brain (EEG signals) [4], eyes (EOG signals) [5] and muscles (EMG signals) [6]. The frequency spectra of these overlaps with that of the ECG, but since they originate much closer to the ear they have significantly higher amplitudes at this location. This means that the base signal-to-noise ratio of ear-ECG recordings is very low, and cannot be easily improved with basic frequency cutoff filtering.

A small number of works have been published in the area. One work did not attempt to extract full time series ECG data from electrode recordings around the ears. Instead, they produced ensemble average single cardiac cycles, averaged using timestamps from separate sensors [3]. Similarly, [7] averaged using ground truth ECG data, and had cross-head ECG signals with very low correlation to their ground truth (mean spectral coherence <0.2). Although an average cycle does contain rich information about the user's cardiac activity, a lot is also lost, including heart rate variability (HRV) data and potential inter-cycle variations in morphology. A recent work took an energy domain approach, based on adaptive Fourier decomposition, which gave good results for R-peak timings but otherwise produced very distorted morphologies [6]. The most successful denoising works in the field so far have used autoencoders trained with noisy ear-ECG as input and simultaneously recorded clean Lead I ECG as output. The first system taking this approach yielded acceptable performance, with large SNR improvements (median 5.9 dB) over the base recording, but was shown in a scatter plot to output signals with very large variations in heart rate from the ground truth, and in some cases failed to recover morphologies [8]. A more recent work refined this method with further processing steps, but at its base still relied on non-explainable signal generation from a black box autoencoder, and focused only on R-peak quality metrics rather than the full ECG morphology [9].

To improve on these results, we propose a multi-modal denoising approach. Heart sounds from inside the ear canal give high-accuracy information about the phase of the ECG signal, and can be recorded using in-ear microphones that already exist in many consumer headphones. We feed these heart sounds into an Extended Kalman Filter (EKF) with a fitted ECG dynamical model for ear-ECG denoising.

II. Background

A. Extended Kalman Filter

The standard Kalman filter is an algorithm which uses noisy measurements to estimate the true state of a dynamic system [10]. A linear system is modelled with state vector $\underline{x}$, state transition matrix A and process noise $\underline{w}$ as:

$$\underline{x}_{k+1} = A\underline{x}_k + \underline{w}_k \tag{1}$$

with observations of $\underline{y}_k$ and measurement noise $\underline{v}_k$. Both the process and measurement noise vectors are assumed to be zero-mean Gaussian vectors with covariance matrices Q_k and R_k respectively.

For every measurement (sample), two steps take place. Firstly, in the **Predict** step, the next states of the system are forecasted using the state transition matrix and process noise estimate as in (1). The state error covariance matrix P_k is also calculated, as:

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q_k \tag{2}$$

Secondly, in the **Update** step, the observed noisy state values are taken into account. The Kalman Gain is given as:

$$K_k = \frac{P_{k|k-1}A^T}{AP_{k|k-1}A^T + R} \tag{3}$$

and is used to adjust the state estimates to take new measurements into account, giving the final estimate for this sample as:

$$\underline{x}_{k|k} = \underline{x}_{k|k-1} + K_k(\underline{y}_k - A\underline{x}_{k|k-1}) \tag{4}$$

It can be seen from (3) that the higher the measurement noise is, the lower the Kalman gain is. This results in the measurement being given less weight when correcting the state prediction in (4), and vice versa when the measurement noise is low. Following this correction, the state error covariance matrix is recalculated as:

$$P_{k|k} = (I - K_k A)P_{k|k-1} \tag{5}$$

after which processing begins for the next sample, with the next states being predicted based on the corrected final estimates for the current states.

For non-linear systems (such as the ECG model introduced below), a modified version of the standard Kalman filter is required: the EKF. In this case the transition matrix doesn't exist. Instead, the transition functions are linearised around the current states at each time step, using a first order Taylor expansion. The linearised transition functions are then evaluated and converted into a matrix in order to calculate the covariance and gain matrices, from which point the filtering process continues as above.

B. ECG Dynamical Model

In order to use the EKF, a model of the ECG signal was required to predict next values based on current states. As introduced in [11], a standard ECG can be approximated as a series of deflections away from the baseline representing the P, Q, R, S and T waves, each of which can be fitted to a scaled Gaussian. The dynamical model introduced uses a three-dimensional state equation, with x and y coordinates representing two-dimensional periodic circulation around a limit cycle and the z coordinate representing the ECG amplitude.

This model was modified in [12] by changing the coordinate system to its polar form, allowing omission of a redundant state equation and inclusion of the signal's phase as an explicit state variable, named θ. This modified version of the model was chosen for our implementation to allow incorporation of recorded audio signals as phase measurements, and has the following discretised form for time step t:

$$\begin{aligned} \theta[k+1] &= \theta[k] + \omega \cdot t \\ z[k+1] &= z[k] - \sum_{i \in \{P,Q,R,S,T\}} t \cdot n_i \end{aligned} \tag{6}$$

where

$$n_i = a_i \cdot \Delta\theta_i \cdot \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) \tag{7}$$

with a_i, b_i, and θ_i representing the Gaussian height, width and location in the cycle for each deflection respectively, and $\Delta\theta_i = (\theta - \theta_i)\mathrm{mod}(2\pi)$.

C. Ear Canal Heart Sounds

For segmenting recorded biopotentials into individual cardiac cycles and to provide phase information about the ECG, we record heart sounds from inside the ear canal. The audio signal consists of S1 heart sounds, corresponding to the QRS complex on the ECG, and S2 sounds, corresponding to the T wave. The sounds can be recorded leveraging a phenomenon called bone conduction: sounds inside the body are conducted through bones, causing the walls of the ear to vibrate [13]. Further to this, when the ear canal is sealed, low frequency sounds inside the canal are amplified through a phenomenon known as the occlusion effect. By occluding the ear canal (for example with a standard earphone tip) and placing a microphone inside, these amplified body sounds can be recorded, including the low frequency heart sounds.

III. Methodology

We recorded heart sounds from inside the ear canal which correspond to the R-peak in the ECG cycle. Initially these heart sounds were used to give timestamps demarcating individual cycles. The corresponding ECG cycles were then averaged together to form a single ensemble average cycle with reduced noise. This single cycle was used to fit the parameters of an ECG dynamical model to an individual user, which was then fed into the state prediction stage of an EKF alongside the continued noisy ear-ECG signal. The EKF also took heart sounds as inputs, giving it low-noise periodic measurements of the ECG phase, to produce a cleaned, accurate time series ear-ECG signal, from a fully explainable filtering process.

A. Hardware

To record ECG data from locations behind the ear we used the OpenEarable ExG, an open-source hardware platform designed to capture a range of biopotential signals from in and around the ear [14]. Standard Ag/AgCl electrodes were chosen for their adhesive properties to minimise motion artifacts. Active electrodes were placed on the mastoid on both sides, with a passive ground electrode on the back of the neck. The in-ear microphone of the OpenEarable 1.4 was used to record heart sounds simultaneously, with these two devices being connected via an I^{2}C bus to synchronise recordings with an interrupt [15]. Finally, a MAX30001G evaluation board was used with Ag/AgCl electrodes to record ground truth Lead I ECG data, with active electrodes on the left and right wrists and a ground electrode on the left wrist.

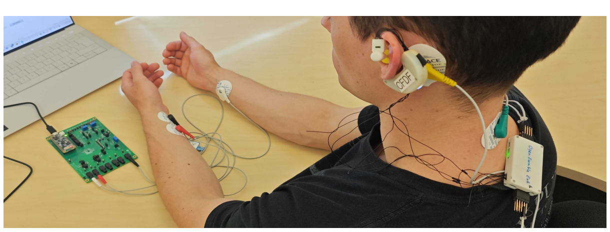

Fig. 2. Experimental Setup: MAX30001G evaluation board recording ground truth Lead I ECG at the wrists, and OpenEarable-based system recording heart sounds from the ear canal and ear-ECG signals

B. ECG Model Fitting

From the ensemble average of the initial recorded ECG signal, the R peak was found first. It was found as the signal maximum closest to the window boundaries, and hence closest to the S1 sounds used to compute the average. From here, the P, Q, S and T peak and trough phases were extracted by searching for maxima and minima working outwards. The amplitudes and Gaussian widths at half-height of all five waves were also extracted, allowing the Gaussians of the dynamical model to be fully fitted to the recorded ensemble average.

C. Extended Kalman Filter

We used an adapted version of the EKF for ECG introduced in [12]. In our system, we defined the ECG phase as zero at the time of an R-peak, which also corresponds to detection of an S1 heart sound. When one of these sounds was detected, a phase measurement of zero was input to the EKF with a low measurement noise setting, giving the EKF periodic corrections and allowing cardiac cycles to be followed by the system despite the high measurement noise in the ECG amplitude state.

D. Data Collection

Eighteen healthy participants (12 males and 6 females, ages 21 to 53) were invited for data collection, which involved participants sitting stationary for 3 minutes wearing the ear-ECG system and electrodes on their wrists connected to the ground truth evaluation board. Care was taken to ensure a good fit on participants' ears, since a good seal between the ear tip and the ear canal was critical for recording good quality heart sounds. A total of 54 minutes of ear-ECG data was collected.

The ear-ECG data was filtered during recording with a fourth-order Butterworth bandpass filter, with cutoff frequencies at 0.5 and 50 Hz. All data was trimmed to remove noise from motion at the beginning and end of recordings. The OpenEarable ExG had some variation in sample rate across the data collection, so all recordings were resampled to a consistent 128 Hz, with the audio recordings also being downsampled from 16 kHz to 128 Hz for alignment in the EKF. Heart sound timings were extracted from the audio signal using peak detection in a moving window.

IV. Results

Our results show promising performance. Fig. 3 displays a sample output of the full system, plotted with the original noisy signal and the ground truth, and Fig. 4 shows the ensemble average denoised signal and corresponding ground truth for one participant. It is evident that the system has extracted the ECG signal and successfully suppressed the noise. To evaluate the effectiveness of the denoising quantitatively, the signal-to-noise ratio (SNR) of the recorded and denoised signals from each participant were calculated using the ”snr” function in MATLAB. The mean input SNR to the filter from recordings was -14.9 dB (σ = 3.3 dB), and the mean output SNR was 6.4 dB (σ = 1.7 dB). This gave a mean improvement of 21.3

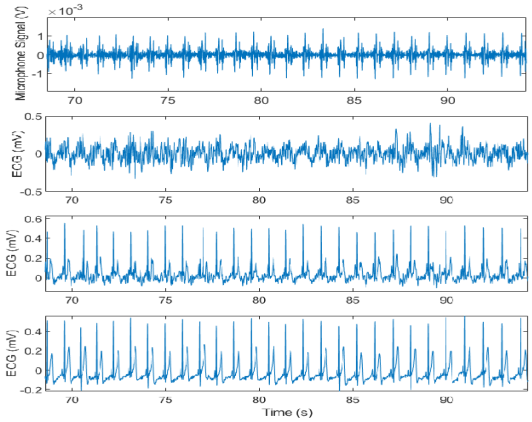

Fig. 3. Example recorded signals for (top to bottom): in-ear heart sound audio, unprocessed ear-ECG, denoised ear-ECG, ground truth ECG

dB (σ = 2.3 dB), significantly higher than any other ear-ECG denoising works. These results are shown in Fig. 5.

To ensure that the extracted ECG signal was correct as well as clean, the heart rate (HR) and QT intervals of the denoised signals were computed and compared to those of the ground truth recording. The mean absolute HR error was 3 BPM (σ = 2 BPM), and the mean absolute QT interval error was 54 ms (σ = 40 ms). For context, normal resting HRs range from 60-100 BPM, and QT intervals below 440 ms are generally considered healthy. These results demonstrate that the system does not simply generate a random clean ECG signal, but accurately denoises the underlying signal in the recording.

In future work we hope to deploy the system in clinical conditions to evaluate its performance on abnormal ECG morphologies. It is expected that some of these morphologies, such as inverted T-waves, should be dealt with straightforwardly by the current system, whereas conditions causing large temporal variations, such as premature ventricular contractions and fibrillation, will be more challenging and require adaptations.

V. Conclusions

This paper introduced a fully explainable approach for recording and denoising time series ear-ECG signals while stationary. Specifically, we employed in-ear microphones to provide information about the phase of individual cardiac cycles, allowing an Extended Kalman Filter to remove noise from other high-amplitude biosignals and retrieve the ECG signal, with a mean SNR improvement of 21.3 dB. These results demonstrate the potential of the system as a base to build passive ear-ECG monitoring upon.

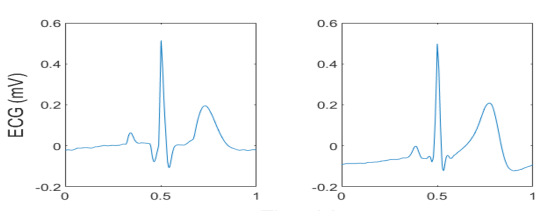

Fig. 4. Ensemble average denoised ear-ECG signal (left), ensemble average ground truth Lead I ECG (right)

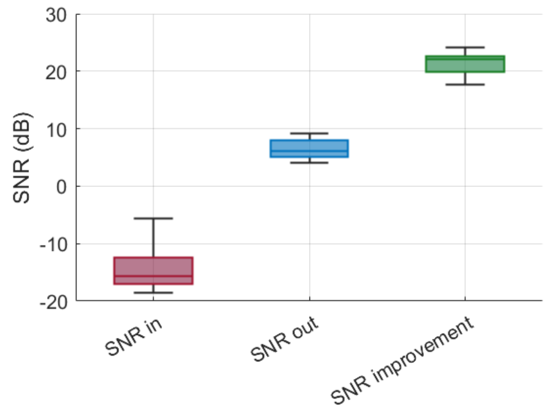

Fig. 5. Boxplots of unprocessed ear-ECG SNR, denoised ear-ECG SNR and improvement in SNR from processing

References

[1] M. B. Conover, *Understanding electrocardiography*. Elsevier Health Sciences, 2002.

[2] D. B. Geselowitz, "On the theory of the electrocardiogram," *Proceedings of the IEEE*, vol. 77, no. 6, 2002.

[3] W. Von Rosenberg, T. Chanwimalueang, V. Goverdovsky, N. S. Peters, C. Papavassiliou, and D. P. Mandic, "Hearables: Feasibility of recording cardiac rhythms from head and in-ear locations," *Royal Society open science*, vol. 4, no. 11, 2017.

[4] P. Kidmose, D. Looney, M. Ungstrup, M. L. Rank, and D. P. Mandic, "A study of evoked potentials from ear-eeg," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, 2013.

[5] T. King, M. Knierim, P. Lepold, C. Clarke, H. Gellersen, M. Beigl, and T. Röddiger, "eareog via periauricular electrodes to facilitate eye tracking in a natural headphone form factor," *arXiv preprint arXiv:2506.07193*, 2025.

[6] S. Park, S.-U. Kim, S.-I. Choi, and H.-J. Hwang, "Systematic investigation of optimal electrode positions and re-referencing strategies on ear biosignals," *International Journal of Human–Computer Interaction*, vol. 41, no. 2, 2025.

[7] A. Adarsh, R. Jhanavi, T. Jayas, A. Biradar, K. Muralidharan, A. Pal, and J. Gubbi, "Design and evaluation of ear-ecg system," in *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2024.

[8] E. Occhipinti, M. Zylinski, H. J. Davies, A. Nassibi, M. Bermond, P. Bachtiger, N. S. Peters, and D. P. Mandic, "In-ear ecg signal enhancement with denoising convolutional autoencoders," *arXiv preprint arXiv:2409.05891*, 2024.

[9] C. Santos, S. Frey, A. Cossettini, L. Benini, and V. Kartsch, "Real-time, single-ear, wearable ecg reconstruction, r-peak detection, and hr/hrv monitoring," *arXiv preprint arXiv:2505.01738*, 2025.

[10] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.

[11] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE transactions on biomedical engineering*, vol. 50, no. 3, 2003.

[12] R. Sameni, M. B. Shamsollahi, C. Jutten, and M. Babaie-Zade, "Filtering noisy ecg signals using the extended kalman filter based on a modified dynamic ecg model," in *Computers in Cardiology, 2005*. IEEE, 2005.

[13] K.-J. Butkow, T. Dang, A. Ferlini, D. Ma, and C. Mascolo, "heart: Motion-resilient heart rate monitoring with in-ear microphones," in *2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2023.

[14] P. Lepold, T. Röddiger, T. King, K. Kunze, C. Maurer, and M. Beigl, "Openearable exg: Open-source hardware for ear-based biopotential sensing applications," in *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2024.

[15] T. Röddiger, J. Stuchbury-Wass, M. Ciliberto, P. Lepold, and M. Beigl, "Openearable 1.4: Dual microphones earpiece to capture in-ear and outer-ear audio signals," in *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2024.