Kalman Filter Based Estimation of Auditory Steady State Response Parameters

Robert Luke and Jan Wouters

Abstract— **Auditory steady state responses (ASSRs) are brain responses to modulated or repetitive stimuli that can be captured in the EEG recording. ASSRs can be used as an objective measure to clinically determine frequency specific hearing thresholds, to quantify the sensitivity of the auditory system to modulation, and have been related to speech intelligibility. However, the detection of ASSRs is difficult due to the low signal to noise ratio of the responses. Moreover, minimizing measurement time is important for clinical applications. Traditionally ASSRs are analyzed using discrete Fourier transform (DFT) based methods. We present a Kalman filter based ASSR analysis procedure and illustrate several benefits over traditional DFT based methods. We show on a data set of 320 measurements that the proposed method reaches valid amplitude estimates significantly faster than the state of the art DFT method. Further, we provide two extensions to the proposed method. First, we demonstrate information can be incorporated from multiple recording electrodes by extending the system model. Secondly, we extend the model to incorporate artifacts from cochlear implant (CI) stimulation and demonstrate electrically evoked auditory steady state responses (EASSRs) can be accurately measured.**

Index Terms— **Auditory steady state responses, cochlear implants, evoked responses, Kalman filter, objective measures.**

I. INTRODUCTION

AUDITORY steady state responses (ASSRs) are neural responses to repetitive auditory stimuli [1]. These evoked potentials are used clinically to objectively determine frequency specific hearing thresholds [2], [3]. Typical ASSR stimuli are amplitude or frequency modulated carriers. Low frequency amplitude modulated ASSRs can be used to probe the auditory systems sensitivity to modulation rates found in the speech envelope. The presence and strength of ASSRs have been related to speech in noise perception [4], [5].

Hearing thresholds can be determined objectively by decreasing the stimulus intensity until no response is detected [2]. Longer measurements provide more accurate ASSR amplitude estimates. Increasing measurement time at each intensity from 5 to 10 min increases the correlation between objective and behavioral thresholds [3]. In clinical conditions the measurement time should be as short as possible while maintaining a high level of accuracy.

To elicit ASSRs, stimuli are presented to the listener at a set repetition or modulation rate (f_m) using clicks, amplitude or frequency modulated signals. Signals recorded using EEG or MEG can then be found to contain increased energy at f_m. Measurements also contain sensor noise and spontaneous neural activity, collectively termed measurement noise. Determining accurate estimates of the evoked response amplitude is difficult due to the small signal-to-noise ratio (SNR). The most commonly used method for determining the amplitude at f_m is with a DFT. The gold standard method is to cut the measurement in to smaller segments (epochs), use a DFT to analyze each epoch, and take the mean value of the bin corresponding to f_m as the signal amplitude. An estimate of the noise can be achieved by analyzing adjacent frequency bins or looking at the variation in the response bin across epochs [6]. The gold standard method to statistically determine the presence of a neural response is to use Hotelling's T^2 test [1].

ASSRs can also be elicited with cochlear implant (CI) users [7]. These are termed electrically evoked auditory steady state responses (EASSRs). In the CI population EASSRs have been used to predict thresholds levels [8] and have been related to modulation sensitivity [9]. EEG recordings performed in CI users while their devices are active contain artifacts from the RF transmission and electrical stimulation that obscures the neural response, the artifact contains energy at both the simulation and modulation rates. These artifacts can be removed by interpolating between artifact free segments of EEG [8]. This is similar to downsampling, but can not be used with a pre filter as that would smear out the CI artifact.

DFT based analysis of ASSR measurements has several shortcomings. First, the measured response is assumed to be stationary, but effects such as attention or adaptation may cause the ASSR to vary over time [10]. Secondly, the DFT based method can not be applied in real time. Lastly, no consensus has been made on how best to integrate information from multiple electrodes. Various approaches have been proposed to utilise information from multiple recording electrodes including averaging the signals in the time domain [9], averaging the output statistic [11], spatial filtering [12] and blind source separation [13]. DFT based EASSR

analysis has additional shortcomings. First, the interpolation based artifact removal technique introduces aliasing in the signal. Secondly, only a small portion of the measured data is used in the analysis. Lastly, artifact free segments are required in the data, which are not available when using clinical stimulation parameters.

Here, we propose a Kalman filter to estimate the ASSR parameters. This approach requires a model of the signal being measured, we propose a simple sinusoid model with unknown amplitude and phase [14]. When the system model is correct and the noise assumptions of the filter are met, the Kalman filter produces optimal estimates of the system variables. We hypothesise that for low SNR signals the Kalman filter will converge to valid amplitude estimates faster than traditional techniques. We propose two extensions to this Kalman filter technique. To integrate information from multiple recording electrodes we propose a multi input multi output model. And for processing recordings from CI users we propose a system model that incorporates CI artifact information.

The Kalman filter approach provides several advantages over traditional DFT based ASSR processing. First, no assumption of signal stationarity is made. Secondly, changes in the response can be analyzed at each sample, opening up new research avenues in to areas such as auditory adaptation or attention. Thirdly, the Kalman filter can be updated in real time during signal acquisition. Lastly, information from multiple recording electrodes can be included in the response estimate. Contrary to DFT based processing, the Kalman filter approach to EASSR processing does not introduce aliasing in the signal and utilizes all the measured data.

Accurate estimates of EEG spectral content has many uses. Information about the amplitude at multiple frequencies, or combined with phase estimates, can be used to statistically assess the presence of a neural response. Beyond ASSRs, accurate spectral estimates may be useful in studies of attention or neurological disorders [5], [15].

The advantages of a state estimation approach over DFT methods in the analysis of nonstationary periodic signals has been previously investigated [16]. A single input single output Kalman filter has been proposed previously for analysis of low SNR EEG signals, specifically the distortion product otoacoustic emissions [17]. But no comparison has been made to current gold standard DFT based methods for measuring ASSRs. Here we compare the proposed method to the current gold standard technique and extend this model to integrate information from multiple electrodes and remove artifacts from EASSRs.

The remainder of this paper is organized as follows. The proposed state space models are defined. The experimental stimuli and recording procedure are described. The statistical tests used to compare the current gold standard DFT procedure and proposed method are described. Results are presented and discussed.

II. METHODS

In this section the proposed state space models are defined. The stimuli, data collection and signal processing is described.

As is the statistical tests to compare results from the current gold standard DFT and proposed methods.

A. Models

Three state space models are introduced to describe: 1) single channel ASSRs, 2) multichannel ASSRs, and 3) single channel EASSRs. A brief description of each model is provided and the model is explicitly defined.

Equations (1) and (2) define the standard discrete state space form all models are placed in [18]. Bold upper case symbols represent matrices, bold lowercase symbols represent vectors and regular lowercase symbols represent numbers. The system state at sample k is represented by x_k. The state transition model is represented by F_k and the process noise is represented by w_k. The observation model is represented by H_k and the measurement noise is represented by v_k

$$x_{k+1} = F_k x_k + w_k \tag{1}$$
$$z_k = H_k x_k + v_k. \tag{2}$$

1) Acoustic Stimulation Model: A sinusoid with unknown amplitude and phase is used to model the ASSR [14]. This model is presented in (3) and (4). The amplitude at each sample (A_k) is given in (5). The phase at each sample (ϕ_k) is given in (6)

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k + w_k \tag{3}$$

$$z_k = \begin{bmatrix} \cos(\omega \Delta t * k) & -\sin(\omega \Delta t * k) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k + v_k \tag{4}$$

$$A_k = \sqrt{x_{1,k}^2 + x_{2,k}^2} \tag{5}$$

$$\phi_k = \arctan(\frac{x_{2,k}}{x_{1,k}}). \tag{6}$$

2) Multichannel Model: A multi input multi output model was designed to integrate information from multiple channels. The number of channels is represented by n. The measurement on channel i and sample k is represented by $z_{i,k}$. This model assumes that each adjacent electrode is measuring the same underlying source with independent noise, this is a common assumption for neighboring electrodes in high density montages [19]. The same state model in (3) is used, but multiple observations are generated as described in (7)

$$\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}_k = \begin{bmatrix} \cos(\omega \Delta t * k) & -\sin(\omega \Delta t * k) \\ \vdots & \vdots \\ \cos(\omega \Delta t * k) & -\sin(\omega \Delta t * k) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k + v_k. \tag{7}$$

3) Electrical Stimulation Model: Measurements of EASSRs are corrupted by artifacts from RF transmission and electrical stimulation. The exact timing and frequency f_a of these artifacts is controlled by the experimenter and thus known a-priori. Fig. 1(a) shows an example EASSR measurement and Fig. 2(a) shows the frequency content of this signal. It should be noted that the artifact is ≈ 2000 times larger than the signal of interest.
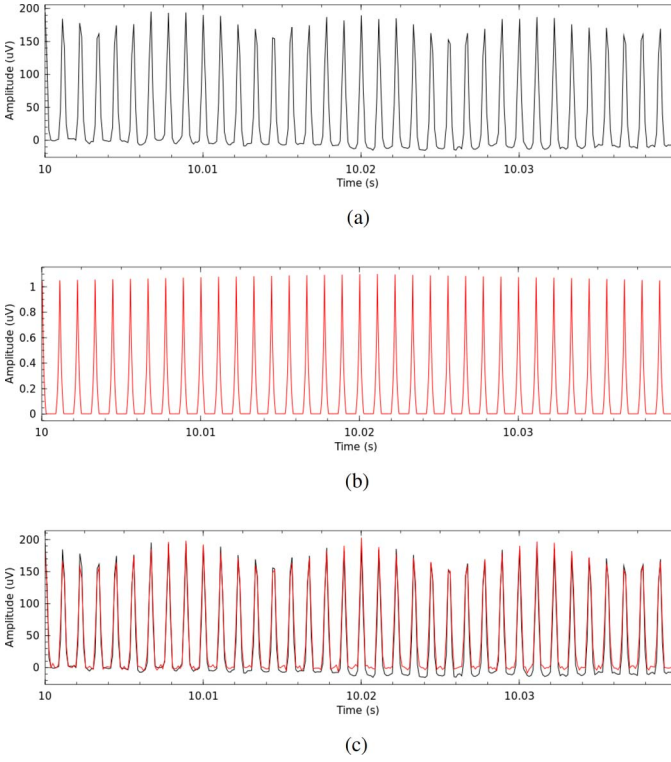
Fig. 1. Example EASSR measurement and model realisation. a) EASSR measurement sampled at 8192 Hz, b) realisation of artifact model sampled at 81920 Hz to capture all features, c) EASSR model realisation (red) overlaid on example measurement (black) both sampled at 8192 Hz.
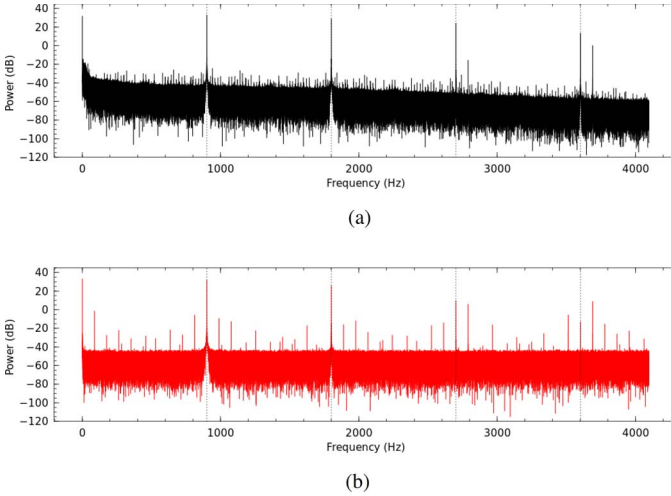

Fig. 2. Spectrum of EASSR measurement and model realization. (a) Measurement (b) Model.

A model of the CI artifact is added to the model in Section II-A1. The artifact is modelled as a linearly increasing and decreasing signal with periodicity equal to the cochlear implant pulse rate ($sw(\Delta t * k, f_a, w)$), where t is the width of the artifact. Fig. 1(b) shows an example of this artifact model. An additional state variable (x_3) is introduced to track the amplitude of the artifact. The complete EASSR model is provided in (8), (9) and (10). An example realization of the EASSR model overlaid on a measurement can be seen in Fig. 1(c) and the frequency content of this realization can be

seen in Fig. 2

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_k + w_k \quad (8)$$

$$z_k = \begin{bmatrix} \cos(\omega \Delta t * k) & -\sin(\omega \Delta t * k) \\ & sw(\Delta t * k, f_a, w) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_k + v_k \quad (9)$$

$$sw(\Delta t * k, f_a, w) = \begin{cases} 1 - \frac{d}{w} & \text{for } -w \leq d \leq w \\ 0 \end{cases}$$

$$\text{where} \quad d = \min(\text{mod}(-t, \frac{1}{f_a}), \text{mod}(t, \frac{1}{f_a})). \quad (10)$$

4) Filter: A standard Kalman filter was applied to the data to determine state estimates at each sample [20]. Initial estimates for x_1 and x_2 were set to 0, representing an amplitude of 0 uV. In the EASSR model the initial value of x_3 was set to the maximum value measured.

The process noise covariance was set to $1e^{-7}$ to represent that little variation was expected in response amplitude. The measurement noise covariance was set to the covariance of the measured signal.

A Rauch–Tung–Striebel (RTS) smoother was used in the Kalman filter based analysis [21]. This consists of a forward pass Kalman filter and a backward pass recursion smoother, it provides better estimates than the forward pass alone.

B. Stimuli

All experiments were conducted in a sound proof and electromagnetically shielded room. EEG measurements were recorded with a BioSemi ActiveTwo 64 channel system. Data was recorded with a sample rate of 8192 Hz and stored for offline processing. Testing was approved by the Medical Ethics Committee of the UZ Leuven (approval number B32220072126) and informed consent was obtained from all participants.

1) Acoustical Data: Twenty Flemish speakers, mean age years 21 (SD = 1.3 years), volunteered for this study. Stimuli consisted of amplitude modulated speech-weighted noise, presented to the participants at 70 dB SPL through ER-3A insert phones. Modulation rates 4, 10, 20, 40, and 80 were each presented for 5 min in a random order. A resting state condition was also measured for 2.5 min. Stimuli were presented binaurally and monaurally to each ear. 16 measurements were taken per participant, in total 320 measurements were recorded. Before further processing the data was downsampled to a sample rate of 256 Hz.

2) Electrical Data: Seven Flemish speakers, mean age 58 years (SD = 19 years), volunteered for this study. All participants were unilateral CI users with Cochlear Nucleus devices. Stimuli consisted of amplitude modulated biphasic pulse trains presented via the NIC research platform. All stimuli were presented in a bipolar mode, both the active and return electrodes were located within the cochlea. In each participant three electrode pairs were stimulated, a basal, middle and apical electrode pair. To match the users daily

processor, a pulse rate of 900 pulses per second was used, the pulses were modulated at a 40 Hz rate. Stimuli were modulated between the threshold and comfort level as described in [9]. Each electrode pair was stimulated twice, using either anodic or cathodic first biphasic pulses. In one participant stimulation of the basal electrode pair was not possible. In total 40 measurements were recorded.

C. Signal Processing

All measurements were processed using the gold standard DFT and proposed Kalman filter methods. Details of each analysis procedure are presented below. Each measurement was analyzed to determine the amplitude at the modulation rate. The resting state data was analyzed to determine the amplitude at 30 Hz, only noise was expected at this rate.

1) DFT Analysis: The DFT analysis begun with a high pass filter from 2 Hz, each signal was then referenced to electrode Cz. The signal was cut in to epochs of 1.024 s, 5% of these were rejected based on peak to peak amplitude. A Fourier transform was conducted on each epoch and the mean amplitude of the bin corresponding to the modulation rate was recorded as the signal amplitude. The standard deviation of this bin divided by the number of epochs was recorded as the noise amplitude.

2) Kalman Analysis: The Kalman analysis consisted of a de-trending procedure with a second order polynomial over a 0.5 s window [22]. De-trending was used instead of low pass filtering as not to smear the artifact in the EASSR analysis. The signal was then referenced to Cz and passed to the Kalman filter and RTS smoother with initial parameters as described in Section II-A4. The amplitude was estimated at each sample using (5), the mean amplitude over all samples was taken for comparison with the DFT analysis.

3) Single Channel Analysis: Single channel analysis was run on electrodes P5, F6, Fp1, and O1. P5 and O1 are commonly used in ASSR analysis as they have a large response amplitude relative to the noise. F6 and Fp1 were analyzed to represent more difficult channels for ASSR detection. Fp1 is commonly contaminated with blink artifacts and both Fp1 and F6 are frontal electrodes with a smaller response amplitude than the electrodes at the back of the head.

4) Multi Channel Analysis: Electrodes P3, P5, PO7, and PO3 were analyzed in the multi channel analysis. These are neighboring electrodes that are commonly used for ASSR analysis. To integrate the information in the traditional processing scheme, the waveforms from multiple channels were averaged after re-referencing to produce a single signal, FFT based analysis was performed on this single channel. In the proposed method, all the channels were passed in to the model as described in (3) and (7).

5) EASSR Analysis: The aim of the EASSR analysis was to determine if the evoked response amplitude could be accurately measured in the presence of the CI artifact. The CI artifact is not limited to the stimulation rate, but also contains energy at the modulation rate [8]. The electrode contra-lateral to the CI was analyzed, either P5 or P6. As a first step in the DFT analysis the signal was blanked 0.0009 s

after each pulse [8]. This value was chosen as [9] showed that the artifact only lasted 0.0007 s with the same stimuli. An anti-aliasing filter was not used prior to interpolation as this smears out the artifact in the time domain, leaving no response dominated samples to interpolate between. For the Kalman analysis (8) and (9) was used.

D. Statistics

Two tests were conducted to ensure the proposed method met or exceeded the performance of the DFT procedure. First, the entire measurements were analyzed to confirm the two measures produced the same amplitude estimates, a regression and Bland Altman analysis was conducted. Secondly, the recordings were split in to smaller segments to determine which measure converged to a valid solution quicker.

1) Regression Analysis: To determine if the two methods produce similar amplitude estimates a linear regression was performed comparing the amplitude from a DFT analysis to the Kalman amplitude estimate. The entire 5 min recording was used in both methods.

2) Bland Altman Analysis: Regression analysis has many shortcomings when comparing two measures [23]. A Bland Altman visual method was used to compare the two measures and define a repeatability coefficient. By definition, 95% of measurements will have a difference in amplitude estimates less than the repeatability coefficient. This analysis was conducted on the entire data set (N = 1536). As the noise level varies per modulation rate, the Bland Altman analysis was also conducted for each modulation rate. The repeatability coefficient was compared to the noise level at each frequency to determine if the differences between measures was related to the noise level or external factors.

3) Convergence Analysis: Analysis was run to determine which method converged to a valid solution faster. Each measurement was cut in to smaller segments from 10 to 250 s in 10 s increments. The amplitude as estimated using a DFT analysis of the full 5 min recording was taken as the true amplitude. The error was defined as the absolute value of the difference between the true amplitude and short segment amplitude. The amplitude estimate was considered valid once the error between it and the 5 min solution was less than the noise level determined on the full 5 min recording using a DFT analysis and remained valid for at least 20 s. The improvement in convergence time was calculated per measurement and for each recording electrode. The mean absolute error was analyzed for each segment from 10 to 150 s.

III. RESULTS

A. Acoustic Stimulation

1) Example: Fig. 3 provides an example of the amplitude estimate over time. Both the DFT and Kalman estimates are visualised. The final DFT estimate is indicated by a dashed black line, the red dashed line indicates the DFT estimate ± the DFT noise amplitude. The blue line indicates the how the DFT estimate is updated over the measurement. The black line represents the Kalman filter estimate, it is initialised with 0 uV amplitude and converges to within the noise level in ≈ 10 s
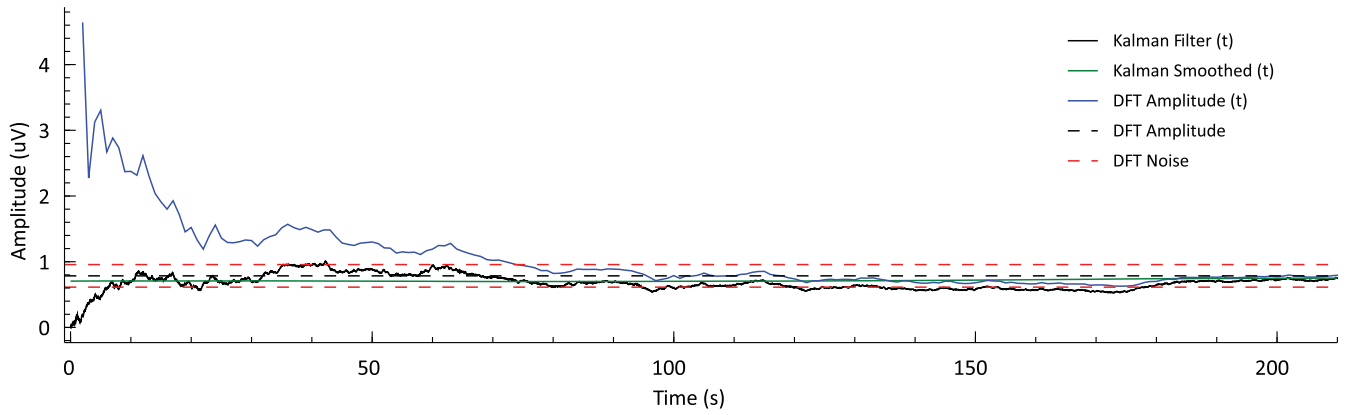
Fig. 3.   Example amplitude estimation for a 4 Hz modulation rate using both Kalman and DFT measures. Time varying estimates are shown as solid lines. The dashed black line indicates the DFT amplitude estimate using the entire 5 min measurement, the dashed red line indicates the amplitude estimate $\pm$ the final noise amplitude.
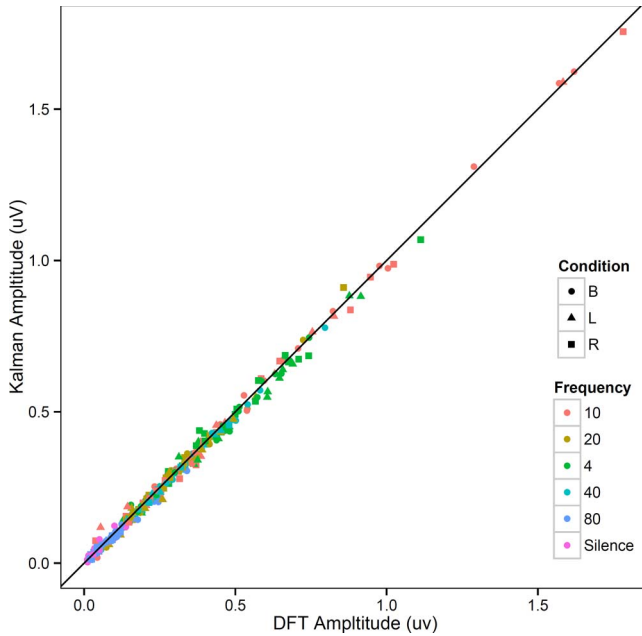

Fig. 4.   Comparison of amplitude estimates for electrode P5 using DFT and Kalman methods. Solid line indicates the line of perfect match.

before fluctuating within the expected noise amplitude range. The green line indicates the RTS smoothed estimate using data from the entire recording.

2) Regression Analysis: Regression analysis showed a significant relation between Kalman and DFT amplitude estimates. Five minute ASSR measurements and 2.5 min resting state conditions were analyzed. A total of 320 amplitude estimates were obtained with the DFT and Kalman methods per electrode. Four electrodes were analyzed to represent both high and low SNR situations. Fig. 4 shows the relation between measures for electrode P5. All points lie along the mid line, indicating good agreement. There was a statistically significant relation between measures for all electrodes. For P5($R^2 = 0.99$, $p < 0.001$, DF $= 318$), for F6($R^2 = 0.95$, $p < 0.001$, DF $= 318$), for Fp1($R^2 = 0.92$, $p < 0.001$, DF $= 318$), for O1($R^2 = 0.99$, $p < 0.001$, DF $= 318$).

3) Bland Altman Analysis: A Bland Altman analysis indicated the difference in measures is less than the EEG noise. The mean difference between the two measures was -0.0035 uV. For the entire data set the repeatability coefficient was 0.088 uV and the mean noise level was 0.11 uV.

Fig. 5 shows a Bland Altman plot. The repeatability coefficient is indicated by dashed lines. The average noise level is plotted as a dotted line.

Fig. 6 plots the average noise per electrode and frequency against the repeatability coefficient. The points generally lie above the solid line indicating the difference in measures is less than the frequency specific measurement noise.

4) Multichannel Analysis: Regression analysis showed that the Kalman and DFT analysis estimated the amplitude similarly when integrating information from multiple channels. DFT analysis averaged the signals in time domain then the standard processing was used. For the Kalman process all channels were fed to the filter. Regression analysis determined the two measures were related ($R^2 = 0.99$, $p < 0.001$, DF $= 318$). The Bland Altman repeatability coefficient was 0.048 uV. The mean noise level was 0.075 uV. Fig. 7 illustrates the similarity between DFT and Kalman estimated amplitudes for the multi channel data.

5) Time Analysis: Fig. 8 illustrates how the mean absolute amplitude error varies with the length of recording for two electrodes and two modulation rates, the dashed horizontal line indicates the average noise level. Results are described for each frequency as the noise level decreases with increasing frequency. For each modulation rate the amplitude was considered valid when the mean absolute error was less than the mean noise level as determined on the full 5 min recording.

The improvement in convergence time was calculated for each measurement on each of the four recording electrodes. On average the proposed Kalman approach converged to a solution 15% faster than the DFT method, a t test indicates this is a significant improvement t(1279) = 5.7, p < 0.001. For 4 Hz conditions the improvement was 27% t(239) = 6.0, p < 0.001. For 10 Hz conditions the improvement was 26% t(239) = 4.8, p < 0.001. For 20 Hz conditions the
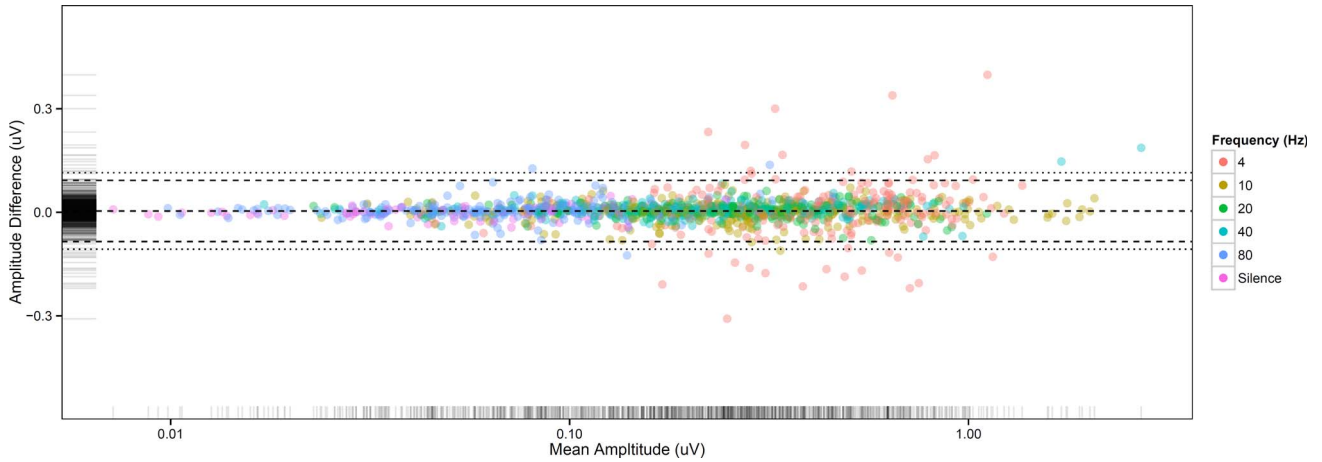
Fig. 5. Bland Altman analysis of difference between DFT and Kalman amplitude estimates for single channel analysis of six modulation rates and four electrodes (N = 1, 280). Dashed lines indicate the repeatability coefficient and dotted lines indicate the mean DFT noise level. The rug indicates the density of distribution for each axis.
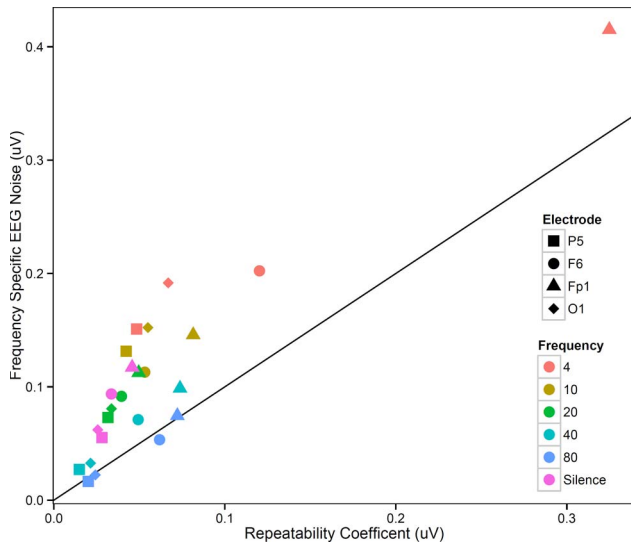

Fig. 6. Relation between repeatability coefficient and noise.

improvement was 3% t(239) = 0.2, p = 0.5. For 40 Hz conditions the improvement was 10% t(239) = 1.5, p = 0.1. For 80 Hz conditions the improvement was 6% t(239) = 0.8, p = 0.4. For resting state data the improvement was 27% t(79) = 4.0, p < 0.05.

When all conditions and modulation rates were combined, the mean absolute error for the Kalman filter achieved a valid amplitude on average between 50 and 60 s, the traditional method between 80 and 90 s. For 4 Hz the Kalman filter achieved a valid amplitude on average between 50 and 60 s, the traditional method between 80 and 90 s. For 10 Hz the Kalman filter achieved a valid amplitude on average between 60 and 70 s, the traditional method between 90 and 100 s. For 20 Hz the Kalman filter achieved a valid amplitude on average between 80 and 90 s, the traditional method between 80 and 90 s. For 40 Hz the Kalman filter achieved a valid amplitude on average between 90 and 100 s, the traditional method between 100 and 110 s. For 80 Hz the Kalman filter

achieved a valid amplitude on average between 30 and 40 s, the traditional method between 70 and 80 s. For resting state data the Kalman filter achieved a valid amplitude on average between 10 and 20 s, the traditional method between 50 and 60 s.

B. Electrical Stimulation

The proposed method determined the EASSR amplitude to be the same as the DFT based blanking method for all participants. Artifact amplitudes varied between participants from 5 uV to 2000 uV. The width of the artifact was chosen by visual inspection of the raw EEG data.

1) Regression Analysis: Regression analysis confirmed the amplitudes from the two methods are related ($R^2 = 0.94$, $p < 0.001$, df = 38).

2) Bland Altman Analysis: Fig. 9 shows the Bland Altman analysis for the EASSR files. The Bland Altman analysis determined the mean difference between methods to be 0.006 uV. The repeatability coefficient was 0.047 uV, the repeatability coefficient for acoustic ASSRs at 40 Hz was 0.051 uV.

IV. DISCUSSION

A method for the real time analysis of ASSR signals was presented. For recordings of 5 and 2.5 min, the proposed method produced the same amplitude estimates as the gold standard DFT method. The variation in measures was shown to be less than the EEG noise level. The proposed Kalman approach reached a valid amplitude significantly faster than the DFT approach. Two extensions of this method were presented. The first, integrated information from multiple sensors. The second, allowed measurement of responses to CI stimulation.

For all modulation rates, the 5 min recordings produced on average essentially no difference in amplitude when compared to the DFT method. This indicates no bias is introduced in the proposed method, and is an essential requirement for any new measure. The variation in results was shown to be less than the EEG noise level. The authors propose this indicates
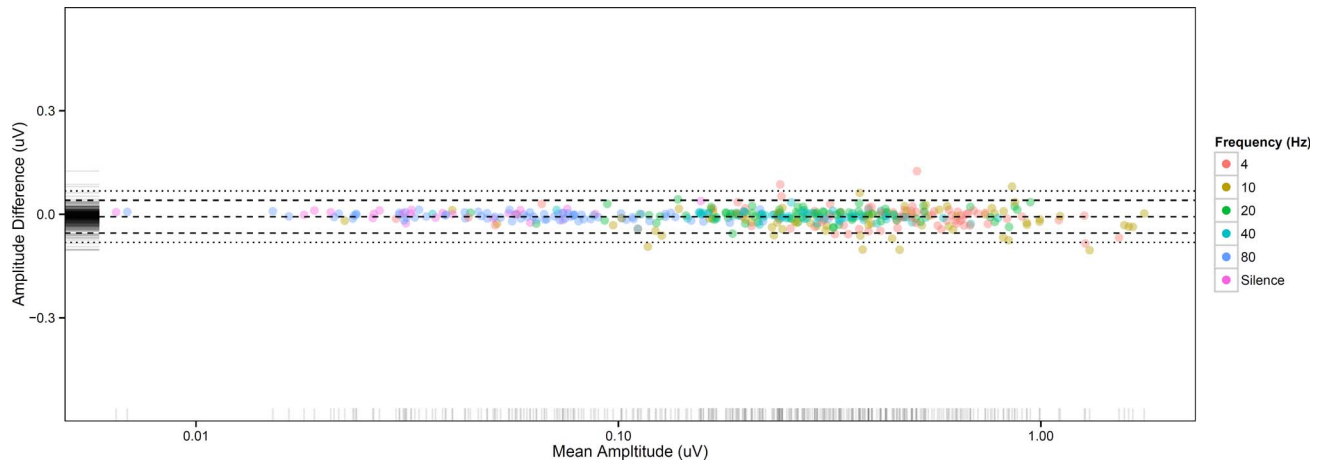
Fig. 7. Bland Altman analysis of difference between DFT and Kalman amplitude estimates for multi channel data (N = 320). Dashed lines indicate the repeatability coefficient and dotted lines indicate the mean DFT noise level. Rug indicates the density of distribution for each axis.
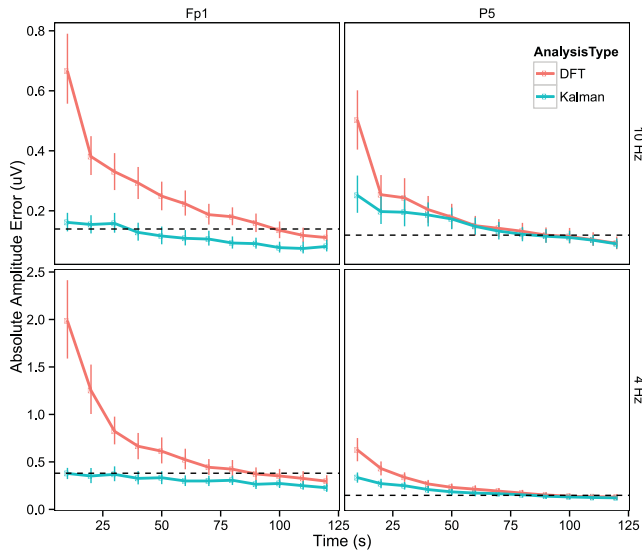

Fig. 8. Error in amplitude estimates as a function of time for two electrodes and two modulation rates. Error bars indicate 95% confidence intervals. Dashed lines indicate the noise amplitude.
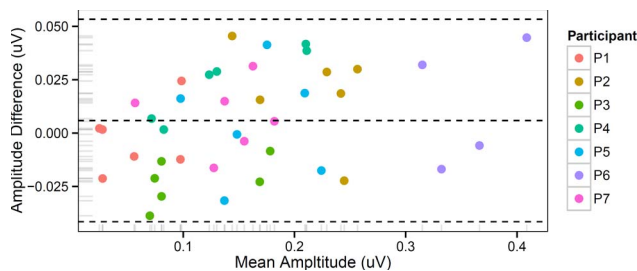

Fig. 9. Bland Altman analysis for EASSR recordings. Dashed lines indicate the mean difference and repeatability coefficient. Rug indicates the density of distribution for each axis.

the two measures can be considered equally accurate for long recordings. The two measures produced the same amplitude estimate for the resting state condition, indicating the Kalman filter approach also works in the absence of a significant neural response to the stimuli. Differences between the two

methods may be due to variation or nonstationarity in the ASSR amplitude over the length of the recording, this variation in the frequency of interest is not usually discussed in DFT based methods.

The recordings were analyzed at different lengths to determine which converged to a valid solution first. A valid solution was defined as when the absolute error in amplitude was less than the noise level for the full recording for at least 20 s. For all modulation rates the Kalman filter matched or outperformed the DFT method, the Kalman filter was significantly faster for three modulation rates. The greatest improvement in convergence time was found in low SNR situations such as frontal electrodes and low modulation rates. The authors propose two reasons for the improvement in detection time. First, the initial condition is set to an amplitude of 0 uV and increases based on the measurements, whereas DFT based methods often over estimate response amplitudes for short measurements. Secondly, the Kalman filter is less likely to be affected by brief periods of noisy data caused by blinking or neck movement. Given the greater reduction in measurement time for frontal electrodes, the authors hypothesise that the proposed method may further outperform DFT methods for stimuli presented closer to hearing threshold.

The proposed Kalman filter was extended to integrate information from multiple electrodes. The model is based on the assumption that each electrode is measuring the same underlying signal but with independent noise, this assumption is not strictly true [24], [25]. However, closely spaced electrodes often show similar amplitude and phase and are combined [11]. For entire scalp analysis, other methods may be more appropriate [13]. Less variability was shown when comparing amplitudes with the DFT method for the multiple electrode data than the single channel data. Indicating a benefit of integrating information from multiple sensors.

A model was developed to include the artifacts introduced when recording EEG with CI stimulation. The artifacts are caused by RF transmission and electrical stimulation. The

artifacts from participants in this study varied in amplitude between 5 uV and 2000 uV. A simple linearly increasing and decreasing model was proposed, this was based on visual inspection of raw EEG data and the signal spectrum. Realisations of the model showed good visual alignment with raw data. Kalman filtering of EASSR data with the extended model showed no difference in amplitudes compared to the current state of the art artifact removal method [8]. If the artifact model was not included in the Kalman filter, the amplitude estimates were inaccurate. The repeatability coefficient between methods for the 40 Hz EASSR was similar to the 40 Hz ASSR, indicating no loss of accuracy is produced relative to measurements in normal hearing participants. The proposed method has the same advantages when processing EASSR data as with the ASSR data, but also has the advantage that it does not require interpolation, which effectively downsamples without pre-filtering, causing aliasing in the signal.

CI stimuli were presented at 900 pulses per second in bipolar mode, these parameters have been previously validated with the blanking technique [8]. The CI artifact is known to vary with stimulation mode and level, and may vary between CI manufacturers. Further investigation is required to determine how the artifact varies with each stimulation parameter. If stimulation parameters dramatically affect the measured artifact, the model in (10) may need to be adjusted accordingly. However, the participants in this study had a representative range of threshold and comfort values, this demonstrated that the Kalman filter was robust to variations in stimulation level.

The DFT analysis methods assume the neural response is stationary, this may not be true as attention and alertness are known to affect auditory neural responses [10]. For the standard Kalman filter to provide an optimal estimate of state variables the model must be correct and noise must be white Gaussian. Fig. 2(a) illustrates that the EEG signal does not have white noise. Despite this limitation the Kalman filter outperforms the DFT based method and does not require the stationarity assumption. Using methods that account for correlated noise may provide further improvements [26].

Initial conditions were set to an amplitude of 0 uV as not to bias the comparison to DFT based methods. In an optimized system, the initial estimate could be set per modulation rate to the expected amplitude based on literature or pilot studies, this should further reduce the time required to converge to a valid amplitude estimate. Further, tracking the response harmonics could be included by simple extension of the model [27].

An alternative approach to validating the proposed Kalman filter, and comparing to the DFT approach, could be to simulate a neural response in noise and compare the results to the simulation parameters. This would allow analysis of the error in different conditions, but would lack the complexity of the experimental data.

V. Conclusion

A new method was proposed to measure auditory steady state responses. The proposed Kalman filter method produced the same amplitude estimates as the gold standard DFT based method for long measurements. The proposed method was shown to converge to a valid solution significantly faster than the current gold standard. The proposed method was extended to include a model of cochlear implant artifacts, and was shown to accurately estimate the amplitude of electrically evoked auditory steady state responses. The authors hope that removing the assumption of stationarity and allowing real time amplitude estimation will open up new avenues in ASSR research.

References

[1] T. W. Picton, *Human Auditory Evoked Potentials*. San Diego, CA: Plural, 2010.

[2] G. Rance, F. W. Rickards, L. T. Cohen, S. De Vidi, and G. M. Clark, "The automated prediction of hearing thresholds in sleeping subjects using auditory steady-state evoked potentials," *Ear Hear.*, vol. 16, no. 5, pp. 499–507, 1995.

[3] H. Luts, C. Desloovere, A. Kumar, E. Vandermeersch, and J. Wouters, "Objective assessment of frequency-specific hearing thresholds in babies," *Int. J. Pediatric Otorhinolaryngol.*, vol. 68, no. 7, pp. 915–926, 2004.

[4] A. Dimitrijevic, M. S. John, and T. W. Picton, "Auditory steady-state responses and word recognition scores in normal-hearing and hearing-impaired adults," *Ear Hear.*, vol. 25, no. 1, pp. 68–84, 2004.

[5] H. Poelmans *et al.*, "Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia," *Ear Hear.*, vol. 33, no. 1, pp. 134–143, 2012.

[6] T. W. Picton, J. Vajsar, R. Rodriguez, and K. B. Campbell, "Reliability estimates for steady-state evoked potentials," *Electroencephalogr. Clin. Neurophysiol./Evoked Potentials Section*, vol. 68, no. 2, pp. 119–131, 1987.

[7] M. Hofmann and J. Wouters, "Electrically evoked auditory steady state responses in cochlear implant users," *J. Assoc. Res. Otolaryngol.*, vol. 11, no. 2, pp. 267–82, Jun. 2010.

[8] M. Hofmann and J. Wouters, "Improved electrically evoked auditory steady-state response thresholds in humans," *J. Assoc. Res. Otolaryngol.*, vol. 13, no. 4, pp. 573–589, 2012.

[9] R. Luke, L. Van Deun, M. Hofmann, A. Van Wieringen, and J. Wouters, "Assessing temporal modulation sensitivity using electrically evoked auditory steady state responses," *Hear. Res.*, vol. 324, pp. 37–45, 2015.

[10] S. I. Kanoh, K.-I Miyamoto, and T. Yoshinobu, "A brain-computer interface (BCI) system based on auditory stream segregation," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, pp. 642–645.

[11] S. Vanvooren, H. Poelmans, M. Hofmann, P. Ghesquière, and J. Wouters, "Hemispheric asymmetry in auditory processing of speech envelope modulations in prereading children," *J. Neurosci.*, vol. 34, no. 4, pp. 1523–1529, 2014.

[12] A. De Cheveigné and J. Z. Simon, "Denoising based on spatial filtering," *J. Neurosci. Methods*, vol. 171, no. 2, pp. 331–339, 2008.

[13] H. M. Bharadwaj and B. G. Shinn-Cunningham, "Rapid acquisition of auditory subcortical steady state responses using multichannel recordings," *Clin. Neurophysiol.*, vol. 125, no. 9, pp. 1878–1888, 2014.

[14] A. A. Girgis and R. G. Brown, "Application of Kalman filtering in computer relaying," *IEEE Trans. Power Apparatus Syst.*, no. 7, pp. 3387–3397, Jul. 1981.

[15] B. Oken, M. Salinsky, and S. Elsas, "Vigilance, alertness, or sustained attention: Physiological basis and measurement," *Clin. Neurophysiol.*, vol. 117, no. 9, pp. 1885–1901, 2006.

[16] R. R. Bitmead, A. C. Tsoi, and P. J. Parker, "A Kalman filtering approach to short-time Fourier analysis," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 34, no. 6, pp. 1493–1501, Dec. 1986.

[17] X. Li, Y. Sokolov, and H. Kunov, "System and method for processing low signal-to-noise ratio signals," U.S. Patent, Patent 7 286 983, Oct. 23, 2007.

[18] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[19] A. De Cheveigné and J. Z. Simon, "Sensor noise suppression," *J. Neurosci. Methods*, vol. 168, no. 1, pp. 195–202, 2008.

[20] B. Anderson and J. B. Moore, "Optimal filtering," in *Prentice-Hall*. Englewood Cliffs, NJ: Prentice-Hall, 1979, vol. 1, Inf. Syst. Sci.

[21] H. E. Rauch, C. Striebel, and F. Tung, "Maximum likelihood estimates of linear dynamic systems," *AIAA J.*, vol. 3, no. 8, pp. 1445–1450, 1965.

[22] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[23] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, 1986.

[24] B. Roß, C. Borgmann, R. Draganova, L. E. Roberts, and C. Pantev, "A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones," *J. Acoust. Soc. Am.*, vol. 108, no. 2, pp. 679–691, 2000.

[25] P. L. Nunez *et al.*, "EEG coherency: I: Statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 5, pp. 499–515, 1997.

[26] K. Wang, Y. Li, and C. Rizos, "Practical approaches to Kalman filtering with time-correlated measurement errors," *IEEE Trans. Aerospace Electron. Syst.*, vol. 48, no. 2, pp. 1669–1681, Apr. 2012.

[27] N. C. Will and R. Cardoso, "Comparative analysis between FFT and Kalman filter approaches for harmonic components detection," in *Proc. 10th IEEE/IAS Int. Conf. IEEE Indust. Appl.*, 2012, pp. 1–7.

Robert Luke received the B.S. degree in physics and an electrical B.Eng. (Hons) degree from the University of Melbourne, Melbourne, Australia. He is a Ph.D. degree candidate at the Department of Neurosciences, University of Leuven, Leuven, Belgium..

His research interests include signal processing, control systems, neuroscience, and auditory processing.

Jan Wouters was born in 1960. He received the M.S. and Ph.D. degrees in physics from the University of Leuven, KU Leuven, Belgium, in 1982 and 1989, respectively, with intermission for officer military service.

From 1989 until 1992, he was a Postdoc Research Fellow with the National Fund for Scientific Research (FWO) at the Institute of Nuclear Physics (UCL Louvain-la-Neuve) and at NASA Goddard Space Flight Center (USA) with a Fullbright-Hays award. Since 1993 he is a Professor at the Department of Neurosciences of the KU Leuven (Full Professor since 2005) where he teaches five physics and audiology courses. His research focuses on audiology, the auditory system and auditory prostheses. He is author of about 250 articles in international peer-reviewed journals and Associate Editor of three international journals.

Dr. Wouters is past President of the European Federation of Audiology Societies EFAS, President of the Belgian Audiology Society B-Audio, Secretary General of the International Collegium of Rehabilitative Audiology ICRA.