BAYESIAN BLIND DECONVOLUTION WITH APPLICATION TO ACOUSTIC FEEDBACK PATH MODELING

Ritwik Giri

Starkey Hearing Technologies
ritwik_giri@starkey.com

Tao Zhang

Starkey Hearing Technologies
tao_zhang@starkey.com

ABSTRACT

Acoustic Feedback Path in a digital hearing aid is not only affected by the user's head and ear, but also by different acoustic environments. But some of these effects are common for a specific style of hearing aid and individual ear, i.e., this part will be invariant to the different acoustic environments and can be interpreted as the effects associated with that specific hearing aid and ear characteristics. In this article we propose a novel Bayesian Blind Deconvolution approach with exponentially decaying kernel and show its application in extracting the invariant part of the feedback path measurements of a digital hearing aid. Efficacy of our proposed approach in extracting the invariant part has been measured by using the extracted invariant part to model unseen test Feedback Path (FBP) measured from the same hearing aid but in a different acoustic environment, over existing methods.

Index Terms— Feedback Paths (FBP), Acoustic Feedback Cancellation (AFC), Blind Deconvolution, Empirical Bayes

1. INTRODUCTION

Acoustic Feedback in digital hearing aids usually occurs because of the coupling between the receiver, i.e., the speaker and the hearing aid microphone, which can result in severe distortion of the desired sound and lead to loud whistling sounds [1]. It has become one of the most common problems associated with the current generation of open fitting digital hearing aids and requires effective strategies to prevent the howling sounds [2].

Among different solutions, recent literatures [3, 4, 5] indicate that the Feedback Cancellation (FBC) algorithms have been particularly successful to counter this problem. FBC algorithms usually estimate the feedback signal and remove it from the hearing aid microphone signal to make sure that only the desired speech signal is amplified in the forward path. Acoustic feedback path is affected by not only the response of user's head and ear, but also by the acoustic environment as well. Because of the dynamic nature of the acoustic surrounding/ environment, Adaptive Feedback Cancellation (AFC) approach has been proposed where the Impulse Response (IR) between the receiver and the hearing aid microphone is estimated using an adaptive filter [3]. In traditional AFC algorithms a Finite Impulse Response (FIR) is used to model the adaptive feedback path. This can often lead to a very long filter to model the FBP. It has been shown in [6, 7] that the convergence speed and the computational complexity of the adaptive filter is determined by the number of adaptive parameters, which makes the above mentioned approach not so effective and motivates to look for solutions which involves far less adaptive parameters to model the feedback path.

This motivated the recent line of work [8, 9, 10, 11], where it has been proposed to model the acoustic feedback path as the convolution of two filters: a time invariant common part which corresponds to the intrinsic properties of a specific hearing aid (transducer characteristics [9]) and also individual ear characteristics [10], and a time-varying variable part which enables us to model the dynamic nature of the acoustic environment (e.g., caused by moving objects around hearing aid [11]). Hence, in order to identify the common part and the variant part from FBP measurements which are measured on a multi-microphone hearing aid and also for different acoustic scenarios, we need to solve a blind deconvolution problem. This modeling approach results in a far shorter adaptive FIR for the time-varying part, contributing to faster convergence and significant reduction in computational load.

The variant part of a FBP can be modeled using finite number of taps, i.e., FIR for an easy and stable adaptation over time using standard adaptive filtering approaches. To model the common invariant part of the FBPs several different modeling techniques can be found in the literature. In [9] authors proposed to model the invariant part using an all pole filter that corresponds to room resonance whereas the variable part was modeled using an FIR leading to popular common-acoustical-pole and zero (CPZ) model. They also proposed an iterative least square search (ILSS) method which does not make any assumption on the pole and zero structure. Their findings showed that ILSS approach with the initialization of CPZ estimate generally produces a useful and robust estimate of the common part. In [10] authors proposed to use a pole-zero filter for the common part and used an Alternating Least Squares (ALS) method to estimate the pole and zero locations of the invariant filter. This approach was further extended in [11] by including a stability constraint on the estimated pole locations, and solving the optimization problem using either a Quadratic Programming (QP) or a Semi Definite Programming (SDP) software.

In this work, we consider the concerned blind deconvolution problem in a Bayesian framework. Bayesian Blind Deconvolution (BBD) [12] problem has been widely investigated in image processing and computer vision community [13], for many relevant applications such as camera shake removal [14], spatially varying blur removal [15] etc. Blind deconvolution problem in general is severely ill-posed and may lead to trivial solution if no constraint is imposed [12]. This is the case for ILSS [9], and hence this algorithm is very sensitive to the initialization. In this work we introduce constraints on the invariant part based on the prior knowledge to regularize the solution space and lessen the sensitivity to the initialization of the algorithm. Though in image processing applications sparsity constraint has been a relevant choice, for our problem it is not so feasible, as it ignores the tail of the invariant part of FBP. Hence as prior information along with sparsity for initial few taps to model any common delay and high nonzero filter coefficients, we also employ an exponentially decaying kernel to model the desired common part of FBP. This specific prior has been exploited in our previous work [16] before in a non-blind deconvolution problem of relative impulse response estimation. In this work we extend our proposed approach to the blind deconvolution case and employ an Empirical Bayes based inference procedure to estimate the concerned filter coefficients.

The rest of the paper is organized as follows. In Section 2, we present the problem of invariant part extraction from FBP measurements. In Section 3, we present the model of our proposed BBD framework, along with an Empirical Bayes based inference procedure to estimate the filter coefficients. We present experimental results of the proposed algorithm and also other existing algorithms using real FBP measurements in Section 4, and finally conclusions and some future directions of this work are presented in Section 5.

2. PROBLEM FORMULATION

Let's assume that L number of Feedback Paths (FBPs) have been measured for the same hearing aid on the same ear but in different acoustic scenarios, which can be denoted as $b_k[n]$ for, $k = 1, ..., L$. Key assumption is that, these FBPs have an invariant part, i.e. a fixed filter which accounts for the invariant properties of each measurement such as, fixed transducer, fixed acoustical couplings and individual characteristics of that particular ear, same for all L measurements. Let $f[n]$ and $e_k[n]$ denote the impulse response of the invariant part and the variant part of the k^{th} FBP $b_k[n]$ respectively. Hence,

$$b_k[n] = f[n] \star e_k[n] \tag{1}$$

In real life we will also consider that the measurement of FBP has some additive noise which can also account for model uncertainty.

Hence,

$$b_k[n] = f[n] \star e_k[n] + \epsilon[n] \tag{2}$$

Estimating the invariant part $f[n]$ from the true measurements of L FBPs, $b_k[n]$, is the main goal of this article.

3. PROPOSED METHOD: BAYESIAN BLIND DECONVOLUTION

Since blind deconvolution problem is fundamentally ill-posed and infinite solutions are possible, some prior knowledge/ information is required to obtain a meaningful solution. Previous work of incorporating pole and zero structure is one way to do that, but the problem with that, is added concern to maintain stability (estimated pole locations) and also sensitivity to any measurement noise [9]. We propose to use an Empirical Bayes based approach with a relevant prior distribution, incorporating sparsity and exponentially decaying kernel to get a robust estimator of the common part of FBPs.

3.1. Model

Concerned deconvolution problem involves finding the common invariant part and then variant parts from L number of FBP measurements ($b_k[n]$), i.e.,

$$b_k[n] = f[n] \star e_k[n] + \epsilon[n] \quad \text{for, } k = 1, ..., L \tag{3}$$

Here both $f[n]$ and $e_k[n]$ are unknown and need to be estimated from the true measurements of FBP, $b_k[n]$ of each length N,

$$\mathbf{b_k} = [b_k[0], ..., b_k[N-1]]^T \in \mathbb{R}^{N\times 1} \tag{4}$$

Let's assume that we can model $f[n]$ using an FIR of length C and each $e_k[n]$ using an FIR of length M, such that $M + C - 1 \leq N$.

$$\mathbf{e_k} = [e_k[0], ..., e_k[M-1]]^T \in \mathbb{R}^{M\times 1} \tag{5}$$

$$\mathbf{f} = [f[0], ..., f[C-1]]^T \in \mathbb{R}^{C\times 1} \tag{6}$$

We also need to truncate the true FBP measurements upto length $M + C - 1$ for the estimation stage, i.e.,

$$\mathbf{b^{tr}}_k = [b_k[0], ..., b_k[M+C-2]]^T \in \mathbb{R}^{M+C-1\times 1} \tag{7}$$

We can rewrite Equation (3) in matrix and vector product using convolution matrix and appending all the truncated FBP measurements $\mathbf{b_k}^{tr}$ together in a long column as,

$$\mathbf{b} = \mathbf{E}\mathbf{f} + \epsilon \tag{8}$$

Where $\mathbf{E}$ is the tall stacked matrix of the convolution matrices $E_k \in \mathbb{R}^{M+C-1\times C}$ constructed from $\mathbf{e_k}$, i.e.,

$$\mathbf{E} = [E_1; E_2; ...E_L] \in \mathbb{R}^{L(M+C-1)\times C} \tag{9}$$

and,

$$\mathbf{b} = [\mathbf{b^{tr}}_1{}^T ....\mathbf{b^{tr}}_L{}^T]^T \in \mathbb{R}^{L(M+C-1)\times 1} \tag{10}$$

Now in our probabilistic framework we will assume that the measurement noise is Gaussian with variance σ^2, which leads to the following likelihood distribution,

$$p(\mathbf{b}|\mathbf{f}, \mathbf{e}_1, ..\mathbf{e}_L; \sigma^2) \sim N(\mathbf{E}\mathbf{f}, \sigma^2) \tag{11}$$

If we assume that the non informative flat priors have been employed over both the common $\mathbf{f}$ and variant part $\mathbf{e}_k$. then the MAP estimate of the unknown filters can be found by solving the following non-linear optimization problem,

$$\hat{\mathbf{f}}, \hat{\mathbf{e_k}} = \arg\min ||\mathbf{b} - \mathbf{E}\mathbf{f}||_2^2 \tag{12}$$

An Iterative Least Square approach has been used in [9] to solve this non-linear problem by alternately estimating $\mathbf{f}$ and $\mathbf{e_k}$ till convergence. As indicated in [9] we will denote this method as ILSS and use it as one of our baseline methods.

As discussed above, blind deconvolution problem is highly ill-posed and there are infinite solutions possible for $\mathbf{f}$ and $\mathbf{e_k}$. This is one of the main reasons why ILSS suffers from severe sensitivity to initialization and often gets stuck to a local minima. To regularize the problem and find a meaningful solution we need to incorporate some prior information in our Bayesian framework by enforcing a prior distribution on the unknown filter coefficients.

3.2. Prior Distribution

In image processing applications of blind deconvolution, sparsity has been a popular regularization strategy to obtain meaningful solutions. But for our problem in hand, sparsity assumption becomes too restrictive to model decaying nature of FBPs and often ignores the tail because of small coefficient values (close to zero). To counter this problem we also employ an exponential decaying kernel to model the tail along with sparsity inducing prior for initial few filter coefficients and any common delay. The prior distribution over $\mathbf{f}$ is proposed to follow:

$$p(\mathbf{f}|\boldsymbol{\gamma}, c_1, c_2) \sim N(0, \Gamma) \tag{13}$$

With:

$$\Gamma = \text{diag}\big[\gamma_1, ..., \gamma_P, c_1 e^{-c_2}, ..., c_1 e^{-c_2 m}, ..., c_1 e^{-c_2 M}\big] \tag{14}$$

Where:

- γ_p corresponds to p^{th} early tap
- $c_1 e^{-c_2 m}$ corresponds to m^{th} tap out of the M exponentially decaying kernel

$\boldsymbol{\gamma} = [\gamma_1, ..., \gamma_P]$, c_1 and c_2 can be interpreted as the hyperparameters of the model, which can be learned from the measurements using an Evidence Maximization approach [17]. Details of this inference procedure will be discussed in the next subsection.

It is not straight forward to see from the above mentioned prior distribution $p(f_i|\gamma_i) = N(f_i; 0, \gamma_i)$ for, $i = 1...P$, how the sparsity is enforced on the initial few taps of $\mathbf{f}$, because the hierarchical nature of the prior disguises its character. To expand on this, let's assume that an Inverse Gamma (IG(α, β)) distribution has been used as the prior over hyperparameters. To find the "true" nature of the prior $p(f_i)$, we integrate out the γ_i and the marginal is obtained as,

$$p(f_i) = \int p(f_i|\gamma_i)p(\gamma_i)d\gamma_i = \frac{\beta^\alpha \Gamma(\alpha + 0.5)}{(2\pi)^{0.5}\Gamma(\alpha)} \left(\beta + \frac{f_i^2}{2}\right)^{-(\alpha+0.5)} \tag{15}$$

This marginal distributions, "true" representation of the behavior of the prior of initial P taps of the common part corresponds to a Student's t-distribution, which is a super Gaussian density (has heavier tails than Gaussian) and has been very popular in sparse recovery literatures because of its ability to promote sparsity [17]. In Figure 1 we present the pdfs of a student's t distribution with degrees of freedom (β) = 0.1, and a Gaussian distribution to show why a student's distribution is suited to promote sparsity. Moreover for our case where a uniform hyperprior $p(\gamma_i)$ has been used (i.e. $\alpha = \beta = 0$), $p(f_i) \propto \frac{1}{|f_i|}$ becomes an improper Jeffrey's prior, which has infinite probability mass at origin.

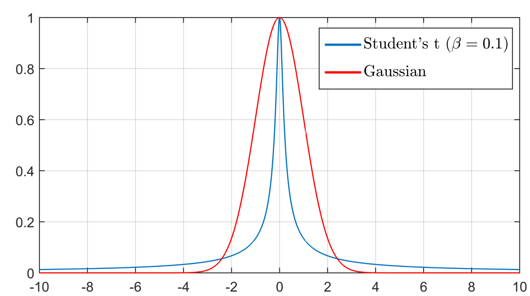

Fig. 1. Tail Behavior: Student's t vs Gaussian

Since the variant part $\mathbf{e_k}$ will be adapted during the Feedback Cancellation stage, we employ a non-informative flat prior on $p(\mathbf{e_k})$ and proceed to the inference stage.

3.3. Inference using Empirical Bayes

It has been shown in [18] that enforcing relevant prior distribution may not be enough to deal with the ill posed nature of the blind deconvolution problem, and discusses that the inference strategy to estimate the concerned parameters, should also be chosen with caution.

Straight forward estimation approach is to look for the Maximum a posteriori (MAP) estimate for both the common part $\mathbf{f}$ and the variant part $\mathbf{e}$ simultaneously, i.e. $\text{MAP}_{\mathbf{f},\mathbf{e}}$ estimate,

$$\hat{\mathbf{f}}, \hat{\mathbf{e}} = \arg\max p(\mathbf{f}, \mathbf{e}|\mathbf{b}) \tag{16}$$

But it has been shown in [12] that there are many problems with this simultaneous MAP estimation approach. One major problem is the presence of many suboptimal local minima which leads to convergence issues and hence sensitivity to initialization. To mitigate some of these issues, as suggested in [12] we will also use an Empirical Bayes based inference procedure also known as Type II/ Evidence maximization for a well conditioned estimate of the common part, $\mathbf{f}$.

We proceed by employing an EM algorithm for inference and treat $\mathbf{e}_k$ as parameters and $\mathbf{f}$ as the hidden random variable. In the E step we need to compute the concerned posterior, $p(\mathbf{f}|\mathbf{b}; \mathbf{E}, \boldsymbol{\gamma}, c_1, c_2)$. Because of the Gaussian nature of the both likelihood and prior, this step leads to the following Gaussian posterior,

$$p(\mathbf{f}|\mathbf{b}; \mathbf{E}, \boldsymbol{\gamma}, c_1, c_2) = N(\mathbf{f}; \mu, \Sigma) \tag{17}$$

Where the mean and covariance are,

$$\hat{\mathbf{f}} = \mu = \sigma^{-2}\Sigma\mathbf{E}^T\mathbf{b} \tag{18}$$

$$\Sigma = (\sigma^{-2}\mathbf{E}^T\mathbf{E} + \Gamma^{-1})^{-1} \tag{19}$$

Note that $\mathbf{E}$ is the stacked convolution matrix following Equation (9). We use the result from the E step to compute the Q function, which is essentially the conditional expectation of the complete data log likelihood with respect to the concerned posterior given in (17).

$$\begin{aligned} &Q(\mathbf{e}_k, \boldsymbol{\gamma}, c_1, c_2) \\ &= \mathbb{E}_{\mathbf{f}|\mathbf{b};\boldsymbol{\gamma}^t, c_1^t, c_2^t, \sigma^2, \mathbf{e}_k}[\log(p(\mathbf{b}|\mathbf{f}; \mathbf{E}, \sigma^2)p(\mathbf{f}|\boldsymbol{\gamma}, c_1, c_2))] \end{aligned} \tag{20}$$

In the Q function expression we will need the following conditional expectation,

$$< f_i^2 >= E_{\mathbf{f}|\mathbf{b};\boldsymbol{\gamma}^t, c_1^t, c_2^t, \sigma^2, \mathbf{e}_k}[f_i^2] = \Sigma_{(i,i)} + \mu_i^2 \tag{21}$$

where $\Sigma_{(i,i)}$ is the i^{th} diagonal element of Σ.

Now in the M step we will maximize the given Q function with respect to $\mathbf{e_k}$, c_1, c_2, and $\boldsymbol{\gamma}$,.

$$\hat{\mathbf{e}_k}, \hat{\boldsymbol{\gamma}}, \hat{c_1}, \hat{c_2} = \arg \max_{\mathbf{e}_k, \boldsymbol{\gamma}, c_1, c_2} Q(\mathbf{e}_k, \boldsymbol{\gamma}, c_1, c_2) \tag{22}$$

After maximizing the Q function we get the following update rules,

$$\gamma_p = \Sigma_{(p,p)} + \mu_p^2 \quad \text{for } p = 1 \ldots P \tag{23}$$

$$c_1 = \frac{1}{M} \sum_{m=1}^{M} e^{c_2 m} < f_{m+P}^2 > \tag{24}$$

$$\sum_{m=1}^{M} m e^{c_2 m} < f_{m+P}^2 > - c_1 \frac{M(M+1)}{2} = 0 \tag{25}$$

$$\hat{\mathbf{e}_k} = \arg\min_{\mathbf{e}_k} ||\mathbf{b}_k^{tr} - \hat{\mathbf{F}}\mathbf{e}_k||^2 + \sum_i w_i e_{k,i}^2 \tag{26}$$

Where, $w_i = \sum_j \Sigma_{i+j,i+j}$.

Note that the convolution matrix $\mathbf{E}$ in the update of $\mathbf{f}$ in Equation (18) will be constructed from the most recent estimates of the variant part. Similarly when we update the variant parts $\mathbf{e}_k$ using (26), we construct the convolution matrix $\hat{\mathbf{F}}$ using the recent estimate of $\mathbf{f}$. We perform this EM based updates for few iterations until a convergence criterion is satisfied. Unlike in [16] we do not learn the noise variance σ^2 in the M step. Instead following [12] an annealing type strategy was employed where after every iteration we update the noise variance, $\sigma^2 \leftarrow \sigma^2/\beta$, where $\beta > 1$ till it reaches a pre specified minimum value (λ_{min}). For all our experiments we have used $\beta = 1.08$ and $\lambda_{min} = 1e-10$. Intuition behind this annealing strategy is that, during initial iterations a high value of σ^2 prevents the algorithm to get stuck to a local minima and as the iteration number grows, decreasing σ^2, i.e., reducing the uncertainty will help our algorithm to converge to the global minima.

4. EXPERIMENTAL EVALUATION

4.1. Setup

Following the setup in [11], four training feedback paths were measured using a two microphone behind-the-ear hearing aid with open-fitting ear-molds on a dummy head with ear canal diameter $(d) = 7$ mm and length $(l) = 15$ mm. To account for the variations in the acoustic conditions a telephone receiver was placed in close distance to the hearing aids for the two FBP measurements. For validating purpose, one test FBP was measured in an unseen (to the algorithms) acoustic scenario, where the dummy head was positioned close to a wall. In Figure 2 we plot the four training FBPs that have been used to extract the common part $\hat{\mathbf{f}}$. In Figure 3 we present the test FBP that has been measured by placing the dummy head near a wall. All the FBPs were sampled using a 16 kHz sampling frequency and were truncated to length, N=100. Extracted common part $\hat{\mathbf{f}}$ from training FBPs is then used to model the unseen test FBP and the resulting Normalized Mean Square Error (NMSE) is computed as a performance metric.

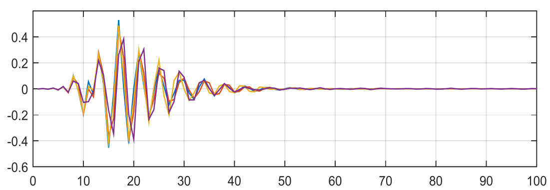

Fig. 2. Training Feedback Paths

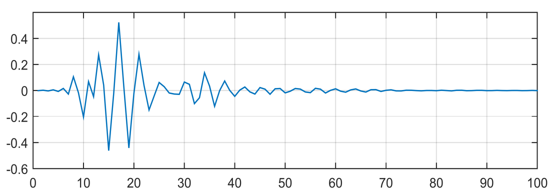

Fig. 3. Test Feedback Path: Wall

4.2. Existing Algorithms

We compare our proposed approach with following existing methods,

- CPZ: All pole model for the invariant part. [9]
- ALS: Pole Zero model for the invariant part, optimized using an Alternating Least Square (ALS) method and initialized with CPZ pole locations. [10]
- W-ALS (QP): Pole Zero model for the invariant part which has been estimated by minimizing weighted equation error (using a Quadratic Programming approach) and a stability constraint has been imposed for the estimated pole locations [11]. Authors have also proposed an SDP based formulation in [11], which has not been included here because of its high computational load.
- ILSS: All zero invariant part, estimated using an Iterative Least Square Search with the initialization of truncated CPZ estimate. [9]
- BBD (proposed): Bayesian Blind Deconvolution method with an exponential kernel, with the choice of $P = C/2$ and with the initialization of truncated CPZ estimate.

4.3. Results

In Figure 4 we show the NMSE measures for different algorithms with respect to the length of the variant part with a fixed $C = 26$. For Pole zero (ALS and W-ALS) based methods, as no specific instruction was provided in the literatures on how to choose the number of poles and zeros, we evenly distribute them, i.e. $C/2$ poles and $C/2$ zeros were used for all our experiments. As expected, in Figure 4 we see that performance improves for all the algorithms as the length of variant part (M) increases, but our proposed approach BBD produces best NMSE for all cases except when $M = 10$. In Figure 5 we show the NMSE measures for different algorithms with respect to increasing C but with fixed variant part length $M = 15$. Again we notice that BBD produces the best NMSE for all cases.

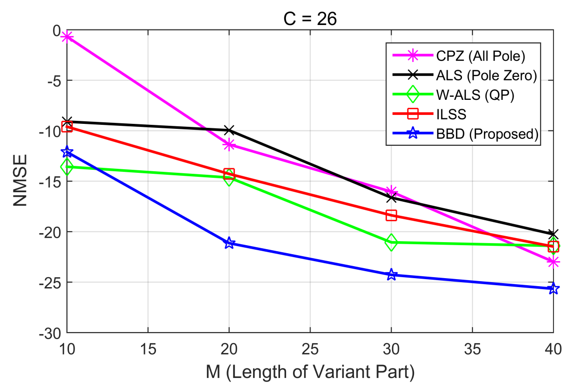

Fig. 4. NMSE comparison with varying M

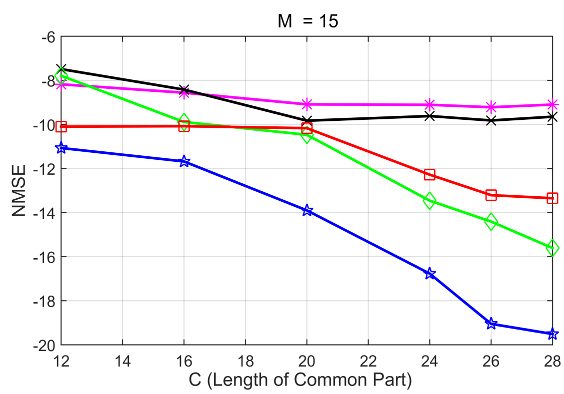

Fig. 5. NMSE comparison with varying C

5. CONCLUSION

In this article we have proposed a novel Bayesian Blind Deconvolution framework with exponentially decaying kernel and have shown its efficacy in extracting the invariant part from a set of FBP measurements in presence of different acoustic variabilities over other existing approaches. Including this modeling approach for acoustic feedback cancellation task will be considered in our future works.

6. ACKNOWLEDGMENT

We would like to thank Henning Schepker and Prof Simon Doclo for providing the code and the measurements that have been used in their paper [11]. We would also like to thank Renato Nakagawa and Karim Helwani for helpful discussion on relevant topics.

7. REFERENCES

[1] Joseph Arlen Maxwell and Patrick M Zurek, "Reducing acoustic feedback in hearing aids," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 4, pp. 304–313, 1995.

[2] Ann Spriet, Geert Rombouts, Marc Moonen, and Jan Wouters, "Adaptive feedback cancellation in hearing aids," *Journal of the Franklin Institute*, vol. 343, no. 6, pp. 545–573, 2006.

[3] Ann Spriet, Ian Proudler, Marc Moonen, and Jan Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Transactions on signal processing*, vol. 53, no. 10, pp. 3749–3763, 2005.

[4] Guilin Ma, Fredrik Gran, Finn Jacobsen, and Finn Thomas Agerkvist, "Adaptive feedback cancellation with band-limited lpc vocoder in digital hearing aids," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 4, pp. 677–687, 2011.

[5] Ann Spriet, Simon Doclo, Marc Moonen, and Jan Wouters, "Feedback control in hearing aids," in *Springer Handbook of Speech Processing*, pp. 979–1000. Springer, 2008.

[6] Simon S Haykin, *Adaptive filter theory*, Pearson Education India, 2008.

[7] Ali H Sayed, *Fundamentals of adaptive filtering*, John Wiley & Sons, 2003.

[8] James Mitchell Kates, "Feedback cancellation apparatus and methods," June 6 2000, US Patent 6,072,884.

[9] Guilin Ma, Fredrik Gran, Finn Jacobsen, and Finn Agerkvist, "Extracting the invariant model from the feedback paths of digital hearing aids," *The Journal of the Acoustical Society of America*, vol. 130, no. 1, pp. 350–363, 2011.

[10] Henning Schepker and Simon Doclo, "Modeling the common part of acoustic feedback paths in hearing aids using a pole-zero model," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 3665–3669.

[11] Henning Schepker and Simon Doclo, "Least-squares estimation of the common pole-zero filter of acoustic feedback paths in hearing aids," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 8, pp. 1334–1347, 2016.

[12] David P Wipf and Haichao Zhang, "Revisiting bayesian blind deconvolution.," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3595–3634, 2014.

[13] Tom E Bishop, S Derin Babacan, Bruno Amizic, Aggelos K Katsaggelos, Tony Chan, and Rafael Molina, "Blind image deconvolution: problem formulation and existing approaches," *Blind image deconvolution: Theory and Applications*, pp. 1–41, 2007.

[14] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T Roweis, and William T Freeman, "Removing camera shake from a single photograph," in *ACM Transactions on Graphics (TOG)*. ACM, 2006, vol. 25, pp. 787–794.

[15] Dilip Krishnan, Terence Tay, and Rob Fergus, "Blind deconvolution using a normalized sparsity measure," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 233–240.

[16] Ritwik Giri, Bhaskar D Rao, Fred Mustiere, and Tao Zhang, "Dynamic relative impulse response estimation using structured sparse bayesian learning," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 514–518.

[17] Ritwik Giri and Bhaskar Rao, "Type i and type ii bayesian methods for sparse signal recovery using scale mixtures," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3418–3428, 2016.

[18] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 2657–2664.