IMPROVED NOISE CHARACTERIZATION FOR RELATIVE IMPULSE RESPONSE ESTIMATION

Tharun Adithya Srikrishnan, Bhaskar D. Rao *

University of California, San Diego
tsrikris, brao@ucsd.edu

Ritwik Giri, Tao Zhang

Starkey Hearing Technologies
Ritwik_Giri, Tao_Zhang@starkey.com

ABSTRACT

Relative Impulse Responses (ReIRs) have several applications in speech enhancement, noise suppression and source localization for multi-channel speech processing in reverberant environments. Noise is usually assumed to be white Gaussian during the estimation of the ReIR between two microphones. We show that the noise in this system identification problem is instead dependent upon the microphone measurements and the ReIR itself. We then present modifications that incorporate this new noise model into three prevalent methods: Least Squares, Non-Stationary Frequency Domain and Sparse Bayesian Learning based approaches. We demonstrated improvements with an experimental study using real-world measurements in various noise environments.

Index Terms— Relative Impulse Response, Speech Enhancement, System Identification, Sparse Estimation

1. INTRODUCTION

Relative Impulse Responses (ReIRs) and their frequency-domain counterparts, the Relative Transfer Functions (RTFs) [1], are important tools in several multichannel audio processing tasks such as noise reduction, speech enhancement and source localization [2, 3]. ReIRs represent the impulse response between two microphones calculated when signals are received from a single source on both. RTF information can be incorporated into beamforming algorithms [2, 4] to produce a noise reference signal used for adaptive interference cancellation and to improve the speech enhancement performance. A solution to the Time Difference of Arrival estimation problem [5] in reverberant environments can also be found using ReIRs. Fast, accurate estimates of the ReIR are desired for optimal and reliable real-time performance for these applications.

ReIRs can be easily computed in a noise-free environment using a traditional Least Squares (LS) formulation as shown in [4]. In [1] the authors have proposed a Frequency Domain (FD) method which exploits the non-stationarity of the target speech signal. This method assumes that the noise and the RTF are much less dynamic, when compared to the target signal. A recently proposed time domain solution, the Structured Sparse Bayesian Learning (S-SBL) algorithm [6] exploits sparsity and the exponential decay in the ReIR during estimation. In this method, sparse early reflections and an exponentially decaying reverberation tail are modeled in a prior distribution as part of an Empirical Bayes formulation.

*The work of Mr. Srikrishnan and Dr. Rao was supported in part by a gift from Starkey Hearing Technologies

In the current state of the art ReIR estimation algorithms [1, 6], the measurement noise is always assumed as a white Gaussian random variable. Section 3 highlights how noise should actually be characterized as a multivariate Gaussian distribution correlated with both the measurements and the unknown estimate.

The effects of the improved noise characterization (Section 3) suggest modifications to the Regularized Least Squares, Frequency Domain methods and the S-SBL algorithm. The approach shown below can be used as a framework for solving any system identification problem where the noise is correlated to measurements and the system itself. Attenuation Rate (ATR) and the Normalized Mean Square Error (NMSE) have been used as performance metrics in this work to show efficacy of this improved noise characterization for ReIR estimation in different noisy environments.

2. PROBLEM FORMULATION

The ReIR estimation problem can be viewed as a system identification problem. Consider a two channel noisy recording of a target speaker (fixed position during duration of measurement) in a reverberant environment. Let h_L and h_R denote the Room Impulse Response between the target and the two microphones (subscript indicating Left (L) or Right (R) microphone). $s[n]$ denotes the target speech while $\epsilon_L[n]$ and $\epsilon_R[n]$ denote the noise components in the microphone measurements $x_L[n]$ and $x_R[n]$ (Index n is dropped in the analysis from now on). Given the exact acoustic channels, we can obtain the ReIR h as $h_R \star h_L^{-1}$. However in practice; the system identification problem is formulated as:

$$x_R = h_R \star s + \epsilon_R \tag{1}$$

$$x_L = h_L \star s + \epsilon_L \tag{2}$$

$$\Rightarrow x_R = x_L * h + \underbrace{(\epsilon_R - \epsilon_L * h)}_{\epsilon} \tag{3}$$

Note that the convolution operation can be replaced by a matrix multiplication; thus mapping it to a linear inverse problem. $\mathbf{x_R}$ and $\mathbf{x_L}$ are $N \times 1$ measurement vectors corresponding to N measurement samples. The unknown ReIR $\mathbf{h}$ of length L (Truncated) is set as a $L \times 1$ vector. The $N \times L$ convolution matrix $\mathbf{X_L}$ designed from $\mathbf{x_L}$ results in $\mathbf{x_R}$ when multiplied with the estimate. We denote the system noise ($\epsilon_R - \epsilon_L * h$) as ϵ.

$$\mathbf{x_R} = \mathbf{X_L}\mathbf{h} + \epsilon \tag{4}$$

This system may have a non-causal behavior. In order to derive a causal ReIR, we delay x_R by D samples such that the ReIR is effectively delayed [7]. In the absence of noise, $h = x_L^{-1}[n] * x_R[n] * \delta[n-D] = x_L^{-1}[n] * x_R[n-D]$. This delay is achieved by zero-padding prior to the delayed signal $x_R[n-D]$.

Most methods currently in use involve formulating the noise as a white Gaussian random variable; i.e $x_R[n] \approx (h \star x_L)[n] + \epsilon_G[n]$, where $\epsilon_G[n] \sim \mathcal{N}(0, \sigma^2)$. However; assuming that the measurement noise ϵ_R and ϵ_L are i.i.d Gaussian random variables, ϵ is clearly not white noise (Equation 3). Note that ϵ is colored irrespective of the correlation between ϵ_R and ϵ_L. ϵ is observed to be correlated with the measurements x_R and x_L; and also the unknown parameter h.

3. NOISE ANALYSIS

We motivate the need for an improved noise model by analyzing the correlations between the noise and the signals. R_{xy} indicates the correlation of the signal x with y. The measurement noise ϵ_R and ϵ_L are modeled as i.i.d colored Gaussian variables with autocorrelation $R_{\epsilon_R \epsilon_R} = R_{\epsilon_L \epsilon_L}$. The non-negative indices of the autocorrelation sequence correspond to $\sigma^2[1, \rho_1, \rho_2, \rho_3 ...]$. We also assume that the noise is independent of the received source signal.

$$
\begin{aligned}
R_{\epsilon\epsilon}[m] &= \mathbb{E}[\epsilon[n+m]\epsilon[m]] && (5)\\
&= \mathbb{E}[(\epsilon_R[n+m] - \sum_a \epsilon_L[a]h[n+m-a]) && (6)\\
&\quad (\epsilon_R[n] - \sum_b \epsilon_L[b]h[n-b])] && (7)\\
&= \mathbb{E}\left[\epsilon_R[n+m]\epsilon_R[n]\right] && (8)\\
&\quad + \mathbb{E}\left[\sum_a \epsilon_L[a]h[n+m-a] \sum_b \epsilon_L[b]h[n-b]\right]\\
&= R_{\epsilon_R\epsilon_R} + \sigma^2 L R_{hh}[m] && (9)\\
&\quad + \sigma^2 L \left[\sum_{i=1}^{\infty} \rho_i (R_{hh}[m-i] + R_{hh}[m+i])\right] && (10)
\end{aligned}
$$

We can derive results similarly for $R_{\epsilon x_L}$ and $R_{\epsilon x_R}$:

$$R_{\epsilon x_L}[m] = -h[m] * R_{\epsilon_L \epsilon_L} \quad (11)$$

$$R_{\epsilon x_R}[m] = R_{\epsilon_R \epsilon_R} \quad (12)$$

$R_{\epsilon\epsilon}$ is not purely a delta function as we would expect if ϵ was white Gaussian noise. The dependence of noise autocorrelation to the ReIR itself allude to iterative methods where a previous estimate is used to compute the next estimate. The later sections attempt to weave in these modifications into existing methods. Modifications to time-domain methods resulted in tedious expressions using the above expressions. The formulation was simplified assuming that the noise at each microphone is white ($\rho_i = 0 \, \forall i$). Note that ϵ is colored even when ϵ_R and ϵ_L are white.

$$R_{\epsilon\epsilon}[m] = \sigma^2 \left[\delta[m] + L R_{hh}[m]\right] \quad (13)$$

$$R_{\epsilon x_L}[m] = -\sigma^2 h[m] \quad (14)$$

$$R_{\epsilon x_R}[m] = \sigma^2 \delta[m] \quad (15)$$

We follow theoretical derivations with experimental results to study the effects of the improved noise modeling.

4. MODIFIED METHODS

4.1. Modified Regularized Least Squares

A rudimentary solution to the ReIR estimation problem (assuming White Gaussian noise) is a maximum likelihood solution obtained using a Least Squares (LS) formulation:

$$\hat{\mathbf{h}}_{LS} = \arg\min_{\mathbf{h}} \|\mathbf{x_R} - \mathbf{X_L h}\|_2^2 \quad (16)$$

The solution to Equation (16) is $\hat{\mathbf{h}}_{\mathbf{LS}} = (\mathbf{X_L}^{\mathbf{T}}\mathbf{X_L})^{-1}\mathbf{X_L^T x_R}$. Possible ill-conditioning of $(\mathbf{X_L}^T \mathbf{X_L})$ is handled through regularization [8]. A maximum likelihood based solution with improved noise characterization (Section (3)) was however tedious to calculate. We simplify the solution by assuming that the noise at each microphone is white Gaussian ($\rho_i = 0 \, \forall i$). An iterative solution is attempted instead; where a previous estimate $\hat{h}_t$ is used to calculate the next estimate $\hat{h}_{t+1}$ until convergence. The noise ϵ given $\hat{h}_t$ is a Gaussian $\mathcal{N}(0, \sigma^2(I + R_H))$; where R_H is a Toeplitz matrix with the first column as LR_{hh} (derived from results in Section 3). An iterative LS solution can then be derived by maximizing:

$$p(x_R | x_L, \hat{h}_t; h_{t+1}) \sim \mathcal{N}(X_L h_{t+1}, \underbrace{\sigma^2(I + R_{H_t})}_{C_t}) \quad (17)$$

$\hat{h}_{t+1}$ is solved for as $(X_L^T C_t^{-1} X_L)^{-1} X_L^T C_t^{-1} x_R$; where C_t is obtained from $\hat{h}_t$ as $\sigma^2(I + R_{H_t})$. Two iterations were found to be sufficient for convergence, beyond which no major performance gain was observed.

4.2. Modified Non-Stationary Frequency Domain (NSFD)

Noise signals are assumed to be stationary and less dynamic when compared to the target speech signal in the NSFD method [1]. The measurement interval is first divided into P frames. Let $\Phi_{AB}^p(\omega)$ denote the cross power spectral density between signals A and B during the p^{th} frame. For the p^{th} frame, we can write:

$$\Phi_{x_R x_L}^p(\omega) = H(\omega)\Phi_{x_L x_L}^p(\omega) + \Phi_{\epsilon x_L}^p(\omega) \quad (18)$$

Since the noise is assumed to be stationary, the NSFD method suggests that we can write $\Phi_{\epsilon x_L}^p = \Phi_{\epsilon x_L}$. These overdetermined set of equations are solved for $p = 1, 2...P$ by a least squares formulation, to estimate the RTF H_{RTF}. The traditional NSFD method implicitly assumes that $\Phi_{\epsilon x_L}$ remains constant. The improved noise framework suggests that $\Phi_{\epsilon x_L}(\omega) = -\Phi_{\epsilon_L \epsilon_L} H(\omega)$, thus justifying the assumption. The Modified NSFD equation for the the p^{th} frame can then be derived as:

$$\Phi_{x_R x_L}^p(\omega) = \Phi_{x_L x_L}^p(\omega)H(\omega) - \Phi_{\epsilon_L \epsilon_L}(\omega)H(\omega) \quad (19)$$

$$\Phi_{x_R x_L}^p(\omega) = \left(\Phi_{x_L x_L}^p(\omega) - \Phi_{\epsilon_L \epsilon_L}(\omega)\right) H(\omega) \quad (20)$$

A possible method of implementation would involve estimating the noise power spectrum during speech-free blocks and estimate the accurate ReIR using the modified NSFD method. On simplifying the formulation by assuming white Gaussian microphone noise:

$$\Phi_{x_R x_L}^p(\omega) = \left(\Phi_{x_L x_L}^p(\omega) - \sigma^2\right) H(\omega) \quad (21)$$

The noise variance σ^2 behaves as a Frequency Domain Regaularizer. This regularization leads to the ReIR peaks being slightly suppressed; along with lower noise-like fluctuations. Setting σ^2 to zero (No Regularization) averages out the Traditional Frequency Domain (FD) [1] solution obtained from the blocks. Frequency Domain methods however were observed to be more sensitive to the noise characterization and thus assuming a flat noise power spectrum (original methods) is not precise. However; such an approximation in low noise environments may not lead to large performance losses.

4.3. Modified Structured Sparse Bayesian Learning Method

The Empirical Bayes based method can be used to estimate the ReIR in time domain by exploiting both its structure and sparsity. Both the sparse early reflections and the reverberant tails are modeled in a unified Bayesian framework as a prior. This formulation is a direct modification of the original Structured Sparse Bayesian Learning (S-SBL) algorithm [6] with a noise component dependent on the previous channel estimate. The algorithm below is iteratively run to calculate this component and update the estimate with another round of Modified S-SBL. We again make the simplifying assumption of white microphone noise. Consider the system model, $\mathbf{x_R} = \mathbf{X_L h} + \epsilon$. The modified Gaussian Likelihood assumption can be shown to be $p(\mathbf{x_R}|\mathbf{h}, \hat{h}) \sim N(\mathbf{X_L h}, \sigma^2(I + R_{\hat{H}}))$. We envisage an iterative approach where a previous estimate can be used for the next estimate. The prior over $\mathbf{h}$ is proposed as:

$$p(\mathbf{h}|\gamma_i, c_1, c_2) \sim N(0, \Gamma) \tag{22}$$

$$\Gamma = diag\left[\gamma_1, ., \gamma_P, c_1 e^{-c_2}, .., c_1 e^{-c_2 M}\right] \tag{23}$$

γ_p is a prior that corresponds to p^{th} early reflection; whereas $c_1 e^{-c_2 m}$ corresponds to the m^{th} tap out of the M-sample exponentially decaying reverberant tail. Note that the proposed approach follows a Relevance Vector Machine (RVM)/Sparse Bayesian Learning (SBL) [9] framework to incorporate the sparse regularization. A Type-II Likelihood/Evidence maximization [10, 11] procedure is used to estimate the Impulse Response, $\mathbf{h}$. The hyper-parameters γ_i (i=1,..,P), c_1, and c_2, are estimated from the data by maximizing the marginal likelihood $p(\mathbf{x}|\gamma_i, c_1, c_2)$. The estimate of the ReIR given the hyperparameters is then:

$$\hat{\mathbf{h}} = \mathbb{E}[\mathbf{h}|\mathbf{x_R}, \hat{\gamma_i}, \hat{c_1}, \hat{c_2}] \tag{24}$$

Let Π denote the covariance of the zero mean Gaussian measurement noise (modified formulation). Π is obtained as $\sigma^2(I + R_H)$. Given the Gaussianity of the prior, the relevant posterior of $\mathbf{h}$ can be computed as:

$$p(\mathbf{h}|\mathbf{x_R}; \gamma, c_1, c_2) \sim N(\mathbf{h}; \mu, \Sigma) \tag{25}$$

$$\mu = \Sigma \mathbf{X_L}^T \Pi^{-1} \mathbf{x_R}; \quad \Sigma = (\mathbf{X_L}^T \Pi^{-1} \mathbf{X_L} + \Gamma^{-1})^{-1} \tag{26}$$

We approximate the true posterior $p(\mathbf{h}|\mathbf{x})$ by $p(\mathbf{h}|\mathbf{x}; \gamma, c_1, c_2)$; a Gaussian distribution whose mean and covariance depend upon the estimated hyperparameters. Following equation (24), we can use $\hat{\mathbf{h}} = \mu$ as the point estimate of the impulse response. The EM algorithm is then used to estimate the hyperparameters from the log-likelihood:

$$\begin{aligned} &Q(\gamma, c_1, c_2, \sigma^2) \\ &= \mathbb{E}_{\mathbf{h}|\mathbf{x_R};\gamma^t, c_1^t, c_2^t, \sigma^2}[\log(p(\mathbf{x_R}|\mathbf{h}; \sigma^2) p(\mathbf{h}|\gamma, c_1, c_2))] \end{aligned} \tag{27}$$

Maximizing this Q-function with respect to the hyperparameters γ, c_1, c_2 and σ^2 [6] results in:

$$\gamma_p^{t+1} = \Sigma_{(p,p)}^t + (\mu_p^t)^2 \quad \text{for } p = 1 \dots P \tag{28}$$

$$c_1^{t+1} = \frac{1}{M} \sum_{m=1}^{M} e^{c_2^t m} \left(\Sigma_{ii}^t + (\mu_i^t)^2\right) \tag{29}$$

$$\sum_{m=1}^{M} m e^{c_2^{t+1} m} < h_{m+P}^2 > -c_1^{t+1} \frac{M(M+1)}{2} = 0 \tag{30}$$

$$\begin{aligned} (\sigma^2)^{t+1} = &\frac{(x_R - X_L h)^T (I + R_H)^{-1} (x_R - X_L h)}{N} \\ &+ \frac{\sigma^{-2} \sum_{i=1}^{(M+P)} \left(1 - \Sigma_{ii}^t / \Gamma_i^t\right)}{N} \end{aligned} \tag{31}$$

The hyperparameters are updated; and the prior mean and covariances are computed until convergence. The final estimate is set as $\mu = \sigma^{-2} \Sigma \mathbf{X_L}^T \Pi^{-1} \mathbf{x_R}$. It can be envisaged that once initial estimates are found through alternative methods; the final pruned solution is obtained through a few iterations of the Modified S-SBL.

4.4. Additional Considerations

The effect of the modified approaches on Time-Domain methods (Section 4.1, 4.3) effectively lead to off-diagonal terms in the noise covariance matrix. We also observe that the matrix is concentrated diagonally; with the correlations far away from the diagonal close to zero due to a decaying R_{hh}. A pure white Gaussian noise characterization as used till now discards these off diagonal terms. Neglecting these terms will lead to simpler computation at the cost of performance. The simplification can also be justified for use in low noise environments. These Improved Noise methods may be run a few iterations given an previous estimate to enhance it.

5. EXPERIMENTAL EVALUATION

We use real-world noise-free speech signals measured in reverberant environments (Experimental setting described in [12] and reverberant recordings generated using measured impulse responses from [13]); combined with different types of noise to evaluate performance. The signal for the target source, a female utterance, has been taken from the task of the online Signal Separation Campaign (SISEC) 2013 [14]. A testing utterance (female talker) 10s long, is divided into intervals of 1024 samples each (128 ms long at 8 kHz). The noise-free speech is combined with different kinds of noise (Ambient Noise (Road), Omni-babble, Directional Machine Noise). The evaluated performance measures are averaged over these intervals.

Two performance metrics are considered here: the Attenuation Rate (ATR) and Normalized Mean Squared Error (NMSE) [6]. The ATR can be evaluated as the ratio between SNR_{out} and SNR_{in} in dB scale, where:

$$SNR_{in} = \frac{\sum_{i=L,R} \sum_n [h_i \star s[n]]^2}{\sum_{i=L,R} \sum_n [\epsilon_i[n]]^2} \tag{32}$$

$$SNR_{out} = \frac{\sum_n [(\hat{h} \star x_L)[n] - x_R[n]]^2}{\sum_n [(\hat{h} \star \epsilon_L)[n] - \epsilon_R[n]]^2} \tag{33}$$

The numerator of SNR_{out} measures the leakage of the target signal whereas the denominator measures the attenuation of the noise signal. A low ATR indicates a good noise reference signal for further processing.

The NMSE measures the normalized square mismatch between the True ReIR and estimated ReIR. The least squares method was used to estimate the true ReIR using a long noise free measurement. Lower NMSE is often an indicator of better signal recovery.

$$NMSE = \frac{||\hat{h} - h^{true}||_2^2}{||h^{true}||_2^2} \tag{34}$$

We evaluate experimental results using performance metrics discussed above. The results are shown in Table 1 and 2 for Least Squares, Table 3 and 4 for NSFD, and Table 5 and 6 for S-SBL respectively. The ATR (Table 1,3,5) and NMSE (Table 2,4,6) suggest that the proposed improved noise characterization consistently helps to improve the ATR and NMSE performance.

Table 1: ATR: Regularized Least Squares Method

Algorithms	Ambient Noise (Road)	Babble Noise	Machine Noise
	ATR (dB)	ATR (dB)	ATR (dB)
Least Squares	-12.38	-4.80	-14.27
Least Squares (Improved Noise)	**-13.46**	**-5.63**	**-14.60**

Table 2: NMSE: Regularized Least Squares Method

Algorithms	Ambient Noise (Road)	Babble Noise	Machine Noise
	NMSE	NMSE	NMSE
Least Squares	2.32	1.26	2.00
Least Squares (Improved Noise)	**1.19**	**0.79**	**0.97**

Table 3: ATR: Frequency Domain Methods

Algorithms	Ambient Noise (Road)	Babble Noise	Machine Noise
	ATR (dB)	ATR (dB)	ATR (dB)
FD	-13.26	-3.10	-18.46
NSFD	-17.06	-6.34	-21.92
Improved Noise NSFD	**-17.86**	**-6.45**	**-22.77**
Improved Noise NSFD ($\sigma^2 = 0$)	**-17.94**	**-6.65**	**-22.84**

Table 4: NMSE: Frequency Domain Methods

Algorithms	Ambient Noise (Road)	Babble Noise	Machine Noise
	NMSE	NMSE	NMSE
FD	1.49	1.07	1.10
NSFD	1.37	0.79	0.72
Improved Noise NSFD	**1.21**	**0.80**	**0.56**
Improved Noise NSFD ($\sigma^2 = 0$)	**1.07**	**0.57**	**0.46**

Table 5: ATR: S-SBL

Algorithms	Ambient Noise (Road)	Babble Noise	Machine Noise
	ATR (dB)	ATR (dB)	ATR (dB)
S-SBL	-16.40	-7.11	-18.77
Improved Noise S-SBL	**-16.62**	**-7.56**	**-20.23**

Table 6: NMSE: S-SBL

Algorithms	Ambient Noise (Road)	Babble Noise	Machine Noise
	NMSE	NMSE	NMSE
S-SBL	1.38	0.79	0.91
Improved Noise S-SBL	**1.06**	**0.48**	**0.38**

In addition, the Improved Noise S-SBL leads to additional sparsity in the estimate as seen in Fig.1. Such a solution will be of relevance for Time Difference of Arrival estimation in reverberant environments using ReIRs [5] since the iterative sparse regularization leads to increased prominence of the peaks.

Improved noise methods can be utilized to improve solutions after an initial estimate is computed using regular noise models. The algorithm can be iterated until desired performance levels are achieved. Fig.2 shows the performance variation with iteration for the Improved S-SBL algorithm when the initial estimate is given through the regular S-SBL algorithm.

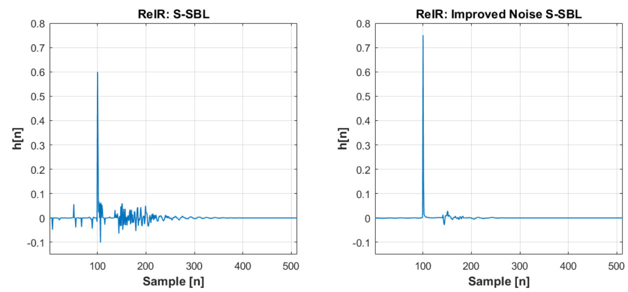

Fig. 1: Increased Sparsity with Improved Noise S-SBL

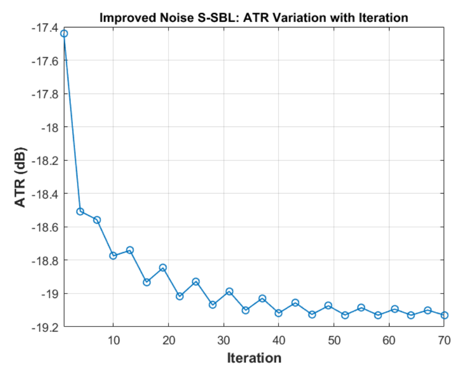

Fig. 2: Algorithm Performance with Iteration

6. CONCLUSION

We formulated an accurate characterization of the noise in the ReIR estimation process. The modifications to existing methods were also studied and the performance gains were evaluated. An improved system identification method when the noise is correlated with measurements and observations was studied using iterative algorithms. We also rationalized current methods as an approximation to the exact noise characterization.

7. REFERENCES

[1] Sharon Gannot, David Burshtein, and Ehud Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *Signal Processing, IEEE Transactions on*, vol. 49, no. 8, pp. 1614–1626, 2001.

[2] Sharon Gannot and Israel Cohen, "Speech enhancement based on the general transfer function generalized sidelobe cancellation and postfiltering," *Speech and Audio Processing, IEEE Transactions on*, vol. 12, no. 6, pp. 561–571, 2004.

[3] Bracha Laufer, Ronen Talmon, and Sharon Gannot, "Relative transfer function modeling for supervised source localization," in *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on*. IEEE, 2013, pp. 1–4.

[4] Alexander Krueger, Ernst Warsitz, and Reinhold Haeb-Umbach, "Speech enhancement with a GSC-like structure employing eigenvector-based transfer function ratios estimation," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 1, pp. 206–219, 2011.

[5] Tsvi G. Dvorkind and Sharon Gannot, "Time difference of arrival estimation of speech source in a noisy and reverberant environment," *Elsevier Signal Process.*, vol. 85, no. 1, pp. 177–204, 2005.

[6] 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.

[7] Zbynek Koldovsky, Petr Tichavsky, and David Botka, "Noise reduction in dual-microphone mobile phones using a bank of pre-measured target-cancellation filters," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 679–683.

[8] Donald W Marquardt and Ronald D Snee, "Ridge regression in practice," *The American Statistician*, vol. 29, no. 1, pp. 3–20, 1975.

[9] Michael E Tipping, "Sparse Bayesian learning and the relevance vector machine," *The journal of machine learning research*, vol. 1, pp. 211–244, 2001.

[10] 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, 2015.

[11] David P Wipf and Bhaskar D Rao, "Sparse Bayesian learning for basis selection," *Signal Processing, IEEE Transactions on*, vol. 52, no. 8, pp. 2153–2164, 2004.

[12] Zbynek Koldovsky, Jiri Malek, and Sharon Gannot, "Spatial source subtraction based on incomplete measurements of relative transfer function," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 23, no. 8, pp. 1335–1347, 2015.

[13] Elior Hadad, Florian Heese, Peter Vary, and Sharon Gannot, "Multichannel audio database in various acoustic environments," in *Acoustic Signal Enhancement (IWAENC), 2014 14th International Workshop on*. IEEE, 2014, pp. 313–317.

[14] Nobutaka Ono, Zbynek Koldovsky, Shigeki Miyabe, and Noboru Ito, "The 2013 signal separation evaluation campaign," in *Machine Learning for Signal Processing (MLSP), 2013 IEEE International Workshop on*. IEEE, 2013, pp. 1–6.