Jointly Leveraging Decorrelation and Sparsity for Improved Feedback Cancellation in Hearing Aids

Kuan-Lin Chen, Ching-Hua Lee, Bhaskar D. Rao, and Harinath Garudadri

Department of Electrical and Computer Engineering
University of California, San Diego

EUSIPCO 2020
December 7, 2020

Outline

Contributions

Kuan-Lin Chen, Ching-Hua Lee, Bhaskar D. Rao, and Harinath Garudadri, "**Jointly leveraging decorrelation and sparsity for improved feedback cancellation in hearing aids**," to appear in *European Signal Processing Conference (EUSIPCO)*, 2020.

- Propose a joint framework for leveraging decorrelation and sparsity to the feedback problem in hearing aids (HAs)
- Using different speech input signals, feedback paths and amplifications, we extensively study the efficacy of AFC using different numbers of subbands and degrees of promoted sparsity
- Both commonly used AFC evaluation criteria and objective evaluations on intelligibility and quality are presented on a large speech corpus to illustrate the effectiveness of the proposed AFC framework
- We show that the benefits of decorrelation and sparsity promoting for AFC are additive and complementary

Acoustic Feedback Problem

- The acoustic feedback or so-called howling effect induces the strong coupling between the receiver (loudspeaker) and the microphone in HAs
- Howling deteriorates the intelligibility, quality and maximum stable gain of the input

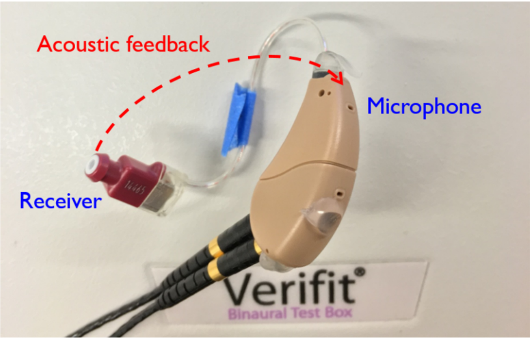


Figure: Illustration of acoustic feedback in the hearing aid.

Challenges and Our Approach

- The classic least mean square (LMS) and normalized LMS (NLMS) [24, 9, 17] both show degraded convergence behaviors when the input signal is colored
- In the AFC literature, many works have been dedicated to either decorrelation [10, 3, 7, 8, 22, 20, 21, 19, 18, 15] or promoting sparsity [12, 14]; a joint exploration on both is lacking

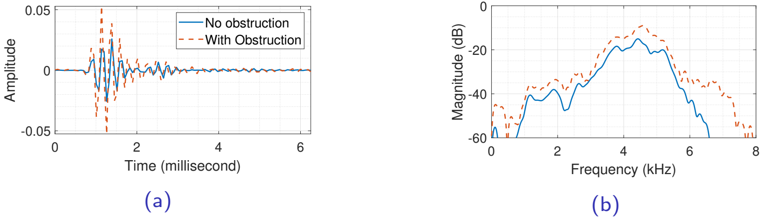


Figure: The truncated FIR filters of different feedback paths were measured from a HA on a dummy head. (a) represents the IRs and (b) shows the magnitudes of the frequency responses.

- In our approach, both decorrelation and sparsity are jointly exploited to eliminate the howling effect

Adaptive Feedback Cancellation (AFC)

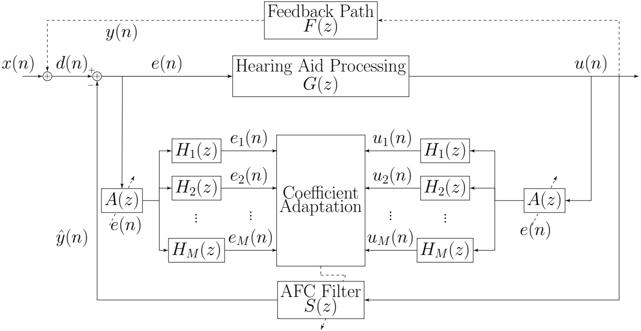


Figure: Block diagram of the proposed AFC framework.

- The prediction-error filter $A(z)$ from the prediction error method (PEM) forms a time-varying analysis filter bank, i.e., $A(z)H_i(z)$.
- The synthesis filters are not required in our proposed framework. The subband error signals are computed and then aggregated together to update the fullband filter taps.
- A generalized update rule is proposed for AFC

Optimization Criterion

We propose the following optimization criterion to jointly exploit sparsity and achieve decorrelation:

$$J(\boldsymbol{s}) = \sum_{i=1}^{M} e_i^2(n) + \tau \|\boldsymbol{s}\|_p^p \tag{1}$$

where

- $\tau \to 0^+$ is a regularization parameter
- $e_i(n) = d_i(n) - \boldsymbol{u}_i^T(n)\boldsymbol{s}$ is the i^{th} subband error scalar
- $d_i(n)$ and $\boldsymbol{u}_i(n)$ are the i^{th} subband desired scalar and the i^{th} subband input vector, respectively
- M is the number of subbands
- optimization variable $\boldsymbol{s} = \begin{bmatrix} s_1 & s_2 & \cdots & s_L \end{bmatrix}^T \in \mathbb{R}^L$ denotes the adaptive filter of length L.
- we have used the p-norm-like diversity measure $\|\boldsymbol{s}\|_p^p = \sum_{i=1}^{L} |s_i|^p$ for promoting sparsity where the parameter $p \in (0, 2]$ controls the degree of sparsity promoting [11, 16]

Jointly Leveraging Decorrelation and Sparsity

$$J(\boldsymbol{s}) = \sum_{i=1}^{M} e_i^2(n) + \tau \|\boldsymbol{s}\|_p^p$$

- Can be used to generalize the proportionate-type normalized subband adaptive filtering (PtNSAF) framework
- Jointly combines decorrelation (first term in (1)) and tunable sparsity exploitation (second term in (1)) in one cost function
- The PEM in our framework can be considered as a way to establish a time-varying analysis filter bank for better decorrelation

Solving the Optimization Problem

We minimize the cost function (1) using the reweighted ℓ_2 framework [11], affine scaling transformation [16] and the regularized Newton's method [2].

Using the reweighted ℓ_2 framework, the criterion (1) becomes

$$J(\boldsymbol{s}) = \sum_{i=1}^{M} \left| e_i(n) \right|^2 + \tau \|\boldsymbol{s}\|^2_{\boldsymbol{W}^{-1}(n)} . \tag{2}$$

To proceed, we perform the affine scaling transform (AST) on the optimization variable $\boldsymbol{s}$:

$$\boldsymbol{q} = \boldsymbol{W}^{-\frac{1}{2}}(n)\boldsymbol{s}. \tag{3}$$

Applying (3) into (2) , we obtain an equivalent optimization problem

$$\min_{\boldsymbol{q}} J(\boldsymbol{q}) = \sum_{i=1}^{M} \left| e_i(n) \right|^2 + \tau \|\boldsymbol{q}\|_2^2 \tag{4}$$

in the $\boldsymbol{q}$ domain.

Newton's Method in the $\boldsymbol{q}$ Domain

We define the *a posteriori* AST variable at time n as $\boldsymbol{q}(n|n) \triangleq \boldsymbol{W}^{-\frac{1}{2}}(n)\boldsymbol{s}(n)$ and the *a priori* AST variable as $\boldsymbol{q}(n+1|n) \triangleq \boldsymbol{W}^{-\frac{1}{2}}(n)\boldsymbol{s}(n+1)$. Now, we consider the regularized Newton's method for the update rule on minimizing $J(\boldsymbol{q})$, i.e.,

$$\boldsymbol{q}(n+1|n) = \boldsymbol{q}(n|n) - \mu \left[\nabla^2_{\boldsymbol{q}} J\left(\boldsymbol{q}(n|n)\right) + 2\delta \boldsymbol{I}\right]^{-1} \nabla_{\boldsymbol{q}} J\left(\boldsymbol{q}(n|n)\right) \tag{5}$$

where $\mu > 0$ is the learning rate or the step size for adaptation and $\delta > 0$ is a regularization parameter. The gradient of $J(\boldsymbol{q})$ is given by

$$\nabla_{\boldsymbol{q}} J(\boldsymbol{q}(n|n)) = -2\boldsymbol{W}^{\frac{1}{2}}(n)\boldsymbol{U}(n)\boldsymbol{e}(n) + 2\tau \boldsymbol{q}(n|n). \tag{6}$$

Next, the Hessian is given by

$$\nabla^2_{\boldsymbol{q}} J(\boldsymbol{q}(n|n)) = 2\boldsymbol{W}^{\frac{1}{2}}(n)\boldsymbol{U}(n)\boldsymbol{U}^T(n)\boldsymbol{W}^{\frac{1}{2}}(n) + 2\tau \boldsymbol{I}. \tag{7}$$

Therefore, the update rule on $\boldsymbol{q}$ domain is given by

$$\boldsymbol{q}(n+1|n) = \left(\boldsymbol{I} - \frac{\mu\tau}{\delta+\tau}\left[\boldsymbol{I} - \Psi(n)\right]\right)\boldsymbol{q}(n|n) + \mu \boldsymbol{W}^{\frac{1}{2}}(n)\boldsymbol{U}(n)\Phi(n)\boldsymbol{e}(n). \tag{8}$$

Update Rule

We have used

$$\Psi(n) = \boldsymbol{W}^{\frac{1}{2}}(n)\boldsymbol{U}(n)\Phi(n)\boldsymbol{U}^T(n)\boldsymbol{W}^{\frac{1}{2}}(n). \tag{9}$$

Notice that the inverse of the regularized weighted subband correlation matrix, i.e.,

$$\Phi(n) = \left[(\delta + \tau)\boldsymbol{I}_M + \boldsymbol{U}^T(n)\boldsymbol{W}(n)\boldsymbol{U}(n)\right]^{-1} \tag{10}$$

is a small matrix inversion which only has M-by-M since we have $L \gg M$ in most cases. Then, by utilizing (3) in (8) to convert $\boldsymbol{q}$ back to the $\boldsymbol{s}$ domain, we have

$$\boldsymbol{s}(n+1) = \left(\boldsymbol{I} - \frac{\mu\tau}{\delta + \tau}\left[\boldsymbol{I} - \Psi(n)\right]\right)\boldsymbol{s}(n) + \mu\boldsymbol{W}(n)\boldsymbol{U}(n)\Phi(n)\boldsymbol{e}(n). \tag{11}$$

Finally, setting $\tau \to 0^+$ leads to the update rule for the GPtNSAF [2]:

$$\boldsymbol{s}(n+1) = \boldsymbol{s}(n) + \mu\boldsymbol{g}(n) \tag{12}$$

where

$$\boldsymbol{g}(n) = \boldsymbol{W}(n)\boldsymbol{U}(n)\left[\delta\boldsymbol{I}_M + \boldsymbol{U}^T(n)\boldsymbol{W}(n)\boldsymbol{U}(n)\right]^{-1}\boldsymbol{e}(n). \tag{13}$$

The Proportionate Matrix

For the proportionate matrix

$$\boldsymbol{W}(n) = \text{diag}\{w_1(n), w_2(n), \cdots, w_L(n)\}, \tag{14}$$

it is given by

$$w_i(n) = \left(\left|s_i(n)\right| + c\right)^{2-p}, i = 1, 2, \cdots, L \tag{15}$$

where $c > 0$ is a regularization constant for avoiding stagnation and instability. The suggested range of the parameter p for sparse, compressible (quasi-sparse) and dispersive solutions are $[1.0, 1.2]$, $(1.2, 1.8)$ and $[1.8, 2.0]$, respectively [12].

Sparsity-promoting Normalized Subband Adaptive Filter (S-NSAF) and Generalization

$$\boldsymbol{s}(n+1) = \boldsymbol{s}(n) + \mu \boldsymbol{W}(n)\boldsymbol{U}(n)\left[\delta \boldsymbol{I}_M + \boldsymbol{U}^T(n)\boldsymbol{W}(n)\boldsymbol{U}(n)\right]^{-1}\boldsymbol{e}(n),$$

$$w_i(n) = \left(\left|s_i(n)\right| + c\right)^{2-p}, i = 1, 2, \cdots, L.$$

In sum, (13) and (15) give the proposed Sparsity-promoting Normalized Subband Adaptive Filter algorithm (S-NSAF).

- $\boldsymbol{W}(n)$ promotes sparsity (induced from the p-norm-like diversity measure)
- $\left[\delta \boldsymbol{I}_M + \boldsymbol{U}^T(n)\boldsymbol{W}(n)\boldsymbol{U}(n)\right]^{-1}$ decorrelates the input signal so that the optimization landscape is not elongated (induced from the subband errors)

	$M = 1$	$M > 1, \boldsymbol{H} \neq \boldsymbol{I}$	$M > 1, \boldsymbol{H} = \boldsymbol{I}$
$p = 2$	NLMS [9]	NSAF [4]	APA [6]
$2 > p > 0$	PtNLMS [23]	PtNSAF [1]	PtAPA [13]

Table: Different cases of S-NSAF. For the correspondence to NSAF and PtNSAF, $\Phi(n)$ needs to be approximated by a diagonal matrix using a proper analysis filter bank.

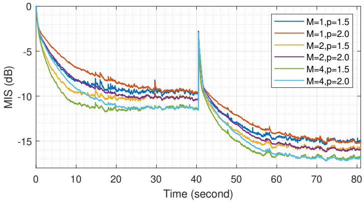
Experimental Setup (1/2)

- The experiments were conducted at 16 kHz with the input speech signal $x(n)$ from the TIMIT dataset [5]
- Two feedback paths were measured from the real-world setup as shown before
- The HA processing, was simulated by $G(z) = gz^{-d}$ where g was the gain in the linear scale and d was the samples of delay corresponding to a fixed latency of 8 milliseconds
- The length $L = 100$ was set to the same size as the truncated FIR filter below and all taps were initialized by 0
- For PEM, the order of the prediction-error filter $A(z)$ was 20 and the filter was updated every 10 milliseconds via Levinson-Durbin recursion with the window length of 160 samples [15]

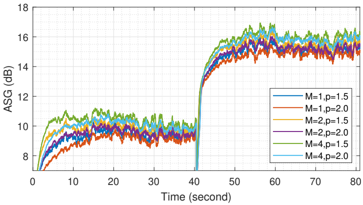
Experimental Setup (2/2)

- The analysis filter bank H is a cosine-modulated pseudo-quadrature mirror filter (QMF) bank. $M = 1, 2, 4$ were chosen to be evaluated. We maintain the same length $N = 16$ of the analysis filters for $M = 2$ and $M = 4$
- The p values which were chosen to be tested are 1.5 [12] and 2.0
- For regularizations, we used $\delta = 10^{-5}$ and $c = 10^{-3}$ for all simulations. The step size is given by $\mu = \frac{1}{M} \times 10^{-3}$ so that the comparison is fair for adaptive filters using different M
- All curves in Fig. 4 and Fig. 5 were ensemble averaged over 100 different speech signals
- During all experiments, a sudden change of the feedback path was introduced at half time where this new path was given by the one with obstruction

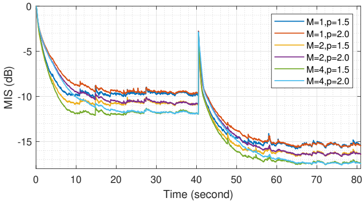
Normalized Misalignment and Added Stable Gain



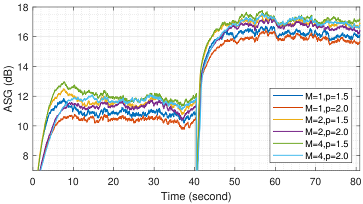
(a) Normalized misalignments (without PEM)



(b) ASGs (without PEM)



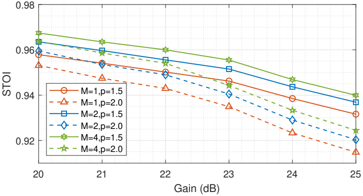
(c) Normalized misalignments (with PEM)



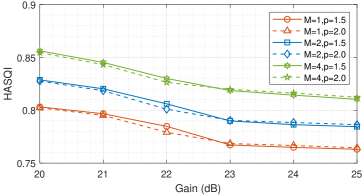
(d) ASGs (with PEM)

Figure: The performance of AFC is better with higher M for a given p; and $p = 1.5$ is better than $p = 2.0$ for a given M, in terms of normalized misalignment and ASG.

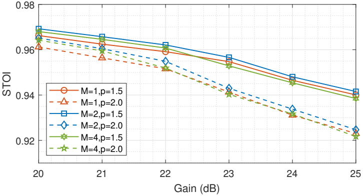
Intelligibility and Quality



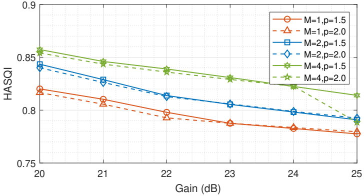
(a) STOI versus gain (without PEM)



(b) HASQI versus gain (without PEM)



(c) STOI versus gain (with PEM)



(d) HASQI versus gain (with PEM)

Figure: In (a), the speech intelligibility is better with higher M for a given p; and $p = 1.5$ is better than $p = 2.0$ for a given M. In (b), the speech quality is improved by choosing higher M; and the p value seems to be irrelevant.

Conclusion

- A new formulation of jointly exploring sparsity promoting and decorrelation is proposed for practical AFC applications
- The effectiveness of using different degrees of sparsity promoting and number of subbands are studied extensively with a large speech corpus and different feedback paths
- Higher number of subbands (up to a certain level) is better
- A proper degree of sparsity promoting gives superior AFC performance
- Commonly used metrics including misalignment, ASG, STOI, and HASQI are better in our proposed method regardless of the incorporation of the PEM

References I

[1] Mohammad Shams Esfand Abadi and Sima Kadkhodazadeh. A family of proportionate normalized subband adaptive filter algorithms. *Journal of the Franklin Institute*, 348(2):212–238, 2011.

[2] Kuan-Lin Chen, Ching-Hua Lee, Bhaskar D. Rao, and Harinath Garudadri. A generalized proportionate-type normalized subband adaptive filter. In *Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, pages 749–753. IEEE, 2019.

[3] Hsiang-Feng Chi, Shawn X Gao, Sigfrid D Soli, and Abeer Alwan. Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids. *Speech Communication*, 39(1-2):147–161, 2003.

[4] Marc De Courville and Pierre Duhamel. Adaptive filtering in subbands using a weighted criterion. *IEEE Transactions on Signal Processing*, 46(9):2359–2371, 1998.

[5] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, N. Dahlgren, and V. Zue. TIMIT acoustic-phonetic continuous speech corpus. *Linguistic Data Consortium*, 1993.

[6] Steven L Gay and Sanjeev Tavathia. The fast affine projection algorithm. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 3023–3026. IEEE, 1995.

[7] Meng Guo, Søren Holdt Jensen, and Jesper Jensen. Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancement. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(9):2549–2563, 2012.

[8] Meng Guo, Søren Holdt Jensen, Jesper Jensen, and Steven L Grant. On the use of a phase modulation method for decorrelation in acoustic feedback cancellation. In *European Signal Processing Conference (EUSIPCO)*, pages 2000–2004. IEEE, 2012.

[9] Simon S Haykin. *Adaptive Filter Theory*. Pearson Education India, 2008.

[10] Johan Hellgren. Analysis of feedback cancellation in hearing aids with Filtered-x LMS and the direct method of closed loop identification. *IEEE Transactions on Speech and Audio Processing*, 10(2):119–131, 2002.

[11] C.-H. Lee, B. D. Rao, and H. Garudadri. Proportionate adaptive filters based on minimizing diversity measures for promoting sparsity. In *Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, pages 769–773. IEEE, 2019.

[12] Ching-Hua Lee, Bhaskar D Rao, and Harinath Garudadri. Sparsity promoting LMS for adaptive feedback cancellation. In *European Signal Processing Conference (EUSIPCO)*, pages 226–230. IEEE, 2017.

[13] Constantin Paleologu, Silviu Ciochina, and Jacob Benesty. An efficient proportionate affine projection algorithm for echo cancellation. *IEEE Signal Processing Letters*, 17(2):165–168, 2009.

[14] Somanath Pradhan, Vinal Patel, Kashyap Patel, Jyoti Maheshwari, and Nithin V George. Acoustic feedback cancellation in digital hearing aids: A sparse adaptive filtering approach. *Applied Acoustics*, 122:138–145, 2017.

[15] Somanath Pradhan, Vinal Patel, Dipen Somani, and Nithin V George. An improved proportionate delayless multiband-structured subband adaptive feedback canceller for digital hearing aids. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(8):1633–1643, 2017.

References II

[16] Bhaskar D Rao and Kenneth Kreutz-Delgado. An affine scaling methodology for best basis selection. *IEEE Transactions on Signal Processing*, 47(1):187–200, 1999.

[17] Ali H Sayed. *Adaptive Filters*. John Wiley & Sons, 2011.

[18] MG Siqueira, R Speece, E Petsalis, A Alwan, S Soli, and S Gao. Subband adaptive filtering applied to acoustic feedback reduction in hearing aids. In *Asilomar Conference on Signals, Systems and Computers (ACSSC)*, pages 788–792. IEEE, 1996.

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

[20] 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*, 53(10):3749–3763, 2005.

[21] Ann Spriet, Geert Rombouts, Marc Moonen, and Jan Wouters. Adaptive feedback cancellation in hearing aids. *Journal of the Franklin Institute*, 343(6):545–573, 2006.

[22] Falco Strasser and Henning Puder. Adaptive feedback cancellation for realistic hearing aid applications. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 23(12):2322–2333, 2015.

[23] Kevin Wagner and Miloš Doroslovački. *Proportionate-type normalized least mean square algorithms*. John Wiley & Sons, 2013.

[24] Bernard Widrow and Samuel D. Stearns. *Adaptive Signal Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1985.