

A ROBUST FREQUENCY-DOMAIN ECHO CANCELLER

Tomas Gänsler

Dept. of Applied Electronics, Signal Processing Group
Lund University, Sweden
tg@tde.lth.se

ABSTRACT

A recursive transfer function estimation algorithm is presented and analyzed. Applications can be found in either electric or acoustic echo cancellation. The proposed algorithm is robust against burst disturbances that are caused by detection misses of double-talk present at the output of the echo path. A frequency-domain technique is used and a robustness function is derived from a criterion that is valid in the application. Analysis of the robust algorithm shows that good performance is to be expected. The performance of the algorithm when operated on real-life speech data in a full duplex communication system is shown by examples. Double-talk detection misses are shown to be well handled by the robust algorithm; yet, convergence rate and variance efficiency are as high as that of a non-robust least squares algorithm.

1. INTRODUCTION

Double-talk, i.e., talkers and listeners interrupting each other in full duplex communication, is limiting the performance of echo cancellers. The reason for this is that the double-talk detector (DTD, cf. Fig. 1) sometimes fails to disconnect the adaptation algorithm. As a consequence, a speech burst at the output of the echo path enters the adaptation algorithm and "moves" the estimate away from "optimum." The algorithm has therefore be made robust against these bursts in order to reduce this problem. A principle of robustness, equivalent to the sign LMS algorithm, was introduced in [1]. This algorithm has the drawback of much slower convergence rate than a standard LMS algorithm, [2, 3].

In this work, a frequency-domain algorithm is proposed that approach the double-talk problem by introducing robustness against deviations in the *distributional assumptions*. Frequency-domain techniques fit well within the framework of robust statistics in the echo canceller application since a contaminated Gaussian model for the residual echo can be adopted. The reasons for this are the following: a frequency transformation converts the burst of speech into outliers, i.e. infrequently appearing large disturbances in the output signal, which are well handled by robust statistics; as a consequence of the central limit theorem, the transformation moves the system noise distribution closer to the Gaussian. Therefore, a Gaussian assumption of the ongoing system noise is well-founded.

A robust M-estimator is derived from the complex contaminated Gaussian model by techniques from classical robust statistical theory, [4]. These techniques are adapted to the complex parameter case of transfer function estimation. Statistical analysis of the algorithm is made and the asymptotic variance of the estimate is derived for the stationary input case.

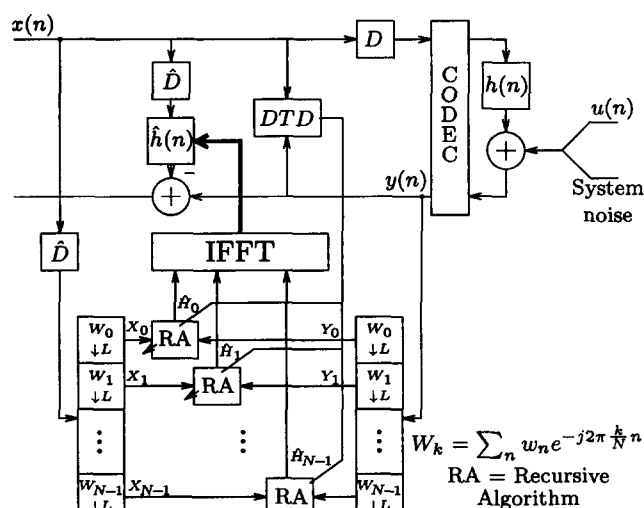


Figure 1. Echo canceller and model of electric echo path. $h(n)$ represents the linear part of the hybrid circuit that performs the transition between 4-wire and 2-wire lines. D is a pure delay. The signals are transformed into the frequency domain where the transfer function, $\hat{H}_k$, is recursively estimated (RA).

2. PRELIMINARIES

The system $h(n)$ is restricted to be FIR of order p and the input, $x(n)$, and output, $y(n)$, are related through,

$$y(n) = \sum_{l=0}^p h(l)x(n-l) + v(n), \quad (1)$$

where $v(n)$ represents the disturbance. Let $Y_k(q)$ be a complex value denoting the windowed discrete Fourier transform (DFT) at frequency k/N of a block of time-domain

data $y(n)$, $n = q \dots q + N - 1$, $q = 0, \dots, m$. Transformed variables of $h(n)$, $x(n)$, $v(n)$ are analogously defined. The frequency-domain model is then given by,

$$\begin{aligned} Y_k(q) &= H_k(q)X_k(q) + \Xi_k(q) + V_k(q) \\ &\approx H_k(q)X_k(q) + V_k(q), \end{aligned} \quad (2)$$

where $\Xi_k(q)$ represents the inherent bias after the approximation. This bias is small and it decreases as the data block size N increases, [5]. For notation of simplicity, the frequency index k is from now on excluded. Denote a residual by $Z(q)$,

$$Z(q) = Y(q) - \hat{H}(q-1)X(q). \quad (3)$$

The transformed system noise is well modeled by a Gaussian probability density function (pdf), $f(Z(q))$, and as a consequence of the possible double-talk, the residual (after convergence of $\hat{H}(q)$) is modeled by a contaminated Gaussian pdf, $f_0(Z(q))$,

$$f_0(Z(q)) = (1 - \epsilon)f(Z(q)) + \epsilon g(Z(q)), \quad (4)$$

where ϵ represents the level of contamination and $g(Z(q))$ the pdf of double-talk. A non-recursive M-estimate of the transfer function, $\hat{H}$, following the principles in [4], is given by the minimization problem,

$$\min_H \sum_{q=0}^m \rho\left(\frac{r(q)}{S}\right), \quad (5)$$

where $r(q)$ is the magnitude of the residual, $Z(q)$, Eq. (3), and S is a scale parameter, in this case a robust estimate of the residual standard deviation. Eq. (5) is equivalent to a least-squares problem if the function ρ is quadratic. Minimization of Eq. (5) with respect to the transfer function gives the implicit equation,

$$\sum_{q=0}^m X^*(q)\Psi\left\{\frac{r(q)}{S}\right\}\text{sign}\{Z(q)\} = 0, \quad \Psi(r) = \frac{d}{dr}\rho(r). \quad (6)$$

Naturally, a good estimator is found by making Eq. (5) a maximum likelihood estimator of the transfer function based on the statistical model, Eq. (4), i.e.,

$$\Psi(r(q)) \sim -\frac{\partial}{\partial r(q)} \ln\{f_0(r(q))\}, \quad (7)$$

This choice results in a robust M-estimator of the transfer function.

Calculus of variations is used to find the worst case candidate for $f_0(r(q))$, [6]. A tractable monotone approximation to the function Ψ resulting from the worst case $f_0(r(q))$ is the two dimensional Huber-function, [4, 6, 7],

$$\Psi(r(q)) = \min\{r(q), r_0\}, \quad r(q) = |Z(q)|. \quad (8)$$

The radius r_0 is a fixed value that determines the level of robustness. The robust estimate of the scale parameter, S , is given by solving the implicit equation, [4],

$$\sum_{q=0}^m \chi\left(\frac{r(q)}{S}\right) = 0, \quad (9)$$

ϵ	r_0	β_0
0.001	2.1615	0.9906
0.002	2.0265	0.9835
0.005	1.8390	0.9660
0.010	1.6892	0.9424

Table 1. Numerical values of ϵ , r_0 and, β_0 .

where χ is an even function and in this paper $\chi\{\cdot\} = \Psi^2\{\cdot\} - \beta_0$. β_0 is chosen so that, $E\{\chi(Z(q))\} = 0$, when $Z(q) \in N\{0, 1\}$, i.e. $\beta_0 = 1 - e^{-r_0^2}$.

The relation between r_0 and ϵ is determined by integration of the contaminated pdf in Eq. (4). The choice of pdf corresponding to Eq. (8) gives,

$$1 + \frac{e^{-r_0^2}}{2r_0^2} = \frac{1}{1 - \epsilon}. \quad (10)$$

This expression can easily be solved numerically. Table I shows r_0 and β_0 for some values of ϵ .

3. ALGORITHMS AND ANALYSIS

In order to derive a block-recursive transfer function algorithm (RA, Fig. 1) that minimizes Eq. (5), the Robbins-Monro technique, [8], is utilized. Tracking is enabled by introducing a forgetting factor, $\lambda \in (0, 1]$, in the resulting algorithm. It is found that, [6],

$$\begin{aligned} \hat{H}(m) &= \hat{H}(m-1) \\ &+ \frac{X^*(m)}{\gamma(m)} \text{sign}\{Z(m)\} S(m) \Psi\left\{\frac{r(m)}{S(m)}\right\}, \\ \gamma_{\text{sub}}(m) &= \lambda \gamma_{\text{sub}}(m-1) + |X(m)|^2, \\ \gamma_{\text{opt}}(m) &= \lambda \gamma_{\text{opt}}(m-1) + |X(m)|^2 \Psi'\left\{\frac{r(m)}{S(m)}\right\}. \end{aligned} \quad (11)$$

The gain factor $\gamma_{\text{sub}}(m)$ is a suboptimal but a practical choice and $\gamma_{\text{opt}}(m)$ is the optimal gain factor leading to the theoretically smallest variance of the estimate. A non-robust least squares estimate, $\hat{H}_{LS}(m)$, is given by a quadratic ρ in Eq. (5), i.e. Ψ is linear. In this case,

$$\hat{H}_{LS}(m) = \hat{H}_{LS}(m-1) + \frac{X^*(m)}{\gamma_{\text{sub}}(m)} Z(m). \quad (12)$$

A robust recursive algorithm for the scale parameter $S(m)$, can be found by the same principle as for the transfer function estimate,

$$\begin{aligned} \hat{S}(m) &= \hat{S}(m-1) \\ &+ \frac{1}{\gamma_2(m-1)} \hat{S}(m-1) \chi\left\{\frac{r(m)}{\hat{S}(m-1)}\right\} \\ \gamma_2(m) &= \lambda \gamma_2(m-1) + \frac{r(m)}{\hat{S}(m)} \chi'\left\{\frac{r(m)}{\hat{S}(m)}\right\}. \end{aligned} \quad (13)$$

The asymptotic variance of the estimate, Eq. (11), can be derived for the case of $\lambda = 1$ and Gaussian system noise. The asymptotic relative efficiency (ARE), compared to the Cramér-Rao bound is defined as,

$$ARE_i = \frac{CRB}{\text{var}\{\hat{H}(\gamma_i)\}}, \quad i = \text{opt}, \text{sub}. \quad (14)$$

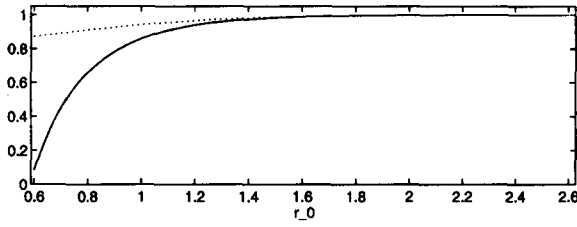


Figure 2. Asymptotic relative (to the CRB) efficiencies of algorithms in Gaussian noise. The robust algorithm, Eq. (11), with suboptimal gain factor (solid line). Result with optimal gain factor (dotted line).

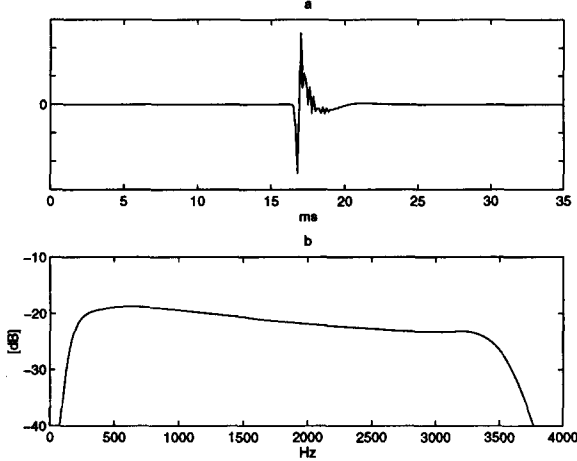


Figure 3. (a) True impulse response of the hybrid circuit; (b) Corresponding amplitude function.

It is found that, [6], the ARE values of Eq. (11) are,

$$ARE_{opt} = \frac{[1 - e^{-r_0^2} + \sqrt{\pi}r_0(1 - \Phi(\sqrt{2}r_0))]^2}{[1 - e^{-r_0^2}]}, \quad (15)$$

$$ARE_{sub} = \frac{[2(1 - e^{-r_0^2} + \sqrt{\pi}r_0(1 - \Phi(\sqrt{2}r_0))) - 1]}{[1 - e^{-r_0^2}]}, \quad (16)$$

where $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{\tau^2}{2}} d\tau$. Figure 2 shows the efficiencies (optimal and suboptimal γ), versus the choice of radius, r_0 . The loss in efficiency when using the suboptimal gain parameter, $\gamma_{sub}(m)$, of Eq. (11) instead of the optimal $\gamma_{opt}(m)$ is insignificant if $r_0 \geq 1.4$.

4. PERFORMANCE

Performance of the robust and non-robust algorithms is presented by estimates of the amplitude function and the misalignment. The misalignment, $M(m)$, is defined as,

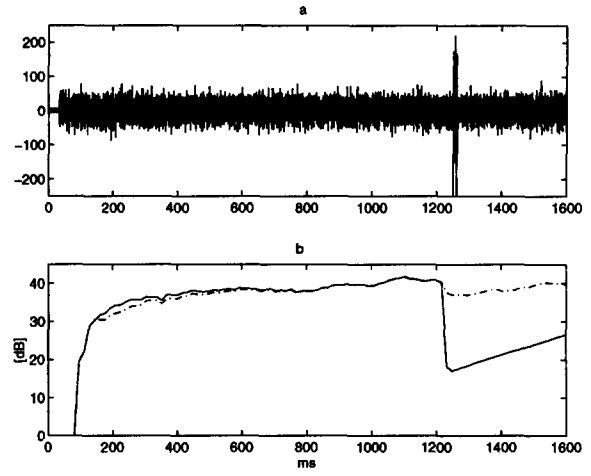


Figure 4. (a) Echo and near-end burst at approximately 1250 ms; (b) Misalignment. The non-robust algorithm (solid line). The robust algorithm (dashed-dotted line). The input signal is white noise in this simulation.

$$M(m) = -10 \log_{10} \left(\frac{\int_{-\frac{1}{2}}^{\frac{1}{2}} |\hat{H}_f(m) - H_f|^2 df}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |H_f|^2 df} \right). \quad (17)$$

In the following examples the system shown in Fig. 3 is used. The attenuation of the hybrid circuit is in this case 20 dB. The transform block length of the algorithm, N , is 256 and ϵ is chosen to 0.002. L , the downsample factor in Fig. 1, is 128.

In the first example, Fig. 4, the input signal is white Gaussian noise and the SNR (input signal to system noise) is 34 dB. A burst of noise (12.5 ms duration) is present at about 1250 ms. A double-talk detector is not implemented in this example. It is shown in Fig. 4b that both algorithms have approximately equal convergence rate. The non-robust algorithm is severely affected by the burst while the robust experiences only minor degradation due to the burst.

The second example shows the performance for speech signals. Figures 5a, b show the data used where the average SNR is 34 dB. A DTD switches off the adaptation when near-end talk occurs. It can be seen in Fig. 5c that the misalignment of the non-robust algorithm (solid line) decreases rapidly when the DTD misses double-talk starting around 7 s. Near-end speech-bursts in this example are clearly visible in Fig. 5b. The estimated amplitude function at 7.08 s (cf. Fig. 5) is shown in Fig. 6. The non-robust estimate deviates considerably from the true amplitude function whereas the robust estimate is practically unaffected by the burst disturbances.

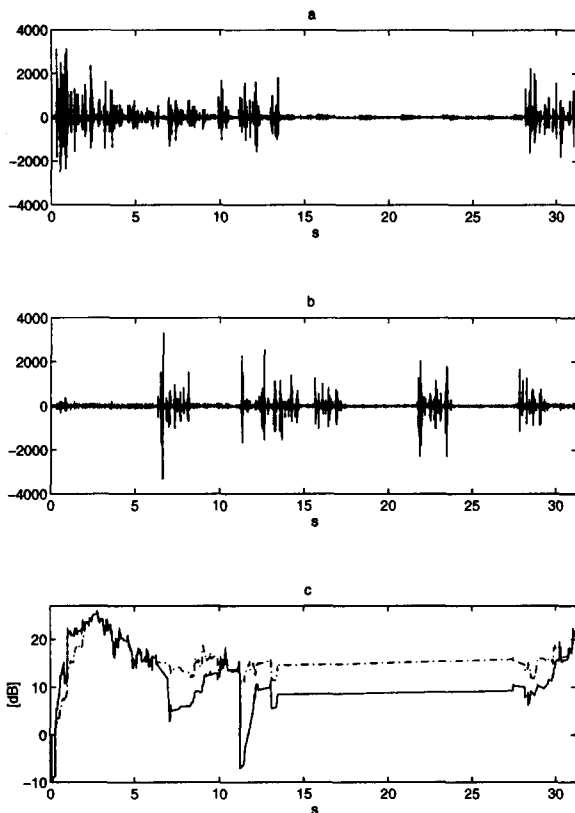


Figure 5. (a) Far-end speech; (b) Near-end speech and echo; (c) Misalignment. The non-robust algorithm (solid line). The robust algorithm (dashed-dotted line).

5. DISCUSSION

The application for the suggested algorithm is echo cancelling where severe burst disturbances can occur. Figure 5 shows the superiority of the robust estimator in double-talk situations.

The block processing contributes to the success of the robust algorithm since bursts are transformed into less frequently occurring outliers and the transformed background noise is well approximated as Gaussian due to the central limit theorem. Thus, the residual obey the ϵ -contaminated Gaussian distribution which form the basis for the robust approach. It is shown that the complex Huber function, Eq. (8), is well-suited for the problem and the level of contamination, ϵ Eq. (4), can be assumed small in the robust algorithm, Eq. (11). The choice of small ϵ makes r_0 large, which means that the asymptotic variance of the estimate approaches the minimum possible, Eq. (15), Fig. 2. In practice, it is possible to choose ϵ small, resulting in high efficiency of the robust algorithm when there are no outliers present.

The robust algorithm is suitable in both electric and acoustic echo cancelling. By employing a robust algorithm a less complex double-talk detector can be used in the echo canceller without degraded performance. This is an advan-

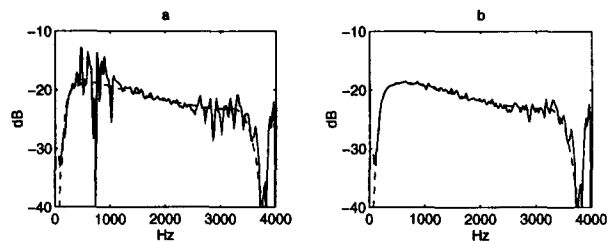


Figure 6. Estimated amplitude function (solid lines). True amplitude function (dashed lines). Estimates at 7.08 s, cf. Fig. 5. (a) Non-robust algorithm; (b) Robust algorithm.

tage, since it is difficult to design a low complexity high performance DTD, especially in acoustic applications. It is also possible to use of an algorithm setting (λ) that leads to faster convergence without the risk of divergence during double-talk.

The ideas presented in this paper can of course be used with general subband echo canceller algorithms where the number of coefficients in the adaptive subband filter is larger than one.

REFERENCES

- [1] M. M. Sondhi, "An adaptive echo canceller," *The Bell System Technical Journal*, vol. XLVI, no. 3, pp. 497–510, March 1967.
- [2] T. A. C. M. Claassen and W. F. G. Mecklenbräuker, "Comparison of the convergence of two algorithms for adaptive fir digital filters," *IEEE Trans. on ASSP*, vol. 29, no. 3, pp. 670–678, June 1981.
- [3] A. Gersho, "Adaptive filtering with binary reinforcement," *IEEE Trans. on Information Theory*, vol. 30, no. 2, pp. 191–199, March 1984.
- [4] P. J. Huber, *Robust Statistics*, Probability and Mathematical Statistics. Wiley, 1981.
- [5] L. Ljung, *System Identification, Theory for the User*, Information and System Sciences Series. Prentice-Hall, Englewood Cliffs, New Jersey 07632, 1987.
- [6] T. Gänslar, "Robust recursive transfer function estimation in echo path identification of telephone channels," Tech. Rep. SPR-34, Signal Processing Group, Lund University, Sweden, April 1996.
- [7] O. Hössjer and M. Mettiji, "Robust multiple classification of known signals in noise-An asymptotic weak signal approach," *IEEE Trans. on Information Theory*, vol. 39, no. 2, pp. 594–608, March 1993.
- [8] D. J. Sakrison, *Advances in Communication Systems*, vol. 2, chapter 2, pp. 51–106, Academic Press, April 1966.