# Echo Cancellation : the generalized likelihood ratio test for double-talk vs. channel change

Jean-Yves Tourneret, Neil J. Bershad, José Carlos Moreira Bermudez

# Echo Cancellation—The Generalized Likelihood Ratio Test For Double-Talk Versus Channel Change

Jean-Yves Tourneret, *Senior Member, IEEE*, Neil J. Bershad, *Fellow, IEEE*, and José Carlos M. Bermudez, *Senior Member, IEEE*

***Abstract*—Echo cancellers (EC) are required in both electrical (impedance mismatch) and acoustic (speaker-microphone coupling) applications. One of the main design problems is the control logic for adaptation. Basically, the algorithm weights should be frozen in the presence of double-talk and adapt quickly in the absence of double-talk. The optimum likelihood ratio test (LRT) for this problem was studied in a recent paper. The LRT requires *a priori* knowledge of the background noise and double-talk power levels. Instead, this paper derives a generalized log likelihood ratio test (GLRT) that does not require this knowledge. The probability density function of a sufficient statistic under each hypothesis is obtained and the performance of the test is evaluated as a function of the system parameters. The receiver operating characteristics (ROCs) indicate that it is difficult to correctly decide between double-talk and a channel change, based upon a single look. However, detection based on about 200 successive samples yields a detection probability close to unity (0.99) with a small false alarm probability (0.01) for the theoretical GLRT model. Application of a GLRT-based EC to real voice data shows comparable performance to that of the LRT-based EC given in a recent paper.**

***Index Terms*—Adaptive filters, adaptive signal processing, adaptive systems, echo cancellation, channel change, double-talk, generalized likelihood ratio test.**

## I. Introduction

ECHO cancellers (EC) have been used in networks for voice quality enhancement for several decades. There are two different kinds of applications for speech EC: network and acoustic echo cancellation [1]. The network or hybrid echo on the public switched telephone network (PSTN) is caused by the four-wire to two-wire impedance mismatch. This mismatch results in unwanted reflection of transmitted energy back to the source. Networks are equipped with EC, known as network or line echo cancellers, to remove these unwanted reflections.

J.-Y. Tourneret is with the University of Toulouse, IRIT—ENSEEIHT—TéSA, 31071 Toulouse cedex 7, France (e-mail: Jean-Yves.Tourneret@enseeiht.fr).

N. J. Bershad is with the Department of Electrical Engineering and Computer Science, University of California at Irvine, Newport Beach, CA 92660 USA (e-mail: bershad@ece.uci.edu).

J. C. M. Bermudez is with the Department of Electrical Engineering, Federal University of Santa Catarina, Florianópolis, SC, Brazil (e-mail: j.bermudez@ieee.org).

10.1109/TSP.2008.2008261

Acoustic echos are caused by the acoustic coupling between the loudspeaker and the microphone, as for instance in hands-free telephones [1], [2].

The two main EC design problems are 1) choice of adaptation algorithm(s), and 2) control logic for adaptation. The latter design problem is caused mainly by double-talk, which happens when both the far-end and the near-end speakers talk simultaneously. The EC observes the channel input vector and the scalar error signal. The error signal can consist of both double-talk and/or the uncancelled outgoing signal due to the far-end speaker. Assume the far-end speech signal is quasi-stationary within a decision interval. Then a significant change in the EC error power suggests either a channel change or double-talk. However, the EC response should be quite different to these two events. The adaptive filter should stop adapting if double-talk is the cause. Otherwise, the near-end speech appears as a large interference and forces the adaptive filter coefficients away from the correct echo channel response. On the other hand, the adaptive algorithm should adapt immediately to track the channel variations if a channel change has occurred. Given the available observations, distinguishing between double-talk and channel change leads to a non-trivial detection problem. Reference [3] presents a good overview of several proposed adaptation control mechanisms to guide the decision making.

The vast majority of the techniques available rely on *ad hoc* statistics to make the decision. Only a few works formulate the problem using a statistical decision framework. For instance, [4] proposes a maximum *a posteriori* (MAP) decision rule using a sliding window of channel output samples as observations, and assumes Bernoulli distributed priors for the different hypotheses formulated. An ARX model is used for the echo channel output. A similar MAP detection approach is used in [5] for a Markov modulated finite impulse response channel model. Reference [6] proposes a generalized likelihood ratio test (GLRT) based on observations composed of windows of samples from both the channel input and output signals. These detection algorithms have all been proposed for the conventional adaptive EC structure [3].

An alternative EC structure has been proposed in [7], which uses a shadow adaptive filter that operates in parallel with the actual echo cancellation filter. The shadow filter coefficients are transferred to the echo cancellation filter when the shadow filter is a better estimate of the unknown channel response than the echo cancellation filter. This structure is shown in Fig. 1. The EC consists of the main echo cancellation filter $\boldsymbol{h}_1$ and the adaptive shadow filter $\boldsymbol{h}_0$. The output of the main filter is subtracted from the echo to obtain the cancelled echo $z_1(n)$. The shadow

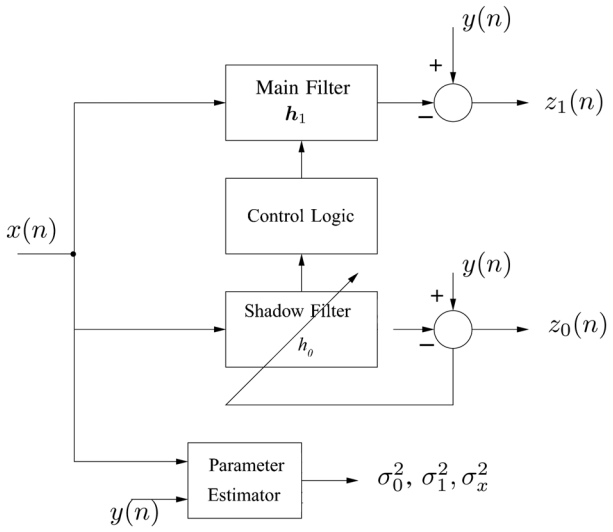


Fig. 1. Basic EC structure.

filter weights are adapted continuously. Control logic decides when copying the shadow filter coefficients to the main filter will improve the EC performance. This logic is based on measurements of the available signals.

Consider the basic behavior of the EC when the cancelled echo power changes significantly. Assume that the system is initially in steady-state so that $\boldsymbol{h}_1 = \boldsymbol{h}_0$ and the short term time-averaged error powers for the two filters are small. Suppose double-talk occurs suddenly at time $n_1$. The two error powers now become large because of the double-talk. The shadow filter (incorrectly) adapts using this large error power and no longer matches the unknown channel. No transfer should occur from the shadow filter to the main filter. On the other hand, suppose a channel change occurs at time $n_1$. The shadow filter now (correctly) adapts on this channel change. After some time, a transfer should occur from the shadow filter to the main filter. The structure in Fig. 1 has interesting potential for optimally deciding between double-talk and channel change. However, the detection scheme proposed in [7] is not based upon a decision-theoretic approach. A recent paper [8] studied the optimum likelihood ratio test (LRT) for this problem. The LRT requires *a priori* knowledge of the background noise and double-talk power levels.

A GLRT is derived in this paper for deciding between double-talk (freeze weights) and a channel change (adapt quickly) using a stationary Gaussian stochastic signal model. The GLRT is then simplified to a sufficient statistic (a function of the observables that depends upon which hypothesis is true) to obtain an optimum test statistic. The probability density function (pdf) of the test statistic under each hypothesis is obtained and the performance of the test statistic is evaluated as a function of the system parameters. This performance is represented through receiver operating characteristics (ROCs) [9, p. 38]. These curves show the probability of detection $(P_{\mathrm{D}})$ (deciding one hypothesis is true when it is actually true) vs. probability of false alarm $(P_{\mathrm{FA}})$ (deciding the same hypothesis is true when it is actually not true). The ROCs have been applied to the double-talk detection problem in [10] without considering the hypothesis of channel change. This paper derives the ROCs for the double-talk/channel-change (DTCC) problem. Finally, different double-talk detectors are compared using Monte Carlo (MC) simulations for both synthetic and real voice data and channels.

The paper is organized as follows. Section II defines a hypothesis test based on the likelihood functions for double-talk versus a channel change. This hypothesis test differs from the one proposed in [8], which involved the channel input ($N$ variates) and the channel output (one variate) for a total of $N+1$ variates. Here only two variates (the instantaneous error signals from the two filters) yield a sufficient statistic for this problem. The LRTs for the one-sample and multiple-sample cases are reinvestigated using these two variates. Finally, the GLRT is introduced for the one-sample and multiple-sample cases. Section III studies the pdf of the sufficient statistics under both hypotheses for both the LRT and GLRT. Section IV presents MC simulations of the ROCs for different sets of parameters for the LRT and GLRT for the one-sample case. The GLRT is then studied for the multiple-sample case. Section V applies the GLRT to full EC implementations for two cases: 1) a synthetically generated data model, and 2) real voice data transmitted over a real channel. Section VI presents some results and conclusions.

## II. Hypothesis Tests

The EC control logic is based upon the error signal $z_1(n)$ (canceller output) and the shadow filter error signal $z_0(n)$ as explained in [8]. Whenever the powers of the error signals increase significantly over some quiescent level, the EC must decide whether the increase is due to double-talk or to a channel change. Either occurrence will cause a significant increase in the error powers. Reference [8] introduced statistical models for these two events. These models are briefly reviewed here.

### A. Signal and Channel Models

The channel input vector $\boldsymbol{x}(n) = [x(n), \ldots, x(n-N+1)]^T$ is of dimension $N \times 1$ with covariance matrix $E[\boldsymbol{x}(n)\boldsymbol{x}^T(n)] = \boldsymbol{\Sigma}_x$ and the channel output is a scalar $y(n)$. Note that the input signal $\boldsymbol{x}(n)$ in [8] was assumed white with covariance matrix $\boldsymbol{\Sigma}_x = \sigma_x^2 \boldsymbol{I}_N$, where $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. However, the analysis presented in this paper will show that the test statistics used for the DTCC problem does not change when the input signal is correlated in time. This is an important result since speech signals are usually correlated.

For mathematical tractability, this paper assumes that the input signal is stationary within the decision periods and that the double-talk signal can be modelled by a white Gaussian process for detection purposes [6]. Also, $[y(n), \boldsymbol{x}^T(n)]^T$ is modelled as a zero-mean Gaussian vector. Consider the following two hypotheses:

$$\begin{aligned} &\mathcal{H}_0 : \text{channel change has happened,} \\ &\mathcal{H}_1 : \text{double-talk is happening.} \end{aligned} \tag{1}$$

Under $\mathcal{H}_0$,

$$y(n) = \boldsymbol{h}_0^T \boldsymbol{x}(n) + n_0(n) \tag{2}$$

where $\boldsymbol{h}_0$ is a known channel (see [8] for motivations), and the additive noise $n_0(n)$ is stationary zero-mean white[1] Gaussian, independent of $\boldsymbol{x}(n)$ with $E[n_0^2(n)] = \sigma_0^2$.

Under $\mathcal{H}_1$,

$$y(n) = \boldsymbol{h}_1^T \boldsymbol{x}(n) + n_0(n) + n_1(n) \tag{3}$$

where $\boldsymbol{h}_1$ can be assumed known (see again [8] for motivations). The second additive noise $n_1(n)$, modeling the double-talk, is zero-mean white Gaussian, and independent of both $\boldsymbol{x}(n)$ and $n_0(n)$ with $E[n_1^2(n)] = \sigma_1^2$.

### *B. The LRT Revisited*

The likelihood ratio detector for the DTCC detection problem was derived in [8]. This section shows that the LRT can be obtained in an easier way by using a different coordinate system. More precisely, two scalars (error signals) are introduced $z_0(n) = y(n) - \boldsymbol{h}_0^T \boldsymbol{x}(n)$ and $z_1(n) = y(n) - \boldsymbol{h}_1^T \boldsymbol{x}(n)$. The equivalence of the LRT derived in [8] and the LRT derived here using $\boldsymbol{z}(n) = [z_0(n), z_1(n)]^T$ can be verified as follows. Consider an $(N-1) \times N$ matrix denoted as $\boldsymbol{M}$ such that 1) there is a one-to-one transformation between $[y(n), \boldsymbol{x}^T(n)]^T$ and $[z_0(n), z_1(n), \boldsymbol{x}^T(n)\boldsymbol{M}^T]^T$, and 2) the vectors $\boldsymbol{z}(n) = [z_0(n), z_1(n)]^T$ and $\boldsymbol{M}\boldsymbol{x}(n)$ are independent[2]. Using the results of [9, p. 35] and noting that the distribution of $\boldsymbol{x}(n)$ does not change under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, the LRTs based on the observation vectors $[y(n), \boldsymbol{x}^T(n)]^T$ and $\boldsymbol{z}(n) = [z_0(n), z_1(n)]^T$ provide the same test statistics. However, this section shows that the LRT for the observation vector $\boldsymbol{z}(n)$ is easier to derive than with $[y(n), \boldsymbol{x}^T(n)]^T$. Another interesting property is that the derivations obtained with $\boldsymbol{z}(n)$ are similar for white or colored input signals $\boldsymbol{x}(n)$.

*1) One-Sample Case:* Under $\mathcal{H}_0$, the error signals $z_0(n)$ and $z_1(n)$ can be written

$$\begin{aligned} z_0(n) &= n_0(n) \\ z_1(n) &= (\boldsymbol{h}_0 - \boldsymbol{h}_1)^T \boldsymbol{x}(n) + n_0(n). \end{aligned} \tag{4}$$

whereas under $\mathcal{H}_1$

$$\begin{aligned} z_0(n) &= (\boldsymbol{h}_1 - \boldsymbol{h}_0)^T \boldsymbol{x}(n) + n_0(n) + n_1(n) \\ z_1(n) &= n_0(n) + n_1(n). \end{aligned} \tag{5}$$

The joint pdf of $\boldsymbol{z}(n)$ is Gaussian under both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ such that

$$p[\boldsymbol{z}(n)|\mathcal{H}_i] \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{i1}), \quad i = 0, 1 \tag{6}$$

where the second subscript in $\boldsymbol{\Sigma}_{i1}$ (1 in this case) indicates the 1-sample case. The covariance matrices under $\mathcal{H}_0$ and $\mathcal{H}_1$ for the $p$-sample case will be denoted, respectively, $\boldsymbol{\Sigma}_{0p}$ and $\boldsymbol{\Sigma}_{1p}$. In (6)

$$\boldsymbol{\Sigma}_{01} = \begin{pmatrix} \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 + c_x^2 \end{pmatrix} \tag{7}$$

$$\boldsymbol{\Sigma}_{11} = \begin{pmatrix} \sigma_0^2 + \sigma_1^2 + c_x^2 & \sigma_0^2 + \sigma_1^2 \\ \sigma_0^2 + \sigma_1^2 & \sigma_0^2 + \sigma_1^2 \end{pmatrix} \tag{8}$$

where

$$c_x^2 = (\boldsymbol{h}_0 - \boldsymbol{h}_1)^T \boldsymbol{\Sigma}_x (\boldsymbol{h}_0 - \boldsymbol{h}_1). \tag{9}$$

The LRT for (1) can then be expressed as

$$\frac{\frac{1}{\sqrt{|\boldsymbol{\Sigma}_{11}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{z}(n)\right]}{\frac{1}{\sqrt{|\boldsymbol{\Sigma}_{01}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T \boldsymbol{\Sigma}_{01}^{-1} \boldsymbol{z}(n)\right]} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_1 \tag{10}$$

or equivalently

$$\boldsymbol{z}(n)^T \left(\boldsymbol{\Sigma}_{01}^{-1} - \boldsymbol{\Sigma}_{11}^{-1}\right) \boldsymbol{z}(n) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_2 \tag{11}$$

where $T_1$ and $T_2$ are threshold settings determined by the probability of detection $P_{\mathrm{D}}$ and the probability of false alarm $P_{\mathrm{FA}}$. The 2×2 matrices $\boldsymbol{\Sigma}_{01}$ and $\boldsymbol{\Sigma}_{11}$ can be inverted easily, yielding the equivalent detection strategy

$$T_{\mathrm{LRT}}(n) = \frac{z_0^2(n)}{\sigma_0^2} - \frac{z_1^2(n)}{\sigma_0^2 + \sigma_1^2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_2. \tag{12}$$

Straightforward computations show that the test statistics $T_{\mathrm{LRT}}(n)$ is equivalent to the one given in [8, Eq. (13)], up to a multiplicative constant[3]. However, the previous analysis applies for white and colored input signals $\boldsymbol{x}(n)$ since the covariance matrix of $\boldsymbol{x}(n)$ only affects $c_x^2$ which does not appear in the test statistics $T_{\mathrm{LRT}}(n)$.

*2) The LRT for Multiple Samples:* The analysis above can be generalized to the case where multiple time samples $\boldsymbol{z}(j)$, for $j = n-p+1, \ldots, n$, are available. This result is interesting and differs from what was obtained when using the vectors $[y(j), \boldsymbol{x}^T(j)]^T$, for $j = n-p+1, \ldots, n$ (see discussion in [8, Sec. V]). The analysis is performed here for two samples (i.e., $p = 2$) for simplicity and is generalized later. When two samples are observed, the error signals $z_0(n), z_0(n-1)$ and $z_1(n), z_1(n-1)$ under $\mathcal{H}_0$ can be written

$$\begin{aligned} z_0(n) &= n_0(n), z_0(n-1) = n_0(n-1) \\ z_1(n) &= (\boldsymbol{h}_0 - \boldsymbol{h}_1)^T \boldsymbol{x}(n) + n_0(n) \\ z_1(n-1) &= (\boldsymbol{h}_0 - \boldsymbol{h}_1)^T \boldsymbol{x}(n-1) + n_0(n-1). \end{aligned} \tag{13}$$

[1]Note here that the whiteness assumption for $n_0(n)$ is not restrictive since it is always possible to whiten the channel outputs by premultiplying consecutive samples by an appropriate matrix. Of course, this operation assumes that the covariance matrix of consecutive noise samples is known or can be estimated.

[2]We can proceed as follows to build a matrix $\boldsymbol{M}$ such that the vectors $\boldsymbol{z}(n) = [z_0(n), z_1(n)]$ and $\boldsymbol{M}\boldsymbol{x}(n)$ are independent Gaussian. Since $\boldsymbol{x}(n)$, $n_0(n)$ and $n_1(n)$ are independent vectors, it is interesting to note that the independence between $\boldsymbol{z}(n)$ and $\boldsymbol{M}\boldsymbol{x}(n)$ is ensured if $\boldsymbol{M}\boldsymbol{x}(n)$ and $(\boldsymbol{h}_0 - \boldsymbol{h}_1)^T \boldsymbol{x}(n)$ are also independent Gaussian. This last property is satisfied provided $\boldsymbol{M}\boldsymbol{x}(n)$ is orthogonal to $\boldsymbol{h}_0^T \boldsymbol{x}(n)$ and $\boldsymbol{h}_1^T \boldsymbol{x}(n)$. Thus the matrix $\boldsymbol{M}$ can be found provided the covariance matrix of $\boldsymbol{x}(n)$ is full-rank and $\boldsymbol{h}_0 \neq \boldsymbol{h}_1$.

[3]In proceeding from [8, Eq. (13) to (14)], the terms in [8, Eq. (13)], that did not change under either hypothesis were ignored for mathematical tractability. However, these terms are not statistically independent of the remaining terms. Hence, [8, Eq. (14)], is only an approximation to the LRT. Equation (12) above results if these ignored terms are included. Hence, we will recompute the ROC curves for the LRT later in this paper.

whereas under $\mathcal{H}_1$

$$\begin{aligned} z_1(n) &= n_0(n) + n_1(n) \\ z_1(n-1) &= n_0(n-1) + n_1(n-1) \\ z_0(n) &= (\boldsymbol{h}_1 - \boldsymbol{h}_0)^T \boldsymbol{x}(n) + n_0(n) + n_1(n) \\ z_0(n-1) &= (\boldsymbol{h}_1 - \boldsymbol{h}_0)^T \boldsymbol{x}(n-1) \\ &\quad + n_0(n-1) + n_1(n-1). \end{aligned} \tag{14}$$

Thus, $\boldsymbol{z}(n) = [z_0(n), z_0(n-1), z_1(n), z_1(n-1)]^T$ is a zero-mean Gaussian vector under both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. Straightforward computations yield the covariance matrices of $\boldsymbol{z}(n)$ under $\mathcal{H}_0$ and $\mathcal{H}_1$. These matrices, denoted respectively as $\boldsymbol{\Sigma}_{02}$ and $\boldsymbol{\Sigma}_{12}$ can be expressed as

$$\boldsymbol{\Sigma}_{02} = \begin{pmatrix} \sigma_0^2 \boldsymbol{I}_2 & \sigma_0^2 \boldsymbol{I}_2 \\ \sigma_0^2 \boldsymbol{I}_2 & \sigma_0^2 \boldsymbol{I}_2 + \boldsymbol{H}_x \end{pmatrix} \tag{15}$$

and

$$\boldsymbol{\Sigma}_{12} = \begin{pmatrix} (\sigma_0^2 + \sigma_1^2)\boldsymbol{I}_2 + \boldsymbol{H}_x & (\sigma_0^2 + \sigma_1^2)\boldsymbol{I}_2 \\ (\sigma_0^2 + \sigma_1^2)\boldsymbol{I}_2 & (\sigma_0^2 + \sigma_1^2)\boldsymbol{I}_2 \end{pmatrix} \tag{16}$$

where

$$\boldsymbol{H}_x = \boldsymbol{D}^T \begin{pmatrix} \boldsymbol{\Sigma}_x & \boldsymbol{R}_x \\ \boldsymbol{R}_x & \boldsymbol{\Sigma}_x \end{pmatrix} \boldsymbol{D} \tag{17}$$

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{h}_0 - \boldsymbol{h}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{h}_0 - \boldsymbol{h}_1 \end{pmatrix} \tag{18}$$

and where $\boldsymbol{R}_x = E[\boldsymbol{x}(n)\boldsymbol{x}^T(n-1)]$ (recall that $\boldsymbol{\Sigma}_x = E[\boldsymbol{x}(n)\boldsymbol{x}^T(n)]$). The determinants and inverses of these block matrices can be computed following [11, p. 572]:

$$|\boldsymbol{\Sigma}_{02}| = \sigma_0^4 |\boldsymbol{H}_x|, \; |\boldsymbol{\Sigma}_{12}| = (\sigma_0^2 + \sigma_1^2)^2 |\boldsymbol{H}_x| \tag{19}$$

and

$$\boldsymbol{\Sigma}_{02}^{-1} = \begin{pmatrix} \frac{1}{\sigma_0^2}\boldsymbol{I}_2 + \boldsymbol{H}_x^{-1} & -\boldsymbol{H}_x^{-1} \\ -\boldsymbol{H}_x^{-1} & \boldsymbol{H}_x^{-1} \end{pmatrix} \tag{20}$$

$$\boldsymbol{\Sigma}_{12}^{-1} = \begin{pmatrix} \boldsymbol{H}_x^{-1} & -\boldsymbol{H}_x^{-1} \\ -\boldsymbol{H}_x^{-1} & \frac{1}{\sigma_0^2 + \sigma_1^2}\boldsymbol{I}_2 + \boldsymbol{H}_x^{-1} \end{pmatrix}. \tag{21}$$

Thus the log-LRT for the 2-sample case rejects hypothesis $\mathcal{H}_0$ if

$$\boldsymbol{z}(n)^T \left( \boldsymbol{\Sigma}_{02}^{-1} - \boldsymbol{\Sigma}_{12}^{-1} \right) \boldsymbol{z}(n) > T_2 \tag{22}$$

where $T_2$ is a threshold depending on $P_{\mathrm{D}}$ and/or $P_{\mathrm{FA}}$. The difference between the two inverse matrices $\boldsymbol{\Sigma}_{02}^{-1}$ and $\boldsymbol{\Sigma}_{12}^{-1}$ can be computed easily using the previous results

$$\boldsymbol{\Sigma}_{02}^{-1} - \boldsymbol{\Sigma}_{12}^{-1} = \begin{pmatrix} \frac{1}{\sigma_0^2}\boldsymbol{I}_2 & 0 \\ 0 & -\frac{1}{\sigma_0^2 + \sigma_1^2}\boldsymbol{I}_2 \end{pmatrix} \tag{23}$$

yielding the following DTCC detection strategy

$$T_{\mathrm{LRT}}(n) = \frac{1}{\sigma_0^2}\|\boldsymbol{z}_0(n)\|^2 - \frac{1}{\sigma_0^2 + \sigma_1^2}\|\boldsymbol{z}_1(n)\|^2 \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_2 \tag{24}$$

where $\|\boldsymbol{z}_i(n)\|^2 = z_i^2(n) + z_i^2(n-1)$ and $T_2$ is the threshold setting. This result can be compared with (12) obtained for the one-sample case. The generalization to more than two samples is straightforward. Indeed, in the $p$-sample case, the matrices $\boldsymbol{\Sigma}_{0p}$ and $\boldsymbol{\Sigma}_{1p}$ are defined as in (15)–(18), except $\boldsymbol{I}_2$ has to be replaced by $\boldsymbol{I}_p$, and $\boldsymbol{H}_x$ is defined differently. However, since $\boldsymbol{H}_x$ cancels from the difference between the two inverses, the LRT for the $p$-sample case is expressed as (24) where $\|\boldsymbol{z}_i(n)\|^2 = \boldsymbol{z}_i(n)^T\boldsymbol{z}_i(n) = \sum_{k=0}^{p-1} z_i^2(n-k)$ is the squared norm of $\boldsymbol{z}_i(n)$. This last result is a proof that the postdetection scheme proposed in [8] is the optimum Neyman-Pearson detector for the $p$-sample case.

### C. Generalized Likelihood Ratio Test

The LRT approach in [8] requires that $\sigma_0^2$, $\sigma_1^2$, and $\boldsymbol{\Sigma}_x$ be known *a priori*. Clearly, this is not the case for echo cancellation as the input signals are nonstationary in general. Thus, these parameters must be estimated from the observed data. The accuracy of these estimates can vary widely as the statistics of the input signals vary over time. Thus, the LRT detector will suffer some degradation in performance. Hence, suboptimal detectors estimating these parameters could, in a practical situation, perform as well as the LRT detector.

When the covariance matrices of the noises $n_0(n)$ and $n_1(n)$ and of the input signal $\boldsymbol{x}(n)$ are unknown, a classical alternative to the Neyman–Pearson detector is the generalized likelihood ratio (GLR) detector. The GLR detector replaces the unknown parameters in the LR by their maximum likelihood (ML) estimates. This section derives the GLR detector for the DTCC problem. As before, the derivations are first conducted for the one-sample case and then generalized to the case of multiple samples.

*1) The GLRT for the One-Sample Case:* The notation $\boldsymbol{z}(n) = [z_0(n), z_1(n)]^T$ is used in the one-sample case. It is first assumed that $c_x^2$ is known (or equivalently that $\boldsymbol{\Sigma}_x$ is known). The GLRT for the DTCC problem rejects hypothesis $\mathcal{H}_0$ if

$$\frac{\sup\limits_{\sigma_0^2, \sigma_1^2} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{11}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{z}(n)\right]}{\sup\limits_{\sigma_0^2} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{01}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T \boldsymbol{\Sigma}_{01}^{-1} \boldsymbol{z}(n)\right]} > T_1 \tag{25}$$

or equivalently,

$$\frac{\sup\limits_{\sigma_0^2, \sigma_1^2} \frac{1}{|c_x|\sqrt{\sigma_0^2 + \sigma_1^2}} \exp\left\{-\frac{[z_0(n) - z_1(n)]^2}{2c_x^2} - \frac{z_1^2(n)}{2(\sigma_0^2 + \sigma_1^2)}\right\}}{\sup\limits_{\sigma_0^2} \frac{1}{|c_x|\sqrt{\sigma_0^2}} \exp\left\{-\frac{[z_0(n) - z_1(n)]^2}{2c_x^2} - \frac{z_0^2(n)}{2\sigma_0^2}\right\}} > T_1. \tag{26}$$

Differentiation of the denominator, with respect to $\sigma_0^2$, yields its supremum for

$$\sigma_0^2 = \frac{1}{2} z_0^2(n). \tag{27}$$

Similarly, the supremum of the numerator of (26) is obtained for

$$\sigma_0^2 + \sigma_1^2 = \frac{1}{2} z_1^2(n). \tag{28}$$

Thus, for a known $c_x^2$, the GLRT for the DTCC problem is

$$T_{\mathrm{GLRT}}(n) = \frac{z_0^2(n)}{z_1^2(n)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_2. \tag{29}$$

The GLRT for the DTCC problem is unchanged when $c_x^2$ is unknown. For this case, the suprema in (25) have to be computed with respect to $(\sigma_0^2, \sigma_1^2, c_x^2)$ and $(\sigma_0^2, c_x^2)$ instead of $(\sigma_0^2, \sigma_1^2)$ and $\sigma_0^2$. However, it is easy to see that the suprema with respect to $c_x^2$ appearing in the numerator and denominator of (25) cancel.

*2) Multiple Samples:* Considering initially the 2-sample case, denote $\boldsymbol{z}_0 = [z_0(n), z_0(n-1)]^T$ and $\boldsymbol{z}_1 = [z_1(n), z_1(n-1)]^T$. Assume first that the matrix $\boldsymbol{H}_x$, defined in (18), is known (or equivalently that $\boldsymbol{\Sigma}_x$ and $\boldsymbol{R}_x$ are known). The GLR detector for the DTCC problem rejects hypothesis $\mathcal{H}_0$ if

$$\frac{\sup\limits_{\sigma_0^2,\sigma_1^2} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{12}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T\boldsymbol{\Sigma}_{12}^{-1}\boldsymbol{z}(n)\right]}{\sup\limits_{\sigma_0^2} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{02}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T\boldsymbol{\Sigma}_{02}^{-1}\boldsymbol{z}(n)\right]} > T_1 \tag{30}$$

or equivalently

$$\frac{\sup\limits_{\sigma_0^2,\sigma_1^2} \frac{1}{\sqrt{\sigma_0^2+\sigma_1^2}} \exp\left[-\frac{1}{2}\frac{\boldsymbol{z}_1^T\boldsymbol{z}_1}{\sigma_0^2+\sigma_1^2}\right]}{\sup\limits_{\sigma_0^2} \frac{1}{\sqrt{\sigma_0^2}} \exp\left[-\frac{1}{2}\frac{\boldsymbol{z}_0^T\boldsymbol{z}_0}{\sigma_0^2}\right]} > T_1. \tag{31}$$

Differentiation of the denominator with respect to $\sigma_0^2$ yields its supremum for

$$\sigma_0^2 = \frac{1}{2}\boldsymbol{z}_0^T\boldsymbol{z}_0 = \frac{1}{2}\|\boldsymbol{z}_0(n)\|^2. \tag{32}$$

Similarly, the supremum of the numerator of (31) is obtained for

$$\sigma_0^2 + \sigma_1^2 = \frac{1}{2}\boldsymbol{z}_1^T\boldsymbol{z}_1 = \frac{1}{2}\|\boldsymbol{z}_1(n)\|^2. \tag{33}$$

Thus, the GLRT with a known matrix $\boldsymbol{H}_x$ for the DTCC problem reduces to

$$T_{\mathrm{GLRT}}(n) = \frac{\|\boldsymbol{z}_0(n)\|^2}{\|\boldsymbol{z}_1(n)\|^2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_2. \tag{34}$$

The GLRT for the DTCC problem is unchanged when $\boldsymbol{H}_x$ is unknown. Indeed

$$\begin{aligned}\frac{\sup\limits_{\sigma_0^2,\sigma_1^2,\boldsymbol{H}_x} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{12}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T\boldsymbol{\Sigma}_{12}^{-1}\boldsymbol{z}(n)\right]}{\sup\limits_{\sigma_0^2,\boldsymbol{H}_x} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{02}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T\boldsymbol{\Sigma}_{02}^{-1}\boldsymbol{z}(n)\right]} \\ = \frac{\sup\limits_{\sigma_0^2,\sigma_1^2} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{12}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T\boldsymbol{\Sigma}_{12}^{-1}\boldsymbol{z}(n)\right]}{\sup\limits_{\sigma_0^2} \frac{1}{\sqrt{|\boldsymbol{\Sigma}_{02}|}} \exp\left[-\frac{1}{2}\boldsymbol{z}(n)^T\boldsymbol{\Sigma}_{02}^{-1}\boldsymbol{z}(n)\right]}.\end{aligned} \tag{35}$$

Thus, the terms depending on $\boldsymbol{H}_x$ in the GLR cancel, which leads to a test statistic equal to that obtained for a known $\boldsymbol{H}_x$. This result indicates that the GLR detectors for the DTCC problem are the same for known and unknown $\boldsymbol{H}_x$, i.e., for known and unknown covariance and correlation matrices $\boldsymbol{\Sigma}_x$ and $\boldsymbol{R}_x$. The generalization to $p$ samples is straightforward and yields the following DTCC detection strategy

$$T_{\mathrm{GLRT}}(n) = \frac{\sum_{i=0}^{p-1} z_0^2(n-i)}{\sum_{i=0}^{p-1} z_1^2(n-i)} = \frac{\|\boldsymbol{z}_0(n)\|^2}{\|\boldsymbol{z}_1(n)\|^2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} T_2. \tag{36}$$

## III. PDFs of the Sufficient Statistics

This section studies the pdfs of the LRT and GLRT sufficient statistics under the two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. Knowledge of these pdfs is important for obtaining the test threshold from the $P_{\mathrm{FA}}$ and for plotting the ROCs for the DTCC problem.

### A. One-Sample LRT

The sufficient statistic $T_{\mathrm{LRT}}(n)$ for the one-sample case ($p = 1$) is defined in (12). $T_{\mathrm{LRT}}(n)$ is the difference between two univariate squared correlated Gaussian random variables, i.e., the difference between two correlated chi-square random variables. The pdf of $T_{\mathrm{LRT}}(n)$ under hypothesis $\mathcal{H}_0$ can then be expressed as follows (see [12, p. 76]):

$$p_0^{\mathrm{LRT}}(t) = \frac{\sqrt{|\boldsymbol{W}_0|}}{2\pi} \exp\left(-\frac{d_w t}{4}\right) K_0\left(\frac{|t|\gamma_0}{4}\right) \tag{37}$$

where $d_w = w_0(1,1) - w_0(2,2)$, $K_0(\cdot)$ is the modified Bessel function of the second kind, $\boldsymbol{W}_0 = [w_0(i,j)]$ is the inverse covariance matrix of the vector $\left[z_0(n)/\sigma_0, z_1(n)/\sqrt{\sigma_0^2+\sigma_1^2}\right]$ under hypothesis $\mathcal{H}_0$, and

$$\gamma_0 = \left\{[w_0(1,1) + w_0(2,2)]^2 - 4w_0^2(1,2)\right\}^{1/2}. \tag{38}$$

Standard calculations lead to

$$\boldsymbol{W}_0 = \frac{1}{\theta_0^2(1-\rho_0^2)} \begin{pmatrix} \theta_0^2 & -\rho_0\theta_0 \\ -\rho_0\theta_0 & 1 \end{pmatrix} \tag{39}$$

with

$$\theta_0^2 = \frac{\sigma_0^2 + c_x^2}{\sigma_0^2 + \sigma_1^2}, \quad \rho_0 = \left[1 + \frac{c_x^2}{\sigma_0^2}\right]^{-1/2}. \tag{40}$$

A similar result is obtained for the pdf of $T_{\mathrm{LRT}}(n)$ under hypothesis $\mathcal{H}_1$. In (34) replace $\boldsymbol{W}_0$ with $\boldsymbol{W}_1$ defined as

$$\boldsymbol{W}_1 = \frac{1}{\theta_1^2(1-\rho_1^2)} \begin{pmatrix} 1 & -\rho_1\theta_1 \\ -\rho_1\theta_1 & \theta_1^2 \end{pmatrix} \tag{41}$$

with

$$\theta_1^2 = \frac{\sigma_0^2 + \sigma_1^2 + c_x^2}{\sigma_0^2}, \quad \rho_1 = \left[1 + \frac{c_x^2}{\sigma_0^2 + \sigma_1^2}\right]^{-1/2}. \tag{42}$$

### B. Multiple-Sample LRT

The LRT test statistics for the $p$-sample case can be expressed as functions of the two vectors $\boldsymbol{z}_0(n) = [z_0(n), z_0(n-1), \ldots, z_0(n-p+1)]^T$ and $\boldsymbol{z}_1(n) = [z_1(n), z_1(n-1), \ldots, z_1(n-p+1)]^T$ as follows

$$T_{\mathrm{LRT}}(n) = \frac{1}{\sigma_0^2}\|\boldsymbol{z}_0(n)\|^2 - \frac{1}{\sigma_0^2+\sigma_1^2}\|\boldsymbol{z}_1(n)\|^2. \tag{43}$$

The distributions of $\boldsymbol{z}_0(n)/\sigma_0$ and $\boldsymbol{z}_1(n)/\sqrt{\sigma_0^2+\sigma_1^2}$ can be determined under both hypotheses. More precisely:

- Under $\mathcal{H}_0$

$$\frac{\boldsymbol{z}_0(n)}{\sigma_0} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p) \tag{44}$$

$$\frac{\boldsymbol{z}_1(n)}{\sqrt{\sigma_0^2+\sigma_1^2}} \sim \mathcal{N}\left(\boldsymbol{0}, \frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}\boldsymbol{I}_p + \frac{1}{\sigma_0^2+\sigma_1^2}\boldsymbol{H}\boldsymbol{\Sigma}_{\widetilde{\boldsymbol{x}}}\boldsymbol{H}^T\right) \tag{45}$$

  where $\boldsymbol{\Sigma}_{\widetilde{\boldsymbol{x}}}$ is the covariance matrix of the vector containing all input samples (i.e., the covariance matrix of $\widetilde{\boldsymbol{x}} = [x(n), x(n-1), \ldots, x(n-p-N+2)]^T$ in the $p$-sample case), and $\boldsymbol{H}$ is a $p \times (p+N-1)$ matrix defined in (46) (see the bottom of the page) where $\boldsymbol{h} = \boldsymbol{h}_0 - \boldsymbol{h}_1$. These results have been obtained by noting that $\boldsymbol{z}_0(n) = \boldsymbol{n}_0(n)$ and $\boldsymbol{z}_1(n) = \boldsymbol{H}\widetilde{\boldsymbol{x}} + \boldsymbol{n}_0(n)$ under hypothesis $\mathcal{H}_0$, where $\boldsymbol{n}_0(n) = [n_0(n), \ldots, n_0(n-p+1)]^T$.
- Under $\mathcal{H}_1$
  Noting that $\boldsymbol{z}_0(n) = \boldsymbol{n}_0(n) + \boldsymbol{n}_1(n) + \boldsymbol{H}\widetilde{\boldsymbol{x}}$ and $\boldsymbol{z}_1(n) = \boldsymbol{n}_0(n) + \boldsymbol{n}_1(n)$ under hypothesis $\mathcal{H}_1$ (with $\boldsymbol{n}_1(n) = [n_1(n), \ldots, n_1(n-p+1)]^T$), the following results are obtained

$$\frac{\boldsymbol{z}_0(n)}{\sigma_0} \sim \mathcal{N}\left(\boldsymbol{0}, \frac{\sigma_0^2+\sigma_1^2}{\sigma_0^2}\boldsymbol{I}_p + \frac{1}{\sigma_0^2}\boldsymbol{H}\boldsymbol{\Sigma}_{\widetilde{\boldsymbol{x}}}\boldsymbol{H}^T\right) \tag{47}$$

$$\frac{\boldsymbol{z}_1(n)}{\sqrt{\sigma_0^2+\sigma_1^2}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p). \tag{48}$$

  The distribution of $\|\boldsymbol{z}_0(n)\|^2/\sigma_0^2$ under hypothesis $\mathcal{H}_0$ (resp. $\|\boldsymbol{z}_1(n)\|^2/(\sigma_0^2+\sigma_1^2)$ under hypothesis $\mathcal{H}_1$) is a central chi-square distribution with $p$ degrees of freedom. However, the distributions of $\|\boldsymbol{z}_1(n)\|^2/(\sigma_0^2+\sigma_1^2)$ under hypothesis $\mathcal{H}_0$ and $\|\boldsymbol{z}_0(n)\|^2/\sigma_0^2$ under hypothesis $\mathcal{H}_1$ cannot be expressed in closed form since they are the distributions of quadratic forms of Gaussian vectors. An alternative is to approximate the distributions of $\|\boldsymbol{z}_0(n)\|^2$ and $\|\boldsymbol{z}_1(n)\|^2$ in (24) by Gaussian distributions using the central limit theorem. However, several experiments have shown that these approximations are valid only for very large values of $p$ (the Gaussian approximations are not in good agreement with the histograms of $T_{\mathrm{LRT}}$, even for $p = 500$). This behavior prevents using these approximations.

### C. One-Sample GLRT

The sufficient statistics for the GLRT can be expressed as follows:

$$T_{\mathrm{GLRT}}(n) = \frac{\|\boldsymbol{z}_0(n)\|^2}{\|\boldsymbol{z}_1(n)\|^2} \tag{49}$$

where $\boldsymbol{z}_0(n) = z_0(n)$ and $\boldsymbol{z}_1(n) = z_1(n)$ in the one-sample case. The joint distribution of $\boldsymbol{z}(n) = [z_0(n), z_1(n)]^T$ is a zero-mean Gaussian distribution under the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ with the following inverse covariance matrices

$$\boldsymbol{V}_0 = \frac{1}{\theta_0^2(1-\rho_0^2)}\begin{pmatrix} \frac{1}{\sigma_0^2+\sigma_1^2} & -\frac{\rho_0\theta_0}{\sigma_0\sqrt{\sigma_0^2+\sigma_1^2}} \\ -\frac{\rho_0\theta_0}{\sigma_0\sqrt{\sigma_0^2+\sigma_1^2}} & \frac{\theta_0^2}{\sigma_0^2} \end{pmatrix} \tag{50}$$

and

$$\boldsymbol{V}_1 = \frac{1}{\theta_1^2(1-\rho_1^2)}\begin{pmatrix} \frac{\theta_1^2}{\sigma_0^2+\sigma_1^2} & -\frac{\rho_1\theta_1}{\sigma_0\sqrt{\sigma_0^2+\sigma_1^2}} \\ -\frac{\rho_1\theta_1}{\sigma_0\sqrt{\sigma_0^2+\sigma_1^2}} & \frac{1}{\sigma_0^2} \end{pmatrix}. \tag{51}$$

The distribution of $Z_{\mathrm{GLRT}}(n) = \sqrt{T_{\mathrm{GLRT}}(n)}$ under hypothesis $\mathcal{H}_i$ (with $i = 0, 1$) can then be determined (see [12, p. 54]):

$$p_i^{\mathrm{GLRT}}(z) = \frac{2\sqrt{|\boldsymbol{V}_i|}}{\pi}\left[\frac{\eta_i(z)}{\eta_i^2(z) - 4v_i^2(1,2)z^2}\right]\mathcal{I}_{\mathbb{R}^+}(z) \tag{52}$$

with $\eta_i(z) = v_i(1,1) + v_i(2,2)z^2$, where $\mathcal{I}_{\mathbb{R}^+}(z)$ is the indicator function defined on $\mathbb{R}^+$.

### D. Multiple-Sample GLRT

Determination of the exact distribution of $T_{\mathrm{GLRT}}(n)$ or $Z_{\mathrm{GLRT}}(n)$ in the general $p$-sample case is complicated. $T_{\mathrm{GLRT}}(n)$ is the ratio of two correlated random variables. The distribution of one of these two random variables is a central chi-square distribution. However, the other random variable is a quadratic form of correlated Gaussian variables (correlated with the first random variable) whose pdf does not have a simple closed-form expression. Note again that approximating the distributions of $\|\boldsymbol{z}_0(n)\|^2$ and $\|\boldsymbol{z}_1(n)\|^2$ in (36) as Gaussians (using the central limit theorem) cannot be used for reasonable values of $p$. Indeed, we have verified that

$$\boldsymbol{H} = \begin{pmatrix} h(1) & h(2) & \ldots & h(N) & 0 & 0 & \ldots & 0 \\ 0 & h(1) & \ldots & h(N-1) & h(N) & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & h(1) & h(2) & \ldots & h(N) & 0 \\ 0 & \ldots & \ldots & 0 & h(1) & \ldots & h(N-1) & h(N) \end{pmatrix} \tag{46}$$

the Gaussian approximations are not in good agreement with the histograms of $T_{\mathrm{GLRT}}$, even for $p = 500$.

## IV. Monte Carlo Simulations

The first set of simulations compares the performances of the LRT and GLRT for the DTCC problem. The channels $\boldsymbol{h}_0$ and $\boldsymbol{h}_1$ are the two one-sided exponential channels studied in [8], i.e.

$$\boldsymbol{h}_i(j) = \begin{cases} k(0.95)^{j-\Delta_i} & j \geq \Delta_i \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

where $\Delta_i$ is the relative delay of the channel $\boldsymbol{h}_i$ and the parameter $k$ is defined by the filter gain $G = \boldsymbol{h}_1^T\boldsymbol{h}_1 = \boldsymbol{h}_0^T\boldsymbol{h}_0$. Two different scenarios are studied here corresponding to $G = -10\,\mathrm{dB}$ (electrical application) and $G = 6\ \mathrm{dB}$ (acoustic application). The performances of the two detectors are determined by the ROCs [9, p. 12]. The ROCs show $P_{\mathrm{D}}$ versus $P_{\mathrm{FA}}$. These two conditional probabilities are defined as follows [9, p. 34]:

$$\begin{aligned} P_{\mathrm{D}} &= P\left[\text{accepting } \mathcal{H}_1 \,|\mathcal{H}_1 \text{ is true}\right] \\ &= \int_{T_2}^{\infty} p_1(u)du \end{aligned} \tag{54}$$

$$\begin{aligned} P_{\mathrm{FA}} &= P\left[\text{accepting } \mathcal{H}_1 \,|\mathcal{H}_0 \text{ is true}\right] \\ &= \int_{T_2}^{\infty} p_0(u)du \end{aligned} \tag{55}$$

where $p_0(u)$ and $p_1(u)$ are the pdfs of the sufficient statistic under the two hypotheses, i.e., $p_i^{\mathrm{LRT}}(u)$ or $p_i^{\mathrm{GLRT}}(u)$ depending on the detector considered, and $T_2$ is the decision threshold.

### A. GLRT ROCs in the One-sample Case

The relation between $P_{\mathrm{FA}}$ and the threshold $T_2$, given in (55), can be inverted for the GLRT, yielding a closed-form expression for the threshold as a function of $P_{\mathrm{FA}}$. More precisely

$$P_{\mathrm{FA}} = 1 - P\left[\frac{\|\boldsymbol{z}_0(n)\|}{\|\boldsymbol{z}_1(n)\|} < Z_2\right] \tag{56}$$

where $Z_2 = \sqrt{T_2}$. Thus

$$P_{\mathrm{FA}} = 1 - P\left[\frac{\sqrt{\sigma_0^2 + c_x^2}}{\sigma_0}\frac{\|\boldsymbol{z}_0(n)\|}{\|\boldsymbol{z}_1(n)\|} < \frac{\sqrt{\sigma_0^2 + c_x^2}}{\sigma_0}Z_2\right] \tag{57}$$

where the ratio in the right hand side of (57) is the ratio of the norms of two correlated Gaussian variables with unity variances and correlation coefficient $\rho_0$ under hypothesis $\mathcal{H}_0$. Using the results of [12, p. 10],

$$P_{\mathrm{FA}} = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{h_0}{\sqrt{1-\rho_0^2}}\right) \tag{58}$$

with

$$h_0 = \frac{1-u_0^2}{2u_0}, \quad u_0 = \frac{\sqrt{\sigma_0^2 + c_x^2}}{\sigma_0}Z_2 = \frac{1}{\rho_0}Z_2. \tag{59}$$

Equation (58) and (59) lead to a quadratic equation in $Z_2$. The quadratic can be solved to yield the threshold as a function of $P_{\mathrm{FA}}$ and $\rho_0$. The probability of detection defined in (54) can be computed analytically using similar calculations:

$$P_{\mathrm{D}} = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{h_1}{\sqrt{1-\rho_1^2}}\right) \tag{60}$$

with

$$h_1 = \frac{1-u_1^2}{2u_1}, \quad u_1 = \frac{\sqrt{\sigma_0^2 + \sigma_1^2}}{\sqrt{\sigma_0^2 + \sigma_1^2 + c_x^2}}Z_2 = \rho_1 Z_2. \tag{61}$$

No numerical problems were encountered for computing the ROCs using this procedure (as opposed to the LRT of [8]).

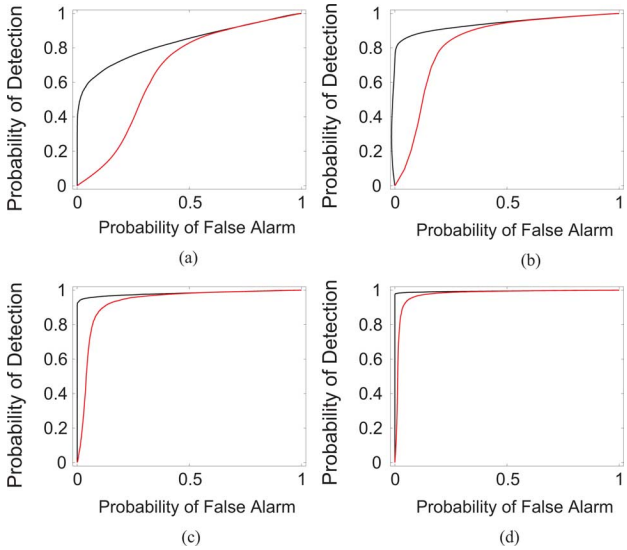


Fig. 2. $P_{\mathrm{D}}$ versus $P_{\mathrm{FA}}$ for the LRT (top curves—black) and GLRT (bottom curves—red). One-sample case. (a) $\sigma_0^2 = 0.1$. (b) $\sigma_0^2 = 0.01$. (c) $\sigma_0^2 = 0.001$. (d) $\sigma_0^2 = 0.0001$.

Fig. 2 shows the well known loss of performance of the GLRT with respect to the LRT for a white input signal with $\sigma_x^2 = 1$ and double-talk variance $\sigma_1^2 = 1$. The filter gain is $G = -10\,\mathrm{dB}$ and corresponds to an electrical application. The two channels $\boldsymbol{h}_0$ and $\boldsymbol{h}_1$ are orthogonal ($\boldsymbol{h}_0^T\boldsymbol{h}_1 = 0$) with lengths $N = 1024$. At a $P_{\mathrm{FA}} = 0.1$ (an acceptable false alarm probability) the GLRT performs very poorly compared to the LRT for $\sigma_0^2 = 0.1$ and $\sigma_0^2 = 0.01$ but is satisfactory for smaller noise powers. This loss of performance is explained by the lack of knowledge of the parameters $\sigma_0^2$, $\sigma_1^2$, and $c_x^2$ for the GLRT. The GLRT ROCs are depicted in Figs. 3 and 4 for different values of $\sigma_0^2$ and $\sigma_1^2$. Fig. 3 supports the previous conclusion that the GLRT does not perform well for $\sigma_0^2 = 0.1$ and $\sigma_0^2 = 0.01$ (without comparison to the LRT). Fig. 4 displays the same crossover behavior of the ROCs as [8, Fig. 4] (discussed in [8]).

### B. GLRT ROCs for Multiple Samples

The next simulations study the ROCs for the $p$-sample case. As explained before, tractable expressions for the distributions of the sufficient test statistics (36) under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are too complicated to derive. Again, an alternative is to approximate the distributions of the numerators and denominators of

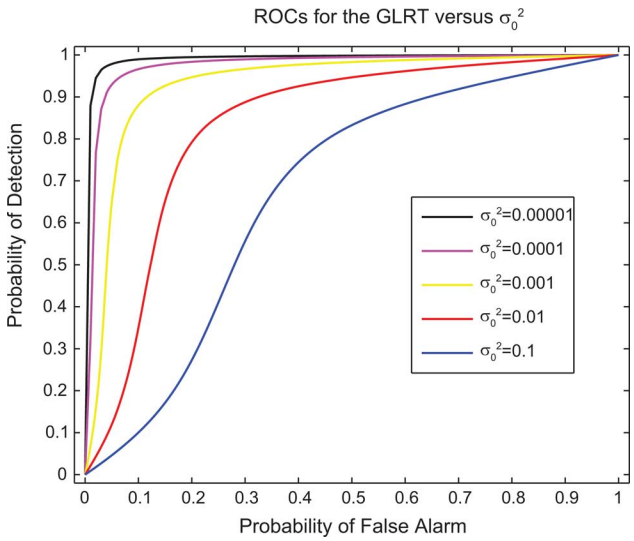


Fig. 3. GLRT ROCs for different values of $\sigma_0^2$ with $\sigma_x^2 = 1, \sigma_1^2 = 1, N = 1024$, and orthogonal channels. One-sample case.

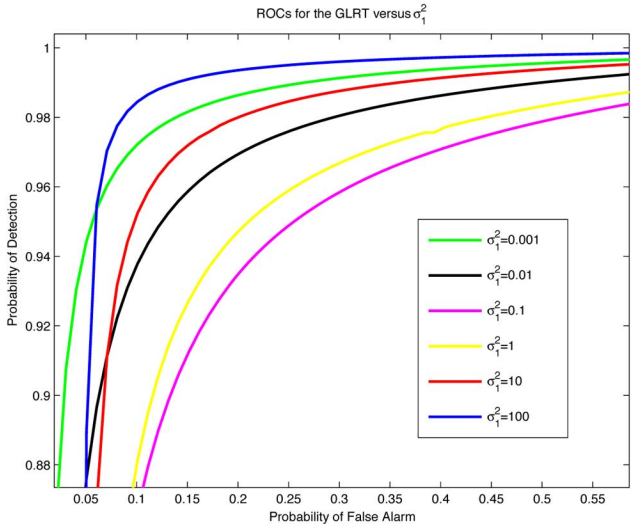


Fig. 4. GLRT ROCs for different values of $\sigma_1^2, \sigma_x^2 = 1, \sigma_0^2 = 0.001, N = 1024$, and orthogonal channels. One-sample case.

(36) by Gaussian distributions using the central limit theorem. However, these approximations are valid only for very large values of $p$ (even for $p = 500$, the Gaussian approximations are not in good agreement with the histograms of $T_{\mathrm{GLRT}}$). Thus the ROCs for the $p$-sample case have been generated from MC simulations instead. Here, 10 000 zero-mean Gaussian vectors have been generated, corresponding to the hypotheses "channel change" and "double-talk" according to (2) and (3). The LRT and GLRT test statistics have then been computed using MC run. The $P_{\mathrm{FA}}$ and $P_{\mathrm{D}}$ have been estimated by counting the relative number of samples exceeding the threshold $T_2$ according to (55) and (54). Note that the threshold value (between the minimum and maximum of the test statistic) has to be selected to yield appropriate values of the pairs $(P_{\mathrm{FA}}, P_{\mathrm{D}})$. Fig. 5 shows the ROCs for the GLRT in the case of orthogonal channels ($\Delta_i = 200$) and channels differing by a small time delay ($\Delta_i = 1$). It is easier to discriminate between channel changes and double-talk for high gain channels, cf. Fig. 5(a) and Fig. 5(b), or Fig. 5(c) and Fig. 5(d). Similarly, it is easier to discriminate between channel changes and double-talk for orthogonal channels, cf. Fig. 5(a) and Fig. 5(c), or Fig. 5(b) and Fig. 5(d).

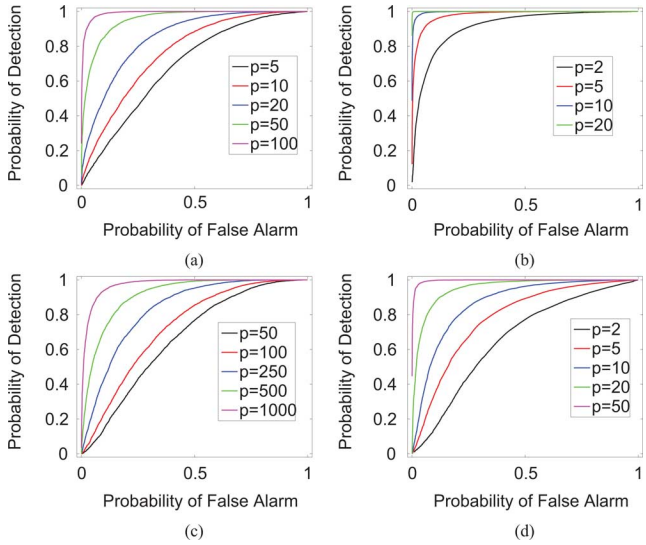


Fig. 5. ROCs for the GLRT in the case of $p$ samples (Top: orthogonal channels, Bottom: channels with $\Delta_i = 1$, left: $G = -10$ dB, right: $G = 6$ dB). $\sigma_0^2 = \sigma_1^2 = \sigma_x^2 = 1$ and $N = 1024$. (a) $G = -10$ dB—Orthogonal channels. (b) $G = 6$ dB—Orthogonal channels. (c) $G = -10$ dB—Channels with relative delay $\Delta = 1$. (d) $G = 6$ dB—Channels with relative delay $\Delta = 1$.

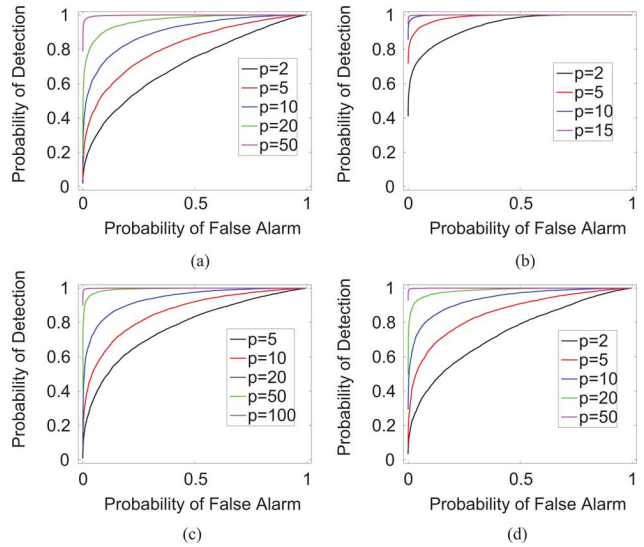


Fig. 6. ROCs for the LRT in the case of $p$ samples. Top: orthogonal channels, Bottom: channels with $\Delta_i = 1$. Left: $G = -10$ dB. Right: $G = 6$ dB. $\sigma_x^2 = \sigma_0^2 = \sigma_1^2 = 1$, $N = 1024$. (a) $G = -10$ dB—Orthogonal channels. (b) $G = 6$ dB—Orthogonal channels. (c) $G = -10$ dB—Channels with relative delay $\Delta = 1$. (d) $G = 6$ dB—Channels with relative delay $\Delta = 1$.

Fig. 6 shows the ROCs for the corrected LRT and replaces [8, Figs. 5 and 6]. Comparison of the two sets of figures indicates improved performance for the corrected LRT. Comparison of the corresponding cases in Figs. 8 and 5 show significant differences in the improvement between the LRT and the GLRT for these cases. Case (a) requires about twice the number of time samples for the GLRT as compared to the LRT. Case (b) requires about the same number of time samples for the GLRT as compared to the LRT. Case (c) requires about 20 times the number of time samples for the GLRT as compared to the LRT.

Case (d) requires about two to three times the number of time samples for GLRT as compared to the LRT. Hence, one cannot generalize the quantitative improvement.

## V. Application To Echo Cancellers

The GLRT theory has been tested for two distinct examples in the full EC implementations of Fig. 1 with transfer logic between the shadow and main filters, modified to use the GLRT statistic in (36). The first example consists of a synthetically generated data set whose channel-change and double-talk parameters are assumed known to the EC. The second example consists of real voice data transmitted over a real channel. The interested reader is invited to consult [8] for more details.

### A. Synthetic Data—Colored Inputs

This section considers the synthetic data generated in [8, Sec. VI-A], where the channels $\boldsymbol{h}_0$ and $\boldsymbol{h}_1$ are described in (53), except the input signal is correlated. The covariance matrix of the input signal has been chosen as follows:

$$\boldsymbol{\Sigma}_x = \sigma_x^2 \begin{pmatrix} 1 & \rho & \dots & \rho^N \\ \rho & 1 & \dots & \rho^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ \rho^N & \dots & \rho & 1 \end{pmatrix} \tag{62}$$

where the parameter $\rho$ controls the input signal correlation. The unknown channel output $y(n)$ consists of four 1 s segments generated as follows:

$$y(n) = \begin{cases} \boldsymbol{h}_1^T \boldsymbol{x}(n) + n_0(n) & n = 1, \dots, 8000 \\ \boldsymbol{h}_0^T \boldsymbol{x}(n) + n_0(n) & n = 8001, \dots, 16000 \\ \boldsymbol{h}_0^T \boldsymbol{x}(n) + n_0(n) + n_1(n) & n = 16001, \dots, 24000 \\ \boldsymbol{h}_0^T \boldsymbol{x}(n) + n_0(n) & n = 24001, \dots, 32000 \end{cases}$$

where $E[n_1^2(n)] = 1$ and $E[n_0^2(n)] = 10^{-3}$. Thus, $y(n)$ consists of channel changes at $n = 1$ and $n = 8001$, double-talk but no channel change at $n = 16001$, and the double-talk disappears at $n = 24001$ without another channel change. Note that $\boldsymbol{h}_1^T \boldsymbol{x}$ and $\boldsymbol{h}_0^T \boldsymbol{x}$ were replaced by the outputs of the main and shadow filters at each iteration. Fig. 7(a)–(c) shows the performance of the LRT for $\rho = 0.5$ for different values of $p$ (i.e., $p = 1$, $p = 10$, and $p = 100$) whereas Fig. 8(a), (b), and (c) is for the GLRT for $\rho = 0.5$ and different values of $p$ (i.e., $p = 100$, $p = 200$, and $p = 500$).

Figs. 7 and 8 have been obtained for orthogonal channels and a gain $G = -10$ dB. The top curves show the EC MSEs (time-average of 100 squared errors), the middle curves indicate the number of transfers from the shadow filter to the main filter every 200 samples and the bottom figures show the LRT and GLRT test statistics compared to the threshold. For the GLRT, the case $p = 100$ results in transfers during the period of double-talk (from 16 k to 24 k). This bad behavior disappears for $p = 200$ and $p = 500$, resulting in the correct drop in MSE during the last 8 k.

Some comments are appropriate about the threshold determination for the GLRT test statistics. The threshold $Z_2$ in (56) was chosen empirically to yield a PFA that resulted in no transfers from the shadow filter to the main filter during double-talk. For $p = 1$, knowledge of the pdfs of the sufficient statistics

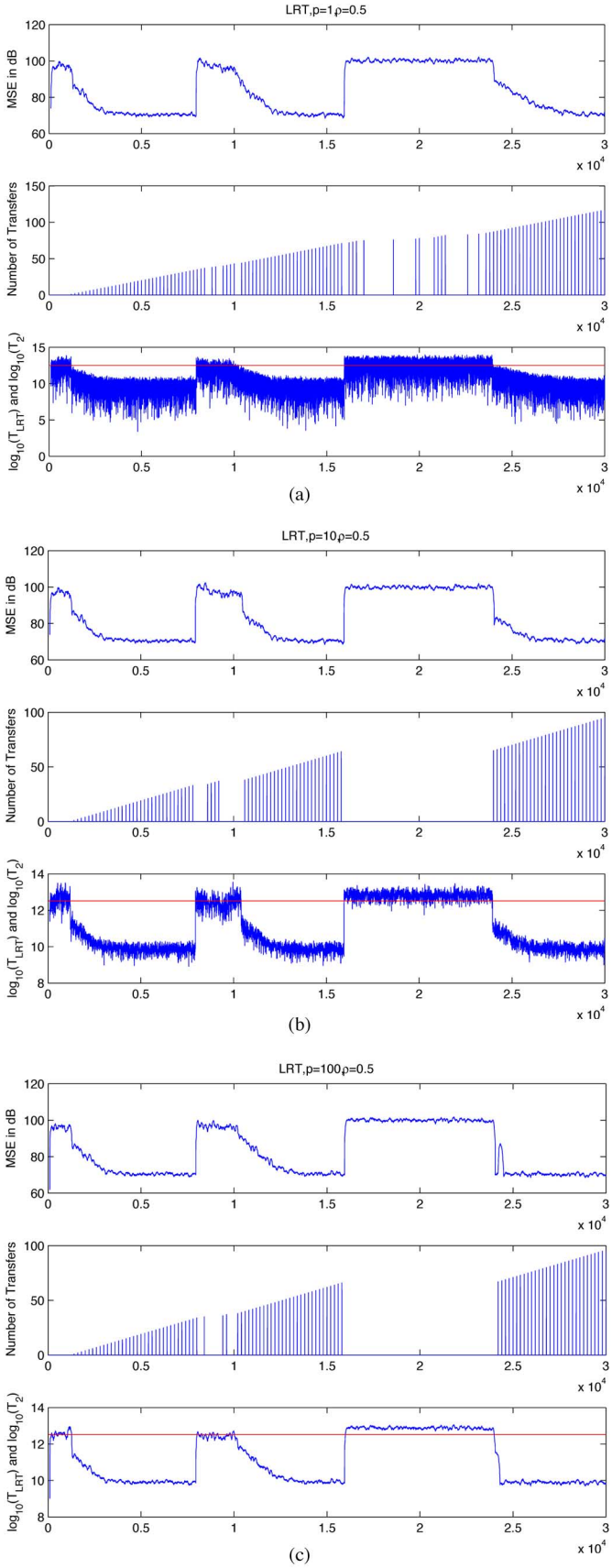


Fig. 7. LRT-based EC performance for synthetic data. (Top) MSE. (Middle) Number of transfers from the shadow filter to the main filter. (Bottom) Time-average sufficient statistics $T_{\mathrm{LRT}}(n)$ and threshold. (a) $p = 1$. (b) $p = 10$. (c) $p = 100$.

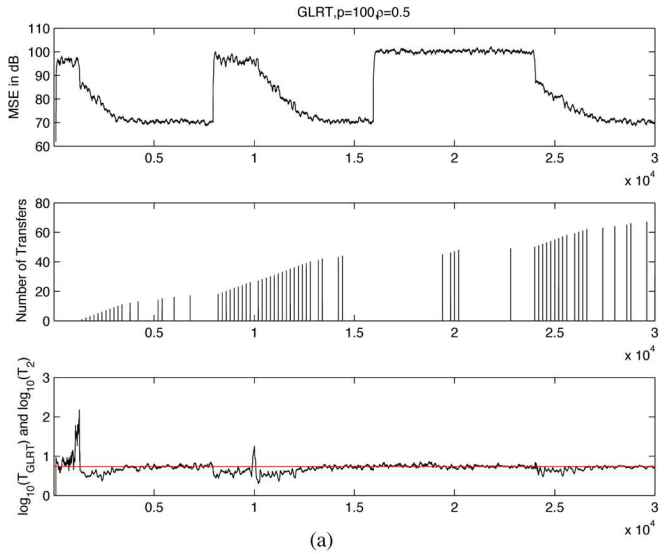


Fig. 8. GLRT-based EC performance for synthetic data. (Top) MSE. (Middle) Number of transfers from the shadow filter to the main filter. (Bottom) Time-average sufficient statistics $T_{\mathrm{GLRT}}(n)$ and threshold. (a) $p = 100$. (b) $p = 200$. (c) $p = 500$.

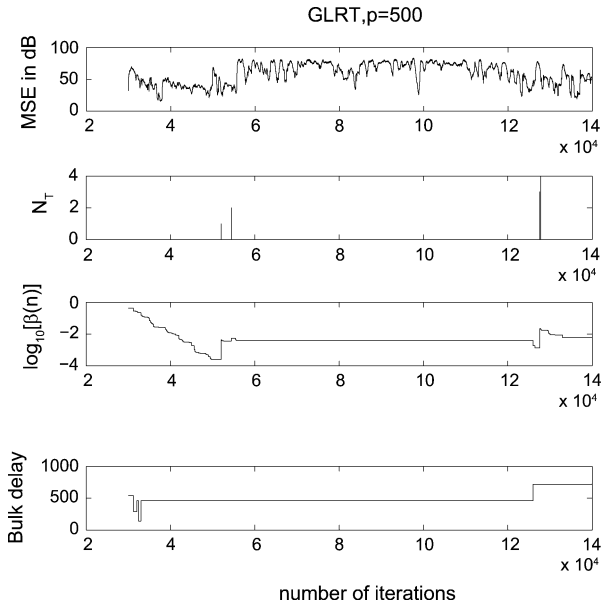


Fig. 9. Performance of the GLRT-based EC for voice data over a real channel. (Top) MSE. (Second) Number of transfers from the shadow filter to the main filter. (Third) $10\ln_{10}(\beta)$. (Bottom) Main filter bulk delay.

determines the test thresholds as a function of $P_{\mathrm{FA}}$. For instance, the GLRT curves have been obtained with a threshold $Z_2 = 0.74$ (corresponding to $P_{\mathrm{FA}} \simeq 0.05$ and $P_{\mathrm{D}} \simeq 0.75$ for $p = 1$). This approach cannot be used for arbitrary $p$, since the pdfs of the GLRT sufficient statistics have no simple closed-form expressions under the hypotheses. However, the GLRT test statistic is the ratio of averaged errors. Thus, the thresholds obtained for $p = 1$ and $p > 1$ should yield similar $P_{\mathrm{FA}}$ and $P_{\mathrm{D}}$. Here, $P_{\mathrm{FA}}$ and $P_{\mathrm{D}}$ have been estimated using the data from the example above to illustrate this point. More precisely, $P_{\mathrm{FA}}$ has been estimated by counting the number of samples exceeding the threshold during the time interval $(8001, 16000)$, according to (55). Likewise, $P_{\mathrm{D}}$ has been evaluated by counting the number of samples exceeding the threshold during the time interval $(16001, 24000)$, according to (54). The values $\widehat{P}_{\mathrm{FA}} \simeq 0.0998$ and $\widehat{P}_{\mathrm{D}} \simeq 0.8281$ for $p = 500$ and $Z_2 = 0.74$ were obtained (better results than for $p = 1$).

### *B. Voice Data Over a Real Channel*

The description of the voice data considered here can be found in [8, Sec. VI-B]. This section concentrates on analyzing the results obtained with the proposed GLRT detector in the multi-sample case. The parameter settings for the EC were $p = 500$ and a threshold $Z_2 = 0.2$ (again, the threshold $Z_2$ was chosen empirically to yield a PFA that resulted in no transfers from the shadow filter to the main filter filter during double-talk). Fig. 9 shows four curves for the EC using the GLRT: the smoothed MSE of the main filter in dB (top), the number of transfers from shadow to main (second), a measure of the adaptive filter weight errors (third) (a small value of the norm of the "delay coefficient" [13], denoted as $\beta$, means that the adaptive filter is at or near convergence), and the bulk delay of the main filter (bottom). Fig. 9 compares favorably to [8, Fig. 11]. Hence, for this case at least, the GLRT-based EC performs as well as the LRT-based EC. This is a very interesting result.

## VI. Results and Conclusion

The LRT statistical hypothesis testing approach in [8] for deciding between double-talk versus channel change for echo cancellation has been extended to the GLRT case. The LRT required *a priori* knowledge of the background noise and double-talk power levels. Here, a GLRT is derived which does not require this knowledge. The GLRT makes joint maximum likelihood estimates of these powers and uses them in the LRT. Since the GLRT uses less *a priori* information than the LRT, it does not perform as well as the LRT in theory.

The probability density function of a sufficient statistic under each hypothesis was studied and the performance of the test was evaluated as a function of the system parameters. The ROCs indicate that it is difficult to correctly decide between double-talk and a channel change based upon a single look. However, post-detection integration of approximately 200 sufficient statistic samples yields a detection probability close to unity (0.99) with a small false alarm probability (0.01) for the theoretical GLRT model. Application of a GLRT-based EC to real voice data showed comparable performance to that of the LRT-based EC given in [8]. Several new and interesting results were obtained.

1) [8] derived the optimum Neyman-Pearson detector for a white data input. Here, it was shown that the same test is valid for colored inputs. This is an important result since speech signals are usually colored.
2) [8] derived the Neyman-Pearson detector for the one-sample case. This result was extended, on an *ad hoc* basis, to the multiple-sample case using post-detection integration. Here, it was proven that this test is also optimum for the multiple-sample case.
3) It was shown that the GLR detectors are the same for both known and unknown input data covariance and correlation matrices. This is a very useful result since the statistics of the input are often changing very rapidly and sometimes cannot be estimated accurately from the input.

## References


[1] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. New York: Springer-Verlag, 2001.

[2] C. Breining, P. Dreiseitel, E. Hänsler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, "Acoustic echo control," *IEEE Signal Process. Mag.*, vol. 16, no. 4, pp. 42–69, Jul. 1999.

[3] A. Mader, H. Puder, and G. Schmidt, "Step-size control for echo cancellation filters—an overview," *Signal Process.*, vol. 80, no. 9, pp. 1697–1719, Sep. 2000.

[4] C. Carlemalm, F. Gustafsson, and B. Wahlberg, "On the problem of detection and discrimination of double talk and change in the echo path," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Atlanta, GA, May 1996, pp. 2742–2745.

[5] C. Carlemalm and A. Logothetis, "On detection of double talk and changes in the echo path using a Markov modulated channel model," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Munich, Germany, Apr. 1997, pp. 3869–3872.

[6] M. Fozumbal, M. C. Hans, and R. W. Schafer, "A decision-making framework for acoustic echo cancellation," in *Proc IEEE Workshop on Appl. Signal Process. Audio Acoust.*, New Paltz, New York, Oct. 1996, pp. 33–36.

[7] K. Ochiai, T. Araseki, and T. Ogihara, "Echo canceller with two echo path models," *IEEE Trans. Commun.*, vol. 25, pp. 589–595, Jun. 1977.

[8] N. J. Bershad and J.-Y. Tourneret, "Echo cancellation: A likelihood ratio test for double talk versus channel change," *IEEE Trans. Signal Process.*, vol. 54, pp. 4572–4581, Dec. 2006.

[9] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York: Wiley, 1968.

[10] J. H. Cho, D. R. Morgan, and J. Benesty, "An objective technique for evaluating doubletalk detectors in acoustic echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 7, pp. 718–724, Nov. 1999.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. 1.

[12] J. Omura and T. Kailath, Some Useful Probability Distributions' Stanford Electronics Laboratories Stanford, CA, Tech. Rep. No. 7050-6, 1965.

[13] E. Haensler and G. Schmidt, *Acoustic Echo and Noise Control*. New York: Wiley, 2004.


**Jean-Yves Tourneret** (SM'08) received the ingénieur degree in electrical engineering from École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique et d'Hydraulique, Toulouse (ENSEEIHT), France, in 1989 and the Ph.D. degree from the National Polytechnic Institute, Toulouse, in 1992.

He is currently a professor with ENSEEIHT and a member of the IRIT Laboratory (UMR 5505 of the CNRS). His research activities are centered around statistical signal processing with a particular interest to Markov chain Monte Carlo methods.

Dr. Tourneret was the program chair of EUSIPCO, held in Toulouse in 2002. He was also a member of the organizing committee for ICASSP'06, held in Toulouse, in 2006. He has been a member of different technical committees including the Signal Processing Theory and Methods (SPTM) Committee of the IEEE Signal Processing Society (2001–2007). He is currently serving as an Associate Editor for the IEEE Transactions on Signal Processing.

**Neil J. Bershad** (F'88) received the B.E.E. degree from Rensselaer Polytechnic Institute, Troy, NY, in 1958, the M.S. degree in electrical engineering from the University of Southern California, Los Angeles, in 1960, and the Ph.D. degree in electrical engineering, also from Rensselaer Polytechnic Institute, in 1962.

He joined the Faculty of the Henry Samueli School of Engineering, University of California, Irvine, in 1966 and is currently an Emeritus Professor of Electrical Engineering and Computer Science. His research interests have involved stochastic systems modeling and analysis. His recent interests are in the area of stochastic analysis of adaptive filters. He has published a significant number of papers on the analysis of the stochastic behavior of various configurations of the LMS adaptive filter. His present research interests include the statistical learning behavior of adaptive filter structures for echo cancellation, active acoustic noise cancellation and variable gain (mu) adaptive algorithms.

Dr. Bershad has served as an Associate Editor of the IEEE Transactions on Communications in the area of phase-locked loops and synchronization. More recently, he was an Associate Editor of the IEEE Transactions on Acoustics, Speech and Signal Processing in the area of adaptive filtering.

**José Carlos M. Bermudez** (S'78–M'85–SM'02) received the B.E.E. degree from Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, the M.Sc. degree in electrical engineering from COPPE/UFRJ, and the Ph.D. degree in electrical engineering from Concordia University, Montreal, Canada, in 1978, 1981, 1985, respectively.

He joined the Department of Electrical Engineering, Federal University of Santa Catarina (UFSC), Florianópolis, Brazil, in 1985, where he is currently a Professor of electrical engineering. During winter 1992, he was a Visiting Researcher with the Department of Electrical Engineering, Concordia University. In 1994, he was a Visiting Researcher with the Department of Electrical Engineering and Computer Science, University of California, Irvine (UCI). His research interests have involved analog signal processing using continuous-time and sampled-data systems. His recent research interests are in digital signal processing, including linear and nonlinear adaptive filtering, active noise and vibration control, echo cancellation, image processing, and speech processing.

Prof. Bermudez served as an Associate Editor for the IEEE Transactions on Signal Processing in the area of adaptive filtering from 1994 to 1996, and from 1999 to 2001, and as the Signal Processing Associate Editor for the *Journal of the Brazilian Telecommunications Society* (2005–2006). He was a member of the Signal Processing Theory and Methods Technical Committee of the IEEE Signal Processing Society from 1998 to 2004. He is currently an Associate Editor for the *EURASIP Journal on Advances in Signal Processing*.