A NEW DOUBLETALK AND CHANNEL CHANGE DETECTION ALGORITHM BASED ON HYPOTHESIS TESTING

Jianming Liu, *Steven L. Grant*[1]

Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, Missouri 65409, Jianming.Liu@mst.edu, sgrant@mst.edu

ABSTRACT

In acoustic and network echo cancellation, the detection of doubletalk and echo path channel changes are important to the control of the echo canceller's adaptive filter. This paper investigates joint doubletalk and channel change detection from an M-hypotheses test perspective. Also, using a stationary Gaussian stochastic input signal model, we propose a doubletalk versus channel change detection algorithm based on the likelihood ratio test. This proposed detection algorithm intuitively has a dynamic threshold which is based on the probabilities of past doubletalk and echo path change detection outputs. Simulation results prove the efficiency of the proposed detection algorithm.

Index Terms— Channel change detection, doubletalk detection, acoustic echo cancellation, hypothesis testing, likelihood ratio test.

1. INTRODUCTION

Acoustic echo cancellation (AEC) is used to remove the undesired echo in hands-free communication, and it is usually done by modeling the echo path impulse with an adaptive filter and subtracting the estimated echo from the microphone output signal. The far-end signal $x(n)$ is filtered through the room impulse response $\mathbf{h}(n)$ to get the echo signal $y(n)$.

$$y(n) = x(n) * h(n) = \mathbf{x}^T(n)\mathbf{h}(n) \tag{1}$$

where

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-L+1) \end{bmatrix}^T, \tag{2}$$

$$\mathbf{h} = \begin{bmatrix} h(0) & h(1) & \cdots & h(L-1) \end{bmatrix}^T, \tag{3}$$

and L is the length of echo path. This echo signal $y(n)$ is added to the near-end signal $v(n)$ and noise $w(n)$ to get the microphone signal,

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n) + w(n) \tag{4}$$

The estimated echo signal $\hat{y}(n)$ is the output of the adaptive filter $\hat{\mathbf{h}}(n)$ as

$$\hat{y}(n) = x(n) * \hat{h}(n) = \mathbf{x}^T(n)\hat{\mathbf{h}}(n) \tag{5}$$

We define the residual echo signal $\hat{e}(n)$ as

$$\hat{e}(n) = y(n) - \hat{y}(n) = y(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n) \tag{6}$$

and considering the near-end signal, the observed error signal $e(n)$ is

$$\begin{aligned} e(n) &= d(n) - \hat{y}(n) + v(n) \\ &= d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n) + v(n) \end{aligned} \tag{7}$$

Traditionally, the doubletalk and channel change detection are used to control the adaption of the filter. The doubletalk detection (DTD) algorithm decides whether there is near-end speech in the microphone signal and freezes the adaptation of the modeling filter when near-end speech is present [1]. However, many doubletalk detectors declare doubletalk during channel changes, and freeze the adaptation of the filter when it most needs to adapt. Thus, another efficient detection algorithm is needed to detect channel change – a channel change detector [2]. According to different adaption strategies during doubletalk and channel change, it is important to differentiate between them [3] [4]. In this paper we will analyze the adaptive control of the filter update from a hypothesis test perspective and then propose a new doubletalk versus channel change (DTCC) detection algorithm based on the likelihood ratio test (LRT).

This paper is organized as follows: Section 2 derives the detection scheme of doubletalk and channel change based on the M-hypothesis testing. Section 3 proposes our doubletalk versus channel detection statistic. Simulation results are presented in Section 4. Finally, we draw conclusions in Section 5.

2. ADAPTIVE CONTROL FROM THE M-HYPOTHESIS TEST PERSPECTIVE

In our simplified model of an acoustic echo canceller there are only three states that effect adaptation: far-end-only-

*This work is performed under the Wilkens Missouri Endowment

[1] Formerly Steven L. Gay

active, doubletalk and channel change in the echo path. It is natural that we could view this as a kind of M hypotheses ($M = 3$) detection problem, in which we define the following hypotheses:

$\boldsymbol{H}_0$: far-end active only,

$\boldsymbol{H}_1$: doubletalk is occurring,

$\boldsymbol{H}_2$: a channel change has happened.

According to the M-Hypothesis Bayes test [5], with an observation vector z, we define the following likelihood ratios:

$$\Lambda_1(z) = \frac{p[z \mid H_1]}{p[z \mid H_0]} \tag{8}$$

$$\Lambda_2(z) = \frac{p[z \mid H_2]}{p[z \mid H_0]} \tag{9}$$

$$\Lambda_3(z) = \frac{p[z \mid H_2]}{p[z \mid H_1]} \tag{10}$$

We denote the cost for the nine possible courses of action as C_{00}, C_{10}, C_{20}, C_{01}, C_{11}, C_{21}, C_{02}, C_{12}, and C_{22}. The first subscript indicates the state chosen, and the second, the state that was true. We have the following decision rules [5]

$$P_1(C_{01} - C_{11})\Lambda_1(z) \underset{H_0\,or\,H_2}{\overset{H_1\,or\,H_2}{\gtrless}} P_0(C_{10} - C_{00}) + P_2(C_{12} - C_{02})\Lambda_2(z) \tag{11}$$

$$P_2(C_{02} - C_{22})\Lambda_2(z) \underset{H_0\,or\,H_1}{\overset{H_2\,or\,H_1}{\gtrless}} P_0(C_{20} - C_{00}) + P_1(C_{21} - C_{01})\Lambda_1(z) \tag{12}$$

$$P_2(C_{12} - C_{22})\Lambda_2(z) \underset{H_1\,or\,H_0}{\overset{H_2\,or\,H_0}{\gtrless}} P_0(C_{20} - C_{10}) + P_1(C_{21} - C_{11})\Lambda_1(z) \tag{13}$$

in which P_0, P_1 and P_2 are the priori probabilities of $\boldsymbol{H_0}$, $\boldsymbol{H_1}$, and $\boldsymbol{H_2}$.

The costs listed above are not all the same. This is due to the different adaptive strategies at doubletalk and channel change. When there is doubletalk, the filter needs to immediately freeze the coefficient update of the filter, meanwhile, when the channel changes, the adaptation should accelerate. So, the cost of declaring doubletalk as channel change should be high, because it might cause the divergence of the filter. The cost of declaring channel change as doubletalk is also high since this kind of error will significantly decrease the tracking ability of the filter.

Next, we will discuss how to define the costs and simplify the detection rule in (11)-(13). According to the M-Hypothesis Bayes test principle, the overall goal is to minimize the total cost. For the application of echo cancellation, the final target of adaptive control is to have an optimal estimate of the real echo path impulse response. In general, according to the normalized least mean square (NLMS) based adaptive algorithm [6], we have

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu(n)\frac{\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{x}(n) + \delta} \tag{14}$$

In practice, we usually have the following variable step size strategy. Based on the detection output of doubletalk and channel change detection, we update the filter in both far-end and channel change state, but we use a smaller step size at far-end than channel change, i.e. we will have a larger step size when channel change is detected. Meanwhile, we use a step size near zero when doubletalk is detected. Therefore, we have

$$0 \le \hat{\mu}_1 < \hat{\mu}_0 < \hat{\mu}_2 \le 1 \tag{15}$$

However, considering our state definitions, there are no channel changes during the “far-end” state. Meanwhile, there may be background noise, so the ideal step size for far-end should be zero as with doubletalk. At the same time, when there is channel change, we’d like to update as quickly as possible, therefore we choose.

$$\mu_1 = \mu_0 = 0,\ \mu_2 = 1 \tag{16}$$

It is reasonable that the cost function should be related to the difference between the used $\hat{\mu}_i$ and correct step size μ_j due to the detection error of doubletalk and channel change. We propose to define the cost C_{ij} as follows

$$C_{ij} = \left|\hat{\mu}_i - \mu_j\right| \tag{17}$$

Therefore, the overall target of the hypothesis test is to minimize the total step size error. The definition of the cost function in (17) is direct and what’s more important is that this definition could simplify the hypothesis test decision rules in (11)-(13). Based on the assumptions in (15) and (16), we could analyze the relation between the costs as follows.

$$C_{10} < C_{00} < C_{20} \tag{18}$$

$$C_{22} < C_{02} < C_{12} \tag{19}$$

$$C_{11} < C_{01} < C_{21} \tag{20}$$

Considering (18), (19), and (20) , we could simplify (11)-(13) as follows,

$$\frac{P_0(C_{00} - C_{10})}{P_2(C_{12} - C_{02})} + \frac{P_1(C_{01} - C_{11})}{P_2(C_{12} - C_{02})}\Lambda_1(z) \underset{H_0\,or\,H_2}{\overset{H_1\,or\,H_2}{\gtrless}} \Lambda_2(z) \tag{21}$$

$$\Lambda_2(z) \underset{H_0\,or\,H_1}{\overset{H_2\,or\,H_1}{\gtrless}} \frac{P_0(C_{20} - C_{00})}{P_2(C_{02} - C_{22})} + \frac{P_1(C_{21} - C_{01})}{P_2(C_{02} - C_{22})}\Lambda_1(z) \tag{22}$$

$$\Lambda_2(z) \underset{H_0\,or\,H_1}{\overset{H_2\,or\,H_1}{\gtrless}} \frac{P_0(C_{20} - C_{00})}{P_2(C_{02} - C_{22})} + \frac{P_1(C_{21} - C_{01})}{P_2(C_{02} - C_{22})}\Lambda_1(z) \tag{23}$$

Meanwhile, due to the definition of the cost in (17), it should be noted that we have

$$\frac{P_1(C_{01} - C_{11})}{P_2(C_{12} - C_{02})} = \frac{P_1(C_{21} - C_{01})}{P_2(C_{02} - C_{22})} = \frac{P_1(C_{21} - C_{11})}{P_2(C_{12} - C_{22})} = \frac{P_1}{P_2} \tag{24}$$

$$\frac{P_0(C_{00}-C_{10})}{P_2(C_{12}-C_{02})}=\frac{P_0(C_{20}-C_{00})}{P_2(C_{02}-C_{22})}=\frac{P_0(C_{20}-C_{10})}{P_2(C_{12}-C_{22})}=\frac{P_0}{P_2} \tag{25}$$

Finally, substituting (24) and (25) into (21)-(23), we observe that that all we need for the so-called doubletalk versus channel change (DTCC) detection statistic is

$$\Lambda_2(z) \underset{H_1}{\overset{H_2}{\gtrless}} \frac{P_0}{P_2}+\frac{P_1}{P_2}\Lambda_1(z) \tag{26}$$

Since $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ have the same effect (μ is made large), we can further simplify the decision in (26) by combining the $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ states into $\boldsymbol{H_1}$ simply by letting $P_0=0$,

$$\Lambda_2(z) \underset{H_1}{\overset{H_2}{\gtrless}} \frac{P_1}{P_2}\Lambda_1(z) \tag{27}$$

It should be noted that this reduces to a binary hypothesis test $\Lambda_3(z)$ with a dynamic threshold P_1/P_2,

$$\Lambda_3(z) \underset{H_1}{\overset{H_2}{\gtrless}} \frac{P_1(C_{21}-C_{11})}{P_2(C_{12}-C_{22})}=\frac{P_1}{P_2} \tag{28}$$

which means that based on the assumption of (15) and (16) we could minimize the step size error (see (17)) based on the DTCC detection in (28). Meanwhile, we automatically get a dynamic threshold which is very useful, since many detection methods suffer from the choice of a fixed threshold in practice.

3. PROPOSED DTCC

There is a DTCC detection algorithm proposed in [3] [4], however it is based on the two-path model. In this section, using a similar approach, we use $\boldsymbol{z}=[d,y]^T$ as the observation vector and derive our proposed LRT detection statistic for doubletalk and channel change detection.

3.1. Multiple-sample LRT

We take the following two-sample observation

$$\boldsymbol{z}(n)=\left[d(n),d(n-1),y(n),y(n-1)\right]^T \tag{29}$$

and assume that there is no channel change between these two samples. Assuming the probability density function (PDF) of far-end and near-end signal are both stationary zero-mean and Gaussian distributed, the joint PDF of $\boldsymbol{z}(n)$ is a zero-mean Gaussian vector such that

$$p[\boldsymbol{z}\,|\,H_i]\sim N(0,\boldsymbol{\Sigma}_{i2}) \tag{30}$$

where $\boldsymbol{\Sigma}_{i2}$, *i=1, 2*, is the 4×4 matrix

$$\boldsymbol{\Sigma}_{12}=\begin{bmatrix}\boldsymbol{H}_{x1}+(\sigma_w^2+\sigma_v^2)\boldsymbol{I}_2 & \boldsymbol{H}_{x1}\\ \boldsymbol{H}_{x1} & \boldsymbol{H}_{x1}\end{bmatrix} \tag{31}$$

and

$$\boldsymbol{\Sigma}_{22}=\begin{bmatrix}\boldsymbol{H}_{x2}+\sigma_w^2\boldsymbol{I}_2 & \boldsymbol{H}_{x2}\\ \boldsymbol{H}_{x2} & \boldsymbol{H}_{x2}\end{bmatrix} \tag{32}$$

where

$$\boldsymbol{H}_{xi}=\boldsymbol{D}_i^T\begin{pmatrix}\boldsymbol{\Sigma}_x & \boldsymbol{R}_x\\ \boldsymbol{R}_x & \boldsymbol{\Sigma}_x\end{pmatrix}\boldsymbol{D}_i,\quad \boldsymbol{D}_i=\begin{pmatrix}\mathbf{h}_i & 0\\ 0 & \mathbf{h}_i\end{pmatrix},\quad i=1,2 \tag{33}$$

$$\boldsymbol{R}_x=E\left[\boldsymbol{x}(n)\boldsymbol{x}^T(n-1)\right],\ \text{and}\ \boldsymbol{\Sigma}_x=E\left[\boldsymbol{x}(n)\boldsymbol{x}^T(n)\right] \tag{34}$$

in which $\mathbf{h}_1$ is the echo path impulse response at doubletalk and $\mathbf{h}_2$ is impulse response at channel change. We define

$$\sigma_{y_1}^2=\mathbf{h}_1^T\Sigma_x\mathbf{h}_1 \tag{35}$$

and

$$\sigma_{y_2}^2=\mathbf{h}_2^T\Sigma_x\mathbf{h}_2\,. \tag{36}$$

The log LRT (LLRT) is expressed as [5]

$$\begin{aligned}\ln\Lambda_3(z)&=\ln\left[\frac{p[\boldsymbol{z}\,|\,H_2]}{p[\boldsymbol{z}\,|\,H_1]}\right]\\ &=\frac{1}{2}\boldsymbol{z}^T\left(\boldsymbol{\Sigma}_{12}^{-1}-\boldsymbol{\Sigma}_{22}^{-1}\right)\boldsymbol{z}+\frac{1}{2}\ln\left(\frac{|\boldsymbol{\Sigma}_{12}|}{|\boldsymbol{\Sigma}_{22}|}\right)\end{aligned} \tag{37}$$

The first part of this LLRT is

$$\begin{aligned}&\frac{1}{2}\boldsymbol{z}^T\left(\boldsymbol{\Sigma}_{12}^{-1}-\boldsymbol{\Sigma}_{22}^{-1}\right)\boldsymbol{z}\\ &=\frac{1}{2}\left(\frac{1}{\sigma_{y_1}^2}-\frac{1}{\sigma_{y_2}^2}\right)\left(y^2(n)+y^2(n-1)\right)\\ &-\frac{1}{2}\left(\frac{1}{\sigma_w^2}-\frac{1}{\sigma_w^2+\sigma_v^2}\right)\Big(\left(d(n)-y(n)\right)^2\\ &+\left(d(n-1)-y(n-1)\right)^2\Big)\end{aligned} \tag{38}$$

Meanwhile, the second part is

$$\begin{aligned}&\frac{1}{2}\ln\left(\frac{|\boldsymbol{\Sigma}_{12}|}{|\boldsymbol{\Sigma}_{22}|}\right)\\ &=\frac{1}{2}\times 2\times\ln\left(\frac{\sigma_w^2+\sigma_v^2}{\sigma_w^2}\right)-\frac{1}{2}\times 2\times\ln\left(\frac{\sigma_{y_2}^2}{\sigma_{y_1}^2}\right)\end{aligned} \tag{39}$$

It should be noted that the above analysis could be generalized to the case where more samples are available. Therefore, in general, the K-sample LLRT is

$$\begin{aligned}&\frac{1}{2}\left(\frac{1}{\sigma_{y_1}^2}-\frac{1}{\sigma_{y_2}^2}\right)\|\boldsymbol{y}\|^2-\frac{1}{2}\left(\frac{1}{\sigma_w^2}-\frac{1}{\sigma_w^2+\sigma_v^2}\right)\|\boldsymbol{d}-\boldsymbol{y}\|^2+\\ &\frac{K}{2}\ln\left(\frac{\sigma_w^2+\sigma_v^2}{\sigma_w^2}\right)-\frac{K}{2}\ln\left(\frac{\sigma_{y_2}^2}{\sigma_{y_1}^2}\right)\underset{H_1}{\overset{H_2}{\gtrless}}\ln\frac{P_1}{P_2}\end{aligned} \tag{40}$$

in which

$$\|\boldsymbol{y}\|^2=y^2(n)+y^2(n-1)+\ldots+y^2(n-K+1) \tag{41}$$

and

$$\|\mathbf{d}-\mathbf{y}\|^2 = \|\mathbf{v}+\mathbf{w}\|^2$$
$$= \left(v(n)+w(n)\right)^2 + \left(v(n-1)+w(n-1)\right)^2 + \ldots \quad (42)$$
$$+ \left(v(n-L+1)+w(n-K+1)\right)^2$$

3.2. Practical implementation

In practice, however, the detection statistic in (40) is not available to us, and we have to consider a practical implementation using estimators. In general, we could use the estimations

$$\|\mathbf{y}\|^2 \approx K\sigma_y^2 \quad (43)$$

and

$$\|\mathbf{v}+\mathbf{w}\|^2 \approx K\left(\sigma_w^2+\sigma_v^2\right) \quad (44)$$

in which σ_y^2 and $\sigma_w^2+\sigma_v^2$ are the power of the real echo and near-end speech and noise. According to the near-end signal energy estimator (NESEE) in [7], we could estimate $\hat{\sigma}_n^2 = \sigma_w^2+\sigma_v^2$ as

$$\hat{\sigma}_n^2 = \sigma_w^2+\sigma_v^2 \approx \hat{\sigma}_e^2 - \frac{\hat{r}_{eX}^T\hat{r}_{eX}}{\hat{\sigma}_x^2} \quad (45)$$

in which $\hat{\sigma}_e^2$ and $\hat{\sigma}_x^2$ are the estimation of error and far-end power, respectively. The statistic, $\hat{r}_{eX}$ is the correlation between far-end and error signal. Meanwhile, use $\sigma_{y_2}^2$ for σ_y^2 and estimate it using the near-end energy as

$$\hat{\sigma}_{y_2}^2 = \hat{\sigma}_y^2 = \sigma_d^2 - \sigma_n^2 \approx \hat{\sigma}_d^2 - \left(\hat{\sigma}_e^2 - \frac{\hat{r}_{eX}^T\hat{r}_{eX}}{\hat{\sigma}_x^2}\right) \quad (46)$$

A simple way to estimate the noise power σ_w^2 is to calculate the noise power $\hat{\sigma}_w^2$ during the silence period of far end and microphone signal,

$$\hat{\sigma}_w^2 = \frac{1}{K-1}\sum_{k=0}^{K-1}\left[d(n-k) - \frac{1}{K}\sum_{j=0}^{K-1}d(n-j)\right]^2 \quad (47)$$

In order to estimate $\sigma_{y_1}^2$, we assume there is no channel change at doubletalk. Then, we estimate the echo from the output of the adaptive filter, which is estimated echo,

$$\hat{\sigma}_{y_1}^2 = \hat{\sigma}_{\hat{y}}^2 = \frac{1}{K-1}\sum_{k=0}^{K-1}\left[\hat{y}(n-k) - \frac{1}{K}\sum_{j=0}^{K-1}\hat{y}(n-j)\right]^2 \quad (48)$$

Then, we have the practical DTCC as follows:

$$\frac{K}{2}\left(\frac{\left(\hat{\sigma}_d^2-\hat{\sigma}_n^2-\hat{\sigma}_{\hat{y}}^2\right)}{\hat{\sigma}_{\hat{y}}^2} - \ln\left(\frac{\hat{\sigma}_d^2-\hat{\sigma}_n^2}{\hat{\sigma}_{\hat{y}}^2}\right)\right) - \frac{K}{2}\left(\frac{\hat{\sigma}_n^2-\hat{\sigma}_w^2}{\hat{\sigma}_w^2} - \ln\left(\frac{\hat{\sigma}_n^2}{\hat{\sigma}_w^2}\right)\right) \underset{H_1}{\overset{H_2}{\gtrless}} \ln\frac{P_1}{P_2} \quad (49)$$

Finally, we could utilize the past detection output to get a dynamic estimation of the priori probabilities of P_1 and P_2.

We argue that our proposed detection statistic in (49) is an ideal detection for doubletalk and channel change as follows. In order to simplify the analysis, we assume $P_1 = P_2$, then the threshold is $\ln P_1/P_2 = 0$. When there is no channel change, we have $\hat{\sigma}_d^2 - \hat{\sigma}_n^2 \approx \hat{\sigma}_{\hat{y}}^2$, and $\hat{\sigma}_d^2 - \hat{\sigma}_n^2 \neq \hat{\sigma}_{\hat{y}}^2$ when channel change occurs. Meanwhile, when there is no doubletalk, we have $\hat{\sigma}_n^2 \approx \hat{\sigma}_w^2$, and $\hat{\sigma}_n^2 > \hat{\sigma}_w^2$ when doubletalk occurs. Therefore, we could easily prove that our detection statistic is larger than zero when there is channel change without doubletalk and smaller than zero when there is doubletalk without channel change. When there is both doubletalk and channel change, considering that the near-end speech energy is much larger than noise, we could still keep our detection statistic smaller than zero even with considerable channel change, which is desirable.

It should be mentioned that, compared to the two-path model method in [3] [4], our proposed method works in the normal single-path echo cancellation scheme and has a dynamic threshold.

4. SIMULATION RESULTS

In our simulation, we use speech sampled at 8 *k*Hz for far-end and near-end speech and a *L*=512-sample room impulse response. The total length of the signal is about 100,000 samples and we simulate the channel change by increasing the gain of channel by 2 at sample 33,000. Near-end speech occurs between samples 50,000 and 70,000. Plots of the far-end, near-end, microphone and error signals are shown in Figure 1.

It should be mentioned that the performance of the practical implementation in (49) depends on the performance of the near end energy estimator. There should be a good match between the estimated and real near-end speech powers even during the period of channel change. Considering that the NESEE and other estimators are not perfect, we have to smooth the detection statistic to make it stable. We plot the smoothed value of our proposed detection statistic and its dynamic threshold in Figure 2. We see that the proposed detection statistic detects doubletalk robustly and is not disturbed by the channel change.

Traditionally, in order to objectively evaluate the performance of the detection statistic, we use the receiver operating characteristic (ROC) [8] [9]. Considering that our proposed detection statistic automatically has a dynamic threshold, we could not get a ROC curve by sweeping a fixed chosen threshold. Therefore, we verified the ROC curve of our proposed detection statistic with different levels of near-end to far-end energy ratio (NFR) as shown in Figure 3.

We see that the probability of declaring doubletalk as channel change increases when the near-end energy becomes small. This is partially because when the near-end speech is small it is easily confused with noise. However, this is not serious because when the near-end energy is small, it will not damage the adaption of filter seriously, which is acceptable. However, when the near-end speech is large, we have very good doubletalk detection with a small error probability of declaring echo path channel change as doubletalk.

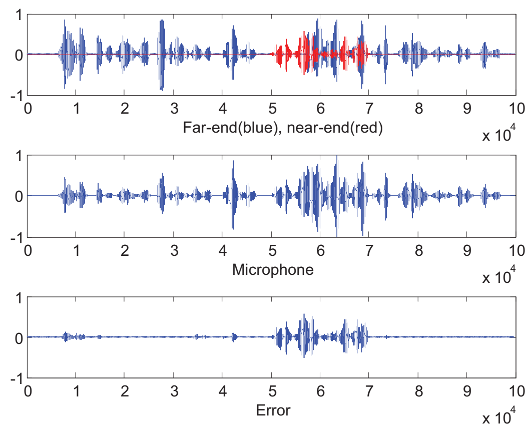

Fig.1 Speech signals in our simulation.

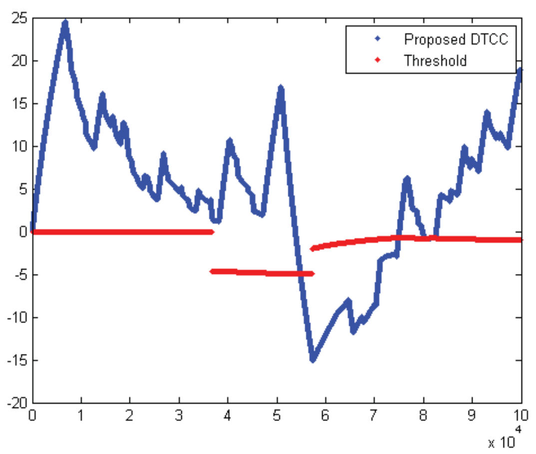

Fig.2 The proposed detection statistic and threshold.

5. CONCLUSION

In this paper, we first analyzed the joint doubletalk and channel change detection from the M-hypothesis testing perspective and then proposed our new doubletalk versus channel change detection statistic which includes a dynamic threshold. Simulation results showed that the proposed detection statistic could detect doubletalk robustly and is not disturbed by the channel change.

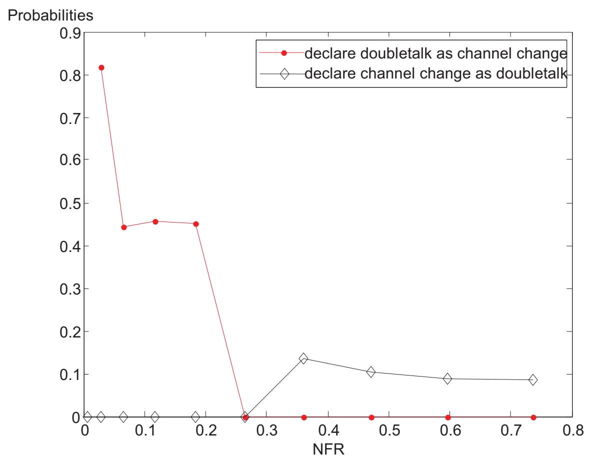

Fig.3 ROC with different levels of near-end energy

6. REFERENCES

[1] J. Benesty, D.R. Morgan, J. H. Cho, “A new class of doubletalk based on cross-correlation,” *IEEE Trans. Speech Audio Process*. 8 (March 2000) 168-172.

[2] Mohammad Asif Iqbal and Steven L. Grant, “A novel normalized cross-correlation based echo-path change detector,” in *Proceedings of the IEEE Region 5 Conference*, Arkansas, April 2007.

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

[4] Jean-Yves Tourneret, Neil J. Bershad, and José Carlos M. Bermudez, “Echo Cancellation—The Generalized Likelihood Ratio Test For Double-Talk Versus Channel Change,” *IEEE Trans. Signal Processing*, vol. 57, pp. 916–926, March. 2009.

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

[6] B. Widrow and S. D. Stearns, Adaptive Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[7] Mohammad Asif Iqbal and Steven L. Grant, “Novel Variable Step Size NLMS Algorithms for Echo Cancellation,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2008. ICASSP-2008.

[8] J. H. Cho, D. R. Morgan, and J. Benesty, “An objective technique for evaluating doubletalk detectors in acoustic echo cancellers,” *IEEE Trans. Speech Audio Processing*, vol.7, pp.718-724, Nov. 1999.

[9] Jianming Liu, Weiqian Liang and Runsheng Liu, “A three-class ROC for evaluating doubletalk detectors in acoustic echo cancellation,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2009. ICASSP-2009. , pp. 205-208, 2009.