# A HYBRID NEAR-FIELD SUPERDIRECTIVE GSC AND POST-FILTER FOR SPEECH ENHANCEMENT

*Parinaz Khayeri*
Speech Processing Research Lab
Electrical and Computer Eng. Dept.
Yazd University
Yazd, Iran
parinaz.khayeri@gmail.com

*Hamid Reza Abutalebi*
Speech Processing Research Lab
Electrical and Computer Eng. Dept.
Yazd University
Yazd, Iran
habutalebi@yazduni.ac.ir

*Vahid Abootalebi*
Electrical and Computer Eng. Dept.
Yazd University
Yazd, Iran
abootalebi@yazduni.ac.ir

**ABSTRACT**

This paper deals with the problem of microphone array speech enhancement using a hybrid Generalized Sidelobe Canceller (GSC), Near-Field Super-Directive (NFSD) beamformer, and post-filter. In this research, we employ a near field compensation block before the blocking matrix (of the GSC) to prevent signal leakage in the reference noise and a Linear Constrained Minimum Variance (LCMV) beamformer instead of blocking matrix to generalize the appropriate performance of the system for different arrival directions of desired speech. We also consider the application of the post-filters on the beamformer output. A modified version of McCowan post-filter is presented by calculating coherence function from microphone-array inputs. Considering similar idea, we also propose a multi-channel version of Short-Time Spectral Amplitude (STSA) post-filter. Our evaluations clearly demonstrate the effectiveness of the proposed modifications in the enhancement of the noisy speech.

***Index Terms-*** speech enhancement, superdirective beamformer, generalized sidelobe canceller, linear constrained minimum variance, post-filter

## 1. INTRODUCTION

Much research has been done in recent years on the use of microphone arrays for the task of speech enhancement by spatial filtering. While adaptive beamformers, such as Generalized Sidelobe Cancellers (GSCs), are especially suitable for the suppression of single, directional noise sources [1], super-directive beamformers have shown superior noise reduction performance in the case of diffuse noise fields [1,2]. Many real-life environments include diffuse, incoherent, and directional noises, simultaneously. To achieve a powerful multi-microphone speech enhancement system for such cases, a hybrid adaptive beamformer and post-filter was proposed that employs a Near Field SuperDirective (NFSD) beamformer as the fixed beamformer in the first path of the GSC [3,4].

However, there is still one issue remained to be solved: The best performance of NFSD occurs in endfire situation (where the desired source is at 0°). On the other side, due to the employed Jim and Griffith's blocking matrix [1], the GSC suffers from the leakage of the desired signal when the signal source is not in broadside situation (or at 90°); This is because of the signal leakage in the interference canceller block of the GSC. As a remedy, in this research we propose a novel blocking matrix to exclude the desired signal for every desired direction, without the signal leakage. Furthermore, the input signals are passed through a near field compensation block before the blocking matrix. The utilization of the compensation block is necessary according to the blocking matrix structure; the input signals of the blocking matrix should be the same (in amplitude and phase characteristics) to prevent signal leakage in the reference noise (output of interference canceller block).

Finally, we also propose some post-filters to be applied on the beamformer output. We modify McCowan post-filter [5,6] by calculating coherence function from microphone-array inputs. Considering the idea behind the modified McCowan post-filter, we also propose a modified version of STSA post-filter [7,8].

The rest of this paper is organized as follows: In Section 2, we explain the basic NFSD-GSC and present the proposed modifications. Section 3 explains post-filters and our proposed methods for post-filtering. In Section 4, we explain the experiments and evaluation results. Finally, Section 5 contains some concluding remarks.

## 2. ADAPTIVE BEAMFORMING

### 2.1. Employing NFSD in GSC Structure (NFSD-GSC)

The basic GSC system has two paths: 1) a standard fixed beamformer with constraints on the desired direction, and 2) an adaptive noise canceller to minimize the noise power at the output. In turn, the second path includes a blocking matrix that removes the desired signal from the noisy inputs, and an Interference Canceller (IC) which is updated using an unconstrained adaptive algorithm, such as the Least Mean Square (LMS) [1]. The so-called NFSD-GSC [3] uses an NFSD beamformer as the fixed beamformer in upper path.

### 2.2. Near-Field Compensation

In far-field conditions, the desired source is far enough from the array so that the received desired signal on microphones can be considered as the same; however, in near-field conditions, the amplitude/ phase of the received desired signal on microphones are not the same Thus, in near-field conditions, even for the broadside signal source, the signal leakage is occurred in the lower path of GSC. As shown in Fig 1, in the case of a near-field desired source, a near-field compensation should be firstly applied on the input signals to align the desired signal on all channels (before the blocking). To ensure full cancellation, we have to compensate for both phase misalignment and amplitude scaling of the desired signal across microphones by [1,3]:

$$\underline{X}^{''} = \underline{X} / \underline{d}_{NF} \,, \tag{1}$$

where $X$ is input vector, $X^{''}$ is input of blocking matrix and $\underline{d}_{NF}$ is steering vector for near-field condition that is given by:

$$\underline{d}_{NF} = \left[ 1 \quad \frac{r_0}{r_1}\exp(-j2\pi f \frac{r_1 - r_0}{c}) \quad \ldots \quad \frac{r_0}{r_{M-1}}\exp(-j2\pi f \frac{r_{M-1} - r_0}{c}) \right]^T \tag{2}$$

where $r_i$ is the space between microphone $i$ and signal source, $f$ is frequency index, and $c$ is the sound velocity.

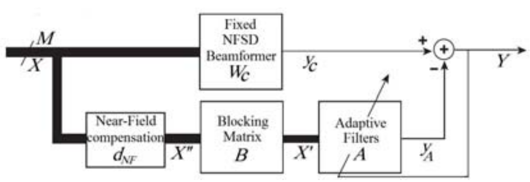


Fig. 1. Block diagram of NFSD-GSC

Fig. 2 shows the directivity pattern at 4*kHz* for the first row of the Jim and Griffiths blocking matrix [1] for an array with 5 microphones with inter-microphone distance of 4 cm.

As shown, in the near-field conditions, the far-field characteristics of the blocking matrix can be approximated by the use of a near-field compensator (equation (1)).

The performance of this system can be summarized as follows. If array has M microphones and X denotes the signals received by array, after time alignment, the output of upper path is given by:

$$y_C = (1/M)\underline{W}_C^H \underline{X} \,, \tag{3}$$

where  is the weighting filters of NFSD that are typically formulated to maximize the array gain as:

$$\underline{W}_C = [C + \varepsilon I]^{-1} \underline{d}_{NF} \left\{ \underline{d}_{NF}{}^H [C + \varepsilon I]^{-1} \underline{d}_{NF} \right\}^{-1} , \tag{4}$$

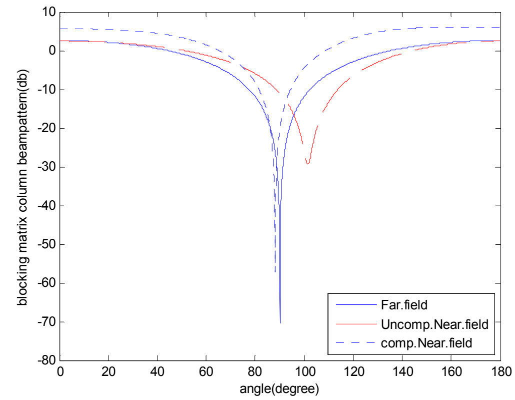


Fig. 2.The beampattern of Jim and Griffiths blocking matrix (first row) in far-, near-, compensated near-field cases.

where $C$ is noise field coherence matrix. The *(i,m)*-th component of $C$ is the coherence between the noise signals that is calculated based on noise PSDs and cross spectral densities between microphones $i$ and $j$ ($\varphi_{i,j}$) [1,5]:

$$C_{i,j} = \varphi_{i,j}(e^{j\Omega}) \Big/ \sqrt{\varphi_{i,i}(e^{j\Omega})\varphi_{j,j}(e^{j\Omega})} \,. \tag{5}$$

In the case of diffuse noise field, the coherence function is expressed by:

$$C_{i,j}(e^{j\Omega})\Big|_{diffuse} = \frac{\sin(\Omega f_s l_{mi} / c)}{\Omega f_s l_{mi} / c} \,, \tag{6}$$

where $l_{mi}$, is microphone spacing, $f_s$ is sampling frequency, $\Omega$ is frequency index and $c$ is the sound velocity.

The output of blocking matrix ($B$) is:

$$\underline{X}^{'} = B^H \underline{X}^{''} \,. \tag{7}$$

If $A$ denotes the adaptive filers of IC block, the output of lower path is given by:

$$y_A = \underline{A}^H . \underline{X}' \,. \tag{8}$$

It should be noted that the adaptive filters are updated using the standard unconstrained LMS algorithm as follows:

$$\underline{A}_{k+1} = \underline{A}_k + \mu y_k \underline{X}'_k \,, \tag{9}$$

where $\mu$ is the adaptation step size and $k$ is the frame index. Finally, the system output is calculated from the outputs of the upper and lower paths as:

$$Y = y_C - y_A \,. \tag{10}$$

### 2.3. Proposed Blocking Matrix

There is an issue in basic NFSD-GSC that should be resolved somehow: The best performance of NFSD occurs in endfire situation. On the other side, due to the employed Jim and Griffith's blocking matrix [1], the GSC suffers from the leakage of the desired signal when the signal source is not in broadside situation.

In this research, we propose the use of Linearly Constrained Minimum Variance (LCMV) beamformer [2,10] to release the above-mentioned limitation of the direction of desired signal.

The weights (filter coefficient) of LCMV are calculated so that minimize the output power (output variance) while imposing multiple linear constraints on the output signal (for preserving desired signal and removing directional undesired signals) [2,10]. Optimal coefficient vector is calculated as [1]:

$$\underline{W}_{opt}=R_x^{-1}C\,(C^H R_x^{-1}C)^{-1}\underline{F}\,, \tag{11}$$

where $R_x$ is input covariance matrix, $\underline{W}_{\mathrm{opt}}$ is weight vector and $C$ and $\underline{F}$ are constraint matrix and vector, respectively.

Here, we exploit the LCMV as a blocking matrix to extract the desired signal coming from a specific direction; thus the constraint matrix and vector should be determined such that fulfill a null towards desired direction and unity value towards directional noise(s). In this way, LCMV provides a signal-free input for the IC block in any arbitrary direction of desired signal. In the following, we refer to the above-mentioned structure as NFSD-Linearly Constrained GSC, or briefly NFSD-LCGSC.

## 3. POST-FILTERING

Post-filtering denotes the processing of beamformer output by a single channel noise suppression filter. The post-filtering can significantly improve the SNR and speech quality [7]. When the required statistical and spectral information of the speech and the noise are present, Wiener post-filter (PF) is the optimum PF that provides a Minimum Mean Squared Error (MMSE) estimation of the desired signal. The transfer function of Wiener filter is estimated by:

$$h_{\mathrm{w}} = \varphi_{\mathrm{ss}} / (\varphi_{\mathrm{ss}} + \varphi_{\mathrm{nn}})\ , \tag{12}$$

where $\varphi_{\mathrm{ss}}$ and $\varphi_{\mathrm{nn}}$ are power spectral densities (PSDs) of the (single-channel) desired signal and noise, respectively.

Obviously, the estimates of the signal and noise PSDs are required to formulate the PF transfer function. There are two main approaches for this purpose: 1) noise PSD is estimated from silent frames of input signal and PF coefficients are calculated by an iterative algorithm [1]. 2) speech and noise PSDs are estimated by the use of the auto- and cross-spectral densities of the multi-channel input signals (after the time alignment module) [1,5,6,7]. The latter is referred to as the *Zelinski PF*. PFs are called *single-channel* when the PF coefficients are calculated from output of a beamformer, and called *multi-channel* when the PF coefficients are estimated from array input signals.

### 3.1. Modified McCowan Post-Filter

Under the assumptions: 1) the speech and noise are uncorrelated ($\varphi_{n_i s} = 0, \forall i$), and 2) the noise power spectrum is the same on all microphones ($\varphi_{n_i n_j} = \varphi_{nn}, \forall i, j$), the auto- and cross-spectral densities of the aligned signals can be calculated and updated using Welsh recursive update formula as [1,5,6]:

$$\hat{\varphi}_{x_i x_j} = \alpha \hat{\varphi}'_{x_i x_j} + (1-\alpha) x_i x_j^*\ , \tag{13}$$

where $\hat{\varphi}'_{x_i x_j}$ and $\hat{\varphi}_{x_i x_j}$ are the spectral estimates for the previous and current frames respectively, and * is the complex conjugate operator. Also, $\alpha$ is a factor close to unity.

If array has $M$ microphones, the speech PSD can thus be estimated as [5,6]:

$$\hat{\varphi}_{\mathrm{ss}}^{(ij)} = \frac{\Re\{\hat{\varphi}_{x_i x_j}\} - \frac{1}{2}\Re\{\hat{C}_{ij}\}(\hat{\varphi}_{x_i x_i} + \hat{\varphi}_{x_j x_j})}{1 - \Re\{\hat{C}_{ij}\}}\,,$$

$$\hat{\varphi}_{\mathrm{ss}} = (2/M(M-1)) \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \hat{\varphi}_{\mathrm{ss}}^{(ij)}\ , \tag{14}$$

where $\hat{\varphi}_{x_i x_j}$ is noise cross spectral density between microphones $i$ and $j$ and $\hat{C}_{ij}$ is coherence function of noise field (equation(5)).

The PF denominator ($\varphi_{\mathrm{ss}} + \varphi_{\mathrm{nn}}$) in equation (12), can in turn be estimated by averaging $\hat{\varphi}_{x_i x_i}$ over all unique microphone combinations. The resulting PF is [5,6]:

$$\hat{h}_{\mathrm{pf}} = \frac{\hat{\varphi}_{\mathrm{ss}}}{(1/M)\sum_{i=1}^{M} \hat{\varphi}_{x_i x_i}}\ . \tag{15}$$

Considering equations (14) and (15), coherence function has a fundamental role in the above equations. McCowan presented a PF by modifying the assumption of zero correlation between the noises in different channels [5,6]. McCowan PF is calculated by considering diffuse noise fields (*sinc* function in equation (6)). Although McCowan PF is appropriate in some conditions, but in realistic noise situations (where simultaneously include diffuse, incoherent, and directional noises), the use of *sinc* coherence function is some misleading.

In this research, we present Modified McCowan (or Mod-McCowan) PF by use of the practical coherence function of noise field, that can be calculated and updated in silent frames of input signals (that are specified by means of a Voice Activity Detector (VAD)). It is expected that the PF performance is improved due to a more accurate estimation of coherence function.

### 3.2. Spectral Amplitude MMSE Post-Filters

A major category of the PFs are based on MMSE estimation of Short-Time Spectral Amplitude (STSA). Considering the major importance of the spectral amplitudes, the STSA of the enhanced signal is estimated and combined with the short-time phase of the noisy (input) speech. By using this PF at the output of beamformer ($T(X)$), the spectral amplitude of the clean speech is estimated by [7,8]:

$$\left|\hat{S}\right|_{MMSE-STSA} = R.\Gamma(1.5)\frac{\sqrt{u}}{\gamma}.\exp(\frac{-u}{2}).\left[(1+u)I_0\left(\frac{u}{2}\right)+uI_1\left(\frac{u}{2}\right)\right], \quad (16)$$

where $R = |T(X)|$, $\Gamma$ denotes the gamma function, and $I_0$ and $I_1$ denote the modified Bessel functions of zero and first orders, respectively. Also, $u$ is defined by $u = (\zeta/(1+\zeta)).\gamma$, where $\zeta$ and $\gamma$ are *a priori* and *a posteriori* signal-to-noise ratio of $T(X)$, respectively:

$$\zeta = \varphi_{ss}/\varphi_{nn}, \gamma = R^2/\varphi_{nn}. \quad (17)$$

Generally, in spectral amplitude MMSE, $\gamma$ is calculated from equation (17) using noise PSD estimation; But $\zeta$ is given by a recursive algorithm as [7]:

$$\zeta = \beta.h_{SA}^2(l-1).\gamma(l-1) + (1-\beta).\max(\gamma(l)-1,0), \quad (18)$$

where $h_{SA}$ denotes PF transfer function, $l$ is frame number, $\beta$ is a constant between zero and one and max is the sign of maximization.

In this research, we have considered a multi-channel version of the STSA PF. Here, we use the algorithm described in sub-section 3.1 to estimate power spectral densities of signal and noise (to be used in the calculation of STSA PF). So, the PF coefficients are estimated from array input signals. Also, $\zeta$ is calculated by equation (17) using desired signal PSD that is specified in equation (14) and noise PSD. The latter is estimated from silent frames of input signals and can be updated. The proposed PF is called Prop. STSA.

## 4. EXPERIMENTS AND RESULTS

For the evaluation of the proposed modifications, we have firstly compared the performance of the NFSD-LCGSC with that of basic NFSD-GSC. Then, we have evaluated the effect of different described PFs been at the output of NFSD-LCGSC. Three objective measures have been considered in these evaluations: SSNR (Segmental SNR), LLR (Log Likelihood Ratio) and PESQ (Perceptual Evaluation of Speech Quality) [10]. For the evaluation of methods, we simulated the sources and microphone array setup shown in Fig. 3 (with *d*=4 cm).

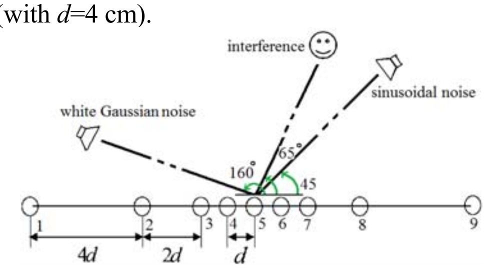


Fig. 3. Array and noise sources situation

According to the broad-band characteristic of speech signal, we have taken advantage of nested sub-array technique [11] for the implementation of the methods. This facilitates appropriate processing on each frequency subband. The implemented sub-array technique divides the broadband signal into three subband signals. A linear uniform array with 5 microphones has been used in each subband as follows:

- $f < 1$ *kHz* : microphones 1, 2, 5, 8, and 9
- 1 *kHz* $< f <$ 2 *kHz* : microphones 2, 3, 5, 7, and 8
- 2 *kHz* $< f <$ 4 *kHz* : microphones 3, 4, 5, 6, and 7

The desired speaker was considered in near-field situation. As shown in Fig. 3, there were three directional noise sources in three different angles with respect to the array axis which are in the far-field: An interfering speaker at the angle of 65°; A white Gaussian noise source at the angle of 160°; and a sinusoidal (3 *kHz*) noise source at 45°. For the desired signal and interferer speaker, we used several segments of speech from the TIMIT database [12]. To make the noisy inputs, we also considered the incoherent and diffuse noises. For modeling incoherent noise, white Gaussian noises with the same amplitude are independently added to each microphone. Also, diffuse noise is generated using equally spaced noise sources, uniformly distributed on the sphere, whose radius is much larger than sensor distance (*d*) [13]. Diffuse, directional, and incoherent noise components were added to each input signal at a proportion of 3, 3 and 1, respectively. All the signals were at the sampling rate of 8 *kHz*.

To examine the effect of the proposed LCMV blocking matrix, the beam-pattern of its first row has been depicted in Fig. 4. The beam-patterns were calculated at the frequency of 1 *kHz* and in the case of different arrival directions of desired signal (0°, 30°, 60°, and 90°).

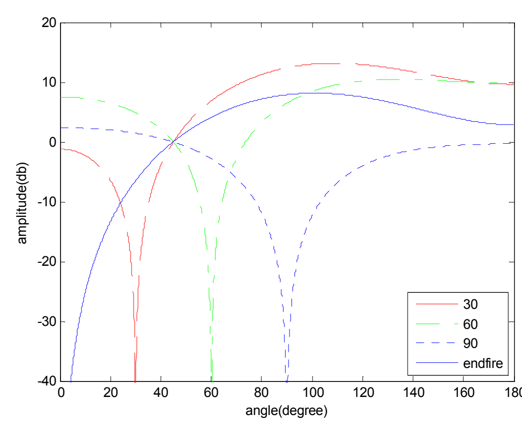


Fig. 4. The beam-pattern of proposed LCMV blocking matrix (first row) in the case of different desired directions

Considering the curves in Fig. 4, the proposed blocking matrix has produced a true null in the direction of desired signal. This guarantees a signal-free input for the IC block, and consequently, no signal cancellation at the output of NFSD-LCGSC. In other word, this has generalized NFSD-GSC for every desired direction without any signal leakage in reference noise.

In the next step, a set of experiments was done to compare the performance of basic NFSD-GSC (in broadside condition) and NFSD-LCGSC (in endfire condition) in various input SNR values (-5 dB, 0 dB, and 5 dB). The comparative results have been listed in Table 1 along with those for the noisy input. The results show that the proposed NFSD-LCGSC is consistently superior to the basic NFSD-GSC in terms of SSNR, LLR and PESQ. This can be justified by considering that in the proposed structure, both upper and lower paths are in their best performance, however, in basic NFSD-GSC only the lower path is in its best performance. It is noted that in this evaluations, both NFSD-GSC and NFSD-LCGSC structures have been in their best working situations.

In the next set of the experiments, we compared the performance of different PFs that were applied to the output of NFSD-LCGSC. Table 2 summarizes the results of these experiments in various input SNR values. The results show that the Mod. McCowan PF is superior to the McCowan PF in terms of SSNR, LLR and PESQ. Furthermore, Mod. McCowan PF has a better performance compared to single-channel Wiener PF in terms of LLR and PESQ; however, it is not better in SSNR measure. Also, while the Prop. STSA PF performs better compared to the STSA PF in terms of SSNR, it is slightly worse in LLR criterion.

## 5. CONCLUSION

In this paper, we introduced a new beamforming structure that incorporates NFSD in upper path of GSC using near-field compensation and generalized this system for different desired directions. The latter was done by replacing the blocking matrix with an LCMV filter. The proposed NFSD-LCGSC obviously outperforms basic NFSD-GSC in the suppression of diffuse, incoherent and directional noises. Furthermore, we proposed some modifications for the post-filters that are applied on the output of NFSD-LCGSC. The evaluation results demonstrate the superiority of the Modified McCowan PF rather than the McCowan PF. Also, the multi-channel Prop. STSA PF showed slightly better performance in comparison with its single-channel version (STSA PF).

## 6. REFERENCES


[1] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*, Berlin, Germany: Springer-Verlag, 2008.

[2] D. Comminiello, M. Scarpiniti, R. Parisi and A. Uncini, “A Novel Affine Projection Algorithm for Superdirective Microphone Array Beamforming,” in *Proc. of 2010 IEEE International Symposium on Circuits and Systems*, Jun. 2010, pp. 2127-2130.

[3] I. McCowan, D. Moore And S. Sridharan, “Speech Enhancement Using Near-Field Superdirectivity with an Adaptive Sidelobe Canceller and Post-Filter,” in *Proc. of 2000 Australian Inernational Coference On Speech Science and Technology*, Dec. 2000, pp. 268-273.

[4] I. McCowan, C. Marro, and L. Mauuary. “Robust Speech Recognition Using Nearfield Superdirective Beamforming with Post-Filtering,” in *Proc. of ICASSP-00*, vol. 3, 2000, pp. 1723–1726.

[5] I. McCowan and H. Bourlard, “Microphone Array Post-Filter Based on Noise Field Coherence,” *IEEE Trans. on Speech and Audio Processing*, vol. 11, no. 6, pp. 709-716, Nov 2003.

[6] S. Lefkimmiatis and P. Maragos, “A Generalized Estimation Approach for Linear and Nonlinear Microphone Array Post-Filters”, *Speech Communication*, vol. 49, pp. 657–666, Feb 2007.

[7] T. Lotter, C. Benien, and P. Vary, “Mutichannel Speech Enhancement Using Bayesian Spectral Amplitude Estimation,” in *Proc. of ICASSP*-03, 2003, pp. I880-I883.

[8] L. H. Kim, M. H. Johnson , and S. Koeng-Mo , “Generalized Optimal Multimicrophone Speech Enhancement Using Sequential Minimum Variance Distortionless Response (MVDR) Beamforming and Post-Filtering,” in *Proc. of ICASSP-06*, vol. 3, July 2006.

[9] M. Souden, J. Benesty, and S. Affes, “A Study of the LCMV and MVDR Noise Reduction Filters,” *IEEE Trans. on Signal Processing,* vol. 58, no. 9, pp. 4925-4935, Sept. 2010.

[10] J. Hansen, B. Pellom, “An Effective Quality Evaluation Protocol for Speech Enhancement Algorithms,” in *Proc. of ICSLP*, Dec. 1998, pp. 2819-2822.

[11] W. Taweesangrungroj, and R. Suleesathira, “Equalizer and Sub-Array Beamformer for Speech Enhancement,” in *Proc. of CACS Automatic Control Conference,* 2005.

[12] J. S. Garofolo, *et al. TIMIT Acoustic-Phonetic Continuous Speech Corpus*, Linguistic Data Consortium, Philadelphia, 1993.

[13] E. Habets, and S. Gannot, “Generating sensor signals in isotropic noise fields”,*url:http://home.tiscali.nl/ehabets/ginfs.html*, 2007.


Table 1: Comparison of basic NFSD-GSC and prop. NFSD-GSC

| | SNR_in = -5 dB | | | SNR_in = 0 dB | | | SNR_in = 5 dB | | |
|---|---|---|---|---|---|---|---|---|---|
| | SSNR | LLR | PESQ | SSNR | LLR | PESQ | SSNR | LLR | PESQ |
| input | -6.41 | 0.69 | 1.95 | -3.24 | 0.52 | 2.02 | -1.01 | 0.31 | 2.22 |
| basic NFSD-GSC | -3.96 | 0.41 | 2.10 | -1.79 | 0.31 | 2.17 | 1.32 | 0.18 | 2.51 |
| NFSD-LCGSC | -2.81 | 0.33 | 2.31 | -0.86 | 0.25 | 2.48 | 2.94 | 0.11 | 2.95 |

Table 2: Evaluation of different post-filters applied to the output of NFSD-LCGSC

| | | NFSD-LCGSC (no PF) | Wiener PF | McCowan PF | Mod. McCowan PF | STSA PF | Prop. STSA PF |
|---|---|---|---|---|---|---|---|
| SNR_in = -5 dB | SSNR | -2.81 | 2.35 | -3.05 | 1.44 | 1.52 | 1.87 |
| | LLR | 0.33 | 0.80 | 0.75 | 0.76 | 0.45 | 0.48 |
| | PESQ | 2.31 | 2.44 | 2.07 | 2.57 | 2.40 | 2.39 |
| SNR_in = 0 dB | SSNR | -0.86 | 3.05 | -0.98 | 3.28 | 3.64 | 3.86 |
| | LLR | 0.25 | 0.74 | 0.58 | 0.59 | 0.41 | 0.43 |
| | PESQ | 2.48 | 2.62 | 2.21 | 2.60 | 2.65 | 2.64 |
| SNR_in = 5 dB | SSNR | 2.94 | 6.82 | 2.31 | 5.67 | 6.33 | 6.45 |
| | LLR | 0.11 | 0.65 | 0.75 | 0.62 | 0.38 | 0.41 |
| | PESQ | 2.95 | 3.11 | 2.65 | 3.09 | 3.14 | 3.10 |