APPROACHES FOR TIME DIFFERENCE OF ARRIVAL ESTIMATION IN A NOISY AND REVERBERANT ENVIRONMENT

Tsvi Gregory Dvorkind and *Sharon Gannot*
Faculty of Electrical Engineering, Technion, Technion City, 32000 Haifa, Israel
e-mail: dvorkind@tx.technion.ac.il, gannot@siglab.technion.ac.il

ABSTRACT

Determining the spatial position of a speaker finds a growing interest in video conference scenario where automated camera steering and tracking are required. As a preliminary step for the localization, microphone array can be used to extract the *time difference of arrival* (TDOA) of the speech signal. The direction of arrival of the speech signal is then determined by the relative time delay between each, spatially separated, microphone pairs. In this work we present novel, frequency domain, approaches for TDOA calculation in a reverberant and noisy environment. Our methods are based on the speech quasi-stationarity property, and on the fact that the speech and the noise are uncorrelated. The proposed methods are supported by an extensive experimental study.

1 Introduction and Model Assumptions

The most common technique for extracting the TDOA is the *generalized cross correlation* (GCC) [1]. However this method assumes that the *acoustical transfer function* (ATF), which relates the source and each of the microphones, is a pure delay. This approximation was shown to be inaccurate in reverberant conditions [2], which frequently occur in enclosed environments. Additionally, in low SNR levels, the GCC method cannot distinguish between the speaker and a directional interference, as it tends to estimate the TDOA of the stronger signal. Approaching our problem from a more realistic point of view, the noisy speech observations are given by:

$$z_m(t) = a_m(t) * s(t) + b_m(t) * n(t);\ m = 1, \ldots, M \tag{1}$$

where $z_m(t)$ is the m-th microphone signal, $s(t)$ is the source signal, and $n(t)$ is a directional interference signal. $a_m(t)$ and $b_m(t)$ are ATF-s relating the desired speech source and the noise interference with the m-th microphone, respectively. M is the number of microphones in the array. We assume that the interfering noise $n(t)$ is stationary and that uncorrelated thermal noise effects can be neglected. Using the more accurate model in (1), several methods were proposed for solving the localization problem under reverberant conditions. Some of these improve the basic GCC method. Others utilize subspace methods for estimating the ATF-s, and then determine the TDOA by finding the difference of their corresponding peaks [3]. In this contribution the ATF-s **ratio** is estimated, rather than the ATF-s themselves.

Let $A_m(\omega)$ and $B_m(\omega)$ be the frequency responses of the m-th ATF $a_m(t)$ and $b_m(t)$, respectively. Define the speech and noise ATF-s ratios by $\mathcal{H}_m(\omega) \triangleq \frac{A_m(\omega)}{A_1(\omega)}$ and $\mathcal{G}_m(\omega) \triangleq \frac{B_m(\omega)}{B_1(\omega)}$ respectively. Denote the inverse Fourier transform of $\mathcal{H}_m(\omega)$ by $h_m(t)$. (Note that since ATF-s are usually not minimum-phase, $h_m(t)$ is non-casual). Experimental study shows that the location of the maximal peak of $h_m(t)$ usually suffices for determining the TDOA. An unbiased method for estimating $H_m(\omega)$, exploiting the speech signal non-stationarity, was presented in [4]. Here, we extend the existing method by imposing decorrelation between the speech and the noise signals.

The organization of the rest of the paper is as follows. In Section 2 we introduce the *decorrelation* criterion as a tool for extracting the TDOA. A method for exploiting the speech *non-stationarity*, contrasted against noise stationarity, is reviewed in Section 3. Batch and recursive solutions for the resulting non-linear set of equations are presented in Sections 4 and 5, respectively. An experimental study is given in Section 6.

2 Decorrelation Criterion

The observations $z_m(t);\ m = 1, \ldots, M$ are a mixture of $s(t)$ and $n(t)$, with the *power spectral density* (PSD) matrix:

$$P_{z_1 z_m}(\omega) \triangleq \begin{bmatrix} \Phi_{z_1 z_1}(\omega) & \Phi_{z_1 z_m}(\omega) \\ \Phi_{z_m z_1}(\omega) & \Phi_{z_m z_m}(\omega) \end{bmatrix} \tag{2}$$

where $\Phi_{z_i z_j} = A_i(\omega)A_j^*(\omega)\Phi_{ss}(\omega) + B_i(\omega)B_j^*(\omega)\Phi_{nn}(\omega)$. $\Phi_{ss}(\omega)$ and $\Phi_{nn}(\omega)$ stand for the speech and noise PSD, respectively. * stands for conjugation. Knowing that $s(t)$ and $n(t)$ are uncorrelated, we wish to construct a decorrelated output. That is, we wish to apply an unmixing matrix $U(\omega)$ which diagonalize the PSD matrix $R(\omega) \triangleq U(\omega)P_{z_1 z_m}(\omega)U^\dagger(\omega)$ ($\dagger$ stands for Hermitian conjugation). Without loss of generality, we can set $U(\omega)$ to the form $U(\omega) = \begin{pmatrix} u_1(\omega) & -1 \\ -u_2(\omega) & 1 \end{pmatrix}$ which yields the (nonlinear) decorrelation criterion:

$$u_2^*(\omega)\left(\Phi_{z_m z_1}(\omega) - u_1(\omega)\Phi_{z_1 z_1}(\omega)\right) = \Phi_{z_m z_m}(\omega) - u_1(\omega)\Phi_{z_1 z_m}(\omega) \tag{3}$$

Eq. (3) is a single nonlinear equation in two unknowns. However, by exploiting the quasi-stationarity of the speech, Eq. (3) becomes a set of equations, obtained by evaluating the PSD-s at different frame indices:

$$u_2^*(\omega)\left(\hat{\Phi}_{z_m z_1}(t,\omega) - u_1(\omega)\hat{\Phi}_{z_1 z_1}(t,\omega)\right) \approx \hat{\Phi}_{z_m z_m}(t,\omega) - u_1(\omega)\hat{\Phi}_{z_1 z_m}(t,\omega)\ ;\ t = 1, \ldots, N \tag{4}$$

where $\hat{\Phi}_{z_i z_j}(t,\omega)$ stands for the appropriate PSD estimation at the t-th frame and N is the number of available frames.

Note that the pair $\{u_2(\omega) = \mathcal{G}_m(\omega), u_1(\omega) = \mathcal{H}_m(\omega)\}$ as well as the pair $\{u_1(\omega) = \mathcal{G}_m(\omega), u_2(\omega) = \mathcal{H}_m(\omega)\}$ solves the equations at hand. This is referred to as the *frequency permutation ambiguity* problem. We will address this problem in the sequel.

Eq. (4), which was derived in [5], constitutes a set of nonlinear equations. In [5] it is suggested to solve the equation set iteratively, by assuming a simplified FIR model for the mixing channels and conducting the solution in the time domain. To maintain simplicity of the solution, we wish to solve the problem in the frequency domain. Thus, the frequency permutation problem, mentioned earlier, have to be addressed.

3 Exploiting Noise Stationarity

The frequency permutation problem can be resolved by proper initialization of (4). This initialization can be obtained by using the method derived in [4], and briefly reviewed here. We try to estimate $\mathcal{H}_m(\omega)$ as the filter relating z_1 and z_m. Consider an analysis interval for which the noise signal can be regarded stationary and the ATF-s time invariant, while the speech signal statistics is changing (quasi-stationarity assumption for the speech signal). The noise signal contributes a bias term. This problem can be avoided by dividing the observation interval into N consecutive frames, resulting an over-determined set of equations for $H_m(\omega)$ and the bias term. This set can be solved by virtue of the *least squares* (LS) method:

$$\hat{\Phi}_{z_m z_1}(t,\omega) = \mathcal{H}_m(\omega)\hat{\Phi}_{z_1 z_1}(t,\omega) + \Phi_{b_1}(\omega) + \xi(k,\omega);\ t = 1,\ldots,N \tag{5}$$

where $\Phi_{b_1}(\omega)$ is a noise-only bias term, independent of the frame index due to its stationarity. Ideally,

$$\Phi_{b_1}(\omega) = (\mathcal{G}_m(\omega) - \mathcal{H}_m(\omega))\,|B_1(\omega)|^2\,\Phi_{nn}(\omega). \tag{6}$$

$\xi(k,\omega)$ are error terms we wish to minimize. The over-determined set in (5) will be called the first form of stationarity (S1 for brevity). Equivalently, a second form of stationarity can be formulated as:

$$\hat{\Phi}_{z_m z_m}(t,\omega) = \mathcal{H}_m(\omega)\hat{\Phi}_{z_1 z_m}(t,\omega) + \Phi_{b_m}(\omega) + \xi_2(k,\omega);\ t = 1,\ldots,N \tag{7}$$

where

$$\Phi_{b_m}(\omega) = (\mathcal{G}_m(\omega) - \mathcal{H}_m(\omega))B_1(\omega){B_m}^*(\omega)\Phi_{nn}(\omega) \tag{8}$$

is a stationary noise-only term as well. Note that Eqs. (6) and (8) are related to the decorrelation equation (3) by setting $u_2^*(\omega) = \frac{\Phi_{b_m}(\omega)}{\Phi_{b_1}(\omega)} = \mathcal{G}_m^*(\omega)$.

4 Batch Solutions for the Nonlinear Decorrelation Equations

Several methods are derived for solving the nonlinear set (4). Here we will address two of them.

4.1 Linear Decorrelation (LD)

Calculate $\Phi_{b_1}(\omega)$ from (5) and $\Phi_{b_m}(\omega)$ from (7). Set $u_2(\omega)^* = \frac{\Phi_{b_m}(\omega)}{\Phi_{b_1}(\omega)}$ and solve the set (4), for evaluating $u_1(\omega)$. This procedure has a twofold advantage. First, using this initialization, the set (4) becomes a **linear** set in $u_1(\omega)$. Thus, LS solution (or, for tracking purposes, a *recursive LS* (RLS) solution) can be applied. Second, by setting $u_2^*(\omega) = \mathcal{G}_m^*(\omega)$, $u_1(\omega)$ is constrained to be $\mathcal{H}_m(\omega)$, thus overcoming the frequency permutation problem.

4.2 Joint Decorrelation and First Form of Stationarity via Gauss iterations (GS1)

A different method for resolving the frequency permutation problem is obtained by solving Eqs. (4) and (5) **simultaneously**. Concatenating these equations we get,

$$\begin{bmatrix} \underline{\hat{\Phi}}_{z_m z_m}(\omega) \\ \underline{\hat{\Phi}}_{z_m z_1}(\omega) \end{bmatrix} \approx \begin{bmatrix} \underline{\hat{\Phi}}_{z_1 z_m}(\omega) & \underline{\hat{\Phi}}_{z_m z_1}(\omega) & -\underline{\hat{\Phi}}_{z_1 z_1}(\omega) & \underline{0} \\ \underline{\hat{\Phi}}_{z_1 z_1}(\omega) & \underline{0} & \underline{0} & \underline{1} \end{bmatrix} \begin{bmatrix} \mathcal{H}_m(\omega) \\ \mathcal{G}_m^*(\omega) \\ \mathcal{H}_m(\omega)\mathcal{G}_m^*(\omega) \\ \Phi_{b_1}(\omega) \end{bmatrix} \tag{9}$$

where $\underline{\hat{\Phi}}_{z_i z_j}(\omega) \triangleq [\hat{\Phi}_{z_i z_j}(1,\omega),\ldots,\hat{\Phi}_{z_i z_j}(N,\omega)]^T$ and $\underline{0}$, $\underline{1}$ stand for column vectors (of proper dimensions) of zeros and ones, respectively. Since the parameter set is nonlinear due to the multiplicative term $\mathcal{H}_m(\omega)\mathcal{G}_m^*(\omega)$, we suggest using several Gauss iterations for obtaining a batch solution.

5 Recursive Solutions for the Nonlinear Decorrelation Equations

In cases where the source is moving, a tracking procedure is required, and a recursive solution for the **nonlinear** LS problem in Eq. (9) is called upon. In Section 5.1 we derive a general procedure for solving the problem, which we denote *recursive Gauss* (RG) procedure. In Section 5.2 we apply this procedure to the problem at hand.

5.1 Recursive Gauss (RG) Procedure

The method starts by resolving the nonlinearities using first-order approximation (as with the original Gauss method), and then deriving a recursion by applying further approximation. This solution will be referred to as *Recursive Gauss* (RG). Consider a nonlinear equation set for the unknown $p \times 1$ parameter vector $\underline{\theta} \in \mathcal{C}^p$:

$$\underline{h}_{1:N}(\underline{\theta}) = \underline{d}_{1:N}$$

where $\underline{h}_{1:N}^T(\underline{\theta}) \triangleq \left[\underline{h}_1^T(\underline{\theta})\cdots\underline{h}_N^T(\underline{\theta})\right]$ and $\underline{d}_{1:N}^T \triangleq \left[\underline{d}_1^T\cdots\underline{d}_N^T\right]$. $\underline{h}_t$ and $\underline{d}_t$ are K nonlinear equations and K measurements, available at time instance t, respectively. Applying first-order approximation around an initial guess $\underline{\theta}^{(0)}$ (as with the Gauss method) we obtain:

$$\underline{h}_{1:N}(\underline{\theta}^{(0)}) + \mathbf{H}_{1:N}(\underline{\theta}^{(0)})\left(\underline{\theta} - \underline{\theta}^{(0)}\right) \approx \underline{d}_{1:N} \tag{10}$$

where $\mathbf{H}_{1:N}$ is the $NK \times p$ gradient matrix:

$$\mathbf{H}_{1:N}^T(\underline{\theta}) \triangleq \left[\mathbf{H}_1^T(\underline{\theta})\cdots\mathbf{H}_N^T(\underline{\theta})\right]$$

where $\mathbf{H}_t(\underline{\theta}) = \nabla_{\underline{\theta}}\underline{h}_t(\underline{\theta})$ is the $K \times p$ gradient matrix of $\underline{h}_t(\underline{\theta})$. According to the Gauss method, the iterative LS solution to the linearized set (10) is:

$$\underline{\theta}^{(l+1)} = \arg\min_{\underline{\theta}} \left\| \underline{d}_{1:N} - \left(\underline{h}_{1:N}(\underline{\theta}^{(l)}) + \mathbf{H}_{1:N}(\underline{\theta}^{(l)})\left(\underline{\theta} - \underline{\theta}^{(l)}\right)\right)\right\|$$

where the superscript denotes the iteration number. Consider the next measurements $\underline{h}_{N+1}(\underline{\theta}) = \underline{d}_{N+1}$ available at time instance $N+1$. In order to estimate $\underline{\theta}$ we use all the available measurements simultaneously. Though we could evaluate all $(N+1)K$ equations at the current estimate

$\underline{\theta}^{(l+1)}$, we do so **only** for the new equations. Namely, instead of minimizing in the LS sense the following residual norm

$$\min_{\underline{\theta}} \left\| \underline{d}_{1:N+1} - \left(\underline{h}_{1:N+1}(\underline{\theta}^{(l+1)}) + \mathbf{H}_{1:N+1}(\underline{\theta}^{(l+1)}) \left(\underline{\theta} - \underline{\theta}^{(l+1)} \right) \right) \right\|$$

we will minimize a modified LS problem

$$\min_{\underline{\theta}} \left\| \begin{matrix} \underline{d}_{1:N} - \left(\underline{h}_{1:N}(\underline{\theta}^{(l)}) + \mathbf{H}_{1:N}(\underline{\theta}^{(l)}) \left(\underline{\theta} - \underline{\theta}^{(l)} \right) \right) \\ \underline{d}_{N+1} - \left(\underline{h}_{N+1}(\underline{\theta}^{(l+1)}) + \mathbf{H}_{N+1}(\underline{\theta}^{(l+1)}) \left(\underline{\theta} - \underline{\theta}^{(l+1)} \right) \right) \end{matrix} \right\|.$$

The reason for this approximation is to keep past solutions intact, thus enabling a recursive solution to be derived. Now, using *stochastic approximation*, i.e. replacing the iteration index by the time index, a sequential algorithm is obtained. To summarize the procedure, an estimate for $\underline{\theta}$ at the current time instance t (denoted by $\underline{\hat{\theta}}(t)$) is obtained by solving the following LS problem sequentially using the *recursive LS* (RLS) procedure:

$$\underline{\hat{\theta}}(t) = \arg\min_{\underline{\theta}} \left\| \begin{bmatrix} \mathbf{H}_1(\underline{\hat{\theta}}(0)) \\ \vdots \\ \mathbf{H}_t(\underline{\hat{\theta}}(t-1)) \end{bmatrix} \underline{\theta} - \underline{y}_{1:t} \right\| \tag{11}$$

where

$$\underline{y}_{1:t} = \begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_t \end{bmatrix} \triangleq \begin{bmatrix} \underline{d}_1 - \underline{h}_1(\underline{\hat{\theta}}(0)) + \mathbf{H}_1(\underline{\hat{\theta}}(0))\underline{\hat{\theta}}(0) \\ \vdots \\ \underline{d}_t - \underline{h}_t(\underline{\hat{\theta}}(t-1)) + \mathbf{H}_t(\underline{\hat{\theta}}(t-1))\underline{\hat{\theta}}(t-1) \end{bmatrix}$$

with $\underline{\hat{\theta}}(0)$ the initial estimate for the parameter set. We note that in many practical situations the parameter set $\underline{\theta}$ might slowly vary with time. In these cases, a common practice is to apply the RLS algorithm with a diagonal weight matrix that uses a forgetting factor $0 < \alpha \leq 1$ to weight past equations.

Another practical issue concerns the computational burden. At each time instance new K equations become available, resulting a $K \times K$ matrix inversion at each RLS iteration. However, by properly varying the forgetting factor α, the computational complexity can be further reduced. This procedure is beyond the scope of this paper.

5.2 Application of the Recursive Gauss to the Decorrelation Equations

For the problem at hand (the GS1 procedure) $\underline{\theta} \triangleq [\mathcal{H}_m(\omega), \mathcal{G}_m^*(\omega), \Phi_{b_1}(\omega)]^T$. Considering the the t-th time instance, for which the gradient matrix $\mathbf{H}_t$ is given by (12) and the measurement $\underline{y}_t$ by (13)

$$\underline{y}_t = \begin{bmatrix} \hat{\Phi}_{z_m z_m}(t,\omega) - \hat{\mathcal{H}}_m(t-1,\omega)\hat{\mathcal{G}}_m^*(t-1,\omega)\hat{\Phi}_{z_1 z_1}(t,\omega) \\ \hat{\Phi}_{z_m z_1}(t,\omega) \end{bmatrix} \tag{13}$$

where $\hat{\mathcal{H}}_m(t-1,\omega)$ and $\hat{\mathcal{G}}_m^*(t-1,\omega)$ are the estimations of $\mathcal{H}_m(\omega)$ and $\mathcal{G}_m^*(\omega)$ available after $t-1$ measurements, respectively. Then $\underline{\theta}$ is evaluated by solving (11) with RLS and a forgetting factor $\alpha < 1$.

6 Experimental Study

In this section we assess the proposed methods (S1, LD, GS1) and compare them with the classical GCC algorithm [1]. Two scenarios are tested. First, batch methods are compared for a static speaker position. Then, the tracking ability of the recursive methods is demonstrated for a moving speaker scenario.

6.1 Static Scenario

For the static scenario we used room dimensions of $[4, 7, 2.75]$ and noise source positioned at $[1.5, 4, 2.08]$ (all dimensions are in meters). PSD estimations are made by the Welch method using 256 samples long frames, with 50% overlap and with the Hann window. The sampling frequency is $F_s = 8000$[Hz]. The speech source is placed at $[2.53, 4.03, 2.67]$ and microphone pair are placed at $[2, 3.5, 1.375]$ and $[1.7, 3.5, 1.375]$. Different reverberation times and different SNR values are tested. The ATF-s are simulated using the image method. Speech segments, 1.76[sec] long, are drawn from the TIMIT database and the overall Monte-Carlo simulation is conducted over more than 5 minutes of speech. The noise source is the speech-like noise drawn from the NOISEX-92 database. The signals are filtered by the simulated ATF-s and summed at different SNR values to construct the measurements $z_1(t)$ and $z_2(t)$. For all the evaluated methods sub-sample TDOA calculation (for $\frac{1}{10}$[sample] resolution) is performed using Shanon interpolation. The percentage of anomalies is calculated (where anomaly is defined as divergence of more than 2 samples from the true TDOA). Only non-anomalous estimates are involved in calculating the *root mean square error* (RMSE). The results are summarized in Table 1. As can be seen, at

T_r**[sec]**	**SNR[dB]**	**S1**	**LD**	**GS1**	**GCC**
		Anomaly percentage			
0.10	5	1	0	0	16
0.50	5	2	5	5	65
0.25	5	0	0	0	20
0.25	0	4	3	3	98
0.25	−5	52	37	24	100
		RMSE [sample]			
0.10	5	0.06	0.06	0.06	0.07
0.50	5	0.15	0.15	0.14	0.13
0.25	5	0.09	0.10	0.10	0.09
0.25	0	0.12	0.12	0.13	0.07
0.25	−5	0.12	0.12	0.15	–

Table 1: Comparison of batch methods.

low reverberation conditions ($T_r = 0.1$[sec]) and high SNR (5[dB]) all methods perform well (this might exclude the GCC method that even at these mild conditions has 16% anomaly). When we test severe (for a moderate size room) reverberation of $T_r = 0.5$[sec], even in the high SNR level, the GCC method, which lacks the reverberant model, rapidly deteriorate in performance. On the other hand, the proposed methods present low anomaly results. This is also the case at mid-range reverberation $T_r = 0.25$[sec] and at lower SNR conditions. Note that at low SNR levels, LD and GS1 outperform S1. Furthermore, at low SNR conditions the GCC becomes useless, since it locks on the directional interference TDOA instead of the speaker TDOA. Evaluation of the RMSE (for the non anomalous experiments) reveals that the TDOA is extracted with high accuracy.

$$\mathbf{H}_t(\hat{\underline{\theta}}(t-1)) = \begin{bmatrix} \hat{\Phi}_{z_1 z_m}(t,\omega) - \hat{\mathcal{G}}_m^*(t-1,\omega)\hat{\Phi}_{z_1 z_1}(t,\omega), & \hat{\Phi}_{z_m z_1}(t,\omega) - \hat{\mathcal{H}}_m(t-1,\omega)\hat{\Phi}_{z_1 z_1}(t,\omega), & 0 \\ \hat{\Phi}_{z_1 z_1}(t,\omega), & 0, & 1 \end{bmatrix} \tag{12}$$

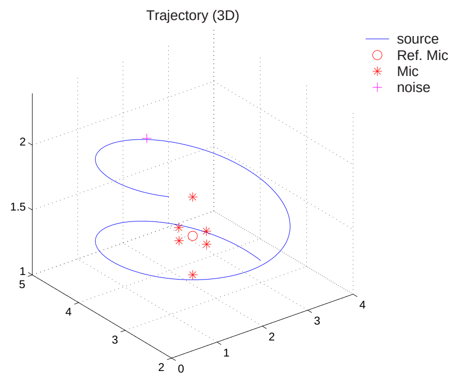

Figure 1: Speaker trajectory

6.2 Tracking Scenario

We proceed by discussing the tracking scenario in which the speaker is moving. Room dimensions and the noise source position are as in the static scenario. The speaker trajectory is set to an helix with radius $R = 1.5$[m] around the reference microphone, at movement speed of 0.5[m/s] and for a total movement time of $T = 30$[sec]. The speaker Cartesian position as a function of time $t \in [0, T]$ is,

$$x(t) = 2 + R\cos(2\pi ft), \quad y(t) = 3.5 + R\sin(2\pi ft)$$
$$z(t) = 1 + \frac{t}{T}$$

with $f = 0.0529[Hz]$. This trajectory is depicted in Fig. 1. The TDOA extraction procedures are the same as in the static scenario. However, for the proposed methods, we now solve the LS problem recursively with a forgetting factor smaller than one. Sampling every $3.75[cm]$ along the speaker trajectory, the ATF-s between the speaker and the microphones are simulated using the image method and used to filter the speech. The mean SNR for the 30[sec] long signal is set to 10[dB]. A forgetting factor of 0.824 is used in the the RLS form of S1, LD and GS1 procedures, to allow tracking of (a slowly changing) $\mathcal{H}_m(\omega)$. TDOA estimation results, with respect to the microphone pair placed at $[2, 3.5, 1.375]$ and $[2.3, 3.5, 1.375]$ are presented in Fig. 2 for the GCC and the recursive forms of LD and GS1. The recursive form of S1 has comparable performance with the LD and GS1. As can be seen from Fig. 2 the GCC method tends to lock on the noise position (note, that the directional noise TDOA is approximately 4[samples]). In contrast, most of the time the proposed methods manage to track the changes in the speaker TDOA. We note however that there are time instances where wrong TDOA selection is made and that the memory of the RLS based algorithms causes slight divergence of the estimated track from the real trajectory. Nevertheless, the obtained performance is significantly superior.

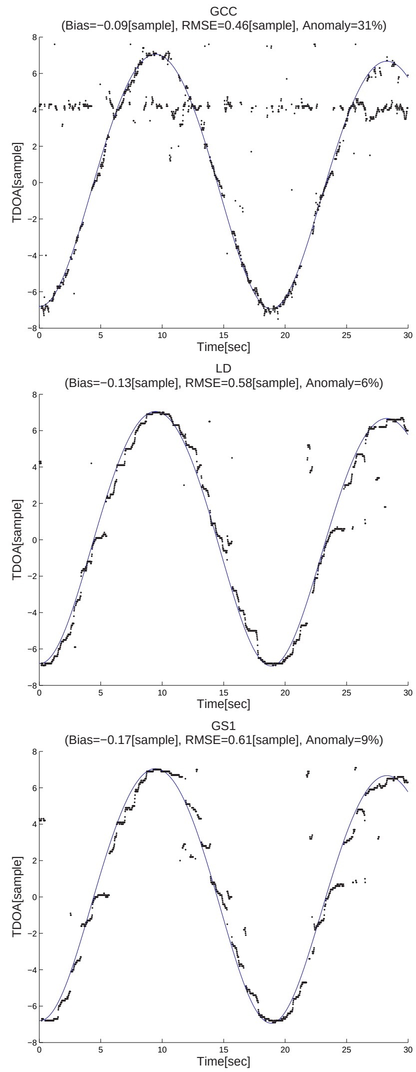

Figure 2: TDOA estimation results. Solid line: True TDOA. Dots: Estimation results. Top to bottom: GCC, LD, GS1.

7 *

References

[1] C.H. Knapp and G.C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[2] B. Champagne S. Bédard and A. Stéphenne, "Performance of time-delay estimation in the presence of room reverberation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 4, no. 2, pp. 148–152, 1996.

[3] S. Doclo and M. Moonen, "Robust time-delay estimation in highly adverse acoustic environments," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 59–62, October 2001.

[4] T. Dvorkind and S. Gannot, "Speaker Localization in a Reverberant Environment," *IEEE proceedings, The 22nd convention of Electrical and Electronics Engineers in Israel.*, pp. 7–9, Dec. 2002.

[5] E. Weinstein, M. Feder, and A. V. Oppenheim, "Multi-Channel Signal Seperation by Decorrelation," *IEEE transactions on Speech and Audio Processing*, vol. 1, no. 4, pp. 405–413, Oct. 1993.