Silence period: While in the silent period of speech, we assume $s_k = 0$. We also assume that the LMS algorithm converges in the silence period, hence the output error function $e_k \to 0$. When the noise sources n_k^1 and n_k^2 are nonzero, from eqn. 1 we have the two equations below:

$$(H_{21} - W_1 H_{22} - W_2 H_{23}) = 0 \tag{2}$$

and

$$(H_{31} - W_1 H_{32} - W_2 H_{33}) = 0 \tag{3}$$

The two transfer functions W_1, W_2, and the acoustic transfer functions can be written in matrix form accordingly

$$\begin{bmatrix} H_{22} & H_{23} \\ H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} H_{21} \\ H_{31} \end{bmatrix} \tag{4}$$

Provided the LHS of eqn. 4 is non-singular ($H_{22}H_{33} - H_{23}H_{32} \neq 0$), the weight vectors W_1 and W_2 have the unique solution

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} H_{22} & H_{23} \\ H_{32} & H_{33} \end{bmatrix}^{-1} \begin{bmatrix} H_{21} \\ H_{31} \end{bmatrix} \tag{5}$$

Of course, in practice a matrix inversion is not required since W_1 and W_2 arise naturally from the convergence of the LMS algorithm.

Speech period: During a speech period, $s_k \neq 0$, and the weight functions W_1, W_2 are no longer updated and are retained unchanged. Substituting the solutions obtained from eqn. 5 into eqn. 1, the output error function becomes

$$e_k = (H_{11} - W_1 H_{12} - W_2 H_{13}) s_k \tag{6}$$

where W_1 and W_2 are the solutions of eqn. 5 obtained from the LMS algorithm. From eqn. 6, it is obvious that the error function e_k has no term in n_k^1 or n_k^2, indicating that the noise terms vanish. It has been found in practice that the transfer function ($H_{11} - W_1H_{12} - W_2H_{13}$) in eqn. 6 does not offer any significant degradation of the perceived speech quality. The importance of such an algorithm is that it adapts to the acoustic transfer functions and not to the signal or noise statistics.

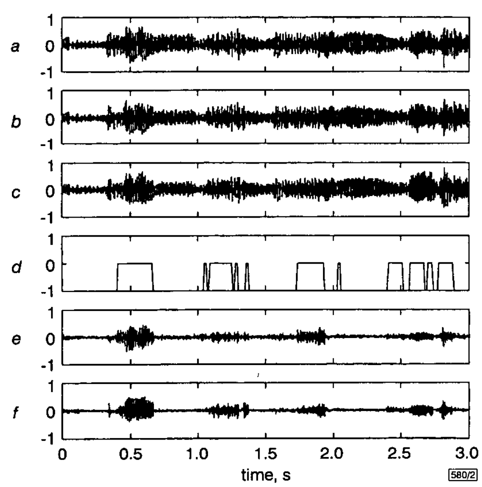

Fig. 2 *Experimental outputs of real data*

a Signal received by microphone a
b Signal received by microphone b
c Signal received by microphone c
d Detected speech period and silence period of desired speech
e Output signal after two-input adaptive algorithm
f Signal after three-input adaptive algorithm

Example of real data experiment: A sentence spoken by one of the authors from the 'active zone' and stereo background music from two speakers from outside the zone were recorded simultaneously using three microphones with the specified geometry introduced in the earlier section. The sampling frequency of all the signal sources was 25.6kHz. The three-channel recorded signals were highpass-filtered with a cutoff frequency of 120Hz in order to remove the embedded 50Hz AC power noise from the amplifier. The SNR average over three microphones of the speech before processing was -4.7dB, whilst the SNRs were 8.9 and 12.4dB, respectively, after the two-input and three-input adaptive algorithms. Fig. 2 shows the speech signals of the three channels before processing, word boundary detection and the results after two-input and three-input adaptive noise cancellation.

Conclusions: This Letter has shown the performance of a three-microphone adaptive algorithm together with word boundary detection. The word boundary detection relies on an active zone where speech and noise are detected. The combination of this word boundary detection algorithm with a more conventional noise-cancellation algorithm based on LMS was then applied to a real data experiment. It was demonstrated that the three-microphone approach has an advantage over the two-microphone case. This dual algorithm approach of three-microphone word boundary detection with noise cancellation should prove useful in a wide range of engineering applications.

28 September 1999
Electronics Letters Online No: 19991357
DOI: 10.1049/el:19991357

W.N. Chen and T.J. Moir (*Department of Electronic Engineering and Physics, University of Paisley, Paisley, PA1 2BE, United Kingdom*)

References

1 CHEN, W.N., and MOIR, T.J.: 'Multiple microphone adaptive noise cancellation using a word boundary speech detector'. ICT'99 Proc. Int. Conf. Telecommun., Cheju, Korea, 15–18 June 1999, Vol. 2, pp. 184–188

2 AGAIBY, H., and MOIR, T.J.: 'Knowing the wheat from the weeds in noisy speech'. Proc. 5th European Conf. Speech Communication and Technology, Rhodes, Greece, 22–25 September 1997, Vol. 3, pp. 1119–1122

3 KNAPP, C.H., and CARTER, G.C.: 'The generalised correlation method for estimation of time delay', *IEEE Trans.*, 1976, **ASSP-24**, (4), pp. 320–327

Angle estimator for signals with known waveforms

Hongbin Li, Guoqing Liu and Jian Li

A large-sample decoupled maximum likelihood angle estimator for signals with known waveforms is presented based on exploiting the *a priori* knowledge that the observation noise can be modelled as spatially white. It is shown that the incorporation of this knowledge significantly improves the angle estimation accuracy over existing angle estimators for signals with known waveforms.

Introduction: Most existing high-resolution angle-of-arrival estimation algorithms, including MUSIC, ESPRIT, and MODE, do not assume any knowledge of the incident signals except for some general statistical properties such as the second-order ergodicity [1]. Recently, there has been growing interest in developing angle estimators that exploit some *a priori* knowledge, e.g. the known waveforms, of the incident signals. Such estimators can be used in various applications including wireless communications where known preamble sequences are often transmitted for training purposes. An interesting algorithm in this category is the decoupled maximum likelihood (DEML) angle estimator for signals with known waveforms [2]. It is an efficient large sample maximum likelihood (ML) method. Although DEML was derived only for uncorrelated signals, an extension of DEML, referred to as coherent decoupled maximum likelihood (CDEML), was made in [3] to handle coherent signals. It has been found that these estimators provide significantly more accurate angle estimation than conventional estimators including MUSIC, ESPRIT, and MODE. Moreover, while most conventional angle estimators require that either the observation noise be spatially white or that the spatial noise covariance matrix be known to within a constant value [1], DEML and CDEML were specifically designed to deal with spatially coloured observation noise.

However, neither DEML nor CDEML are optimal when the observation noise can be reasonably modelled as spatially white. In this Letter, we present a new angle estimator for signals with known waveforms (but unknown gains) which are corrupted by spatially white observation noise. The resulting estimator is referred to as white decoupled maximum likelihood (WDEML) since, similarly to DEML, it is a decoupled large-sample ML method that assumes spatially white noise. We show that WDEML achieves a similar performance to the optimal exact ML estimator, but at a significantly reduced computational complexity.

Problem formulation: We consider the problem of estimating the arrival angles of K signals impinging on an array of M sensors. The array output vector can be written as [1 – 4]

$$\mathbf{x}(n) = \mathbf{A}(\theta)\mathbf{s}(n) + \mathbf{n}(n) \qquad n = 1, ..., N \qquad (1)$$

where $\mathbf{x}(n) \in \mathbb{C}^{M\times 1}$ is the array output vector, $\mathbf{s}(n) \in \mathbb{C}^{K\times 1}$ is the signal vector, $\mathbf{n}(n) \in \mathbb{C}^{M\times 1}$ is the noise vector, and N denotes the number of snapshots. The matrix $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{M\times K}$ is the array manifold matrix corresponding to the angle vector $\boldsymbol{\theta} \in \mathbb{R}^{K\times 1}$. The noise vector $\mathbf{n}(n)$ is circularly symmetric zero-mean complex Gaussian with $E[\mathbf{n}(n_i)\mathbf{n}^H(n_j)] = \sigma^2\mathbf{I}_M\delta_{i,j}$, where $\mathbf{I}_M$ denotes the $M \times M$ identity matrix and $\delta_{i,j}$ denotes the Kronecker delta. It is assumed that the waveforms of $\mathbf{s}(n)$ are known but with unknown complex gains. Specifically, $\mathbf{s}(n) = \boldsymbol{\Gamma}\mathbf{y}(n)$, where $\mathbf{y}(n) = [y_1(n), ..., y_L(n)]^T$ denotes the known waveform vector and $\boldsymbol{\Gamma} \in \mathbb{C}^{K\times L}$ denotes the unknown gain matrix, which has the form: $\boldsymbol{\Gamma} = \mathrm{diag}\{\boldsymbol{\Gamma}_1, ..., \boldsymbol{\Gamma}_L\}$, with $\boldsymbol{\Gamma} = [\gamma_{l,1}, ..., \gamma_l,K_l]^T$. Here, K_l denotes the number of the incident signals corresponding to the lth waveform $\{y_l(n)\}_{n=1}^N$ and $\Sigma_{l=1}^L K_l = K$. Hence, $K_l > 1$ implies coherent multipaths for the lth waveform. We next partition $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, ..., \boldsymbol{\theta}_L^T]^T$, where $\boldsymbol{\theta}_l = [\theta_{l,1}, ..., \theta_l,K_l]^T$, each element denoting the arrival angle of one multipath component corresponding to the lth waveform. Similarly, we let $\boldsymbol{\gamma} = [\boldsymbol{\gamma}_1^T, ..., \boldsymbol{\gamma}_L^T]^T$. We assume that the L waveforms $\{y_l(n)\}_{l=1}^L$ are uncorrelated so that the covariance matrix, $\mathbf{R}_{yy} \triangleq \lim_{N\to\infty}(1/N)\Sigma_{n=1}^N \mathbf{y}(n)\mathbf{y}^H(n)$, is diagonal. We also assume that the waveform and noise vectors are uncorrelated. The problem of interest is to determine the arrival angles $\boldsymbol{\theta}$ and the complex gains $\boldsymbol{\gamma}$ from the measurements $\{\mathbf{x}(n)\}_{n=1}^N$.

WDEML estimator: It is straightforward to show that maximising the likelihood function of $\{\mathbf{x}(n)\}_{n=1}^N$ with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ is equivalent to minimising

$$L_1(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \frac{1}{N}\mathrm{tr}\left\{\sum_{n=1}^{N}[\mathbf{x}(n) - \mathbf{B}\mathbf{y}(n)][\mathbf{x}(n) - \mathbf{B}\mathbf{y}(n)]^H\right\} \qquad (2)$$

where $\mathbf{B} \triangleq \mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma} = [\mathbf{A}(\boldsymbol{\theta}_1)\boldsymbol{\gamma}_1, ..., \mathbf{A}(\boldsymbol{\theta}_L)\boldsymbol{\gamma}_L]$. We let $\hat{\mathbf{R}}_{yx} \triangleq (1/N)\Sigma_{n=1}^N \mathbf{y}(n)\mathbf{x}^H(n)$, and $\hat{\mathbf{R}}_{yy}$ and $\hat{\mathbf{R}}_{xx}$ be similarly defined. We can show that minimising L_1 with respect to $\mathbf{B}$ yields the unstructured estimate $\hat{\mathbf{B}} = \hat{\mathbf{R}}_{yx}^H \hat{\mathbf{R}}_{yy}^{-1}$. We next rewrite the cost function L_1 as

$$L_1(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \mathrm{tr}\{\hat{\mathbf{R}}_{\mathbf{xx}} - \hat{\mathbf{B}}\hat{\mathbf{R}}_{\mathbf{yy}}\hat{\mathbf{B}}^H\} + \mathrm{tr}\{(\mathbf{B} - \hat{\mathbf{B}})\hat{\mathbf{R}}_{\mathbf{yy}}(\mathbf{B} - \hat{\mathbf{B}})^H\} \qquad (3)$$

Since the first term of eqn. 3 is independent of $\mathbf{B}$, minimising L_1 reduces to minimising the second term of eqn. 3. We observe that $\hat{\mathbf{R}}_{yy}$ and $\hat{\mathbf{B}}$ are consistent estimates of $\mathbf{R}_{yy}$ and $\mathbf{B}$, respectively. Thus, replacing $\hat{\mathbf{R}}_{yy}$ by $\mathbf{R}_{yy}$ in the second term of eqn. 3 will not affect its asymptotic properties. Minimising L_1 is asymptotically (for large N) equivalent to minimising

$$L_2(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \mathrm{tr}\{\mathbf{R}_{\mathbf{yy}}(\mathbf{B} - \hat{\mathbf{B}})^H(\mathbf{B} - \hat{\mathbf{B}})\} \qquad (4)$$

Since $\mathbf{R}_{yy}$ is diagonal, the minimisation of L_2 decouples into L minimisation problems:

$$\{\hat{\theta}_l, \hat{\gamma}_l\} = \arg\min_{\theta_l, \gamma_l} \|[\mathbf{A}(\theta_l)\gamma_l - \hat{\mathbf{b}}_l]\|^2 \qquad l = 1, ..., L \quad (5)$$

where $\hat{\mathbf{b}}_l$ is the lth column of $\hat{\mathbf{B}}$ and $\|\cdot\|$ denotes the Euclidean norm. The solution to eqn. 5 is easy to obtain:

$$\hat{\gamma}_l = [\mathbf{A}^H(\hat{\theta}_l)\mathbf{A}(\hat{\theta}_l)]^{-1}\mathbf{A}^H(\hat{\theta}_l)\hat{\mathbf{b}}_l \qquad (6)$$

$$\hat{\theta}_l = \arg\min_{\theta_l} \left\|\{\mathbf{I} - \mathbf{A}(\theta_l)[\mathbf{A}^H(\theta_l)\mathbf{A}(\theta_l)]^{-1}\mathbf{A}^H(\theta_l)\}\hat{\mathbf{b}}_l\right\|^2 \qquad (7)$$

The minimisation of eqn. 7 can be performed by, for example, searching over a K_l-dimensional parameter space. If a uniform linear array (ULA) is employed, then the K_l-dimensional optimisation is asymptotically equivalent to solving a polynomial rooting problem [2]. To avoid matrix singularity, we must have $N \geq \max\{M, K\}$ for DEML [2] and $N \geq \max\{M, K\}$ for CDEML [3]. For WDEML, however, we only need $N > L$ (so that $\hat{\mathbf{R}}_{yy}$ has full rank).

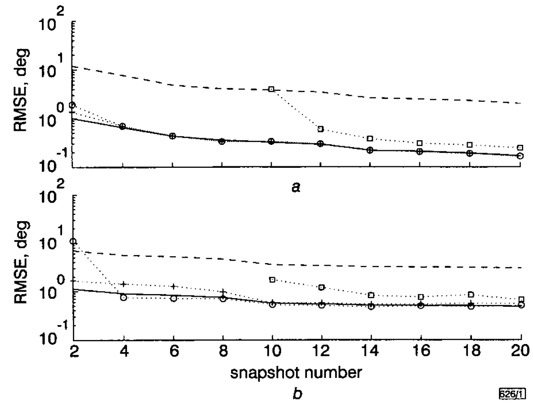

Fig. 1 *RMSEs and CRBs against N*

SNR = 5dB and $M = 10$
a $\theta_3 = 5°$, uncorrelated
□ CDEML/DEML
+ WDEML
○ EXACT-ML
——— CRB-known waveform
– – – – CRB-unknown waveform
b $\theta_1 = 0°$, coherent
□ CDEML
+ WDEML
○ EXACT-ML
——— CRB-known waveform
– – – CRB-unknown waveform

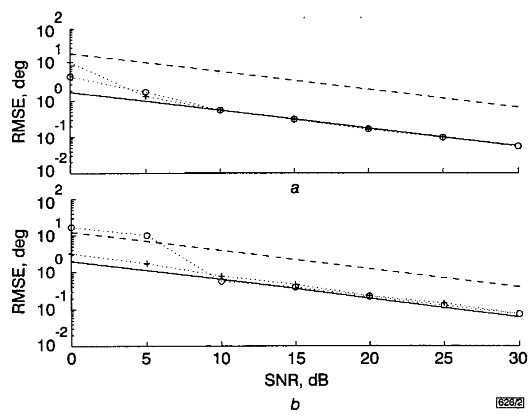

Fig. 2 *RMSEs and CRBs against SNR*

$N = 2$ and $M = 10$
a $\theta_3 = 5°$, uncorrelated
b $\theta_1 = 0°$, coherent
+ WDEML
○ EXACT-ML
——— CRB-known waveform
– – – – CRB-unknown waveform

Numerical examples: We considered three signals with equal power from $\theta_1 = 0°$, $\theta_2 = 10°$, and $\theta_3 = 5°$ impinging on a ULA of $M = 10$ sensors separated by a half wavelength. The first two signals were assumed to be coherent and the third signal was assumed to be uncorrelated with the first two. The two known waveforms were zero-mean complex Gaussian sequences that were temporally white and uncorrelated with each other. Four angle estimators for signals with known waveforms, namely WDEML, DEML, CDEML, and the exact ML method (EXACT-ML), were examined. The EXACT-ML method was implemented in an iterative manner, as in [4]. Since all methods approach the corresponding Cramér-Rao bound (CRB) for sufficiently large N, we considered the more difficult cases of small N and/or low SNR.

Fig. 1 shows the results when SNR = 5dB and N varied from 2 to 20. The CRB with and without the knowledge of the waveforms was plotted to demonstrate the merits of incorporating the knowledge of the waveforms. For θ_3 (the uncorrelated signal), DEML and CDEML are identical, whereas for θ_1 (the coherent signal), DEML fails and was not considered. As explained before, the CDEML/DEML methods produce valid angle estimates only when $N \geq 10$ in this example. While the CDEML/DEML angle estimates deviate from the CRB with known waveforms noticeably at $N = 20$, WDEML is observed to approach the CRB starting from $N = 2$. For θ_1, WDEML appears to work even better than EXACT-ML at $N = 2$. This is because the iterative EXACT-ML failed to converge to the global minimum of its cost function.

We then let $N = 2$, the smallest snapshot number required by WDEML for this problem, and let the SNR vary from 0 to 30dB. All the other parameters were the same as in the preceding example. Both DEML and CDEML failed with $N = 2$, so only WDEML and EXACT-ML were considered. The results are shown in Fig. 2. Note that both methods approach the CRB with known waveforms for SNR ≥ 10dB. Note also that the EXACT-ML estimates could be worse than the WDEML estimates due to the local convergence of EXACT-ML at low SNR.

29 September 1999

Electronics Letters Online No: 19991389
DOI: 10.1049/el:19991389

Hongbin Li (*Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA*)

E-mail: hli@stevens-tech.edu

Guoqing Liu and Jian Li (*Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA*)

References

1 OTTERSTEN, B., VIBERG, M., STOICA, P., and NEHORAI, A.: 'Exact and large sample ML techniques for parameter estimation and detection in array processing' *in* HAYKIN, S., *et al.* (Eds.): 'Radar array processing' (Springer-Verlag Inc., New York, NY, 1993) Chap. 4

2 LI, J., HALDER, B., STOICA, P., and VIBERG, M.: 'Computationally efficient angle estimation for signals with known waveforms', *IEEE Trans. Signal Process.*, 1995, **43**, pp. 2154–2163

3 CEDERVALL, M., and MOSES, R.L.: 'Efficient maximum likelihood doa estimation for signals with known waveforms in the presence of multipath', *IEEE Trans. Signal Process.*, 1997, **45**, pp. 808–811

4 LI, J., and COMPTON, R.T.: 'Maximum likelihood angle estimation for signals with known waveforms', *IEEE Trans. Signal Process.*, 1993, **41**, pp. 2850–2862

Comparison of DDE and ETDGE for time-varying delay estimation

H.C. So

The performances of DDE and ETDGE, which are two recently proposed methods for the direct estimation of time delay between signals received at two spatially separated sensors, are compared. Although both algorithms are computationally efficient, it is shown that the ETDGE method generally outperforms the DDE method for tracking nonstationary delays with different source signals.

Introduction: The problem of estimating and tracking the time delay between signals received at two spatially separated sensors arises in many application fields such as sonar, radar and seismology [1]. Let the two sensor outputs be represented by

$$\begin{aligned} x(k) &= s(k) + n_1(k) \\ y(k) &= s(k-D) + n_2(k) \end{aligned} \tag{1}$$

where $s(k)$ is the unknown source signal, $n_1(k)$ and $n_2(k)$ are the uncorrelated white Gaussian noise sources which are statistically independent of $s(k)$, and D is the differential delay to be determined. Without loss of generality, it is assumed that the signal and noise spectra are bandlimited between –0.5 and 0.5Hz while the sampling period is 1s.

Based on the property that a time-shifted version of a bandlimited signal can be expressed as the convolution of a sinc function and the signal itself, Chan *et al.* [2] introduced a parameter estimation approach to model the time delay as an FIR filter, $W(z) = \Sigma_{i=-P}^{P} w_i z^{-i}$, in one of the receiver channels. Once the filter coefficients are estimated, the time difference of arrival is found by interpolating their values. For time-varying delay estimation, this approach can be made adaptive by adjusting the filter weights according to Widrow's least mean square (LMS) algorithm [3]. Recently, two LMS-style algorithms, the direct delay estimator (DDE) [4] and the explicit time delay and gain estimator (ETDGE) [5], were proposed to provide direct delay measurements and their computational complexities are much lower than the computational complexity of the algorithm presented in [3] because they do not involve the interpolation of filter weights. Basically, the DDE uses the largest filter weight and one of its adjacent coefficients to compute the delay estimate while the filter coefficients are expressed as a function of the delay estimate and a gain factor in the ETDGE. In this Letter, we compare the DDE and ETDGE in terms of delay convergence rates and variances as well as computational requirements.

Performances of DDE and ETDGE: In the DDE, the filter weights $\{w_i(k)\}$, $i = -P, -P+1, ..., P$, are adapted iteratively to minimise the mean square output error $E\{e^2(k)\}$ subject to $\Sigma_{i=-P}^{P} w_i^2(k) = 1$ as follows [4]:

$$w_i(k+1) = \left(1 + \mu_d \hat{\sigma}_n^2(k)\right) w_i(k) + \mu_d e(k) x(k-i) \tag{2}$$

where $e(k) = y(k) - \Sigma_{i=-P}^{P} w_i(k)x(k-i)$ and $\hat{\sigma}_n^2(k) = \beta\hat{\sigma}_n^2(k-1) + (1-\beta)(x^2(k) - y(k)\Sigma_{i=-P}^{P} w_i(k)x(k-i))$. The positive scalar μ_d controls the convergence rate and stability of the algorithm while $\hat{\sigma}_n^2(k)$ represents the estimate of σ_n^2 which is the power of $n_1(k)$ or $n_2(k)$ and $\beta \in (0, 1]$ is a smoothing factor. We denote the peak weight at time k by $w_L(k)$; the delay estimate of the DDE, $\hat{D}_d(k)$, is calculated as

$$\hat{D}_d(k) = L + (A - L)\frac{w_A(k)}{w_L(k) + w_A(k)} \tag{3}$$

where $w_A(k) = \max\{w_{L-1}(k), w_{L+1}(k)\}$. Assuming that $s(k)$ is a white process with variance σ_s^2 and taking the expectation of eqn. 2, we obtain

$$E\{w_i(k)\} = \text{sinc}(i-D) + (w_i(0) - \text{sinc}(i-D))(1-\mu_d\sigma_s^2)^k \tag{4}$$

where $\{w_i(0)\}$ represent the initial values of the filter weights. By substituting eqn. 4 into eqn. 3, we can obtain the learning trajectory of $\hat{D}_d(k)$. For $\mu_d\sigma_s^2 << 1$, the delay variance of the DDE, $\text{var}(\hat{D}_d)$, is given by

$$\text{var}(\hat{D}_d) \simeq \frac{\mu_d\sigma_s^2(2(1-2|D_i|+2D_i^2)(1+SNR)+((1-|D_i|)^2 w_A^{o2} + D_i^2 w_L^{o2}))}{2(w_L^o + w_A^o)^2 SNR^2} \tag{5}$$

where $D_i = (L-D) \in (-0.5, 0.5)$, $w_i^o = \text{sinc}(i-D)$, $i = A, L$, and $SNR = \sigma_s^2/\sigma_n^2$. Fig. 1 shows $\text{var}(\hat{D}_d)$ against $|D_i|$ at different SNRs with $\mu_d\sigma_s^2 = 0.004$. It can be seen that $\text{var}(\hat{D}_d)$ decreases monotonically with increasing $|D_i|$ or SNR. For each sampling interval, $(6P+7)$ additions and $(6P+10)$ multiplications are required for the DDE algorithm.

In the ETDGE, the filter weights $\{w_i(k)\}$ are expressed as $\{\hat{\alpha}(k)\text{sinc}(i - \hat{D}_e(k))\}$ where $\hat{D}_e(k)$ represents the delay estimate while $\hat{\alpha}(k)$ is a variable gain for optimal filtering. The ETDGE algorithm is given by [5]

$$\hat{D}_e(k+1) = \hat{D}_e(k) - \mu_e e(k) \sum_{i=-P}^{P} f(i - \hat{D}_e(k))x(k-i) \tag{6}$$

$$\hat{\alpha}(k+1) = \hat{\alpha}(k) + \mu_e e(k) \sum_{i=-P}^{P} \text{sinc}(i - \hat{D}_e(k))x(k-i) \tag{7}$$