Window Length Effect on Blind Symbol Estimation

Qi Cheng*
School of Engineering
University of Western Sydney
Kingswood NSW 2747, Australia
Email: qcheng@uws.edu.au

Huimin Zhang
Department of Electronic & Information Engineering
Zhengzhou Railway Vocational Technical College
Zhengzhou, Henan 450052, China

Abstract: In this paper, we study the window length effect on the estimation accuracy of symbols given by an existing approach. Cramer-Rao bounds under the norm and linear constraints are discussed, derived and compared. The norm-constrained and the linear-constrained estimation errors have also been studied.

I Introduction

Consider a communication system consisting of M channels. The i-th channel is characterised by a finite impulse response, denoted by $h_i(0), ..., h_i(L)$, $i = 1, 2, ..., M$. L is the maximum order of all M channels, in the sense that $h_i(L) \neq 0$ for at least one i and $h_j(0) \neq 0$ for at least one j. All channels are driven by the same input symbols s_0, ..., s_{N+L-1} transmitted from a single user. If we let $y_i(n), n = 0, ..., N-1$, be the output samples at N different time instants, for the i-th channel, the output samples of the i-th channel can be written as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{s}, \quad i = 1, 2, ..., M \tag{1}$$

where

$$\mathbf{y}_i = [y_i(0), y_i(1), ..., y_i(N-1)]^T \tag{2}$$

$$\mathbf{H}_i^T = \begin{bmatrix} h_i(L) & & \\ \vdots & \ddots & \\ h_i(0) & & h_i(L) \\ & \ddots & \vdots \\ & & h_i(0) \end{bmatrix}_{(N+L)\times N} \tag{3}$$

$$\mathbf{s} = [s_0, ..., s_{N+L-1}]^T. \tag{4}$$

In the context of channel equalization, channel response $h_1(0), \cdots, h_1(L)$, $\cdots, h_M(0), \cdots, h_M(L)$ is first estimated using a short length of samples and the channel response estimate is then used to carry out equalization (computation of symbols $s_0, ..., s_{N+L-1}$) based on the subsequent samples. Techniques for channel estimation can be found in [2,3,4] and the references therein. To have a reliable symbol estimate, channels should remain unchanged during the period of equalization. For fast changing channels, this scheme clearly does not work, and direct symbol estimation is most appreciated.

For single user case, the approach in [6] is a direct symbol estimation method. For multiple user case, techniques in [9,10,11,12] can provide symbol estimates without prior knowledge of channel response. Another advantage of direct symbol estimation is that inversion of an ill-conditioned channels can be avoided.

The approach in [6] applies a sliding window to the original data to generate a number of data sets for the construction of a data covariance

*Corresponding author. This work is supported by the University of Western Sydney Research Grant and DRG Grant Scheme.

matrix satisfying a desired rank condition. In this paper, we will investigate in which way the estimation accuracy of symbols varies with the window length. Note that the channel order L is assumed known. In practice, this number is unknown and can be estimated using available techniques, e.g., the least square technique in [5].

II Blind Symbol Estimation

The approach in [6] relies on the following averaged data matrix

$$\mathbf{Y}_{i,K} = \begin{bmatrix} y_i(0) & \cdots & y_i(N-K) \\ y_i(1) & \cdots & y_i(N-K+1) \\ \vdots & \ddots & \vdots \\ y_i(K-1) & \cdots & y_i(N-1) \end{bmatrix}. \quad (5)$$

In $\mathbf{y}_{i,K}$, K is called window length. As seen later, this parameter plays an important role in the accuracy of the symbol estimates. Let

$$\mathbf{S}_K = \begin{bmatrix} s_0 & \cdots & s_{N-K} \\ s_1 & \cdots & s_{N-K+1} \\ \vdots & \ddots & \vdots \\ s_{L+K-1} & \cdots & s_{N+L-1} \end{bmatrix} \quad (6)$$

and $\mathbf{H}_{i,K}$ be a $K \times (K+L)$ Sylvester matrix having the same structure as $\mathbf{H}_i$. Then

$$\mathbf{Y}_{i,K} = \mathbf{H}_{i,K}\mathbf{S}_K \quad (7)$$

and one can construct the following covariance matrix

$$\mathbf{R}_Y = \sum_{i=1}^{M} \mathbf{Y}_{i,K}^H \mathbf{Y}_{i,K} = \mathbf{S}_K^H \left[\sum_{i=1}^{M} \mathbf{H}_{i,K}^H \mathbf{H}_{i,K} \right] \mathbf{S}_K. \quad (8)$$

For the algorithm [6] to apply, it is required that $\mathbf{R}_Y$ has $L+K$ positive eigenvalues. To this end, it is sufficient that (1) $\mathbf{S}_K$ should have full row rank and (2) $\sum_{i=1}^{M} \mathbf{H}_{i,K}^H \mathbf{H}_{i,K}$ has a full rank. Requirement (1) is satisfied if $\mathbf{s}$ has more than $L+K$ modes, $L+K < N-K+1$. Requirement (2) is satisfied if all channels do not share any common zeros and $K \geq L/(M-1)$. Note that the requirement on K is

$$L/(M-1) \leq K < (N-L+1)/2. \quad (9)$$

Define the collection of eigenvectors of $\mathbf{R}_Y$ in (8) corresponding to zero eigenvalues by $\mathbf{U}$, then it is easy to see $\mathbf{S}_K\mathbf{U} = \mathbf{0}$. Based on the aforementioned conditions, $\mathbf{S}_K\mathbf{U} = \mathbf{0}$ leads to a unique (up to a scalar) solution of $\mathbf{s}$. The reason is that $\mathbf{S}_K^H$ and $\mathbf{U}$ are orthogonal complements and the Hankel structure of $\mathbf{S}_K$ as outlined in [6]. Note that

$$\mathbf{S}_K\mathbf{U} = \mathbf{0} \Longrightarrow \mathbf{S}_K\mathbf{u}_i = \mathbf{0} \Longrightarrow \mathbf{U}_i\mathbf{s} = \mathbf{0}, \quad (10)$$
$$i = 1, 2, ..., N-L-2K+1,$$

where $\mathbf{u}_i$ is the i-th column of $\mathbf{U}$,

$$\bar{\mathbf{U}}_i = \begin{bmatrix} u_{i,1} & \cdots & u_{i,d} & & \\ & \ddots & & \ddots & \\ & & u_{i,1} & \cdots & u_{i,d} \end{bmatrix}_{(L+K)\times(N+L)} \quad (11)$$

with $d = N-K+1$ and $u_{i,j}$ being the (i,j)-th element of $\mathbf{U}$. Then $\mathbf{s}$ is solved as the solution to the following over-determined linear equations

$$\begin{bmatrix} \bar{\mathbf{U}}_1 \\ \vdots \\ \bar{\mathbf{U}}_{N-L-2K+1} \end{bmatrix} \mathbf{s} = 0. \quad (12)$$

Note that for the unique (up to a scalar) determination of $\mathbf{s}$, there should be more equations that parameters. This imposes a further condition on K:

$$(N-L-2K+1)(L+K) \geq N+L \quad (13)$$

For noisy data, only an estimate $\hat{\mathbf{U}}$ of $\mathbf{U}$ is obtained, and $\mathbf{s}$ is determined as the eigenvector associated with the smallest eigenvalue of the following matrix

$$\sum_{i=1}^{N-L-2K+1} \hat{\bar{\mathbf{U}}}_i^H \hat{\bar{\mathbf{U}}}_i \quad (14)$$

where $\hat{\bar{\mathbf{U}}}_i$ has the same structure as $\bar{\mathbf{U}}_i$ and computed using noisy elements of $\hat{\bar{\mathbf{U}}}$.

III Cramer-Rao Bounds

To evaluate Cramer-Rao bounds, we assume that independent and identically distributed (complex) Gaussian noise has been added to the output samples $\mathbf{y}_i, i = 1, 2, ..., M$. Each noise variable has zero mean and variance σ_w^2. Since symbol vector estimate is inherently not unique, various constraints are needed to obtain a unique estimate (for the maximum likelihood algorithms). Hence Cramer-Rao bound evaluation also depends on what constraint is used. In this paper, we consider two widely used constraints: (1) $\|\mathbf{s}\| = 1$ plus $\Im m(s_1) = 0$ [7] termed as norm constraint and (2) $s_0 = 1$ termed as linear constraint.

In [8], Cramer-rao bound evaluation has been discussed for the channel response under the norm and linear constraints. In what follows, we will present the CRB expressions for input symbols, following the principle in [8]. σ_w^2 is assumed known, because the CRB for the unknown parameter set $\boldsymbol{\theta}$ and the CRB for σ_w^2, are decoupled. $\boldsymbol{\theta} = [\Re e(\mathbf{h}_1)^T, \cdots, \Re e(\mathbf{h}_M)^T, \Im m(\mathbf{h}_1)^T, \cdots, \Im m(\mathbf{h}_M)^T, \Re e(\mathbf{s})^T, \Im m(\mathbf{s})^T]^T$ where $\mathbf{h}_i = [h_i(0), \cdots, h_i(L)]^T$,

Let

$$\bar{\mathbf{S}} = \begin{bmatrix} s_L & \cdots & s_0 \\ s_{L+1} & \cdots & s_1 \\ \vdots & \ddots & \vdots \\ s_{N+L-1} & \cdots & s_{N-1} \end{bmatrix} \tag{15}$$

$$\breve{\mathbf{S}} = diag[\bar{\mathbf{S}}, ..., \bar{\mathbf{S}}] \tag{16}$$

$$\breve{\mathbf{H}} = [\mathbf{H}_1^T, ..., \mathbf{H}_M^T]^T \tag{17}$$

$$\mathbf{Q} = [\breve{\mathbf{S}}, j\breve{\mathbf{S}}, \breve{\mathbf{H}}, j\breve{\mathbf{H}}]. \tag{18}$$

According to [4], the Fisher information matrix is

$$\mathbf{J} = E[\frac{\partial f}{\partial \boldsymbol{\theta}} \frac{\partial f}{\partial \boldsymbol{\theta}^T}] = \frac{2}{\sigma_w^2} \Re e[\mathbf{Q}^H \mathbf{Q}] \tag{19}$$

where f is the maximum likelihood function for a given measurement. Note that $\mathbf{J}$ is rank deficient by 2, because 2 unknowns are not free parameters.

Let $\mathbf{U}_l = diag(\mathbf{I}_{2M(L+1)}, \bar{\mathbf{I}}_{N+L}, \bar{\mathbf{I}}_{N+L})$ where $\bar{\mathbf{I}}_{N+L}$ is the identity matrix $\mathbf{I}_{N+L}$ with the first column deleted. The CRB under the linear constraint can be determined from the matrix $\mathbf{CRB}_l = \mathbf{U}_l(\mathbf{U}_l^T \mathbf{J} \mathbf{U}_l)^{-1} \mathbf{U}_l^T$. Let $\mathbf{G} = [\dot{\mathbf{s}}, \mathbf{e}]$ where $\dot{\mathbf{s}} = [\Re e(\mathbf{s})^T, \Im m(\mathbf{s})^T]^T$ and $\mathbf{e}$ is the $(N+L+1)$-th column of $\mathbf{I}_{2N+2L}$, and construct $\mathbf{P}_n = \mathbf{I}_{2N+2L} - \mathbf{G}(\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H$. Denote the eigendecomposition of $\mathbf{P}_n$ by $\mathbf{P}_n = \mathbf{V}\mathbf{V}^T$. The CRB under the norm constraint can be shown to be $\mathbf{CRB}_n = \mathbf{U}_n(\mathbf{U}_n^T \mathbf{J} \mathbf{U}_n)^{-1} \mathbf{U}_n^T$ where $\mathbf{U}_n = diag[\mathbf{I}_{2m(L+1)}, \mathbf{V}]$.

IV Accuracy Comparison

The norm and linear constraints have also been used by other estimation techniques for the uniqueness of solution. It is now a natural question as to which constraint gives rise to a more accurate estimate. The comparison may not be possible in general if norm-constrained and linear-constrained estimates are obtained by using different approaches and/or under different conditions. In the following, we will compare the accuracy of the two estimates inter-related by

$$\hat{\boldsymbol{\rho}} = \frac{1}{\sqrt{1 + \hat{\boldsymbol{\gamma}}^H \hat{\boldsymbol{\gamma}}}} \begin{bmatrix} 1 \\ \hat{\boldsymbol{\gamma}} \end{bmatrix} \tag{20}$$

where $\hat{\boldsymbol{\rho}}$ is the estimate of $\mathbf{s}$ under the norm constraint, and $[1, \hat{\boldsymbol{\gamma}}^T]^T$ the estimate under the linear constraint.

The relationship (20) holds for various kinds of estimates, such as norm-to-linear and linear-to-norm converted estimates as defined in [7]. For the problem of blind symbol estimation, norm-constrained estimates are generated by many approaches in the first instance and then converted to the linear-constrained ones for the comparison against the CRB under the linear constraint (e.g., see [6]). However, as shown below, norm-to-linear converted estimates are less accurate than norm-constrained ones and the norm-to-linear conversion should not be encouraged.

Let the noiseless estimate of $\mathbf{s}$ under the norm constraint, be $\boldsymbol{\rho}$ and the estimate under the linear constraint be $[1, \boldsymbol{\gamma}^T]^T$. Denote $\hat{\boldsymbol{\rho}} - \boldsymbol{\rho} \stackrel{def}{=} \triangle\boldsymbol{\rho}$ and $\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \stackrel{def}{=} \triangle\boldsymbol{\gamma}$ and assume they are small enough[1]. Applying the first order approximation to (20), one can represent errors in norm-constrained estimate in terms of those in linear-constrained one,

$$\triangle\boldsymbol{\rho} = \frac{1}{\sqrt{1 + \boldsymbol{\gamma}^H \boldsymbol{\gamma}}} \left\{ \begin{bmatrix} 1 \\ \triangle\boldsymbol{\gamma} \end{bmatrix} - \frac{1}{2} \frac{\triangle(\boldsymbol{\gamma}^H \boldsymbol{\gamma})}{1 + \boldsymbol{\gamma}^H \boldsymbol{\gamma}} \begin{bmatrix} 1 \\ \boldsymbol{\gamma} \end{bmatrix} \right\}.$$

[1] This requirement can be met in general if noise is small and/or the amount of data is large.

Then

$$\triangle\rho^H\triangle\rho = \frac{1}{1+\gamma^H\gamma}\{\triangle\gamma^H\triangle\gamma - \frac{1}{4}\frac{\triangle(\gamma^H\gamma)\triangle(\gamma^H\gamma)}{1+\gamma^H\gamma}\}$$

and we have

$$E\{\triangle\rho^H\triangle\rho\} \leq E\{\frac{1}{1+\gamma^H\gamma}\triangle\gamma^H\triangle\gamma\}. \quad (21)$$

The inequality (21) means that the variance of the total estimation errors under the norm constraint is less than that under the linear constraint, asymptotically. This finding tells us that we should always try to obtain an estimate under the norm constraint or convert a linear-constrained estimate to a norm-constrained one. The finding easily extends to other linear constraints with $s_k = 1, k > 1$, and holds for any unbiased estimates with arbitrary statistical distribution. Another consequence of this finding is that the CRB under the norm constraint is always no larger than the CRB under the linear constraint if the regularity conditions are satisfied.

V Performance Evaluation

In this section, we will investigate the effect of the window length K on the accuracy of the estimated symbols. We considered a two-channel system, where each channel impulse response equal to

$$[h_i(0), h_i(1), h_i(2)] = [1, -2\cos(\theta_i), 1],$$
$$i = 1, 2, \quad \theta_1 = \pi/10, \theta_2 = \pi/5.$$

The system is driven by one realization of a white binary random process of variance $\sigma_s^2 = 1$, along with additive Gaussian noise. The number of output samples from each channel was chosen to be $N = 30$. The symbol sequence generated is $1, -1, -1, 1, 1,\ 1, 1, -1, 1, -1,\ -1, 1, 1, 1, 1,\ 1, 1, 1, 1, 1,\ 1, 1, -1, 1, -1,\ -1, -1, -1, -1, -1,\ 1, -1$. The signal-to-noise ratio of this 2-channel system is defined as

$$SNR(dB) = 20\log_{10}(\frac{||\mathbf{h}||\sigma_s}{\sigma_w})$$

where $\mathbf{h} = [h_1(0), h_1(1), h_1(2), h_2(0), h_2(1), h_2(2)]^T$. The range of K is $2 \leq K \leq 13$ according to (9) and (13).

Denote by $\hat{\boldsymbol{\rho}}_k$ the norm-constrained estimate computed in the k-th run, with the first element having the same sign as s_0. The performance of the symbol estimation under the norm constraint is measured by the mean squared error

$$10\log_{10}(\frac{1}{N_s}\sum_{k=1}^{N_s}||\hat{\boldsymbol{\rho}}_k - \frac{\mathbf{s}}{||\mathbf{s}||}||^2)$$

where $N_s = 100$ is the number of simulation runs, and that under the linear constraint is measured by the mean squared error

$$10\log_{10}(\frac{1}{N_s}\sum_{k=1}^{N_s}||\frac{\hat{\boldsymbol{\rho}}_k}{|\hat{\boldsymbol{\rho}}_k(1)|} - \frac{\mathbf{s}}{||\mathbf{s}||}||^2).$$

CRB under the norm constraint was computed as $\sum_{i=2M(L+1)+1}^{2M(L+1)+2(N+L)} \mathbf{CRB}_n(i,i)/||\mathbf{s}||^2$. CRB under the linear constraint was computed as $\sum_{i=2M(L+1)+1}^{2M(L+1)+2(N+L)} \mathbf{CRB}_l(i,i)/||\mathbf{s}||^2$.

MSEs and CRBs are shown in Figure 1 for varying window length, at $SNR = 50dB$. It is easy to see that the optimum value of K is $K \approx N/3$. This observation is consistent with that for frequency estimation. For other larger window lengths, the approach gives a poor performance and requires more computations.

VI Conclusions

Window length effect for an existing symbol estimation approach has been investigated and an optimum window length is obtained via simulation. Theoretic support is highly desired and currently under investigation. Effect of the norm and linear constraints has also been studied. the norm-constrained estimate is more accurate than the linear-constrained one.

Acknowledgement

We like to thank Dr. W. Yu at the University of Melbourne for providing manuscript.

REFERENCES

[1] Tong, L, Xu, G., and Kailath, T., "Blind identification and equalisation of multipath

channels," *Proc. International Conference on Communications*, pp. 1513-1517, June 1992.

[2] Xu, G., Liu, H., Tong, L., and Kailath, T., "A least-square approach to blind channel identification," *IEEE Transactions on Signal Processing*, Vol. 43, No. 12, pp. 2982-1993, Dec 1995.

[3] Moulines, E., Duhamel, P, Cardoso, J., and Mayrargue, S., "Subspace methods for blind identification of multichannel FIR filters," *IEEE Transactions on Signal Processing*, Vol. 43, No. 2, pp. 516-525, Feb 1995.

[4] Hua, Y., "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Transactions on Signal Processing*, Vol. 44, No. 3, pp. 661-672, March 1996.

[5] Cheng, Q., "Channel order estimation using least square error function," *Proc. IASTED International Conference on Signal and Image Processing*, pp. 372-375, Las Vegas, Nevada, USA, November 19-23, 2000.

[6] Liu, H., and Xu, G., "A deterministic approach to blind symbol estimation," *IEEE Signal Processing Letters*, Vol. 1, No. 12, pp. 205-207, 1994.

[7] Stoica, P., and Ng, B. C., "Chapter 7: Performance bounds for blind channel estimation," *Signal Processing Advances in Wireless Communications*, Vol. 2, P. Stoica, G. Giannakis, Y. Hua, and L. Tong, Eds. Upper Saddle river, NJ:Prentice Hall, 2000.

[8] Stoica, P., and Ng, B. C., "On the Cramer-Rao bound under parametric constraints," *IEEE Signal Processing Letters*, Vol. 5, No. 7, pp. 177-179, 1998.

[9] Yellin, D., and Porat, B., "Blind identification of FIR systems excited by discrete-alphabet inputs," *IEEE Transactions on Signal Processing*, Vol. 41, No. 3, pp. 1331-1339, 1993.

[10] Talwar, S., Viberg, M., and Paulraj, A., "Blind estimation of multiple co-channel

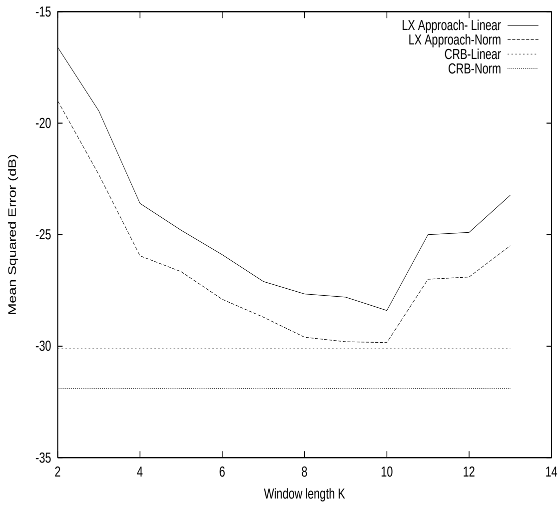

Figure 1: Performance versus Window Length.

signals using an antenna array," *IEEE Signal Processing Letters*, Vol. 1, No. 2, pp. 29-31, 1994.

[11] Talwar, S., Viberg, M., and Paulraj, A., "Blind separation of synchronous co-channel signals using an antenna array-Part I: algorithms," *IEEE Transactions on Signal Processing*, Vol. 44, No. 5, pp. 1184-1197, 1996.

[12] van der Veen, A., Talwar, S., and Paulraj, A., "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Transactions on Signal Processing*, Vol. 45, No. 1, pp. 173-190, 1997.