AN OPTIMAL TRAINING SIGNAL STRUCTURE FOR FREQUENCY OFFSET ESTIMATION IN AWGN CHANNEL

Hlaing Minn, *Member, IEEE* and Shaohui Xing
Department of Electrical Engineering, University of Texas at Dallas
Email: Hlaing.Minn@utdallas.edu

***Abstract*— This paper addresses training signal design for frequency offset estimation in additive white Gaussian noise (AWGN) channel. Based on the Cramer-Rao lower bound (CRB), the optimal training signal structure is obtained with constraints on the total energy and peak sample energy within a fixed training block length. This training signal structure consists of two non-zero parts of the same length and each non-zero sample energy is determined by the allowable peak sample energy. The optimality proof of the proposed training signal structure is provided. Numerical and simulation results are presented to show the advantage of the proposed training signal structure.**

I. INTRODUCTION

Frequency offset is unavoidable at the receiver due to oscillators' inaccuracies and the Doppler shift of the mobile wireless channel. To eliminate or reduce the performance degradation caused by the frequency offset, the receiver usually performs frequency synchronization or frequency offset compensation based on the frequency offset estimation. Usually training signals are used for frequency offset estimation. Proper training signals can not only improve the estimation accuracy, but also reduce the estimation complexity. Hence, in this paper we investigate the training signal design for frequency offset estimation in AWGN channel and present an optimal training structure.

II. THE CRAMER-RAO LOWER BOUND (CRB)

We consider frequency offset estimation based on the training signal vector $\boldsymbol{s} = [s_0,\ s_1,\ \ldots,\ s_{N-1}]^T$, where the superscript T denotes the transpose. Let us consider a signal model with an arbitrary carrier phase ψ given by

$$\boldsymbol{r} = e^{j\psi}\ \boldsymbol{W}(v)\boldsymbol{s} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{r}$ is the received training signal vector of length N, $\boldsymbol{W}(v)$ is a diagonal matrix with diagonal elements $\{exp(j2\pi 0v/N)$, $exp(j2\pi 1v/N)$, $\ldots$, $exp(j2\pi(N-1)v/N)\}$ corresponding to the normalized frequency offset v (normalized by $1/(NT)$ where T is the sample duration), and $\boldsymbol{n}$ is a zero-mean complex Gaussian noise vector with a covariance matrix $\boldsymbol{C_n} = \sigma_n^2 \boldsymbol{I}$ where $\boldsymbol{I}$ is an identity matrix.

For the parameter vector $\boldsymbol{\alpha} = [v,\ \psi]^T$, the received vector $\boldsymbol{r}$ has a complex Gaussian pdf, $p(\boldsymbol{r};\boldsymbol{\alpha})$, with the mean vector and the covariance matrix respectively given by $\boldsymbol{\mu_r} = e^{j\psi}\ \boldsymbol{W}(v)\ \boldsymbol{s}$ and $\boldsymbol{C_r} = \boldsymbol{C_n}$. The $[i,j]$ element of the Fisher information matrix $\boldsymbol{J}$ is given by

$$J_{i,j} = E\left[\frac{\partial\ ln\ p(\boldsymbol{r};\boldsymbol{\alpha})}{\partial \alpha_i}\ \frac{\partial\ ln\ p(\boldsymbol{r};\boldsymbol{\alpha})}{\partial \alpha_j}\right] \tag{2}$$

$$= tr\left[\boldsymbol{C_r}^{-1}\ \frac{\partial \boldsymbol{C_r}}{\partial \alpha_i}\ \boldsymbol{C_r}^{-1}\ \frac{\partial \boldsymbol{C_r}}{\partial \alpha_j}\right] + 2\,\Re\left[\frac{\partial \boldsymbol{m_r}^H}{\partial \alpha_i}\ \boldsymbol{C_r}^{-1}\ \frac{\partial \boldsymbol{m_r}}{\partial \alpha_j}\right].$$

After some calculation, we have

$$J_{1,1} = \frac{8\pi^2}{N^2\ \sigma_n^2}\ \boldsymbol{s}^H \Lambda^2 \boldsymbol{s},\ \ J_{2,2} = \frac{2}{\sigma_n^2}\boldsymbol{s}^H\ \boldsymbol{s} \tag{3}$$

$$J_{1,2} = J_{2,1} = \frac{4\pi}{N\sigma_n^2}\ \boldsymbol{s}^H \Lambda \boldsymbol{s} \tag{4}$$

where Λ is a diagonal matrix with diagonal elements $\{0,1,\ldots,N-1\}$. Then the CRB of the frequency offset estimation for the signal model (1) is given by

$$\text{CRB} = [\boldsymbol{J}^{-1}]_{1,1} = \frac{N^2\sigma_n^2}{8\pi^2}\ \frac{\boldsymbol{s}^H\boldsymbol{s}}{\boldsymbol{s}^H\boldsymbol{s}\boldsymbol{s}^H\Lambda^2\boldsymbol{s} - (\boldsymbol{s}^H\Lambda\boldsymbol{s})^2}. \tag{5}$$

III. TRAINING SIGNAL DESIGN

We investigate the training signal design that minimizes the CRB. The best training signal vector $\boldsymbol{s}^\dagger$ is given by

$$\boldsymbol{s}^\dagger = arg\ min_{\boldsymbol{s}}\ \{\text{CRB}\}\ ,\ \text{constraint: } \boldsymbol{s}^H\boldsymbol{s} = E_s \tag{6}$$

where E_s is the total energy of the training signal vector. After some manipulation, we have the best training signal given by

$$\boldsymbol{s}^\dagger = arg\ min_{\boldsymbol{s}}\ \left((\boldsymbol{s}^H\Lambda\boldsymbol{s})^2 - \boldsymbol{s}^H\boldsymbol{s}\boldsymbol{s}^H\Lambda^2\boldsymbol{s}\right) \tag{7}$$

$$= arg\ min_{\boldsymbol{s}}\ \left(\sum_{k=0}^{N-1}\sum_{n=0}^{N-1}|s_k|^2|s_n|^2(nk - n^2)\right) \tag{8}$$

with the constraint $\boldsymbol{s}^H\boldsymbol{s} = E_s$. Equation (8) indicates that it is only the training energy allocation among the training samples that has effect on the CRB. Define the following

$$\boldsymbol{y} = [|s_0|^2,\ |s_1|^2,\ \ldots,|s_{N-1}|^2]^T \tag{9}$$

$$G_{k,n} = n(k-n)\ ,\ 0 \le k \le (N-1),\ 0 \le n \le (N-1) \tag{10}$$

where $G_{k,n}$ is the $[k,n]$ element of an $N \times N$ matrix G. $\boldsymbol{y}$ can be viewed as the training energy vector or the training energy allocation. Then the best training vector is determined by the best training energy allocation given by

$$\boldsymbol{y}^\dagger = arg\ min_{\boldsymbol{y}}\ \left\{\boldsymbol{y}^T G \boldsymbol{y}\right\},\ \text{constraint: } \boldsymbol{y}^T\boldsymbol{1} = E_s,\ 0 \le y(i) \le P \tag{11}$$

This work was supported, in part, by the Clark Foundation Grant, Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, USA.

where $\mathbf{1}$ is an all-one vector of length N, P is a design value representing the allowable peak training sample energy (to avoid non-linear distortion). The optimization of (11) can be solved and the general result is given by

$$y_k^{\dagger} = \begin{cases} min\{P, E_s/2\} & \text{if } k=0 \text{ or } N-1 \\ min\left\{P, E_s/2 - \sum_{l=0}^{m-1} y_l^{\dagger}\right\} & \text{if } k=m \text{ or } k=N-1-m \\ & \text{and } 1 \le m \le N/2-1 \\ E_s - 2\sum_{l=0}^{k-1} y_l^{\dagger} & \text{if } k=(N-1)/2 \\ & \text{and } N \text{ is an odd integer.} \end{cases} \tag{12}$$

IV. Optimality of the Training Signal Structure

This section provides a proof for the optimality of the training signal structure of (12). The optimization objective function from (11), which is to be minimized, can be given by the following objective function, which is to be maximized:

$$f(\boldsymbol{y}) = \boldsymbol{y}^T \boldsymbol{H} \boldsymbol{y}, \text{ constraint: } \boldsymbol{y}^T\mathbf{1} = E_s,\ 0 \le y(i) \le P \tag{13}$$

where

$$\boldsymbol{H}_{k,n} = \begin{cases} -(G_{k,n}+G_{n,k}) = (k-n)^2, & k>n \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

We start from an arbitrary training signal structure. We apply on the training signal structure an algorithm which moves training sample energy from one sample to another in a prescribed way. Each move will be shown to increase the objective function (13), which is to be maximized. The algorithm will finally reach to the training signal structure of (12). Since any arbitrary training signal structure can finally be transformed into the training structure of (12) by the algorithm while increasing the objective function, the training structure of (12) is optimal and unique.

Let us consider a length-N training energy vector defined by

$$\boldsymbol{y} = [P\cdot \mathbf{1}_{1\times M},\ y_M,\ y_{M+1}, \dots,\ y_{N-L-1},\ P\cdot\mathbf{1}_{1\times L}]^T \tag{15}$$

where $0 \le M \le N_1$, $0 \le L \le N_1$, $M+L+2 \le N$, $0 \le y_m \le P$, $0 \le m \le N-1$, $\sum_{m=0}^{N-1} y_m = E_s$, $\sum_{m=M}^{N-L-1} y_m = E_u$, and $N_1 = \lfloor E_s/(2P) \rfloor$. If there are more than N_1 consecutive samples of energy P at either side, the corresponding value of M or L is set to N_1. The above training structure satisfies the total energy and the peak energy constraints. The algorithm is described in the following. If $M \le L$, then the algorithm moves energy ε from the training energy sample y_k to y_M where y_k with $M+1 \le k \le N-L-1$ is the nearest non-zero sample to y_M and ε is given by

$$\varepsilon = \begin{cases} min(P-y_M,\ y_k) & \text{if } (E_u \ge 2P) \\ & \text{or } (E_u < 2P, M \ne L) \\ min(E_u/2 - y_M,\ y_k) & \text{if } (E_u < 2P,\ M=L, \\ & \text{and } y_M < E_u/2) \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

If $M \ge L$, the algorithm moves energy ε from the training energy sample y_k to y_{N-L-1} where y_k with $M \le k \le N-L-2$ is the nearest non-zero sample to y_{N-L-1} and ε is given by

$$\varepsilon = \begin{cases} min(P-y_{N-L-1},\ y_k) & \text{if } (E_u \ge 2P) \\ & \text{or } (E_u < 2P, M \ne L) \\ min(E_u/2 - y_{N-L-1},\ y_k) & \text{if } (E_u < 2P,\ M=L, \\ & \text{and } y_{N-L-1} < E_u/2) \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Then the values of M and L are updated and the energy moving is repeated. Finally, the resulted training signal will become that of (12). In the following, we will prove that each move increases the objective function which is to be maximized.

Each move changes the training structure $\boldsymbol{y} = [y_0,\ y_1,\ \dots, y_l,\ \dots, y_n,\ \dots, y_{N-1}]^T$ into $\boldsymbol{y}' = [y_0, y_1,\ \dots, y_l', \dots, y_n', \dots, y_{N-1}]^T$. The only differences between $\boldsymbol{y}$ and $\boldsymbol{y}'$ are at the indexes l and n. The objective function (13) can be given by (18) where Q is not a function of y_l and/or y_n.

$$\begin{aligned} f(y) =& y_l P \Big[M l^2 + M(M-1)(\frac{2M-1}{6} - l) \\ &+ L(L+1)(\frac{2L+1}{6} - N + l) + L(N-l)^2 \Big] \\ &+ y_n P \Big[M n^2 + M(M-1)(\frac{2M-1}{6} - n) \\ &+ L(L-1)(\frac{2L-1}{6} + N - L - n) + L(N-L-n)^2 \Big] \\ &+ y_l \Big[\sum_{i=M}^{l-1} y_i\ (l-i)^2 + \sum_{i=l+1}^{n-1} y_i\ (l-i)^2 \\ &+ \sum_{i=n+1}^{N-L-1} y_i\ (l-i)^2 \Big] \\ &+ y_n \Big[\sum_{i=M}^{l-1} y_i\ (n-i)^2 + \sum_{i=l+1}^{n-1} y_i\ (n-i)^2 \\ &+ \sum_{i=n+1}^{N-L-1} y_i\ (n-i)^2 \Big] + y_l y_n (n-l)^2 + Q \end{aligned} \tag{18}$$

Define $\Delta f = f(\boldsymbol{y}') - f(\boldsymbol{y})$. Then we just need to show that $\Delta f > 0$. Now, consider the case with $M \le L$ where the energy ε is moved from y_k to y_M. Then we have $y_k = y_k - \varepsilon$ and $y_M = y_M + \varepsilon$. After substituting M and k in place of l and n in (18) and manipulating, we have

$$\Delta f = g_1 + g_2 \tag{19}$$

where g_1 and g_2 are respectively given by

$$\begin{aligned} g_1 =& \varepsilon P\ (k-M)L\{2N - k - M - L - 1 - (k+1)M/L\} \\ =& \{\varepsilon P\ (k-M)L(2N - 2k - M - L - 2)\} \\ &+ \{\varepsilon P(k-M)(k+1)(L-M)\} \\ \ge& \{\varepsilon P(k-M)(k+1)(L-M)\} \end{aligned} \tag{20}$$

$$\begin{aligned} g_2 =& \varepsilon(k-M)\{ \sum_{i=k}^{N-L-1} y_i\ (k-M+2(i-k))\} \\ &- \varepsilon(k-M)^2(y_M+\varepsilon) \\ \ge& \varepsilon(k-M)^2\{(\sum_{i=k}^{N-L-1} y_i) - (y_M+\varepsilon)\}. \end{aligned} \tag{21}$$

Then we obtain

$$\begin{aligned} \Delta f \ge& \{\varepsilon P(k-M)(k+1)(L-M)\} \\ &+ \varepsilon(k-M)^2\{(\sum_{i=k}^{N-L-1} y_i) - (y_M+\varepsilon)\}. \end{aligned} \tag{22}$$

When $(E_u > 2P)$ or $(E_u < 2P$ and $M = L)$, we have $(\sum_{i=k}^{N-L-1} y_i) > (y_M + \varepsilon)$ and the first term of (22) is greater than or equal to zero. Hence, $\Delta f > 0$. When $(E_u < 2P$ and $M < L)$, i.e., $(E_u < 2P$ and $M = L-1)$, we have

$(\sum_{i=k}^{N-L-1} y_i) - (y_M + \varepsilon) > -P$ and the first term of (22) is $\varepsilon P\ (k-M)(k+1)$. Hence, we still have $\Delta f > 0$.

Now, consider the case with $M \geq L$ where the energy ε is moved from y_k to y_{N-L-1}. Then we have $y_k = y_k - \varepsilon$ and $y_{N-L-1} = y_{N-L-1} + \varepsilon$. After substituting k and $(N-L-1)$ in place of l and n in (18) and manipulating, we have (19) with g_1 and g_2 respectively given by

$$\begin{aligned} g_1 =& \varepsilon P\ (N-L-k-1)\{M(N-L-M+k) - L(N-k)\} \\ =& \varepsilon P\ (N-L-k-1)\{M(2k-M-L) + (M-L)(N-k)\} \\ \geq& \varepsilon P\ (N-L-k-1)\{(M-L)(N-k)\} \end{aligned} \quad (23)$$

$$\begin{aligned} g_2 =& \varepsilon(N-L-k-1)\{\sum_{i=M}^{k} y_i\ (N-L-k-1+2(k-i))\} \\ & -\varepsilon(N-L-k-1)^2(y_{N-L-1}+\varepsilon) \\ \geq& \varepsilon(N-L-k-1)^2\{(\sum_{i=M}^{k} y_i) - (y_{N-L-1}+\varepsilon)\}. \end{aligned} \quad (24)$$

Consequently, we obtain

$$\begin{aligned} \Delta f \geq& \varepsilon P\ (N-L-k-1)\{(M-L)(N-k)\} \\ &+ \varepsilon(N-L-k-1)^2\{(\sum_{i=M}^{k} y_i) - (y_{N-L-1}+\varepsilon)\}. \end{aligned} \quad (25)$$

When $(E_u > 2P)$ or $(E_u < 2P$ and $M = L)$, we have $(\sum_{i=M}^{k} y_i) > (y_{N-L-1}+\varepsilon)$ and the first term of (25) is greater than or equal to zero. Hence, $\Delta f > 0$. When $(E_u < 2P$ and $M > L)$, i.e., $(E_u < 2P$ and $M = L+1)$, we have $\{(\sum_{i=M}^{k} y_i) - (y_{N-L-1}+\varepsilon)\} > -P$ and the first term of (22) is $\varepsilon P\ (N-L-k-1)(N-k)$. Hence, we still have $\Delta f > 0$.

The above proof assumes $M+L+2 \leq N$. The only possibility left is $M+L+1 = N$ which is already in the form of the training structure (12). This proves that the training structure of (12) is optimal and unique.

V. Simulation Results and Discussions

The simulation parameters are as follows. The length of training signal structure is $N = 1024$, and the normalized frequency offset is $v = 0.4$. The S-BLUE frequency offset estimation method proposed in [1] is used in the simulation. The frequency offset estimator is

$$\hat{v} = \boldsymbol{\theta}^T \boldsymbol{w} \quad (26)$$

where $\boldsymbol{\theta}$ is a column vector containing $\{\theta_k\}$ given by

$$\theta_k = \frac{N}{2\pi D_k} angle\{R(D_k)\} \quad (27)$$

and $R(D_k)$ is sample-wise autocorrelation of the received training signals with a correlation distance D_k. The weighting vector $\boldsymbol{w}$ is given by [2]

$$\boldsymbol{w} = \frac{\boldsymbol{C}_\theta^{-1}\,\mathbf{1}}{\mathbf{1}^T\,\boldsymbol{C}_\theta^{-1}\,\mathbf{1}} \quad (28)$$

where the calculation of $\boldsymbol{C}_\theta$, the covariance matrix of $\boldsymbol{\theta}$, is lengthy and the results are given in [1].

Different training structures are evaluated. Conventional training signal structure has N samples with sample energy E_s/N. Structure K_128 and K_1 are the proposed structure with $K = 128$ and $K = 1$, respectively. In practice, the value of K is determined by the total training energy and the allowable peak training sample energy. From Fig.1 (b), it is observed that K_1 can achieve a 5 dB SNR advantage over Conventional structure. K_128 has a 4 dB SNR advantage over Conventional structure. Strcuture K_1 has 1 dB SNR advantage over Strcuture K_128 but K_1 has a larger PAR than K_128. The CRBs of these structures are shown in the Fig.1 (a).

TABLE I

Computation Load

Structure	K_1	K_128	Conventional
#RM	5	$8K^2 + 2K - 8$	$2N^2 - 2$
#RA	2	$8K^2 - 7K + 1$	$2N^2 - 3N$
$angle\{\cdot\}$	1	$3K - 2$	$N - 1$

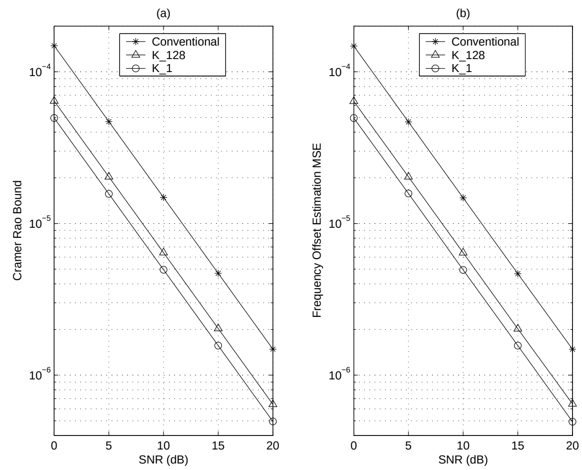

Fig. 1. The frequency offset estimation performance comparison

Based on the S-BLUE frequency offset estimation method and assuming $N \geq 3K - 1$, the required computational load is shown in Table I. For the considered parameters, Conventional structure needs approximately 16 times the real multiplications (RM), 16 times the real additions (RA) and 2.7 times the angle operations if compared to those required for the structure K_128.

VI. Conclusions

An optimal training signal structure for frequency offset estimation is developed by minimizing the CRB with constraints on the total energy E_s and the peak sample energy P within a fixed training block length. The proposed training signal structure consists of two non-zero parts, each having K samples determined by E_s and P. The first part is at the beginning of the training block and the second is at the end. The proposed training signal structure can not only improve the estimation performance but also reduce the estimation complexity over the conventional training structure.

References

[1] H. Minn, and S. Xing, "An Optimal Training Signal Structure For Frequency Offset Estimation," Submitted to *IEEE Trans. Commun.*

[2] S. M. Kay, "Fundamentals of Statistical Signal Processing: Estimation Theory," *Prentice Hall PTR*, 1993.

[3] H. Minn, P. Tarasak and V. K. Bhargava, "OFDM frequency offset estimation based on BLUE principle," *IEEE VTC'02 (Fall)*, Vancouver, Canada, Sept. 2002, pp.1230-1234.