# Bounds on Joint Phase and Frequency Estimation for Pilot-assisted PSK/QAM Modulation

Feng Rice
*CRC for Sensor Signal and Information Processing*
*Department of Electrical and Electronic Engineering*
*University of Adelaide, Australia 5005, feng@eleceng.adelaide.edu.au*

**Abstract**

*This paper derives the true Cramér-Rao lower bounds (CRLB) on the individual phase estimation and the coupled carrier phase and frequency estimation for transmission bursts with pilot-assisted PSK/QAM modulation. The reduction in frequency and phase CRLBs due to introduction of embedded reference symbols increases as the modulation order increasing. It is found that the better CRLB may be obtained if the reference symbols are distributed symmetrically in the burst, especially using a preamble and postamble. At low SNR, 5% pilot symbols in the burst reduces the bound on the phase standard deviation by a factor of 10 for QPSK or higher order modulated signals. The new CRLBs for distributed pilot symbols are also presented. The comparison of the true CRLB and the CRLB of the pilot symbols has revealed that using only pilot symbols for estimation is an effective approach only if more than 5% of pilot symbols in the burst and SNR$\leq 0$ dB.*

## 1. Introduction

In packet transmission for Time Division Multiple Access (TDMA) systems, the data symbols alone may be insufficient for effective synchronization of the burst. This has motivated the use of synchronization overhead in the form of inserted reference symbols. These are sometimes grouped into a preamble, postamble or midamble, or alternatively distributed throughout the packet as pilots. In the most general case the reference symbols are interleaved with data symbols in some predetermined order. Given that reference symbols are inserted, the synchronization performance may be enhanced where both data and reference symbols are used to estimate received signal parameters. The previous CRLB results [1], [2], [3], [4] may be used to approximate the best case estimator performance with either known symbols at low SNR or unknown data symbols at high SNR. It is known that estimator performance may be enhanced by combining algorithms for both known and unknown data, especially at moderate SNR (practical operating range). However, the CRLB for these cases has not been well understood. Reference [5] presents the phase and frequency CRLBs mainly for a special decoupled case - symmetrically distributed pilot symbols in transmission burst. In this papers, the more general CRLBs of the coupled frequency and phase estimation are provided for an arbitrary mixture of interleaved reference and data symbols. The paper investigates the synchronizer CRLBs with various design burst structures, such as number of the reference symbols and their location in the burst. A static AWGN channel is assumed.

The paper is constructed as following. Section 2 derives the phase CRLB for the interleaved data and reference vector assumed known frequency. The new bounds with various percentage pilot symbols in the burst are compared to the well known CW CRLB (100% pilot symbols) and blind estimation CRLB (0% pilot symbols), which verifies the new CRLB. Section 3 presents the CRLBs for joint frequency and phase estimation with an arbitrary pilot distribution. Since the location of the reference symbols is important for the frequency CRLB the effective burst structures are investigated. Section 4 provides the new frequency and phase CRLBs for the interleaved known symbols in the burst. The conventional CW CRLB no longer applies to those bursts since the reference symbols are in sub-groups and far apart. Brief conclusions are presented in Section 5.

## 2. Phase CRLB for Reference and Data Interleaved Burst

Figure 1 illustrates a general reference and data interleaved transmission burst. The length of the burst is $N$ symbols, $D$ is the length of data symbols and $P$ is the length of reference symbols in the burst, $n$ is the number of sub-groups of reference symbols and $m$ is the number of the sub-groups of data symbols.

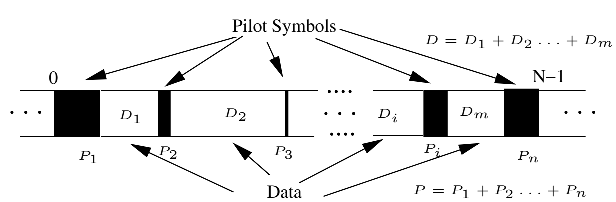



**Fig. 1:** Arbitrary pilot-assisted transmission burst

We initially assume that a received modulated signal has an unknown fixed phase offset $\phi$, and that the signal has been ideally filtered and sampled at the optimum sampling instant. In this case, the received samples are

$$x_k = a_k e^{j\phi} + w_k, \qquad k = 0, 1, \ldots, N-1, \qquad (1)$$

where $a_k$ are the transmitted symbols of modulated symbols (known or unknown) of a PSK/QAM constellation of a unit average energy. For example, with BPSK, **C**=$\{$-1, 1$\}$ and for Square 16QAM **C**=$(1/\sqrt{10})\{\pm1\pm j, \pm3\pm3j, \pm3\pm j, \pm1\pm3j\}$. $w_k$ is the k-th noise sample whose real and imaginary parts are independent zero-mean Gaussian random variables, each with variance $\sigma^2$ and the $w_k$'s are mutually independent. The symbol SNR is $E_s/N_o = 1/(2\sigma^2)$ [10].

The CRLB on the variances of an unbiased estimator of $\phi$, namely $\hat{\phi}$, is given by [6],

$$CRLB(\hat{\phi}) = \frac{1}{-E[\frac{\partial^2 \ln p(\mathbf{X}|\phi)}{\partial \phi^2}]} \qquad (2)$$

where **X**= $(X_0X_1 \ldots X_{N-1})$, $X_k$ is iid and E[·] denotes statistical expectation with respect to the probability density function $p(\mathbf{X} \mid \phi)$.

$$\text{CRLB}^{-1}(\hat{\phi}) = 2\frac{E_s}{N_o}\left[P + F(\frac{E_s}{N_o})D\right] \qquad (3)$$

where $F_i(\frac{E_s}{N_o})$ is the ratio of the CRLB for PSK/QAM signals to the CRLB for an unmodulated carrier of the same power [1].

The derivation reveals that the CRLB of individual phase estimate is independent of the location of reference symbols. Figure 2 shows a comparison of the phase CRLBs for the pilot-assisted QPSK bursts. The length of the burst is 100 symbols with the different length of preamble and pilot symbols. If the burst consists of all known data i.e., $N = P$, the phase CRLB is equal to the well-known CW CRLB [9]. If the burst does not contain any known symbols, i.e., $P = 0$, then the bound is the same as the CRLB of PSK/QAM which is derived in [1], [4]. At low SNR, the use of 5% reference symbols in a burst improves estimation accuracy significantly. At SNR = -10 dB the CRLB (standard deviation) decreases by a factor of 10; while at high SNR it converges to the CW CRLB.

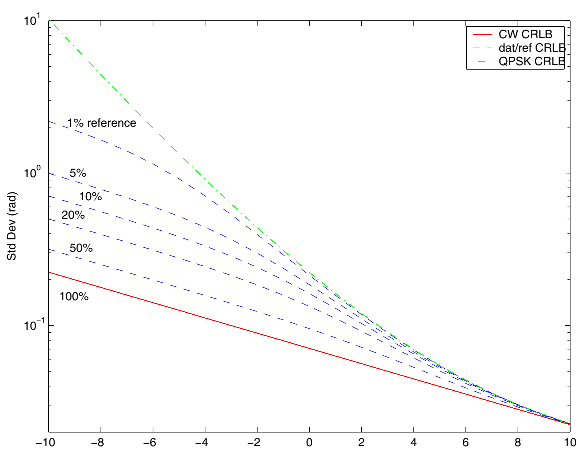



**Fig. 2:** CRLBs for Pilot-assisted QPSK modulation

## 3. CRLBs FOR JOINT FREQUENCY AND PHASE ESTIMATION

Section 2 has discussed the CRLB for individual phase estimation. The CRLBs for the joint frequency and phase estimation can be easily extended. We assume that the received signal has perfect symbol timing and the frequency offset is relatively small so that the received signal is free of intersymbol interference. The received signal is given by

$$x_k = a_k e^{j(k\omega+\phi)} + w_k, \qquad k = \begin{cases} 0, 1 \ldots, N-1 & \text{or} \\ -\frac{N-1}{2} \ldots, \frac{N-1}{2} & \end{cases}$$

where $\omega$ is the frequency offset in radians per symbol period, $\phi$ is a fixed, arbitrary value. The range of $k$ affects the estimate accuracy, it will be discussed later.

Based on the previous derivation the Fisher information matrix (FIM) is given by [6]

$$\mathbf{I} = \begin{bmatrix} I_\omega & I_{\omega\phi} \\ I_{\omega\phi} & I_\phi \end{bmatrix} = \frac{1}{\sigma^2}\begin{bmatrix} \sum_{k\in\text{ref}} k^2 & \sum_{k\in\text{ref}} k \\ \sum_{k\in\text{ref}} k & \sum_{k\in\text{ref}} 1 \end{bmatrix} + \frac{F(\sigma^2)}{\sigma^2}\begin{bmatrix} \sum_{k\in\text{dat}} k^2 & \sum_{k\in\text{dat}} k \\ \sum_{k\in\text{dat}} k & \sum_{k\in\text{dat}} 1 \end{bmatrix}. \qquad (4)$$

where the first matrix is related to the reference symbols and the second matrix is related to the data symbols.

The CRLBs are given by

$$\text{CRLB}(\hat{\omega}) = \frac{I_\phi}{I_\omega I_\phi - I_{\omega\phi}^2}, \qquad \text{CRLB}(\hat{\phi}) = \frac{I_\omega}{I_\omega I_\phi - I_{\omega\phi}^2} \qquad (5)$$

Note that the elements $I_\omega$ and $I_{\omega\phi}$ are dependent on the range of $k$ and the location of the reference symbols. We define $k$ from 0 to $N-1$ as range $\mathcal{K}_N$ and $k$ from $-\frac{N-1}{2}$ to $\frac{N-1}{2}$ as range $\mathcal{K}_S$ (symmetrical range).

### A. Range $\mathcal{K}_N$ from 0 to $N-1$

$$I_{\omega\phi} = 2\frac{E_s}{N_o}\Big\{F_1(\frac{E_s}{N_o})\frac{A_1(A_1-1)}{2} + \sum_{i=2}^{n+m} F_i(\frac{E_s}{N_o})\Big[\frac{A_i(A_i-1)}{2} + A_i\Big(\sum_{j=1}^{i-1} A_j\Big)\Big]\Big\} \qquad (6)$$

$$I_\omega = 2\frac{E_s}{N_o}\Big\{F_1(\frac{E_s}{N_o})\frac{A_1(A_1-1)(2A_1-1)}{6} + \sum_{i=2}^{n+m} F_i(\frac{E_s}{N_o})\Big[\frac{A_i(A_i-1)(2A_i-1)}{6} + A_i(A_i-1)\Big(\sum_{j=1}^{i-1} A_j\Big) + A_i\Big(\sum_{j=1}^{i-1} A_j\Big)^2\Big]\Big\}, \qquad (7)$$

$$I_\phi = 2\frac{E_s}{N_o}\left[P + F_i(\frac{E_s}{N_o})D\right], \qquad (8)$$

where $i$ standards for the sub-group number, $A_i$ is the length of the $i$-th sub-group either the known symbol $P_i$ or the data symbol $D_i$, $\sum_{j=1}^{i-1} A_j$ is the sum of the total symbol length before the $i$-th sub-group and

$$F_i(\frac{E_s}{N_o}) = \begin{cases} 1, & \text{when} \quad A_i = P_i \quad \text{(reference symbols)}, \\ F_i(\frac{E_s}{N_o}), & \text{when} \quad A_i = D_i \quad \text{(data symbols)}. \end{cases}$$

where for the data symbols $D_i$, $F_i(\frac{E_s}{N_o})$ depends on the sub-group modulation scheme and it can be found in [1].

For example, a preamble burst with the range of $\mathcal{K}_N$ from 0 to $N-1$. The corresponding elements of FIM are:

$$I_{\omega\phi} = \frac{2E_s}{N_o}\left[\frac{P(P-1)}{2} + F(\frac{E_s}{N_o})\left(\frac{D(D-1)}{2} + PD\right)\right]$$

$$I_{\omega} = \frac{2E_s}{N_o}\left[\frac{P(P-1)(2P-1)}{6} + F(\frac{E_s}{N_o})\left(\frac{D(D-1)(2D-1)}{6} + D(D-1)P + DP^2\right)\right]$$

$I_\phi$ is shown in Eqn (8). Figure 3 (a) and (b) show the CRLBs for the joint frequency and phase estimation of the preamble burst. The data symbols are 16QAM and the burst length N=1000. The solid black line and the dashed blue line represent 10% and 20% preamble in the burst respectively. The circle line 'o-' is the CW CRLB (100% reference symbols) and the star line is the CRLB for the 16QAM (100% data symbols). The CRLBs with interleaved bursts are bounded between these two lines. At high SNR the interleaved burst CRLBs will converge to the CW CRLB for equal average power. At low SNR the reference symbols dominate the estimation, and the interleaved burst CRLBs asymptotically approach the CW CRLBs of the preamble. Note that the frequency CRLB with 10% preamble does not improve much compared to the 16QAM CRLB. If the same length preamble is divided into preamble and postamble the estimation will be improved significantly [5].

### B. Range $\mathcal{K}_S$ ($k = -\frac{N-1}{2}, \ldots, \frac{N-1}{2}$)

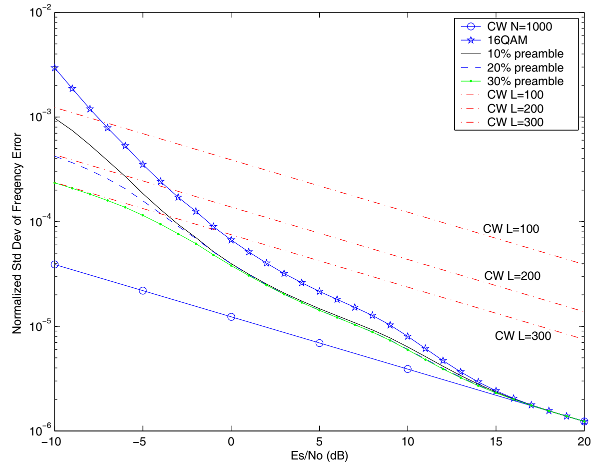
In FIM Eqn (4), the core of element $I_{\omega\phi}$ is given by $\sum_k k$. If $k$ is selected so that the sampling times are symmetrically located about zero and also the reference symbols are distributed symmetrically about the center of the burst then this term equals to zero $\sum_{-\frac{N-1}{2}}^{\frac{N-1}{2}} k = 0$, where $N$ is an odd number.

Under this condition, the FIM is diagonal matrix, which means that the joint frequency and phase estimation are *decoupled*. The frequency and phase estimators become independent. In general, the estimator variance will be reduced.

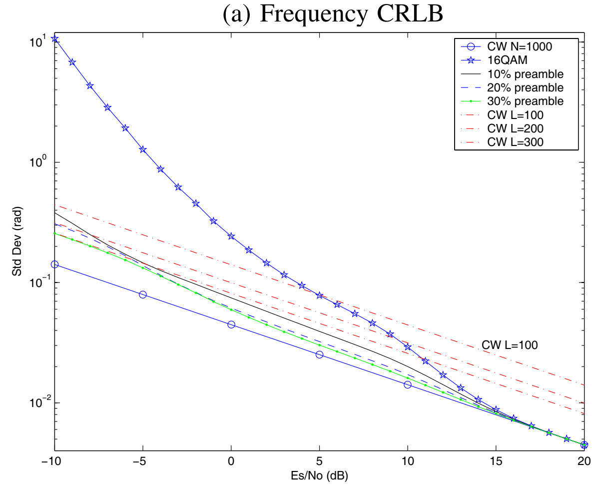
Fig 4 presents two structures of transmission burst with different $k$ ranges. Fig 4 (a) and (b) show the same structure but different $k$ ranges, (a) $k$ from 0 to $N-1$ and (b) $k$ is with symmetrical range.

Figure 5 shows the effect of the $k$ ranges on the joint frequency and phase CRLBs. The midamble burst with 16QAM modulation is used as an example, where $N = 101, P = 13$. The solid red line is with $\mathcal{K}_N$ and the dashed green line is with $\mathcal{K}_S$. The ranges $\mathcal{K}_N$ or $\mathcal{K}_S$ does not impact on the frequency CRLB (can be approved from the calculation). However, the phase CRLB of $\mathcal{K}_N$ is worse than that of $\mathcal{K}_S$. The reason is that the frequency estimate degrades the phase CRLB of $\mathcal{K}_N$ range. The degradation for the CRLBs at high SNR is $\frac{4N-2}{N+1}$ [1]. For the decoupled case $\mathcal{K}_S$ the phase error decreases compared with that of $\mathcal{K}_N$, especially at low SNR.

There are many difference arrangement of data and reference symbols in symmetrical distributed burst. In general, we can divide them into two categories: (1) the reference symbols placed in the middle of the burst are called *midamble transmission burst* (MTB); (2) the data symbols placed in the middle of the burst are called *data centered transmission*

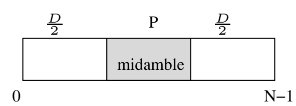



(a) Frequency CRLB

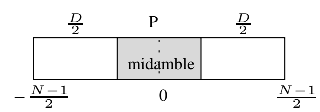



(b) Phase CRLB

**Fig. 3:** Joint frequency and phase CRLBs for preamble burst, N=1000, 16QAM with various P




(a) Range $\mathcal{K}_N$




(b) Range $\mathcal{K}_S$

**Fig. 4:** Examples of interleaved transmission bursts

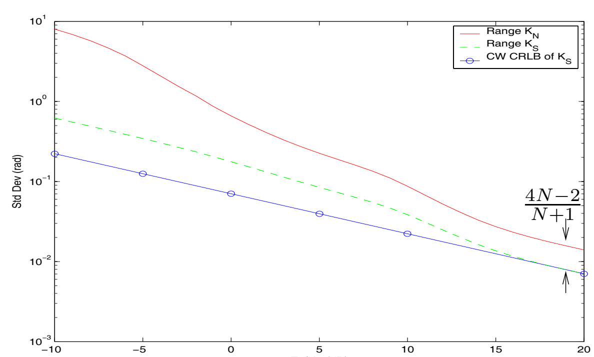



**Fig. 5:** Comparison of the phase CRLBs for different $k$ ranges, 16QAM modulation, midamble burst, N=101, P=13

*burst* (DCTB). For this case it does not matter where the rest reference symbols distributed in the packet as long as they are symmetrically placed about the center of the burst. The detailed discussion of effect of the location of reference symbols on frequency accuracy was given in [5].

## 4. CRLB of Pilot Symbols

Since the location of the reference symbols is critical in frequency estimation the conventional CW frequency CRLB [9] is not directly applicable to the distributed reference symbols. The CRLB for the symmetrical inserted known symbols is the first curly brackets (reference part) of the Eqns (9) and (10) in [5]. The FIM of the range $\mathcal{K}_N$ is the pilot symbol related terms in Eqns (6), (7) and (8) ($F_i = 1$), disregarding the data symbol related terms.

Comparison of the CRLBs of interleaved reference and data transmission burst and the CRLBs of the distributed reference symbols only is given in Fig 6. The DCTB is with N=301, P=30, D=271 and different number of sub-group reference symbols $n$. When $n = 2$ the burst contains the preamble and postamble with 15 symbols each. The crossed line "x-" is the CRLB for the burst with reference and data symbols. The dashed line is the CRLB of preamble and postamble only. When $n = 6$, the burst is inserted with 6 groups of reference symbols separated by 45 data symbols from the central data symbol. When $n = 30$, the 30 pilot groups are inserted in the burst with 10 data symbols apart. At the low SNR the CRLBs of reference symbols are good approximations of the true CRLBs. However, at moderate and high SNR they deviate from the true CRLBs of the interleaved reference and data bursts.

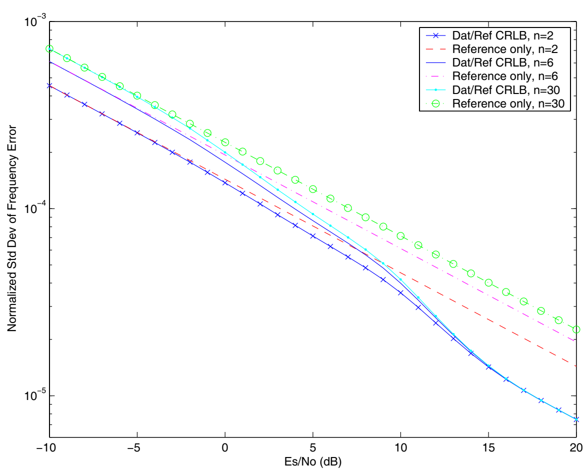



**Fig. 6:** Frequency CRLBs for 16QAM, P=30, N=301,DCTB

Figure 7 shows the frequency CRLBs for 16QAM modulation with the preamble and postamble burst, $P = 10$ and various N. The crossed line "x-" is the true CRLB for N=101 and the circle line "o-" is the CRLB of preamble and postamble only. In the rest of plot $P = 10$ and $N$ varies. To obtain good approximation more than 5% of the reference symbols are needed in the burst for estimation process using the reference symbols only at SNR $\leq 0$ dB.

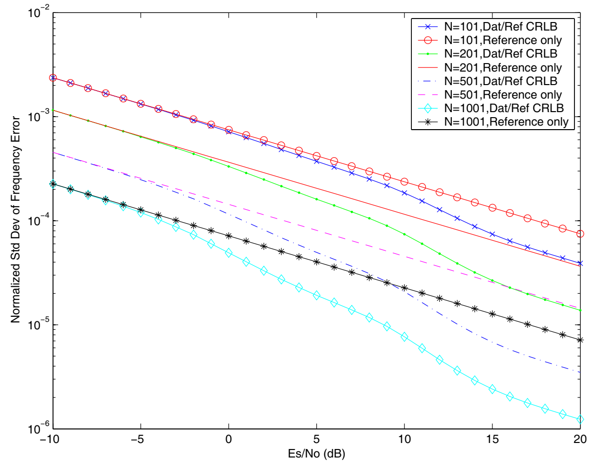



**Fig. 7:** Frequency CRLBs for 16QAM modulation with preamble and postamble burst, $P = 10$

## 5. Conclusion

Joint carrier phase and frequency CRLBs for transmissions with pilot-assisted burst have been derived. The reduction in frequency and phase CRLBs due to introduction of pilot symbols increases as the modulation order increasing. Symmetrical burst may have better CRLBs than that of non-symmetrical burst.

The pilot symbols CRLB has also been presented. The comparison of the true CRLB and the CRLB of the pilot symbols has revealed that using only pilot symbols for estimation is an effective approach only if more than 5% of pilot symbols in the burst and SNR$\leq$ 0 dB.

## Acknowledgments

The author would like to acknowledge helpful discussions with Dr Mark Rice that motivated the work.

## References

[1] F. Rice, B. Cowley, B. Moran, and M. Rice, "Cramér-Rao lower bounds for QAM phase and frequency estimation," *IEEE Trans. Commun.*, vol. 49, NO. 9 pp. 1582–1591, September 2001.
[2] F. Rice, M. Rice, and B. Cowley, "A New Bound and Algorithm for Star 16QAM Carrier Phase Estimation," *IEEE Trans. Commun.*, vol. 51, NO. 2 pp. 161–165, February 2003.
[3] F. Rice, *Bounds and Algorithms for Carrier Frequency and Phase Estimation*. Ph.D Thesis, University of South Australia, [on-line] http://www.library.unisa.edu.au/adt-root/public/, 2002.
[4] W. Cowley, "Phase and frequency estimation for PSK packets: Bounds and algorithms," *IEEE Trans. Commun.*, vol. 44, pp. 1–3, January 1996.
[5] F. Rice,"Carrier Phase and Frequency Estimation Bounds for Transmissions with Embedded Reference Symbols" accepted by *IEEE Trans. Commun.*.
[6] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
[7] U.Mengali and A. D'Andrea, *Synchronization Techniques for Digital Receivers*. Plenum Press, 1997.
[8] M. Moeneclaey, "On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Trans. Commun.*, vol. 46, pp. 1536–1544, November 1998.
[9] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 591–598, September 1974.
[10] W.T.Webb and L.Hanzo, *Modern Quadrature Amplitude Modulation Principles and Applications for Fixed and Wireless Communications*. IEEE Press and Pentech Press, 1994.