Hybrid Data-Aided and Code-Aided Carrier Synchronization Schemes for Burst-Mode Communications

Jin-Hua Sun, Jing Cheng

State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an, China
E-mail: jhsun@xidian.edu.cn, xdjcheng@126.com

Abstract-**For burst communication systems, high estimation accuracy and large estimation range cannot be achieved at the same time at low signal to noise ratio(SNR), a hybrid data-aided(DA) and code-aided(CA) carrier synchronization scheme is proposed. Firstly, the synchronization parameters coarse estimation is performed based on maximization of periodogram and cross-correlation algorithm utilizing known pilot symbols. Secondly, the fine estimation is obtained based on expectation-maximization (EM) algorithm exploiting the pilot symbols and soft information from the extended Turbo decoder. On the basis, the cascade of periodogram maximization and multiple cross-correlations has been proposed, which improves the estimation performance with low complexity. Simulation results show that the proposed algorithm can satisfy the requirements of frequency estimation range and accuracy, BER performance of Turbo decoding is very close to perfect synchronization and the estimation range can reach half of the symbol rate at a relatively efficient pilot cost.**

Keywords- synchronization; pilot; maximization periodogram; turbo code; expectation maximization

I.   INTRODUCTION

The applications of burst-mode communications have been more and more extensive in modern engineering practice, such as military communication and the deep-space communication. However, there is a relatively high-speed movement between both sides of communication, which will produce a large Doppler shift and thus cause synchronous receiver unable to work normally and finally result in a serious decline in the quality of communication. In such case, design an efficient synchronization algorithm is particular importance, which closely related to three indicators: estimation range, accuracy and complexity. So, our target algorithm should provide wide range, high precision, at the same time as well as moderate complexity.

Traditional data-aided carrier synchronization estimation algorithms such as Kay, Fitz, L&R, ML[1][2], etc, generally improve the estimation accuracy by increasing the length of the pilot symbols. However, that does not meet the requirements of burst communication systems. In addition, utilizing only pilot sequence cannot meet the estimation accuracy of precise synchronization required in a Shannon limit coding system such as Turbo code or LDPC (low-density parity-check) code. In order to reduce the cost of pilot symbols and obtain accurate synchronization, it's an effective way that introducing unknown data symbols to increase the total number of symbols for synchronization parameters estimation. This is the so called hybrid DA and CA synchronization scheme[3-10]. Obviously, the CA algorithms substantially

improve the estimation accuracy and achieve the accurate synchronization with limited pilot symbols at low SNR.

In order to obtain larger estimation range and higher estimation accuracy, a carrier synchronization scheme based on the pilot cross-correlation and EM algorithm is presented for Preamble-Postamble (PP) and Preamble-Middle (PM) structure in [3]. However this method can only correct small frequency offset $\left(\Delta fT_s \le 6\times10^{-4}\right)$. Reference [4] proposes a maximum likelihood iterative soft decision directed (ML-ISDD) carrier synchronization algorithm that exploits periodogram peaks and turbo decoding soft information. It is shown, via simulation, that this algorithm can improve the performance of synchronization parameters estimation at low SNR. Especially for $\Delta fT_s \le 7\times10^{-4}$, the CRB (Cramer-Rao bound) can be achieved by using only 22 pilot bits. However, due to the limited number of pilots, when $\Delta fT_s > 7\times10^{-4}$ the remaining frequency offset after coarse estimation can't meet the requirement of Turbo decoding. After iterative decoding, the reliability of soft information is still unable to be improved. The algorithm based on ML combines DA cross-correlation with turbo decoding information to estimate synchronization parameter[8]. Compared with the algorithm in [4], this algorithm exploits short pilot sequence but drastically improves the estimation accuracy and extends the estimation range to $\Delta fT_s \le 1.25\times10^{-3}$. However, the estimation range is relatively narrow in contrast to the periodogram-based algorithm. Therefore, how to estimate large frequency offset without reducing the estimation accuracy is a question we need to explore.

In burst communication systems, pilot position has an important influence on the estimation range and accuracy[11,12]. In order to obtain high estimation accuracy and wide estimation range, we adopt a three-block pilot frame structure based on [8]. Since the wide range FFT-based algorithm cannot obtain high estimation accuracy at limited pilot symbols, we consider utilizing the high accuracy of cross correlation with far pilot block spacing to improve the estimation performance. And code-aided EM synchronization utilizing data is also considered as an improved scheme.

The work of the paper is organized as follows. Firstly, the synchronization parameters coarse estimation is performed based on maximization of periodogram and cross-correlation algorithm utilizing known pilot symbols, and the frequency offset of coarse estimation is compensated to the receive signal. Secondly, the fine estimation is obtained based on EM algorithm exploiting the pilot sequence and soft information from the extended Turbo decoder. On the basis, the cascade of periodogram maximization and multiple cross-correlations has

been proposed. Finally, the periodogram maximization-cross-correlation with EM algorithm scheme and the periodogram maximization with multiple cross-correlation schemes are simulated separately in order to select ideal carrier synchronization estimation method. Simulation and analysis show that our proposed scheme greatly expands the frequency offset estimation range and improves estimation accuracy compared with the algorithms in [4,8].

II. SYSTEM MODEL

The system model considered in this paper is shown in Fig.1. Information bits are grouped into blocks and Turbo encoded. The length-N Turbo coded sequence is formed into a burst frame with a length-L pilot sequence as shown in Fig.2. The burst frame is then BPSK modulated and transmitted over an AWGN channel. Assuming perfect symbol timing recovery (which can be determined by the pilot symbols and data symbols), the complex baseband representation of the received signal is

$$r_k = s_k e^{j(2\pi \Delta f T_s k)} + w_k, k = 1,\ldots,N+L \tag{1}$$

where s_k is the transmitted signal, T_s is the symbol duration, Δf is the frequency offset and $\Delta f T_s \leq 1$, w_k is the kth complex AWGN sample with real and imaginary components having variance σ_w^2. L is the total length of the three-block pilot. The total length of the frame is $N+L$, where $N = N_1 + N_2$.

In [8], for the PMPM structure (the pilot sequence is equally divided and inserted into the data block with the same length), the spacing between two adjacent pilot blocks is D ($D = L/n + N/n$, n is the number of pilot blocks, and the data also has n blocks). The more blocks, the wider the estimation range, but the less estimation accuracy. The estimation range of the cross-correlation algorithm is $|\Delta f T_s| < 1/2D$. Since the data stream is divided into blocks with the same length in PMPM structure, the estimation range is determined by the spacing between two adjacent pilot blocks. The estimation accuracy usually is not very high due to the short distance between two adjacent pilot blocks. So, given the length of the data burst and the cost of pilot symbols ,we select this frame structure and consider unequal spacing between two adjacent pilot blocks in our frame structure. Generally, $N_1 < N_2$ as shown in Fig.1, smaller N_1 can guarantee large estimation range and larger N_2 can guarantee high estimation accuracy for cross-correlation algorithm. However, the estimation range of cross-correlation is still restricted by N_1, which is not larger than the estimation range of FFT-based algorithm. So we consider that maximization of periodogram estimation is done before cross-correlation estimation in the coarse estimation. And the data block length, N_1 and N_2, can be designed according to the estimation accuracy of maximization periodogram and the estimation

range of cross-correlation algorithm, both satisfying the requirements of estimation accuracy and obtaining higher spectrum efficiency. In addition, the code-aided EM algorithm and cross-correlation using pilot blocks with far distance are also considered in the fine estimation to improve the estimation accuracy further.
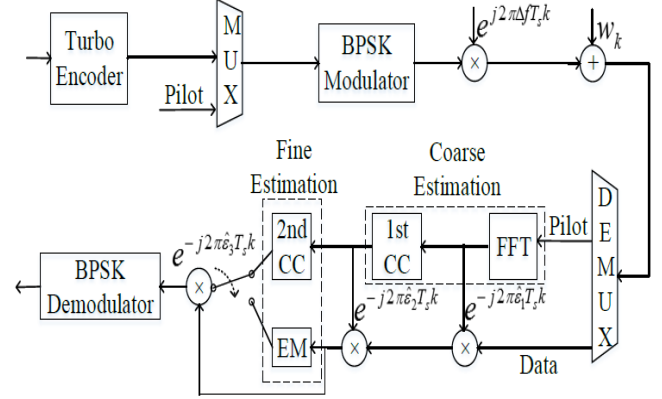

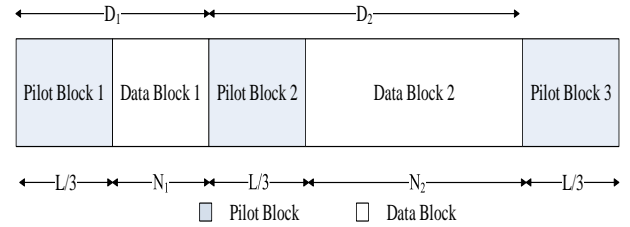Figure 1.   System model.

Figure 2.   Frame structure.

III. CARRIER SYNCHRONIZATION SCHEME

A. Coarse Estimation

For DA estimation, the dependence can be removed by multiplying r_k by s_k^* corresponding to the three pilot blocks. We have

$$z_k^1 = r_k s_k^*, k = 1,\ldots,L/3 \tag{2}$$

$$z_k^2 = r_k s_k^*, k = L/3 + N_1 + 1,\ldots,2L/3 + N_1 \tag{3}$$

$$z_k^3 = r_k s_k^*, k = 2L/3 + N + 1,\ldots,L+N \tag{4}$$

Then three frequency offsets are obtained according to the maximization of periodogram of three pilot blocks, i.e.

$$\hat{f}_i = \arg\max_m \left\{ FFT\left(z_k^i, M\right)\right\}/T_s M,$$

$$m = 0,1,......M-1, i=1,2,3 \tag{5}$$

with M being FFT points.

Averaging on $\hat{f}_i$, we get a preliminary frequency offset estimate,

$$\hat{\varepsilon}_1 = \left(\hat{f}_1 + \hat{f}_2 + \hat{f}_3 \right)/3 \tag{6}$$

According to the simulation results, with limited pilot symbols, the estimation performance based on FFTs of these three pilot blocks still cannot meet the residual frequency offset requirement of turbo decoding and guarantee convergence of turbo decoding. If too many pilot symbols are inserted into the data stream, the spectrum efficiency will be reduced. Therefore, we consider a cascade of cross-correlation algorithm after FFT with limited pilot symbols. However, how to perform cross-correlation needs further consideration, which will be detailed in section IV.

According to [3], the summation of the cross-correlation of two pilot blocks with distance D is used to estimate the frequency offset,

$$\hat{\varepsilon}_2 = \frac{1}{2\pi DT_s} \arg \left\{ \sum_{k=1}^{L/3} \sum_{m=1}^{L/3} z_{m+D} z_k^* \right\} \tag{7}$$

As it will be seen later, we have two schemes. One is that we can choose cross-correlation based on pilot block 1 and pilot block 2 with further code aided estimation (such as EM algorithm). The other one is further cross-correlation based on pilot block 2 and pilot block 3 is performed after cross-correlation based on pilot block 1 and pilot block 2, which is called multiple cross-correlations. So the coarse estimation value is that $\hat{\varepsilon} = \hat{\varepsilon}_1 + \hat{\varepsilon}_2$ and D= L/3+N1.

B. Fine Estimation

When the EM algorithm[5][6] is applied to the case of synchronization, the carrier synchronization parameters are iteratively estimated from the incomplete data $\mathbf{r}$ (the received signal vector) when there exists the nuisance parameter vector $\mathbf{s}$ (the transmitted symbols $(s_1, s_2, ..., s_{N+L})$). In the fine estimation, we consider introducing code-aided synchronization to further improve the estimation performance.

We define the value of frequency offset as $\hat{\eta}^n$,

$$\hat{\eta}^{n+1} = \arg \max_{\tilde{\eta}} \left\{ Q\left(\tilde{\eta}, \hat{\eta}^n \right) \right\} \tag{8}$$

$$Q\left(\tilde{\eta}, \hat{\eta}^n \right) = \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{r}, \hat{\eta}^n \right) \ln p\left(\mathbf{r} \mid \mathbf{s}, \tilde{\eta} \right) \tag{9}$$

where $\tilde{\eta}$ is a trial value of η , $p(\mathbf{s}\mid \mathbf{r}, \hat{\eta}^n)$ is posterior probability of the transmitted symbols given $\mathbf{r}$ and $\hat{\eta}^n$.

Taking (1) into account, we obtain

$$\ln p\left(\mathbf{r} \mid \mathbf{s}, \tilde{\eta} \right) \propto \left| \sum_{k=1}^{N+L} \left(r_k - s_k e^{j2\pi\tilde{\eta}T_s k} \right) \right|^2 / 2\sigma_w^2$$

$$\propto \Re \left[\sum_{k=1}^{N+L} \left(r_k s_k^* e^{-j2\pi\tilde{\eta}T_s k} \right) \right] \tag{10}$$

where $\Re$ represents the real part.

Since $\mathbf{s}$ and η are independent, (9) can be written as
$$Q\left(\tilde{\eta}, \hat{\eta}^n \right) = \sum_{\mathbf{s}} p\left(\mathbf{s} \mid \mathbf{r}, \hat{\eta}^n \right) \ln p\left(\mathbf{r} \mid \mathbf{s}, \tilde{\eta} \right)$$

$$= \Re \left\{ \sum_{k=1}^{N+L} \left(\mu_k^* \left(\mathbf{r}, \hat{\eta}^n \right) r_k e^{-j2\pi\hat{\eta}^n T_s k} \right) \right\} \tag{11}$$

$$\mu_k \left(\mathbf{r}, \hat{\eta}^n \right) = \sum_s s_k p(s_k = s \mid \mathbf{r}, \hat{\eta}^n)$$
where   .
For the known pilot symbols, the a posterior average values are $\mu_k \left(\mathbf{r}, \hat{\eta}^n \right) = s_k$. For the unknown data symbols, supposing the symbols are independent of each other and uniformly distributed in the constellation alphabet, the a posterior average values are given by

$$\mu_k \left(\mathbf{r}, \hat{\eta}^n \right) = \sum_{s \in \Lambda} P\left(s_k = s \mid r, \hat{\eta}^n \right) s_k \tag{12}$$

where $P\left(s_k = s \mid r, \hat{\eta}^n \right)$ is the posterior probability of the decoded words, Λ is the constellation alphabet. In summary, the frequency offset estimation based on the EM algorithm is described as follows:

- step 1: set $n = 0$, $\eta^0 = \hat{\varepsilon}$, compute the **a** posterior average values $\mu_k \left(\mathbf{r}, \hat{\eta}^0 \right)$ from (12)
- step 2: calculate $\hat{\eta}^{n+1}$ from (11) and (8)
- step 3: calculate $\mu_k \left(\mathbf{r}, \hat{\eta}^{n+1} \right)$, $n = n+1$
- step 4: if n is less than the initialized number of iterations go back to step 2, else go to step 5
- step 5: $\hat{\varepsilon}_3 = \hat{\eta}^{n+1}$

If we consider further cross-correlation based on pilot block 2 and pilot block 3 in the fine estimation ,

$$\hat{\varepsilon}_3 = \frac{1}{2\pi DT_s} \arg \left\{ \sum_{k=1}^{L/3} \sum_{m=1}^{L/3} z_{m+D} z_k^* \right\}, D = L/3 + N_2 \tag{13}$$

IV. PERFORMANCE RESULTS

This section is devoted to the study of the system performance. We consider a rate-1/3, Turbo coded BPSK system with generating polynomial $(5,7)_8$ and iteration number of six. For the symbol rate of 10Mbps, there are 10

Turbo encoded blocks with each block of $1098bit$ in one $1.098\,\text{ms}$ burst. The pilot symbols are divided into three blocks and inserted into the data stream. The length of pilot symbols is detailed as follows.

A   Determination of Frame Parameters

First, the coarse estimation based on periodogram maximization(FFT) and cross-correlation(CC) utilizing pilot block 1 and 2 (2FFT-1CC) is considered. We find through Fig.3 that even the length of each pilot block is increased to $732bit$ and FFT point is increased to 4096, the RMSE (root mean square error) can only reach about 3.5×10^{-4}, which does not meet the requirements of turbo decoding. That is why we introduce three pilot blocks in the frame structure. It can be demonstrated through simulation that the RMSE can reach 1×10^{-4} when each pilot block has $366bit$ and FFT point is 2048 based on 3FFTs. And the cascade of cross-correlation can improve the estimation accuracy further.
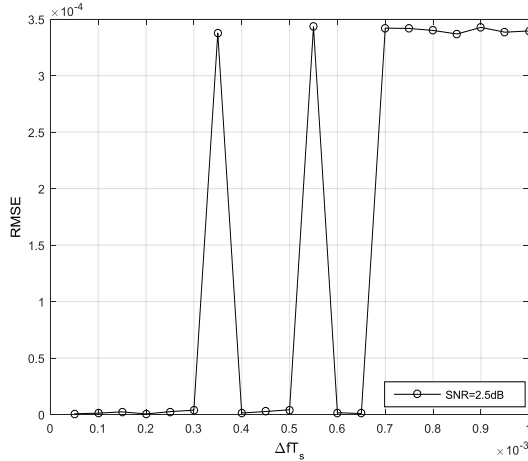
(a) 2FFT-1CC (L/3=732bits, M=4096)
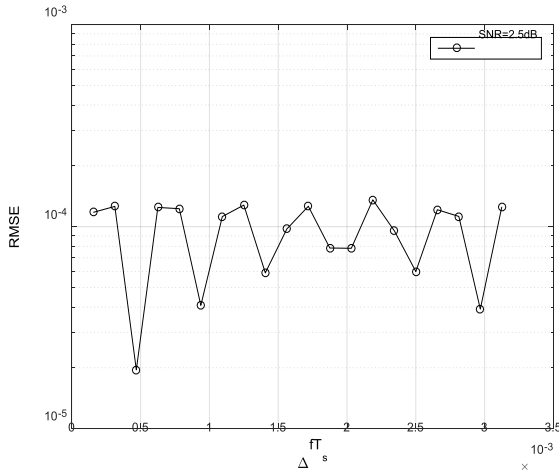
(b) 3FFT(L/3=366bits, M=2048)

Figure 3.   RMSE of 2FFT-1CC and 3FFT.

According to the estimation range $1/2D$ of the cross-correlation algorithm, the residual frequency offset should be with the estimation range of $1CC^{\left(1/2D_1 > 1\times10^{-4}\right)}$, so we can determine that $D_1 < 5000bit$. However, RMSE is the statistics average of root square error. In order to guarantee each single estimation meet the above point, we choose $N_1 = 1098bit$, so $D_1 = 1464bit$, $N_2 = 9882bit$, $D_2 = 10248bit$, the ratio of the pilot symbols to the data symbols is 10%.

B   Coarse Estimation

Fig.4 illustrates the performance of turbo decoding with 10 Turbo encoded blocks of $10980bits$ in total under different frequency offsets. As can be seen from Fig.4, when BER is in the range of $10^{-1} \sim 10^{-4}$ and $\Delta fT_s = 7\times10^{-6}$, 6×10^{-6} and 5×10^{-6}, the SNR degradations are about: 0.75dB, 0.5dB and 0.25dB compared with perfect synchronization. When $\Delta fT_s < 7\times10^{-6}$, the BER performance loss is relatively small at low SNR. Therefore $\Delta fT_s = 6\times10^{-6}$ is determined as a RMSE requirement of coarse estimation in this paper, which can guarantee convergence of turbo decoding and SNR degradation to be less than 0.5dB.
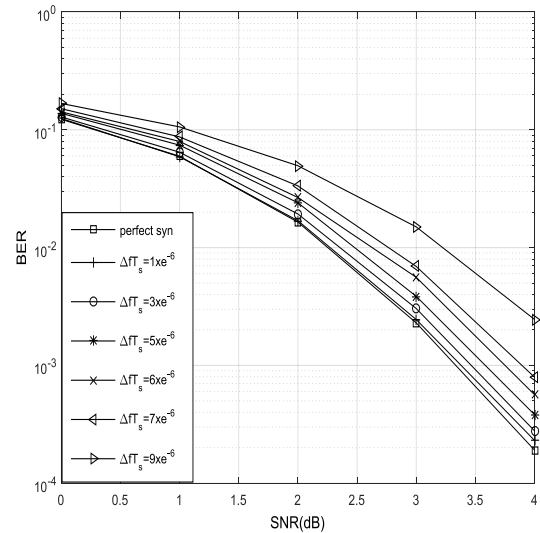

Figure 4.   BER performance for different frequency offsets.

According to the frame structure shown in Fig.2, there are three pilot blocks which can be used to carry out cross-correlation after 3FFTs. In the coarse estimation, cross-correlation using pilot block 1 and pilot block 2 (denominated as 1CC) is performed after 3FFTs. The residual normalized frequency offset for 3FFT-1CC is presented in Fig.5(a). As mentioned in section III, we also consider perform cross-correlation using pilot block 2 and pilot block 3 after 1CC (denominated as 2CC) in the fine estimation, the residual normalized frequency offset for 3FFT-2CC is also given in Fig.5(b). As can be seen, the remaining normalized frequency offset of 3FFT-1CC is about 7×10^{-6} at SNR of 0dB, which cannot meet the requirement of $\Delta fT_s < 6\times10^{-6}$. The remaining

normalized frequency offset can reach about 1×10^{-6} when exploiting 3FFT-2CC, which can make the BER performance of Turbo decoding close to theoretical BER Performance.

C   Fine Estimation

As stated above, the estimation accuracy of coarse estimation exploiting 3FFT-1CC cannot meet the requirements. So we consider the cascade of code-aided EM algorithm after 3FFT-1CC (3FFT-1CC-EM) and the scheme of 3FFT-2CC.
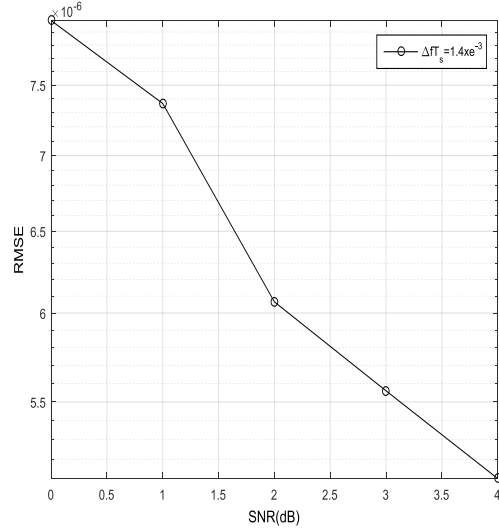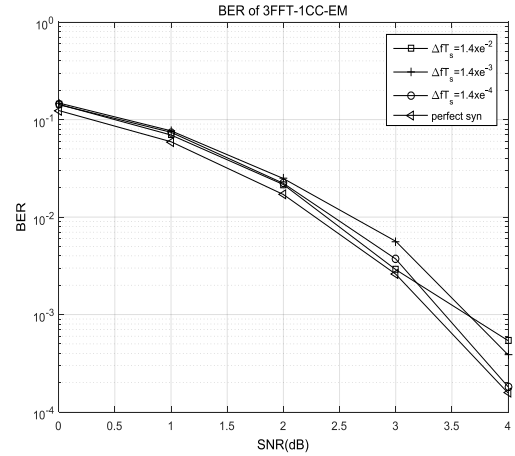

(a) 3FFT-1CC
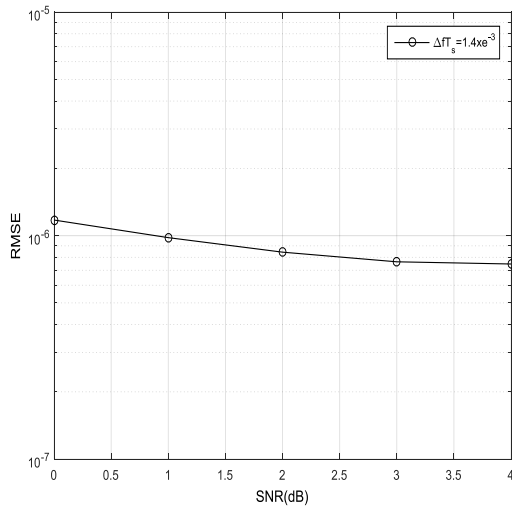
(b) 3FFT-2CC

Figure 5.   The remaining normalized frequency offset of 3FFT-1CC and 3FFT-2CC.

Fig.6 shows BER performance when $\Delta f T_s$ is 1.4×10^{-2}, 1.4×10^{-3} and 1.4×10^{-4} respectively compared with theoretical curve. As can be seen the two schemes can effectively estimate and correct large or small frequency offsets. Compared with 3FFT-1CC-EM, the BER performance of 3FFT-2CC is almost close to that under ideal synchronization. The SNR degradation is about 0.125dB.

Since there introduces too much more computation of posterior probabilities in code aided EM algorithm, and the complexity is very high due to the iteratively synchronization parameters estimation and Turbo decoding, 3FFT-2CC scheme is better with low complexity.

D   Estimation Range

Fig.7 shows the estimation range of our two schemes and CC-ML algorithm[8] at SNR of 3dB. It is clear that the estimation range of 3FFT-2CC and 3FFT-1CC-EM scheme is determined by the estimation range of FFT algorithm which can reach half of the symbol rate, and the estimation range of CC-ML is only 2×10^{-3}. Our scheme can estimate large frequency offsets effectively.
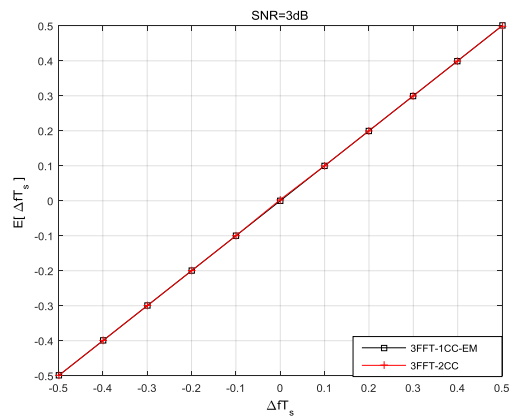

(a) 3FFT-1CC-EM
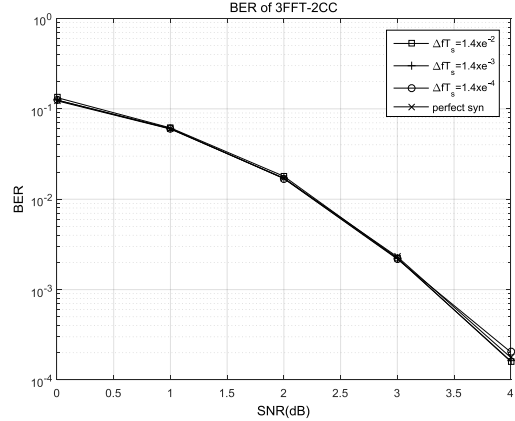
(b) 3FFT-2CC

Figure 6.   BER performance.

V.   CONCLUSIONS

In burst communication systems, carrier synchronization at low SNR and large frequency offsets is a question which needs to be solved. We propose 3FFT-2CC and 3FFT-1CC-EM synchronization schemes based on three-pilot-block frame structure. It can be demonstrated that the proposed schemes can correct large frequency offsets at 10% pilot cost, and make the BER performance of Turbo decoding approach that of perfect synchronization .

(a) 3FFT-1CC-EM and 3FFT-2CC
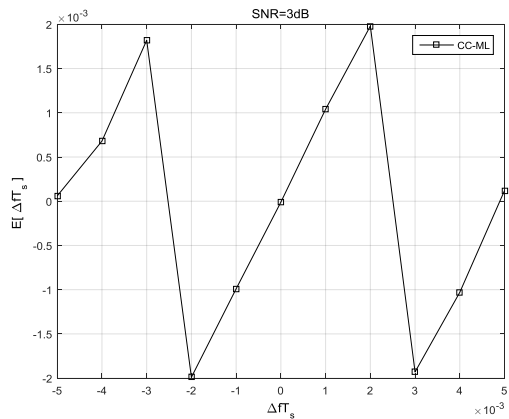
(b) CC-ML

Figure 7. Estimation range.

ACKNOWLEDGMENT

REFERENCES

[1] U. Mengali and A.N.D'Andrea, Synchronization Techniques for Digital Receiver, New York and London: Plenum Press, 1997.

[2] Y. Zhu, F. Yao, Y. Liu, "ML-based carrier frequency offset estimation and its low SNR threshold," Systems Engineering and Electronics, vol. 33,no.2,pp.427-431, Feb. 2011.

[3] H. Yan, F. Tang, Z. Zhang, "A joint assisted carrier synchronization algorithm with code and pilot based on low rate LDPC code," Journal of Electronics & Information Technology, vol.33,no.2,pp. 470-474, Feb. 2011.

[4] Y. Rahamim, A. Freedman, A Reichman, "ML Iterative Soft-Decision-Directed (ML-ISDD): A Carrier Synchronization System for Short Packet Turbo Coded Communication," IEEE Transactions on Communications, vol.56, no.7, pp.1169-1177, July 2008.

[5] N. Noels, C. Herzet, A. Dejonghe, et al. "Turbo synchronization: an EM algorithm interpretation,Communications," in Proc. IEEE International Conference on Communications. IEEE, May 2003, vol.4, pp.2933-2937.

[6] H. Wymeersch, M. Moeneclaey, "Iterative code-aided ML phase estimation and phase ambiguity resolution," EURASIP Journal on Advances in Signal Processing, no.6,pp.1-8, May 2005.

[7] D. J. Jakubisin, R. M. Buehrer, "Performance complexity, and receiver design for code-aided frame synchronization in multipath channels," IEEE Transactions on Communications,vol.63,no.9,pp. 3363-3376. Sep. 2015.

[8] J. Sun, X. Wang, X. Wu, "Joint pilot and iterative decoding carrier synchronization for short burst transmission system," Journal of Xidian University, vol.41,no.1,pp. 23-28, Feb. 2014.

[9] D. Jakubisin, C. I. Phelps, R. M. Buehrer, "Iterative joint detection, decoding, and synchronization with a focus on frame timing," 2014 IEEE Wireless Communications and Networking Conference (WCNC), pp.446-451, 2014.

[10] F. Bellili, A. Methenni, S. Affes, "Closed-form CRLBs for CFO and Phase estimation from turbo-coded square-QAM-modulated transmissions," IEEE Transactions on Communications,vol.14,no.5,May 2015.

[11] S. Godtmann, A. Pollok, N. Hadaschik, el at, "On the Influence of Pilot Symbol and Data Positioning on Turbo Synchronization," IEEE 65th Vehicular Technology Conference, VTC-2007 Spring, pp.1723-1726, 2007.

[12] S. Godtmann, N. Hadaschik, W. Steinert, G. Ascheid, "A concept for data-aided carrier frequency estimation at low signal-to-noise ratio," IEEE International Conference on Communications, ICC'08,pp.463-467, 2008.