Partial FFT demodulation scheme based on fast convolution structure

Xu Lin, Student Member, IEEE, Lin Mei, Member, IEEE, Fabrice Labeau, Senior Member, IEEE, Xuejun Sha, Member, IEEE and Yusi Zhang

Abstract—In systems that aim to mitigate the inter-carrier interference (ICI) caused by doubly selective (DS) channels, the partial Fast Fourier Transform (PFFT) has emerged as an interesting alternative to the conventional FFT. In this letter, we propose a novel PFFT demodulation scheme based on fast convolution (FC) structure. The proposed scheme uses the overlap-save operation of FC to further suppress the residual interference (RI) to improve performance over DS channels, including in terms of bit error rate (BER), while, at the same time, allowing for an improvement in spectral efficiency by avoiding the use of a cyclic prefix (CP). We provide a theoretical analysis of RI and numerical simulations to demonstrate the superiority of the proposed scheme in terms of BER performance.

Index Terms—Partial FFT, fast convolution, doubly selective channels.

I. INTRODUCTION

Future wireless communication systems must be designed to cope with doubly selective (DS) fading, due to the combined effect of high Doppler and multi-path channels [1]. High mobility communications [2], underwater acoustic (UWA) communications [3] and V2X communications [4] are good examples of such systems that are affected by DS channels. These DS channels lead to severe inter-carrier interference (ICI) for orthogonal frequency-division multiplexing (OFDM) based systems. Thus, equalization techniques [1] and ICI suppressing methods such as interference cancelation methods [5] and pulse shaping schemes [6] are proposed to combat ICI.

In recent years, a novel partial Fast Fourier Transform (PFFT) demodulation scheme [7] has shown its effectiveness for OFDM based systems to mitigate ICI caused by the highly distorted DS channels. In [8][9], Li et al. applied a PFFT demodulation scheme to the hybrid carrier (HC) communication system and showed the combined advantage of PFFT and HC modulation over DS channels. In [10], PFFT was applied to the differential coherent multichannel detection of acoustic OFDM signals. Then, Han et al. [11] extended the scheme to the multiple-input multiple-output (MIMO) case. Besides, an eigendecomposition-based PFFT demodulation scheme over UWA channels has been proposed in [12].

Similar to PFFT, fast convolution (FC) is also a kind of segmentation processing method, which usually used to implement linear convolution and has been exploited to try and alleviate the need for a cyclic prefix (CP) and hence increase throughput [13]. Recently, the FC structure has been applied in waveform design, in order to implement filter banks flexibly with low complexity [14], [15]. Besides, FC has also been used to improve the performance of banded equalization to form a novel hybrid carrier modulation successive transmission scheme [16]. Here, we use FC to improve the PFFT scheme.

In this letter, we propose to further use the overlap-save method from FC, in conjunction with PFFT, as a means to reduce residual interference (RI) from ICI, and hence improve BER performance. The main contribution of this paper is in the demodulation of the PFFT scheme, where we further segment received block through overlap-save, carry out MMSE weight compensation on each block, and discard the edges of each overlapping block, in which our processing has concentrated the effect of interference. This allows us to increase receiver performance without the need for a CP.

Acronyms: For the sake of clarity, the conventional OFDM scheme (including cyclic prefix) is designated as CP-OFDM; the conventional Discrete Fourier Transform-spread-OFDM scheme is designated as CP-DFT-S-OFDM. Similarly, equivalent schemes without a cyclic prefix are designated as NCP-OFDM and NCP-DFT-S-OFDM respectively.

Notation: Transpose, conjugate and conjugate transpose are denoted as $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$, respectively. $E[\cdot]$ denotes the statistical expectation of a random variable. $I_N$ is the $N \times N$ identity matrix. $(\cdot)^*$ denotes modulus-N calculation. $\lceil \cdot \rceil$ denotes the operation of rounding up, diag{$\{a\}$} denotes a diagonal matrix with a on its diagonal, $[\cdot]_{k,n}$ extracts the $k^{th}$ row and $n^{th}$ column entry from a matrix.

II. PFFT DEMODULATION SCHEME BASED ON FC

A. Signal model

The modulated system with PFFT demodulation based on FC is illustrated in Fig. 1, where the precoding module with dashed line can be inserted or not to generate two kinds of transmitted symbols in the absence of CP. The transmitted data is a continuous stream of NCP-DFT-S-OFDM or NCP-OFDM symbols when the precoding module is present or not. If the precoding module is omitted, an FFT module
at the end of the receiver is needed; the continuous stream of information data is split into symbols denoted as $d_j = [d_j(0), d_j(1), \ldots, d_j(N-1)]^T$, where $d_j(n)$ is the modulated symbol and $N$ is the information symbol length. Thus, the transmitted symbol is

$$\mathbf{x}_j = \begin{cases} \mathbf{F}^H \mathbf{d}_j, & \text{NCP-OFDM} \\ \mathbf{d}_j, & \text{NCP-DFT-S-OFDM} \end{cases},$$

(1)

where $\mathbf{F}$ is the normalized $N$-point DFT matrix. Define the total number of symbols as $J$, so the continuous stream yields

$$\mathbf{x} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \ldots, \mathbf{x}_{J-1}^T]^T = [x(0), x(1), \ldots, x(NJ-1)]^T.$$

(2)

Receiver with overlap-save fast convolution: As shown in Fig. 2, the received signal $y$ is processed according to the overlap-save operation of FC because the system transmits the symbols without CP. The basic principles of FC can be found in [15] and references therein. First, $y$ is divided into non-overlapping blocks of length $M$, the $i$th block being denoted as $y_i$, $i = 0, 1, \ldots, \left\lfloor \frac{JN}{M} \right\rfloor - 1$. Then, the $i$th overlapping block $\bar{y}_i$ can be obtained by taking $N = M + G$ middle values from $3$ adjacent blocks. The overlapping operation of FC can be described as follows:

$$\bar{y}_i = \mathbf{O}[y_{i-1}^T, y_i^T, y_{i+1}^T]^T,$$

(3)

where $\mathbf{O} = [0_{N \times (M-G/2)} \mathbf{I}_N 0_{N \times (M-G/2)}]$, is the overlapping operation matrix. For ease of processing, $G$ is chosen to be an even number and $N$ to be power of 2. Practically, $G/2$ should be larger than the maximum delay spread $L$. If $i = 0$, we can insert an $M$-point zero vector before $[y_0^T, y_1^T]^T$ to complete the aforementioned overlapping operation.

In order to analyze the effect of the channel, the transmitted signal $x$ can also be divided into a sequence of $M$-point non-overlapping blocks $x_i$. The corresponding $i$th overlapping block before passing the DS channel can be noted as $\bar{x}_i = \mathbf{O}[x_{i-1}^T, x_i^T, x_{i+1}^T]^T$. The $N$-point block data in front of $\bar{x}_i$ on the continuous stream, which causes the inter-block interference (IBI) to affect the transmission of $x_i$, when passing the DS channel, is denoted as $\mathbf{g}_i = \left[0_{N \times (2M-G/2-N)} \mathbf{I}_N 0_{N \times (2M-G/2-N)} \right]^T$.

Define the time-domain circular convolution channel matrix as $\mathbf{H}_{i,t}$, where $\mathbf{H}_{i,t}[n,(n-l)]_N = h_t(n, l)$, $n = 0, 1, \ldots, N-1$, $l = 0, 1, \ldots, L-1$. $h_t(n, l)$ is the $l$th tap response at time $n$ of the $i$th block, and the other elements of $\mathbf{H}_{i,t}$ are set to 0. Thus, the received overlapping block can be expressed as

$$\bar{y}_i = (\mathbf{H}_{i,t} - \mathbf{H}_{i,IBI})\bar{x}_i + \mathbf{H}_{i,IBI}\mathbf{g}_i + \mathbf{z}_i = \mathbf{H}_{i,t}\bar{x}_i + \mathbf{H}_{i,IBI}(\mathbf{g}_i - \bar{x}_i) + \mathbf{z}_i,$$

(4)

where $\mathbf{z}_i$ is the additive white Gaussian noise with zero mean and variance $N_0$, and the equivalent IBI $v_i = \mathbf{H}_{i,IBI}(\mathbf{g}_i - \bar{x}_i)$ is caused by the interference channel, shown as

$$[\mathbf{H}_{i,IBI}]_{n,(n-l)} = \begin{cases} h_t(n, l), & \text{if } n \leq l \\ 0, & \text{otherwise} \end{cases}.$$

(5)

B. PFFT demodulation

The discrete form of PFFT demodulation without CP is described in this subsection. As shown in Fig. 2, the overlapping block $\tilde{y}_i$ is divided into $U$ non-overlapping intervals\(^1\) as

$$\tilde{y}_i = [\bar{y}_{i,0}, \bar{y}_{i,1}, \ldots, \bar{y}_{i,U-1}]^T,$$

(6)

and $\bar{y}_{i,u} = W_u\tilde{y}_i$, where $W_u = \text{diag}\{w_u\}$ with the rectangular window $w_u$ is

$$w_u = [0, 0, \ldots, 0, 1, 1, \ldots, 1, 0, 0, \ldots, 0].$$

(7)

Then, the output of the PFFT demodulation is expressed as

$$\mathbf{m}_u = \mathbf{F}W_u\tilde{y}_i.$$

(8)

Substituting (4) into (8) yields

$$\mathbf{m}_u = \mathbf{F}W_u[H_{i,t}\bar{x}_i + v_i + \mathbf{z}_i] = \mathbf{F}W_u[H_{i,t}\mathbf{F}\bar{x}_i + \mathbf{F}W_u\mathbf{v}_i + \mathbf{F}W_u\mathbf{z}_i] = \mathbf{G}_u\mathbf{F}\bar{x}_i + \mathbf{V}_{IBI} + \mathbf{Z}_u,$$

(9)

where

$$\mathbf{m}_u = [m_u(0), \ldots, m_u(k), \ldots, m_u(N-1)]^T, k = 0, 1, \ldots, N-1 \text{ and } \mathbf{G}_u = \mathbf{F}W_u\mathbf{F}^H,$$

shown as

$$[\mathbf{G}_u]_{k,n} = \frac{1}{N} e^{\frac{2\pi i (n-k)(2n+1)}{2N}} \sin\left(\frac{\pi(n-k)}{N}\right).$$

(10)

Note that the block index $i$ is omitted for simplicity, and as the following analysis is based on a single overlapping block.

Assuming that the channel parameters vary slowly in each interval of length $N/U$, the channel impulse responses can be approximated by the ones at time of the midpoint in each interval. Thus the current $u$th interval channel matrix $\mathbf{W}_u\mathbf{H}_{i,t}$ can be approximated as $\mathbf{W}_u\mathbf{H}_{i,t}$, where $\left[H_{i,t}(n-l)\right]_{N}$ = \begin{cases} h_t(n, l), & \text{if } n \leq l \\ 0, & \text{otherwise} \end{cases}.$
\[ h_i((2u + 1)N/(2U), l), \] and the other entries of \( H_{\text{IBI}} \) are set to 0. The received signal (9) can now be simplified as
\[
m_u \approx G_u F H_i^i F^H X + V_{\text{IBI}} + Z_u = G_u A_u X + V_{\text{IBI}} + Z_u,
\]
where \( A_u = F H_i^i F^H = \text{diag}\{ [H_u(0), H_u(1), ..., H_u(N - 1)] \} \) and the entry of \( m_u \) can be shown as
\[
m_u(k) = \sum_{n=0}^{N-1} X(n) H_u(n) [G_u]_{k,n} + V_u(k) + Z_u(k).
\]
By extracting the \( k^{th} \) subcarrier from each \( m_u \), \( m_{\text{k}} = [m_0(k), m_1(k), ..., m_{U-1}(k)]^T \). The corresponding coefficient vector can be expressed as
\[
\Psi_{n,k} = \{ G_0[k,n], G_1[k,n], ..., G_{U-1}[k,n]\}^T.
\]
Then, \( m_{\text{k}} \) can be obtained by
\[
m_{\text{k}} = \sum_{n=0}^{N-1} X(n) H_u(k,\Psi_{n,k} + V_k + Z_k,
\]
where \( H_u = \text{diag}\{ [H_0(n), ..., H_{U-1}(n)] \} \) contains the channel frequency response of the \( n^{th} \) subcarrier in each interval, \( V_k = [V_0(k), ..., V_{U-1}(k)]^T \) denotes the equivalent IBI of the \( k^{th} \) subcarrier, and \( Z_k = [Z_0(k), ..., Z_{U-1}(k)]^T \) denotes the noise of the \( k^{th} \) subcarrier.

C. Weight compensation

Define \( \varphi_k = [\varphi_0(k), \varphi_1(k), ..., \varphi_{U-1}(k)]^H \) as the combiner weights of the \( k^{th} \) subcarrier, then the estimation of the \( k^{th} \) subcarrier after PFFT and weight compensation yields
\[
\hat{X}(k) = \varphi_k^H m_{\text{k}}.
\]

The \( i^{th} \) overlap-saving block after compensation, IFFT, and the saving operation shown in Fig. 2 can be expressed as
\[
\hat{x}_i = S\hat{x}_i = SF^H H_i [\varphi_0^H m_{\text{0}}, \varphi_1^H m_{\text{1}}, ..., \varphi_{U-1}^H m_{\text{U-1}}] = SX(k) = \varphi_k^H m_{\text{k}},
\]
where \( S = [0_{M \times G/2}, I_M, 0_{M \times G/2}] \) is the saving operation matrix used to complete FC, and \( \hat{x}_i \) is the estimation of \( x_i \).

If the MMSE estimation criterion is adopted, the optimal weights \( \varphi_{i,\text{opt-MMSE}} \) should minimize \( E[|X(k) - \varphi_k^H m_{\text{k}}|^2] \). Through matrix derivation, we have
\[
\varphi_{i,\text{opt-MMSE}} = \{ E[|m_{\text{k}}|^2] \}^{-1} E[|m_{\text{k}}|X^*(k)].
\]

After the PFFT, the equivalent IBI is divided into subcarriers. So the effect of \( V_k \) is limited and can be ignored. Bring \( m_{\text{k}} = \sum_{n=0}^{N-1} X(n) H_u[k,n] + Z_k \) into the derivation, we get
\[
R_{\text{m}_k} = E[|m_{\text{k}}|^2] = \sum_{n=0}^{N-1} H_n[k] \Psi_{n,k}^H H_n + N_0 \frac{1}{U} I_U
\]
and
\[
R_{\text{m}_k, X(k)} = E[|m_{\text{k}}|X^*(k)] = H_k \Psi_{k,k}, \text{where } I_U \text{ is the } U \times U \text{ identity matrix and } \Psi_{k,k} = \frac{1}{U} [1, 1, ..., 1]^T \text{ according to (13) and } [G_u]_{k,k} = \frac{1}{U} \sqrt{2}. \text{ Then, we have}
\]
\[
\varphi_{i,\text{opt-MMSE}} = R_{\text{m}_k}^{-1} H_k \Psi_{k,k}.
\]

The derivation of the optimal weights is consistent with that proposed in [7]. But derivations are all in discrete form in this letter. Now, the optimal combined PFFT outputs in (16) can be rewritten as
\[
\hat{x}_i = SF^H \{ (R_{\text{m}_k}^{-1} H_0 \Psi_{0,0}) H_{\text{0}}, (R_{\text{m}_k}^{-1} H_1 \Psi_{1,1}) H_{\text{1}}, \ldots, (R_{\text{m}_k}^{-1} H_{N-1} \Psi_{N-1,N-1}) H_{N-1} \}^T.
\]

According to (18) and (20), the optimal estimation can be obtained. Then, the continuous stream \( \hat{x} \) after compensation and segmentation shown in Fig. 1 can be expressed as \( \hat{x} = [\hat{x}_0^T, \hat{x}_1^T, ..., \hat{x}_{JN}^T, ...]^{-1} = [\hat{x}_0^T, \hat{x}_1^T, ..., \hat{x}_{JN-1}^T]^T \), where \( \hat{x}_j = \hat{x}(jN) \), \( \hat{x}(jN + 1), ..., \hat{x}(jN + N - 1) \) is the estimation of the transmitted symbol \( \hat{x}_j \). Finally, the estimation of \( d_j \) can be obtained by
\[
d_j = \left\{ \begin{array}{ll}
F \hat{x}_j, & \text{NCP-OFDM} \\
\hat{x}_j, & \text{NCP-DFT-S-OFDM}
\end{array} \right..
\]

D. Interference analysis

Define the deviation of \( \hat{x}_i \) after compensation as
\[
\Delta \hat{x}_i = \hat{x}_i - x_i = (C_i - I_N) \hat{x}_i + P_i (v_i + z_i),
\]
where \( C_i \) and \( P_i \) are expressed as
\[
C_i = \sum_{k=0}^{N-1} f_k w_k^H \Psi_{k,\text{opt-MMSE}} \sum_{u=0}^{U-1} b_u^H w_u F \Psi_{u} H_{u,i},
\]
and \( P_i = \sum_{k=0}^{N-1} f_k w_k^H \Psi_{k,\text{opt-MMSE}} \sum_{u=0}^{U-1} b_u^H w_u F \Psi_{u} \). Note that \( f_k \) and \( w_k \) denotes the \( k^{th} \) column of \( F^H \) and \( I_N \), respectively, and \( b_u \) denotes the \( u^{th} \) column of \( I_U \). \( (C_i - I_N) \hat{x}_i \) indicates the RI after the PFFT, weight compensation, and IFFT. \( P_i v_i \) indicates the RI from the equivalent IBI. The expectations of them are demonstrated in Fig. 3, which are performed over a wide-sense stationary uncorrelated scattering (WSSUS) channel described in Section IV and done for index \( i \) with 3000 channel realizations. Here, SNR is the ratio between total signal power of \( x \) and noise power \( N_0 \). As shown in Fig. 3, most of \( (C_i - I_N) \hat{x}_i \) is localized near both edges of the compensated overlapping block, so is \( P_i v_i \). Therefore, the saving operation of FC can discard the compensated data near both edges with most of them when \( G \) is much greater than \( 2L \). The total power of the RI can be reduced, and then the BER performance of the proposed scheme is improved.

E. Computational complexity

For a length-\( N \) overlapping block, the computational complexity mainly depends on the computation of PFFT and the calculation of weight compensation, which is the same as that of the conventional PFFT demodulation scheme with a length-\( N \) block. PFFT can be implemented by \( U \) FFTs, so it requires \( O(U \log N) \) operations. And the weight compensation requires \( O(N U^3) + O(N^2 U^2) \) operations, from the computation of one inverse matrix and two covariance matrices in (17). As for \( J \) symbols, the approximation of total computational complexity is shown in Table I. So the proposed scheme requires \( O(\lceil \frac{JN}{\sqrt{N}} \rceil U \log N) + O(\lceil \frac{JN}{\sqrt{N}} \rceil N U^3) + O(\lceil \frac{JN}{\sqrt{N}} \rceil N^2 U^2) \). Compared with the conventional PFFT demodulation scheme, the overlap-save operation of the proposed scheme increases \( \lceil \frac{JN}{\sqrt{N}} \rceil - J \) times, so the increase in complexity is very tolerable.
III. SIMULATIONS AND DISCUSSION

The CP-DFT-S-OFDM system and the CP-OFDM system are two kinds of typical CP based systems used for the transmission to resist channel fading. Different from these CP based systems described in [17], the FC is introduced in the proposed scheme to allow to discard data with more RI on both edges of the overlapping block after compensation. Numerical simulations are presented to compare the performance of the proposed scheme with the ones of these two CP based systems under the same compensation criterion and simulation conditions.

Here, the modulation mode is 4QAM; no channel coding is used. Let $N = 512$ and $J = 5000$, a continuous stream of symbols is transmitted without the insertion of CP for the proposed scheme. The DS channel is modeled by the WS-SUS Rayleigh-fading channel with an exponential multipath intensity profile [18] and $L = 36$. The normalized Doppler frequency is $f_d T_d = 0.32$. Such parameters correspond to high mobility communications [2] with bandwidth $B = 2$MHz, center frequency $f_c = 2.7$GHz, subcarrier spacing $\Delta f = B/N = 3.9$kHz, symbol duration $T_d = 250$us, delay spread 18us and velocity 500kmph. For the CP-OFDM and CP-DFT-S-OFDM systems, CP $T_p = T_d/8 = 32$us is inserted over the same DS channel. It is assumed that perfect channel state information (CSI) is known at the receiver.

Fig. 4 shows the BER versus SNR for the proposed PFFT demodulation scheme and the conventional PFFT demodulation scheme of the CP-DFT-S-OFDM system or the CP-OFDM system. It can be seen that the performance of the CP-DFT-S-OFDM system is much better than that of the CP-OFDM system under this DS channel, so only the NCP-DFT-S-OFDM data stream is considered for the proposed scheme. The conventional ones with CP can better eliminate the influence of IBI, but the proposed scheme is greatly affected by IBI when $G = 0$. The BER performance of the proposed scheme is much better than those of the conventional ones with the same $N$ and $U$, when $G/2$ is large enough.

When $G/2$ is greater than the maximum delay spread $L$, the overlap-save operation removes the influence of most IBI between blocks, and the introduction of the PFFT algorithm can better eliminate the ICI effect. Meanwhile, the proposed scheme performs PFFT demodulation and weight compensation on the overlap-saving block, then keep the middle part and discard the data at both edges after the IFFT. As analyzed in subsection D of Section II, the RI caused by the approximation of the channel frequency response is mainly localized in both edges of data blocks after the compensation and converting to the time domain. When $G/2$ is large enough, the operation of overlap-save can eliminate the influence of IBI and discard the data of poor performance, so the proposed scheme has the best performance.

As shown in Fig.4, the performance of $G = 200$ is almost the same as that of $G = 240$ with the same $U = 16$. It implies that the performance of the proposed scheme tends to be the same with the increase of $G$ when $G/2$ is large enough. As for the number of intervals $U$ of the PFFT, it can improve the performance of the system with the increase of complexity.

Another advantage of the proposed scheme is that the system does not rely on CP, so the spectral efficiency can be increased by $(T_d + T_p)/T_d - 1 = 0.125$ times.

IV. CONCLUSION

In this letter, we propose a novel PFFT demodulation scheme based on FC structure in the absence of CP to further suppress the ICI by discarding the data of poor compensation performance through the overlap-save operation of FC. The proposed scheme has better BER performance and higher spectral efficiency than the conventional PFFT scheme at a moderate complexity. Moreover, the analysis of RI and numerical simulations are presented to demonstrate the superiority of the proposed scheme.
REFERENCES


