Single Carrier Frequency Domain Compressed Training Adaptive Equalization

Baki Berkay Yilmaz
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0250, US
Email:b.berkayyilmaz@gatech.edu

Alper Tunga Erdogan
Department of Electrical and Electronics Engineering
Koc University
Istanbul, 34450, Turkey
Email:alper@stanfordalumni.org

***Abstract*—Compressed training approach offers to improve the spectral efficiency by significantly reducing the required training length. It is based on a convex optimization setting (SCFDE-CoTA) which combines two cost functions: 1) Least squares based training reconstruction performance, and 2) Infinity norm of equalizer outputs, which exploits magnitude boundedness assumption of digital communication symbols. In this article, we provide an extension of this framework to single carrier frequency domain equalizer based transceivers, where we show that the compressed training algorithms can be effectively implemented for training frequency domain equalizers. We provide examples to demonstrate the potential gain of the algorithm in terms of training length reduction and performance improvement.**

***Index Terms*—Single Carrier; Frequency Domain Equalizers; Adaptive Equalizers; Adaptive Signal Processing**

I. Introduction

The use of linear equalization scheme is a low complexity approach to counteract against the Inter-Symbol-Interference (ISI) effects caused by the multipath propagation environment [1]. The main challenge in this scheme is to adapt the coefficients of the equalizer filter for the unknown/changing propagation channel.

In the conventional communication systems, the adaptation is achieved by employing training/pilot symbols in the transmit packets which are also known by the receiver. These training symbols are utilized to adjust equalizer coefficients. Unfortunately, embedding training symbols into communication packets reduces the precious bandwidth available for the information symbols, especially for the channels with short coherence time. In this respect, blind algorithms are proposed to eliminate the training phase. The most conventional method, which completely removes the training symbol concept, is known as Constant Modulus Algorithm (CMA) [2], [3]. Despite the success of CMA on removing training symbols, its cost function is non-convex and has some ill/mis-convergence issues. Some papers [4]–[7] propose convex blind frameworks to remove these drawbacks. Even though these algorithms solve ill/mis-convergence behaviour of CMA, they require relatively large number of data, therefore, their performance in a highly mobile environment could not be satisfying. Furthermore, these algorithms could not handle the unresolved ambiguity on the delay and the phase of the equalizer outputs. To overcome these problems, semi-blind algorithms, which get benefit from a small number of training symbols embedded into data blocks, are proposed in [8]–[10]. These algorithms exploit both training symbols and a prior knowledge about the information symbols.

As an alternative to the time domain equalization (TDE) approaches, frequency domain equalization (FDE) based schemes have received more attention due to their low complexity structures enabling the real-time applications for high data rates. Moreover, the adaptive FDE algorithms were reported to converge faster than TDE algorithms [11].

A time domain equalization technique which utilizes the infinity norm of receiver outputs was proposed in [12]. It utilizes the infinity norm of the equalizer outputs for the whole packet as the regularizer to reduce the amount of training data required for training. This regularization cost function reflects the sparsity of the combined equalizer-channel impulse response where its minimization together with training reconstruction error leads to the reduction of the ISI. In this article, we propose the extension of this framework for single carrier FDE based transceivers, which we refer as **S**ingle **C**arrier **F**requency **D**omain **E**qualization - **Co**mpressed **T**raining **A**daptive method (SCFDE-CoTA). The framework is a combination of two convex optimization settings:

* Minimization of training symbol reconstruction error by employing least-squares cost function,
* Minimization of maximum magnitude of equalizer outputs by employing infinity norm cost function based on the pre-knowledge about the boundedness of communication symbols

whose goal is to reduce the number of training symbols to increase the efficiency of the available bandwidth and decreasing the complexity of the equalization process.

Fig. 1. Transmitted Data Block to Eliminate Inter-Block Interference

II. Adaptive Equalization Setting

In this section, we introduce the set-up for the single carrier FDE receiver with multiple diversity branches [13], [14]. The

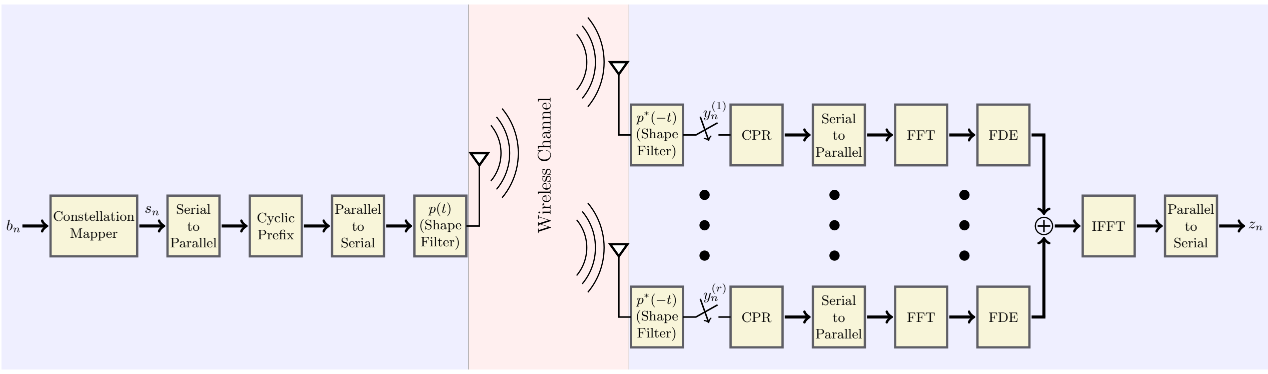

Fig. 2. Single-Carrier Frequency-Domain System

baseband equivalent communication system is provided in Fig. 2 and can be described as follows:

- Mapping of the binary input sequence $\{b_n\}$ to transmit symbols $\{s_n\}$ is performed by constellation mapper where the transmitter uses PAM or QAM constellation schemes.
- A sequence, generated by a known Unique Word (UW), $\{x_n \mid n \in \{1, \cdots, L_T\}\}$, followed by data symbols, $\{d_n \mid n \in \{1, \cdots, L_D\}\}$, is fed to serial-to-parallel (S/P) converter to produce data block of length $L_P = L_D + L_T$.
- To create a circular convolution channel and eliminate inter-block interference (IBI), UW is also placed at the end of the data block as illustrated in Fig. 1.
- The size of UW is selected to be greater than the presumed channel spread because it serves as the cyclic-prefix block as well as the training sequence.
- The channel is assumed to be finite and time invariant (FIR) whose output is corrupted by noise.
- A receiver with r diversity branches is assumed (i.e. separate antennas and/or an oversampling phase).
- To minimize self ISI, the front-end receiver contains a matched filter with impulse response $p^*(-t)$ (i.e. $p(t)$ is the transmit pulse shape filter).
- The discrete time equivalent channel model under symbol-space sampling for the k^{th} receiver branch is

$$y_n^{(k)} = \sum_{l=0}^{L_C-1} h_l^{(k)} s_{n-l} + v_n^{(k)}, \quad k = 1, \ldots r \qquad (1)$$

  where $\{y_n^{(k)}\}$ is the sampled receiver sequence, $\{h_n^{(k)} \mid n \in \{0, \ldots, L_C - 1\}\}$ represent the discrete-time equivalent FIR channel coefficients, which include the communication channel and transmit-receiver filters and $v_n^{(k)}$ represents zero-mean complex Gaussian noise samples with variance σ_v^2.
- We do not assume that $\{h_n^{(k)}\}$ is sparse. On the contrary, we remove all constraints which impose sparse structure to channels.
- The pulse shape $p(t)$ is chosen to have zero auto-correlation at lag, therefore, $\{v_n^{(k)}\}$ is assumed to be an independent and identically distributed (i.i.d.) sequence.
- The received data block at the k^{th} antenna can be written as

$$\mathbf{y}_n^{(k)} = \begin{bmatrix} y_n^{(k)} & y_{n-1}^{(k)} & \ldots & y_{n-L_T-L_P+1}^{(k)} \end{bmatrix}^T \qquad (2)$$

  for $k = 1, \ldots, r$.
- After prefix removal, time domain receiver samples can be written as

$$\mathbf{y}^{(k)} = \mathrm{CPR}\left\{\mathbf{y}_n^{(k)}\right\} = \begin{bmatrix} y_n^{(k)} & y_{n-1}^{(k)} & \ldots & y_{n-L_P+1}^{(k)} \end{bmatrix}^T$$

  for $k = 1, \ldots, r$ where $\mathrm{CPR}\{\bullet\}$ stands for the cyclic-prefix removal operation.
- $\mathbf{y}^{(k)}$ is converted to frequency domain vector $\mathbf{Y}^{(k)}$ through L_P-point FFT operation such that $\mathbf{Y}^{(k)} = \mathbf{F}^H \mathbf{y}^{(k)}$. Here, $\mathbf{F}$ is the normalized DFT basis matrix with

$$\mathbf{F}_{lm} = e^{\frac{j2\pi(l-1)(m-1)}{L_P}} / \sqrt{L_P}. \qquad (3)$$

- The frequency domain equalizer coefficient vector for the k^{th} branch is $\mathbf{W}^{(k)}$ where $k = 1, \ldots, r$.
- Elementwise-multiplied frequency domain vectors corresponding to the equalizing operation is defined as

$$\mathbf{Z} = \sum_{k=1}^{r} \mathbf{W}^{(k)} \odot \mathbf{Y}^{(k)} \qquad (4)$$

  where $\odot$ is the elementwise multiplication operator. The equalizer output converted to the time domain is given by $\mathbf{z} = \mathbf{F}\mathbf{Z}$.

III. Adaptive Single-Carrier Frequency-Domain Equalization

In this section, we introduce the SCFDE-CoTA as an approach for adaptively training the the frequency domain equalizers described in the previous section. As the first step, a circulant convolution channel is obtained by embedding cyclic-prefix. As illustrated by Fig. 1, a predetermined Unique-Word (UW) sequence $\{x_n\}$ is placed to the beginning and to the end of the data sequence $\{d_n\}$ to form the transmit block.

The main goal of our discussion is to survive in highly mobile environment with shorter channel coherence time, in other words, the adaptive setting is constraint to be trained using only one single block. Actually, this is a desired property which is compelled by the wireless systems. In this respect, our framework is a very favourable candidate because the framework aims to adapt equalizer coefficients with relatively small number of training symbols to provide more space for data symbols. This phenomena attracts more attention as shorter data packets are mandated due to the shorter channel coherence time (the presented approach can be easily extended to multi-block-based training). For the adaptive compressed training-based SCFDE, we propose the following optimization setting:

$$\underline{\textit{Setting-I:}} \qquad \begin{aligned} \underset{\mathbf{W}}{\text{minimize}} \quad & \|\mathbf{E}_d\mathbf{z} - \mathbf{x}\|_2 \\ \text{subject to} \quad & \mathbf{W} \in \mathcal{F}_{L_E} \\ & \|\tilde{\mathbf{z}}\|_\infty \leq \delta \\ & \mathbf{z} = \mathbf{F}\sum_{k=1}^{r} \mathbf{W}^{(k)} \odot \mathbf{Y}^{(k)} \end{aligned}$$

where

- The vector containing the UW (training) symbols is $\mathbf{x} = \left[\begin{array}{ccc} x_1 & x_2 & \ldots, x_{L_T} \end{array}\right]$,
- $\tilde{\mathbf{z}} = [\Re e\{\mathbf{z}\}^T \quad \Im m\{\mathbf{z}\}^T]^T$ where $\Re e\{\bullet\}$ and $\Im m\{\mathbf{z}\}$ give the real and imaginary part of their arguments respectively.
- d is the target equalization delay.
- $\mathbf{E}_d$ extracts the parts of the $\mathbf{z}$ vector corresponding the UW based on the selected choice of delay as
$$\mathbf{E}_d = \left[\begin{array}{cc} \mathbf{0} & \mathbf{I}_{L_T} \end{array}\right]\mathbf{A}_d \tag{5}$$
where multiplication of $\mathbf{z}$ with $\mathbf{A}_d$ is equivalent to applying the circular d-advance operation on $\mathbf{z}$ to position the elements of $\mathbf{z}$, which corresponds to UW, to the last L_T rows, i.e., the compensation of the equalization delay. Multiplying by $\left[\begin{array}{cc} \mathbf{0} & \mathbf{I}_{L_T} \end{array}\right]$ extracts the UW region from the circularly shifted $\mathbf{z}$.
- $\mathcal{F}_{L_E}$ is the set of frequency domain equalizers whose impulse response length less than or equal to L_E, which can be written as
$$\mathcal{F}_{L_E} = \left\{\mathbf{W} \in C^{L_P} : \left[\begin{array}{cc} \mathbf{0}_{L_P - L_E \times L_E} & \mathbf{I}_{L_P - L_E} \end{array}\right]\mathbf{F}\mathbf{W} = \mathbf{0}\right\}.$$
- δ is chosen based on the prior information about the input constellation.

Setting-I minimizes the reconstruction error of training symbols. On the other hand, the optimization setting can be modified by replacing the objective function and the second constraint of *Setting-I* to obtain the minimum size hypercube that contains all outputs as follows :

$$\underline{\textit{Setting-II:}} \qquad \begin{aligned} \underset{\mathbf{W}}{\text{minimize}} \quad & \|\tilde{\mathbf{z}}\|_\infty \\ \text{subject to} \quad & \mathbf{W} \in \mathcal{F}_{L_E} \\ & \|\mathbf{E}_d\mathbf{z} - \mathbf{x}\|_2 \leq \epsilon \\ & \mathbf{z} = \mathbf{F}\sum_{k=1}^{r} \mathbf{W}^{(k)} \odot \mathbf{Y}^{(k)} \end{aligned}$$

Algorithm 1 Compressed Training-Based Adaptive SCFDE Algorithm Update

1 $\texttt{while}\left(\sum_l \|\mathbf{U}^{(l)}[k-1] - \mathbf{U}^{(l)}[k-2]\|_2 \geq \epsilon\right)$ {
2 `// Equalizer output in frequency domain`
3 $\mathbf{Z}[k] \leftarrow \sum_{l=1}^{r} \mathbf{W}^{(l)}[k] \odot \mathbf{Y}^{(l)}$
4 `//Equalizer output in time domain`
5 $\mathbf{z}[k] \leftarrow \texttt{ifft}_{L_P}(\mathbf{Z}[k])$
6 `//TD training symbols reconstruction error`
7 $\mathbf{e}[k] \leftarrow \mathbf{E}_d\mathbf{z}[k] - \mathbf{x}$
8 `//FD training symbols reconstruction error`
9 $\mathbf{E}[k] \leftarrow \texttt{fft}_{L_P}\left(\left[\begin{array}{c} \mathbf{0} \\ \mathbf{e}[k] \end{array}\right]\right)$
10 `//Real peak outputs in time domain`
11 $\mathbf{p}_{re}[k] \leftarrow \sum_{l=1}^{L_P} 1_{|\tilde{z}_l[k]| \geq \alpha\|\tilde{\mathbf{z}}[k]\|_\infty} \texttt{sign}(\tilde{z}_l[k])\mathbf{e}_l$
12 `//Imaginary peak outputs in time domain`
13 $\mathbf{p}_{im}[k] \leftarrow \sum_{l=L_P+1}^{2L_P} 1_{|\tilde{z}_l[k]| \geq \alpha\|\tilde{\mathbf{z}}[k]\|_\infty} \texttt{sign}(\tilde{z}_l[k])\mathbf{e}_{l-L_P}$
14 `//Transform peaks into frequency domain`
15 $\mathbf{P}[k] \leftarrow \texttt{fft}_{L_P}\left(\underbrace{\mathbf{p}_{re}[k] + i\mathbf{p}_{im}[k]}_{\mathbf{p}[k]}\right)$
16 `//Subgradient projection`
17 $\mathbf{G}^{(l)}[k] \leftarrow \mathcal{P}_{\mathcal{F}_{L_E}}\left(\left(\frac{\mathbf{E}[k]}{\|\mathbf{e}[k]\|_2} + \frac{\lambda}{\|\tilde{\mathbf{p}}[k]\|_1}\mathbf{P}[k]\right) \odot \bar{\mathbf{Y}}^{(l)}\right)$
18 `//Subgradient update of the equalizer`
19 $\mathbf{U}^{(l)}[k] \leftarrow \mathbf{W}^{(l)}[k] - \mu[k]\mathbf{G}^{(l)}[k]$
20 `//Nesterov Step`
21 $\mathbf{W}^{(l)}[k+1] = \mathbf{U}^{(l)}[k] + \frac{k-1}{k+2}(\mathbf{U}^{(l)}[k] - \mathbf{U}^{(l)}[k-1])$
22 `//Increase the number of iteration`
23 `k = k + 1`
24 }

Another setting can be obtained by combining the objective function and the second constraint of *Setting-I* to form a Lagrangian function. Hence, we can propose the following optimization setting:

$$\underline{\text{SCFDE CoTA:}} \qquad \begin{aligned} \underset{\mathbf{W}}{\text{minimize}} \quad & \|\mathbf{E}_d\mathbf{z} - \mathbf{x}\|_2 + \lambda\|\tilde{\mathbf{z}}\|_\infty \\ \text{subject to} \quad & \mathbf{W} \in \mathcal{F}_{L_E} \\ & \mathbf{z} = \mathbf{F}\sum_{k=1}^{r} \mathbf{W}^{(k)} \odot \mathbf{Y}^{(k)} \end{aligned}$$

where λ is the regularization constant.

The iterative algorithm to solve the optimization problem SCFDE-CoTA is given in the Algorithm 1:

- Lines 2 through 5 construct the equalizer outputs. Equalizer outputs are first generated in frequency domain, then, transformed into time domain.
- In lines 6 through 9, reconstruction error vector of training symbols is created at each iteration and converted to its frequency domain counterpart.
- Lines 10 through 13 contain the operations which are required to determine the potential peaks of the equalizer outputs. These operations are very crucial to generate the

subgradient in Line 19 [6]. In these lines, $\tilde{\mathbf{z}}$ represents cascaded real and imaginary components of time domain equalizer outputs, where $\tilde{\mathbf{z}} = [\Re e\{\mathbf{z}\}^T \quad \Im m\{\mathbf{z}\}^T]^T$, and $\|\tilde{\mathbf{z}}\|_\infty$ gives the maximum magnitude of these components.

Due to existence of additive noise in the communication channel, the sample peak, which yields the value of $\|\tilde{\mathbf{z}}\|_\infty$, may not coincide with the actual peak sample(s). Therefore, absence of actual peak sample(s) can cause decrease in algorithm performance. To increase the probability of including the true peak sample(s) of the system to the subgradient vector, we cover all subgradient vectors corresponding to cascaded equalizer outputs such that the magnitudes of the selected samples are within α factor range of $\|\tilde{\mathbf{z}}\|_\infty$. Here, α is an empirical algorithm parameter. In this respect, $1_{|\tilde{z}_l[k]| \geq \alpha \|\tilde{\mathbf{z}}[k]\|_\infty}$ is defined to choose the potential peaks where $1_{(\bullet)}$ is the indicator function which returns one or zero if its input condition is true or false respectively. After selection of possible peak samples, l^{th} component of $\mathbf{p}_{re}[k]$ ($\mathbf{p}_{im}[k]$) is set to zero or sign of the l^{th} component of $\tilde{\mathbf{z}}$ according to output of the indicator function.

- In line 15, time domain subgradient of the infinity norm cost function is transformed into its frequency domain representative.
- Line 17 finds the overall subgradient of the objective function of SCFDE-CoTA in frequency domain. Here, $\tilde{\mathbf{p}}[k] = [\Re e\{\mathbf{p}[k]\}^T \quad \Im m\{\mathbf{p}[k]\}^T]^T$ and $\mathcal{P}_{\mathcal{F}_{L_E}}$ is the projection defined as

$$\mathcal{P}_{\mathcal{F}_{L_E}}(\mathbf{a}) \triangleq \texttt{fft}\left\{\left[\begin{array}{cc} \mathbf{0}_{N-L_E \times L_E} & \mathbf{I}_{N-L_E} \end{array}\right] \texttt{ifft}\left\{\mathbf{a}\right\}\right\} \tag{6}$$

  which transforms its input vector to time domain, zeros out all components except the first L_E ones in time domain and then transforms back to the frequency domain.
- Lines 18 through 21 accelerate the convergence behaviour of the algorithm by exploiting "Nesterov Method" [15], where $\mathbf{U}[k]$ is intermediate variable of the algorithm and $\mu[k]$ is the step-size parameter.

In addition to steps provided in Algorithm 1, the structure of the algorithm is also suitable to work on a decision directed mode. In that respect, the performance of the algorithm can be improved further by expanding the training region with reliable decisions after achievement of first eye opening.

The main goals of SCFDE-CoTA are to reduce the required number of training symbol by exploiting the duality between ℓ_1 and ℓ_∞ norm based cost functions [16], and decrease the computational complexity of equalization process. Complexity reduction is achieved by proposing a convex optimization setting function in frequency domain rather than complex convolution operations in time domain. Likewise, the proposed algorithm offers robust convergence behaviour. In terms of complexity of Algorithm 1, $\texttt{fft}(\bullet)/\texttt{ifft}(\bullet)$ operations dominate the algorithm's complexity. Therefore, the proposed scheme requires on the order of $L_P \log(L_P)$ operations. Although the complexity of the proposed algorithm is low, with the improvement in machine learning and big data based large-scale sparsity driven optimization settings, it is possible to increase the convergence speed or decrease the complexity of the algorithm further [17].

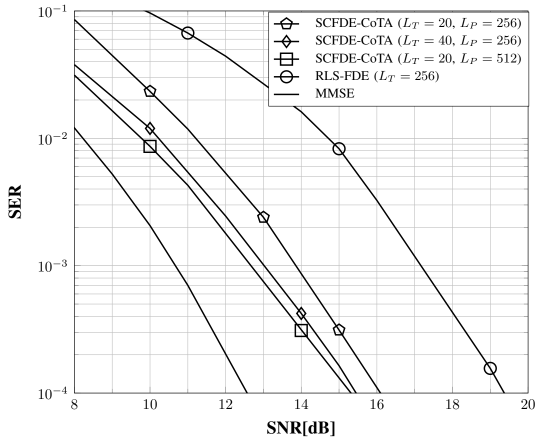

Fig. 3. Performance of SCFDE-CoTA for Varying Packet Length and Training Length

IV. Numerical Experiments

In this section, we provide examples to demonstrate the strength of the proposed algorithm to decrease the required number of training symbols as well as its fast convergence feature. The setup for the experiment:

- 4-QAM input constellation scheme is selected.
- The scenario with packet length $L_P = 256$ is considered. The performance of the proposed algorithm is compared with the one given in [18] and Minimum Mean Square Error (MMSE) equalizer. Furthermore, to clarify the effect of packet length, we also provide the performance of the algorithm when $L_P = 512$.
- α is set to 0.8 and decision directed mode is activated after 30^{th} iteration.
- We consider discrete time equivalent channel whose length is 16 and the tabs of the channel are assumed to be identical and independent complex (circular) Gaussian.

Fig. 3 provides the performance of the SCFDE-CoTA for varying training length, packet length and, SNR. We also plot the performance of adaptive RLS-FDE [18] which exploits the entire first packet as training. Hence, to be able to calculate the Symbol Error Rate (SER) performance of RLS-FDE, we assume coherence time of the channel allows to send two data packets where the first packet is filled with training symbols while the second packet is full of information symbols. Therefore, we can conclude that RLS-FDE algorithm uses 256 training symbol. On the other hand, for the proposed scheme, we assume the coherence time of the channel allows to send only one data packet for the purpose of clarifying the strength of the proposed algorithm for a highly mobile environment. We observe that even with $L_T = 20$, we can achieve better results than RLS-FDE. When we increase the

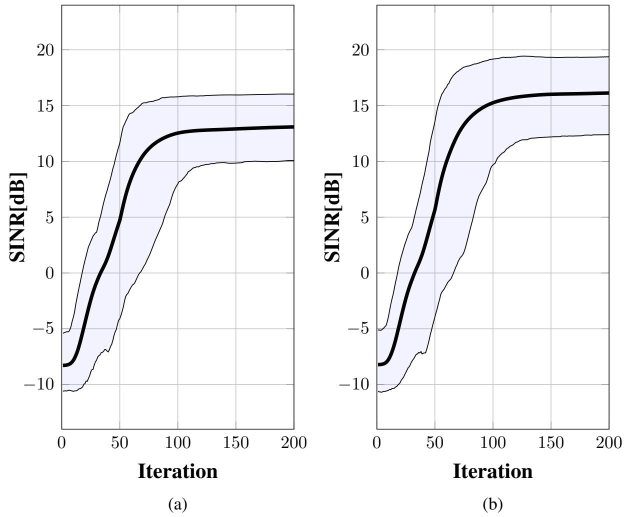

Fig. 4. Convergence Performances of SCFDE-CoTA for a) 16dB b) 20dB. Shaded areas are %95 confidence intervals and the thick black curves are the mean SINR calculated at each iteration.

number of training symbols, the performance of the algorithm gets better. However, the packet length increases the performance of SCFDE-CoTA more. The reason behind that comes from the duality between ℓ_1 and ℓ_∞ norms as shown in [16]. As the number of output samples increases, the probability of the condition that the duality between these norms holds increases. Hence, enlarging data packet contributes more to the performance of the algorithm. Moreover, we provide the performance behavior of the proposed framework for different SNR values as a function of iteration in Fig. 4. We can see that the algorithm converges around 80 iteration which demonstrates the low complexity nature of the algorithm.

V. Conclusion

In this article, we provide single carrier frequency domain based adaptive equalization scheme. The approach utilizes training symbols and the magnitude boundedness of digital communication symbols. The framework also exploits training symbols to overcome the dispersion caused by the multipath propagation environment. Because it combines the least square cost function to exploit reconstruction error of training symbols, and infinity norm cost function to use the prior information about the input constellation structure, the framework can be categorized in semi-blind algorithms.

The algorithm is proposed to decrease the required number of training symbols to open the eye and decrease the complexity by performing in frequency domain. Since the framework is based on convex optimization setting, and the connection with sparse optimization literature [16] is also available, further improvement in computational complexity can be done by merging the ideas in big data or machine learning.

The numerical examples provided in the article demonstrate the remarkable performance of the proposed framework in terms of reduction in the number of training samples required to achieve a given target SNR level for the noisy scenarios. The significant reduction achieved in training length through the proposed framework has three major implications:

- Increase in efficiency of bandwidth utilization by decreasing the required number of training symbols, in other words, providing more space for information symbols.
- Increasing the effective data rates on a highly mobile environment with lesser training symbols.
- Single packet adaptation to support applications with low latency requirements.

References

[1] J. G. Proakis, *Digital Communications*. New York, NY: McGraw Hill, 2000.

[2] A. J. van der Veen and A. Paulraj, "An analytical constant modulus allgorithm," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[3] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1927–1950, 1998.

[4] S. Vembu, S. Verdu, R. Kennedy, and W. Sethares, "Convex cost functions in blind equalization," *IEEE Transactions on Signal Processing*, vol. 42, pp. 1952–1960, Aug. 1994.

[5] Z. Ding and Z. Luo, "A fast linear programming algorithm for blind equalization," *IEEE Transactions on Communications*, vol. 48, pp. 1432–1436, Sep. 2000.

[6] A. T. Erdogan and C. Kizilkale, "Fast and low complexity blind equalization via subgradient projections," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2513–2524, Jul. 2005.

[7] A. T. Erdogan, "A fractionally spaced blind equalization algorithm with global convergence," *Signal Processing*, vol. 88, no. 1, pp. 200–209, 2008.

[8] E. D. Carvalho and D. T. M. Slock, "Maximum-likelihood blind fir multichannel estimation with gaussian prior for the symbols," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP' 97)*, vol. 5. Munich, Germany: IEEE, Apr. 1997, pp. 3593–3596.

[9] V. Zarzoso and P. Comon, "Blind and semi-blind equalization based on the constant power criterion," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4363–4375, Nov. 2005.

[10] S. Chen, "Semi-blind fast equalization of qam channels using concurrent gradient-newton cma and soft decision-directed scheme," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 6, pp. 467–476, 2010.

[11] M. Morelli, L. Sanguinetti, and U. Mengali, "Channel estimation for adaptive frequency-domain equalization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2508–2518, 2005.

[12] B. B. Yilmaz and A. T. Erdogan, "Compressed training adaptive equalization," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP' 16)*. IEEE, Mar. 2016, pp. 4920–4924.

[13] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Communications Magazine*, vol. 40, no. 4, pp. 58–66, Apr. 2002.

[14] F. Pancaldi, G. M. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization," *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 37–56, September 2008.

[15] Y. Nesterov, "A method of solving a convex programming problem with convergence rate $o(1/k^2)$." *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.

[16] B. B. Yilmaz and A. T. Erdogan, "Compressed training adaptive equalization: Algorithms and analysis," *IEEE Transactions on Wireless Communications*, vol. 65, no. 9, pp. 3907–3921, 2017.

[17] V. Cevher, S. Becker, and M. Schmidt, "Convex optimization for big data: Scalable, randomized, and parallel algorithms for big data analytics," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 32–43, 2014.

[18] M. V. Clark, "Adaptive frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1385–1395, 1998.