Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 88 (2008) 200–209
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

# A fractionally spaced blind equalization algorithm with global convergence ☆

Alper T. Erdogan*

*EE Department, Koc University, Sariyer, 34450 Istanbul, Turkey*

Received 1 September 2006; received in revised form 4 April 2007; accepted 23 April 2007
Available online 6 May 2007

**Abstract**

Two different fractionally spaced extensions of the SubGradient based Blind equalization Algorithm (SGBA) are provided. The first one is the direct extension of the linearly constrained SGBA for the symbol spaced setting. The second extension is the weighted and the 2-norm constrained fractionally spaced SGBA (FS-SGBA) algorithm. It is proven that the latter algorithm is globally convergent to a perfect equalization point under the well-known equalizability conditions for the fractionally spaced setting. The simulation results provided illustrates the relative merit of the proposed algorithm in comparison to the state of the art algorithms.


*Keywords:* Adaptive filtering; Blind equalization; Fractionally spaced; Subgradient

## 1. Introduction

In the area of blind equalization, convex cost functions play an important role due to their surface structures which are free of local and false minima and saddle points. The pioneering work in this field is due to Vembu and Verdu [1] who cast the blind equalization problem as a convex infinity norm minimization problem under a linear constraint. This approach exploits the magnitude bounded structure of the PAM constellations used in digital communications. They also proposed the use of large-$p$ norm approximation of the proposed cost function to obtain a gradient search based iterative algorithm.

In [2], Ding and Luo posed the infinity norm minimization of the affine function corresponding to blind equalization as a linear programming problem. The same reference proposes the modification of the infinity norm based cost function for handling complex QAM constellations. The extension of this work for the fractionally spaced equalizers is proposed in [3].

Although the linear programming based approaches have better performance than $p$-norm based approximation in [1], they are computationally expensive. Recently, a subgradient based framework for direct iterative minimization of infinity norm based blind equalization cost function was proposed [4] as an alternative. The proposed framework enables the development of iterative algorithms with low complexity for the convex problem posed in [1] and its variations.

☆This work was supported in part by TUBITAK Career Award, Contract No.: 104E073.
*Tel.: +90 212 338 1490; fax: +90 212 338 1548.
*E-mail addresses:* alperdogan@ku.edu.tr, alper@stanfordalumni.org (A.T. Erdogan).

0165-1684/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2007.04.017

In this article, we present the extension of the approach in [4] for the fractionally spaced equalization problem. We will first provide the simple generalization of the Sub Gradient based Blind equalisation Algorithm (SGBA) algorithm based on the linear constraint. We later replace the linear constraint on the equalizer with the quadratic constraint and provide an algorithm for this case. In particular, we show that the resulting algorithm is globally convergent to a perfect equalizer point under the well-known equalizability assumptions for the fractionally spaced case and under a generically true assumption about the initial search vector.

The organization of the article is as follows: In Section 2.2, we provide the blind equalization setup and the convex formulation proposed in [1]. Section 3, is the main part of the article where we provide the fractionally spaced SGBA algorithms. Section 4 focuses on the convergence of the proposed algorithms. The simulation examples illustrating the performance of these algorithms are provided in Section 5. Finally, Section 6 is the conclusion.

## 2. Notation and blind equalization setup

### 2.1. Notation

We use the following notation throughout the paper:

| *Symbol* | *Use* |
|---|---|
| $\mathbf{A}$ (bold-capital letters) | matrices |
| $\mathbf{x}$ (bold-lowercase letters) | vectors |
| $s$ (normal-lowercase letters) | scalars |
| $\bar{\mathbf{A}}$ | $\mathbf{A}$ with conjugated elements |
| $\mathbf{A}^{\mathrm{T}}$ | transpose of $\mathbf{A}$ |
| $\mathbf{A}^{\mathrm{H}}$ | Hermitian transpose of $\mathbf{A}$ |
| $\mathbf{A}^{\dagger}$ | pseudoinverse of $\mathbf{A}$ |
| $\mathbf{A}_{:,n}$ | $n$th column of $\mathbf{A}$ |
| $\mathbf{A}_{m,:}$ | $m$th row of $\mathbf{A}$ |
| $\mathscr{R}e\{.\}$ | real Part operator |
| $\mathscr{I}m\{.\}$ | imaginary Part operator |

### 2.2. Blind equalization setup

The setup for fractionally spaced equalization is shown in Fig. 1 where

- $\{x_n\}$ is the transmitted digital communication sequence. We assume that $x_n$ takes its values from a square QAM constellation where

$$\max(\mathscr{R}e\{x_n\}) = -\min(\mathscr{R}e\{x_n\}), \quad (1)$$

$$= \max(\mathscr{I}m\{x_n\})$$

$$= -\min(\mathscr{I}m\{x_n\}) = Q. \quad (2)$$

Although we will assume complex QAM constellation for the rest of the article, the presented algorithms are trivially applicable to real baseband case with PAM constellations. We assume a uniform distribution for the constellation points with a variance $\sigma_x^2$.

- $M$ is the oversampling factor of the fractionally spaced equalization. We will assume that $M = 2$ without loss of generality.
- $\{h_n; n \in \{0, \ldots, N_H - 1\}\}$ is the effective impulse response of the overall communication channel, where we assume $N_H$ to be even without loss of generality. If we define $H(z) = \sum_{n=0}^{N_H-1} h_n z^{-n}$ and write

$$H(z) = H^{(e)}(z^2) + H^{(o)}(z^2)z^{-1} \quad (3)$$

then $H^{(e)}(z)$ and $H^{(o)}(z)$ represent the $Z$-transforms of the even and odd subsamples of the channel impulse response.

- $\{y_n\}$ is the oversampled received sequence at the receiver. If we define $Y(z)$ as the $Z$-transform of $\{y_n\}$ and write

$$Y(z) = Y^{(e)}(z^2) + Y^{(o)}(z^2)z^{-1} \quad (4)$$

then $Y^{(e)}(z)$ and $Y^{(o)}(z)$ represent the $Z$-transforms of the even and odd subsamples of $\{y_n\}$. Since

$$Y(z) = H(z)X(z^2) = \underbrace{H^{(e)}(z^2)X(z^2)}_{Y^{(e)}(z^2)} + \underbrace{H^{(o)}(z^2)X(z^2)}_{Y^{(o)}(z^2)} z^{-1}$$

we can write

$$\underbrace{\begin{bmatrix} Y^{(e)}(z) \\ Y^{(o)}(z) \end{bmatrix}}_{\mathbf{Y}(z)} = \underbrace{\begin{bmatrix} H^{(e)}(z) \\ H^{(o)}(z) \end{bmatrix}}_{\mathbf{H}(z)} X(z), \quad (5)$$

which is nothing but the multichannel representation of the oversampled setup.

For the adaptive implementation we assume that a window of channel output samples $\{y_k : k = 1, \ldots, \Omega\}$, where $\Omega$ is the length of the window, is available for the adaptation of the equalizer.

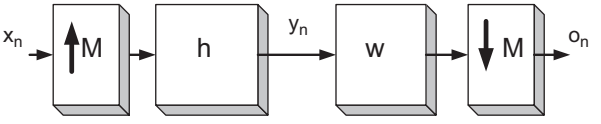


Fig. 1. The fractionally spaced equalization setup.

- $\{w_n; n \in \{0,\ldots,N_W-1\}\}$ is the impulse response of the equalizer, where we assume $N_W$ is even without loss of generality.
- $\{o_n\}$ is the overall output which can be written as

$$o_n = \underbrace{\begin{bmatrix} w_0 & w_1 & w_2 & w_3 & \cdots & w_{N_W-2} & w_{N_W-1} \end{bmatrix}}_{\mathbf{w}^{\mathrm{T}}} \times \underbrace{\begin{bmatrix} y_{2n} \\ y_{2n-1} \\ y_{2n-2} \\ y_{2n-3} \\ \vdots \\ y_{2n-N_W+2} \\ y_{2n-N_W+1} \end{bmatrix}}_{\mathbf{y}_n} \tag{6}$$

$$= \underbrace{\mathbf{w}^{\mathrm{T}}\mathscr{T}(\mathbf{H})}_{\mathbf{g}^{\mathrm{T}}} \underbrace{\begin{bmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-\frac{N_W+N_H}{2}+2} \end{bmatrix}}_{\mathbf{x}_n}, \tag{7}$$

where $\mathscr{T}(\mathbf{H})$ is the $N_W \times L$ block convolution matrix corresponding to the transfer function $\mathbf{H}(z)$ where $L = (N_H + N_W)/2 - 1$.
For equalizability, we assume [5],

- $H^{(e)}(z)$ and $H^{(0)}(z)$ do not have any common zeros,
- $N_W/2 + 1 \geqslant N_H/2$.

Under these assumptions, $\mathscr{T}(\mathbf{H})$ would be a full-rank tall (or square) matrix such that we can find a vector $\mathbf{w}$ for which

$$\mathbf{w}^{\mathrm{T}}\mathscr{T}(\mathbf{H}) = \mathbf{e}_d^{\mathrm{T}},$$

where $\mathbf{e}_d$ is a standard basis vector with single non-zero entry located at the index $d$.

## 3. Fractionally spaced SGBA

In [2], the fractionally spaced blind equalization problem is posed as the optimization problem

$$\begin{aligned} \text{minimize} \quad & \max_n |\mathscr{R}e\{o_n\}| \\ \text{subject to} \quad & \mathscr{R}e\{w_0 + w_1\} + \mathscr{I}m\{w_0 + w_1\} = 1, \\ & w_{2k-1} = 0, \;\; k = \frac{N_H}{2},\ldots,\frac{N_W}{2}, \end{aligned}$$

where the linear programming is proposed to obtain the solution of the problem. The linear programming solution is computationally involved. Instead, we propose a simple modification to SGBA proposed in [4], which yields the following iterations:

$$\underline{\mathbf{w}}^{(k+1)} = \mathbf{w}^{(k)} - \mu^{(k)} \operatorname{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\bar{\mathbf{y}}_{n^{(k)}}, \tag{8}$$

$$\mathbf{w}^{(k+1)} = \mathscr{P}\{\underline{\mathbf{w}}^{(k+1)}\}, \tag{9}$$

where

- $n^{(k)}$ is the index for which the maximum magnitude $\mathscr{R}e\{o_n^{(k)}\}$ is achieved,
- $\mu^{(k)}$ is the step size,
- $\mathscr{P}$ is the minimum distance projection operator which projects its argument to the affine set defined by the constraints on the equalizer coefficients. The relation between $\mathbf{w}^{(k+1)}$ and $\underline{\mathbf{w}}^{(k+1)}$ is simply given by

$$w_l^{(k+1)} = \begin{cases} \dfrac{\mathscr{R}e\{3\underline{w}_0^{(k+1)} - \underline{w}_1^{(k+1)}\} - \mathscr{I}m\{\underline{w}_0^{(k+1)} + \underline{w}_1^{(k+1)}\} + 1}{4} \\ \quad + j\dfrac{\mathscr{I}m\{3\underline{w}_0^{(k+1)} - \underline{w}_1^{(k+1)}\} - \mathscr{R}e\{\underline{w}_0^{(k+1)} + \underline{w}_1^{(k+1)}\} + 1}{4}, & l = 0, \\ \dfrac{\mathscr{R}e\{3\underline{w}_1^{(k+1)} - \underline{w}_0^{(k+1)}\} - \mathscr{I}m\{\underline{w}_0^{(k+1)} + \underline{w}_1^{(k+1)}\} + 1}{4} \\ \quad + j\dfrac{\mathscr{I}m\{3\underline{w}_1^{(k+1)} - \underline{w}_0^{(k+1)}\} - \mathscr{R}e\{\underline{w}_0^{(k+1)} + \underline{w}_1^{(k+1)}\} + 1}{4}, & l = 1, \\ 0, & l = N_H - 1, N_H + 1, N_H + 3, \ldots, N_W - 1, \\ \underline{w}_l^{(k+1)} & \text{otherwise.} \end{cases}$$

Note that since the corresponding optimization problem is convex (with convex cost function and convex constrained set), the algorithm defined by (8) and (9) converges to the global optimal point. However, this approach requires the exact knowledge of the channel length $N_H$, which limits the practicality of the algorithm.

We obtain an alternative version of this algorithm by removing (9) and using a weighting

$$\underline{\mathbf{w}}^{(k+1)} = \mathbf{w}^{(k)} - \mu^{(k)} \operatorname{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\bar{\mathbf{\Pi}}_{\mathbf{y}}^{\dagger}\bar{\mathbf{y}}_{n^{(k)}}, \tag{10}$$

where

$$\mu^{(k)} = \frac{|\mathscr{R}e\{o_{n^{(k)}}^{(k)}\}| - (\sqrt{2}Q/\sigma_x\sqrt{\Omega})\sqrt{\sum_{l=1}^{\Omega}|\mathscr{R}e\{o_l^{(k)}\}|^2}}{2\mathbf{y}_{n^{(k)}}^{\mathrm{H}}\mathbf{\Pi}_{\mathbf{y}}^{\dagger}\mathbf{y}_{n^{(k)}}} \tag{11}$$

and $\mathbf{\Pi}_{\mathbf{y}}$ is the covariance of $\mathbf{y}$ which is equal to

$$\mathbf{\Pi}_{\mathbf{y}} = \sigma_x^2\mathscr{T}(\mathbf{H})\mathscr{T}(\mathbf{H})^{\mathrm{H}}. \tag{12}$$

Note that in applications, this covariance matrix should actually be estimated from the observations. However, to simplify our analysis later, we'll assume that our estimate of the covariance is equal to the true covariance.

In order to avoid all zeros solution and to fix a numerical range for $\mathbf{w}$ we introduce the normalization

$$\mathbf{w}^{(k+1)} = \frac{\underline{\mathbf{w}}^{(k)}}{\|\underline{\mathbf{w}}\|_{\bar{\mathbf{\Pi}}_{\mathbf{y}}/\sigma_x^2}}, \tag{13}$$

where

$$\|\underline{\mathbf{w}}\|_{\bar{\mathbf{\Pi}}_{\mathbf{y}}/\sigma_x^2} \triangleq \sqrt{\underline{\mathbf{w}}^{(k)\mathrm{H}}\frac{\bar{\mathbf{\Pi}}_{\mathbf{y}}}{\sigma_x^2}\underline{\bar{\mathbf{w}}}^{(k)}}, \tag{14}$$

which is a scaling by a weighted norm of $\underline{\mathbf{w}}^{(k)}$ such that the average equalizer output power is fixed as $\sigma_x^2$.

One clear advantage of this variation over the linearly constrained version is that, the exact knowledge of the channel length is not required.

## 4. Convergence analysis of the fractionally spaced blind equalization algorithm

Among the two alternative fractionally spaced SGBA algorithms presented in the previous section, the former algorithm corresponds to a conventional subgradient search algorithm for a convex cost function with a convex constraint. The corresponding cost function was shown to have a perfect equalization point as its optimal point in [3] and a smart choice of step size rule satisfying zero-limit-divergent-sum (ZLDS) rule, i.e.,

$$\lim_{k\to\infty}\mu^{(k)} \to 0 \quad \text{zero limit}, \tag{15}$$

and

$$\lim_{\Omega\to\infty}\sum_{k=0}^{\Omega}|\mu^{(k)}| \to \infty \quad \text{divergent sum} \tag{16}$$

would guarantee convergence of the algorithm to this perfect equalization point [4,6,7]. Note that the choice

$$\mu^{(k)} = \frac{\mu^{(0)}}{k+1} \tag{17}$$

would satisfy the requirements of the ZLDS rule.

The convergence of the latter algorithm is less trivial due to the fact that the iterations correspond to the non-convex optimization problem

$$\begin{aligned} &\text{minimize} \quad \|o\|_{\infty}, \\ &\text{subject to} \quad \|\mathbf{w}\|_{\bar{\mathbf{\Pi}}_{\mathbf{y}}/\sigma_x^2} = 1, \end{aligned}$$

for which the cost function is convex but the constraint set is not. Therefore, a direct assessment of convergence based on subgradient optimization literature is not possible. In this section, we are going to show that the second fractionally spaced SGBA algorithm introduced above is globally convergent under some mild assumptions.

We start by multiplying both sides of (10) by $\mathscr{T}(\mathbf{H})^{\mathrm{T}}$ from left

$$\underbrace{\mathscr{T}(\mathbf{H})^{\mathrm{T}}\underline{\mathbf{w}}^{(k+1)}}_{\underline{\mathbf{g}}^{(k+1)}} = \underbrace{\mathscr{T}(\mathbf{H})^{\mathrm{T}}\mathbf{w}^{(k)}}_{\mathbf{g}^{(k)}} - \mu^{(k)}\operatorname{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\mathscr{T}(\mathbf{H})^{\mathrm{T}}\bar{\mathbf{\Pi}}_{\mathbf{y}}^{\dagger}\bar{\mathbf{y}}_{n^{(k)}}, \tag{18}$$

where $\mathbf{g}^{(k)}$ would be the overall effective impulse response at the $k$th step. If we write down a (full) singular value decomposition of $\mathscr{T}(\mathbf{H})$ as

$$\mathscr{T}(\mathbf{H}) = \mathbf{U}\begin{bmatrix}\mathbf{\Sigma}_r \\ \mathbf{0}\end{bmatrix}\mathbf{V}^{\mathrm{H}}, \tag{19}$$

where

$$\mathbf{\Sigma}_r = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix}$$

with $\sigma_1, \sigma_2, \ldots, \sigma_r > 0$, and $\mathbf{U}, \mathbf{V}$ are unitary matrices. Since $\mathbf{y}_n = \mathscr{T}(\mathbf{H})\mathbf{x}_n$ it is easy to show that

$$\mathbf{\Pi}_y = \mathscr{T}(\mathbf{H})\mathscr{T}(\mathbf{H})^{\mathrm{H}}\sigma_x^2 \tag{20}$$

$$= \sigma_x^2 \mathbf{U}\begin{bmatrix} \mathbf{\Sigma}_r^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{U}^{\mathrm{H}}, \tag{21}$$

and consequently,

$$\bar{\mathbf{\Pi}}_{\mathbf{y}}^{\dagger} = \sigma_x^{-2}\bar{\mathbf{U}}\begin{bmatrix} \mathbf{\Sigma}_r^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{U}^{\mathrm{T}}. \tag{22}$$

Therefore, from (18) we can write

$$\underline{\mathbf{g}}^{(k+1)} = \mathbf{g}^{(k)} - \mu^{(k)}\,\mathrm{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\bar{\mathbf{V}}[\mathbf{\Sigma}_r \;\; \mathbf{0}] \times \mathbf{U}^{\mathrm{T}}\sigma_x^{-2}\bar{\mathbf{U}}\begin{bmatrix} \mathbf{\Sigma}_r^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{U}^{\mathrm{T}}\bar{\mathbf{U}}\begin{bmatrix} \mathbf{\Sigma}_r \\ \mathbf{0} \end{bmatrix}\mathbf{V}^{\mathrm{T}}\bar{\mathbf{x}}_{n^{(k)}} \tag{23}$$

$$= \mathbf{g}^{(k)} - \mu^{(k)}\,\mathrm{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\sigma_x^{-2}\bar{\mathbf{x}}_{n^{(k)}}. \tag{24}$$

If we define

$$\tilde{\mathbf{g}} = \begin{bmatrix} \mathscr{R}e\{\mathbf{g}\} \\ \mathscr{I}m\{\mathbf{g}\} \end{bmatrix} \tag{25}$$

and

$$\check{\mathbf{x}} = \begin{bmatrix} \mathscr{R}e\{\mathbf{x}\} \\ -\mathscr{I}m\{\mathbf{x}\} \end{bmatrix}, \tag{26}$$

then we can rewrite (24) as

$$\underline{\tilde{\mathbf{g}}}^{(k+1)} = \tilde{\mathbf{g}}^{(k)} - \mu^{(k)}\,\mathrm{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\sigma_x^{-2}\check{\mathbf{x}}_{n^{(k)}}. \tag{27}$$

Note that

$$\mathscr{R}e\{o_n\} = \tilde{\mathbf{g}}^{\mathrm{T}}\check{\mathbf{x}}_n \tag{28}$$

and its maximum magnitude is achieved for

$$\check{\mathbf{x}}_{n^{(k)}} = \pm Q\,\mathrm{sign}(\tilde{\mathbf{g}}^{(k)}) + \dot{\mathbf{x}}_{n^{(k)}}, \tag{29}$$

where $\dot{\mathbf{x}}_{n^{(k)}}$ is the vector that has zero values for the indexes where $\tilde{\mathbf{g}}^{(k)}$ is non-zero, and has arbitrary values from the set of real and imaginary components of the constellation points, i.e., the $l$th component of $\dot{\mathbf{x}}_{n^{(k)}}$ is given by

$$\begin{cases} 0 & \text{if } \tilde{g}_l^{(k)} \neq 0, \\ \mathscr{A}_l^{(k)} & \text{otherwise} \end{cases} \tag{30}$$

for all $l \in [1, 2L]$, where $\mathscr{A}_l^{(k)}$ is a random value from the set of real and imaginary components of the constellation. Note that since $\dot{\mathbf{x}}_{n^{(k)}}$ is orthogonal to $\mathbf{g}^{(k)}$, it has no effect at the output. Therefore, we have

$$\mathrm{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\check{\mathbf{x}}_{n^{(k)}} = Q\,\mathrm{sign}(\tilde{\mathbf{g}}^{(k)}) + \mathrm{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\dot{\mathbf{x}}_{n^{(k)}}. \tag{31}$$

As a result, by plugging (31) in (27),we obtain

$$\underline{\tilde{\mathbf{g}}}^{(k+1)} = \tilde{\mathbf{g}}^{(k)} - \frac{\mu^{(k)}Q}{\sigma_x^2}\mathrm{sign}(\tilde{\mathbf{g}}^{(k)}) - \frac{\mu^{(k)}}{\sigma_x^2}\mathrm{sign}(\mathscr{R}e\{o_{n^{(k)}}^{(k)}\})\dot{\mathbf{x}}_{n^{(k)}}. \tag{32}$$

Based on (32), the magnitude of the $v$th element of $\underline{\tilde{\mathbf{g}}}^{(k+1)}$ would be equal to

$$|\underline{\tilde{g}}_v^{(k+1)}| = \begin{cases} \left|\,|\tilde{g}_v^{(k)}| - \dfrac{\mu^{(k)}Q}{\sigma_x^2}\right| & \text{if } \tilde{g}_v^{(k)} \neq 0, \\ \dfrac{\mu^{(k)}|\mathscr{A}_v^{(k)}|}{\sigma_x^2} & \text{otherwise} \end{cases} \tag{33}$$

for all $v$. Here

$$\mu^{(k)} = \frac{|\mathscr{R}e\{o_{n^{(k)}}^{(k)}\}| - (Q\sqrt{2}/\sigma_x\sqrt{\Omega})\sqrt{\sum_{l=1}^{\Omega}|\mathscr{R}e\{o_l^{(k)}\}|^2}}{2\mathbf{y}_{n^{(k)}}^{\mathrm{H}}\mathbf{\Pi}_{\mathbf{y}}^{\dagger}\mathbf{y}_{n^{(k)}}} \tag{34}$$

$$\leqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 Q - \|\tilde{\mathbf{g}}^{(k)}\|_2 Q}{2\mathscr{L}_{\tilde{g}}^{(k)}Q^2\sigma_x^{-2}} \tag{35}$$

$$= \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \|\tilde{\mathbf{g}}^{(k)}\|_2}{2\mathscr{L}_{\tilde{g}}^{(k)}Q\sigma_x^{-2}}, \tag{36}$$

where $\mathscr{L}_{\tilde{g}}^{(k)}$ is the number of non-zero elements of $\tilde{\mathbf{g}}$. Here going from (34) to (35), we used:

- the fact that
$$|\mathscr{R}e\{o_{n^{(k)}}^{(k)}\}| = |\check{\mathbf{x}}_{n^{(k)}}^{\mathrm{T}}\tilde{\mathbf{g}}^{(k)}|, \tag{37}$$
and since $\check{\mathbf{x}}_{n^{(k)}} = \pm Q\,\mathrm{sign}(\tilde{\mathbf{g}}^{(k)}) + \dot{\mathbf{x}}_{n^{(k)}}$,
$$|\mathscr{R}e\{o_{n^{(k)}}^{(k)}\}| = \|\tilde{\mathbf{g}}^{(k)}\|_1 Q \tag{38}$$
easily follows;
- the approximation
$$\frac{1}{\Omega}\sum_{l=1}^{\Omega}|\mathscr{R}e\{o_l^{(k)}\}|^2 \approx \|\tilde{\mathbf{g}}^{(k)}\|_2^2\frac{\sigma_x^2}{2}; \tag{39}$$
- and the inequality
$$\mathbf{y}_{n^{(k)}}^{\mathrm{H}}\mathbf{\Pi}_{\mathbf{y}}^{\dagger}\mathbf{y}_{n^{(k)}} = \mathbf{x}_{n^{(k)}}^{\mathrm{H}}\mathscr{T}(\mathbf{H})^{\mathrm{H}}\mathbf{\Pi}_{\mathbf{y}}^{\dagger}\mathscr{T}(\mathbf{H})\mathbf{x}_{n^{(k)}} \tag{40}$$
$$= \mathbf{x}_{n^{(k)}}^{\mathrm{H}}\sigma_x^{-2}\mathbf{x}_{n^{(k)}} \tag{41}$$
$$= Q^2\|\,\mathrm{sign}(\tilde{\mathbf{g}}^{(k)})\|_2^2\sigma_x^{-2} + \sigma_x^{-2}\|\dot{\mathbf{x}}_{n^{(k)}}\|_2^2 \tag{42}$$
$$\geqslant Q^2\mathscr{L}_{\tilde{g}}^{(k)}\sigma_x^{-2}. \tag{43}$$

From (36), we can write

$$\mu^{(k)} \leqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \|\tilde{\mathbf{g}}^{(k)}\|_2}{2\mathscr{L}_{\tilde{g}}^{(k)} Q\sigma_x^{-2}} \tag{44}$$

$$\leqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \|\tilde{\mathbf{g}}^{(k)}\|_\infty}{2\mathscr{L}_{\tilde{g}}^{(k)} Q\sigma_x^{-2}} \tag{45}$$

$$= \frac{\|\acute{\mathbf{g}}^{(k)}\|_1}{2\mathscr{L}_{\tilde{g}}^{(k)} Q\sigma_x^{-2}} \tag{46}$$

$$\leqslant \frac{\|\acute{\mathbf{g}}^{(k)}\|_\infty}{2Q\sigma_x^{-2}}, \tag{47}$$

where we define $\acute{\mathbf{g}}^{(k)}$ as the vector obtained by deleting the maximum magnitude element of $\tilde{\mathbf{g}}^{(k)}$. Here (45) follows from the norm inequality $\|\mathbf{g}\|_2 \geqslant \|\mathbf{g}\|_\infty$, (46) follows directly from the definition of $\acute{\mathbf{g}}^{(k)}$ and (47) follows from the norm inequality

$$\frac{\|\acute{\mathbf{g}}^{(k)}\|_1}{\mathscr{L}_{\tilde{g}}^{(k)}} \leqslant \frac{\|\acute{\mathbf{g}}^{(k)}\|_1}{(\mathscr{L}_{\tilde{g}}^{(k)} - 1)} \tag{48}$$

$$\leqslant \|\acute{\mathbf{g}}^{(k)}\|_\infty, \tag{49}$$

where $\mathscr{L}_{\tilde{g}}^{(k)} - 1$ is the number of non-zero elements of $\acute{\mathbf{g}}^{(k)}$. As a result,

$$\underbrace{\frac{Q\mu^{(k)}}{\sigma_x^2}}_{\eta^{(k)}} \leqslant \frac{\|\acute{\mathbf{g}}^{(k)}\|_\infty}{2}. \tag{50}$$

Note that $\|\acute{\mathbf{g}}^{(k)}\|_\infty$ is nothing but the second magnitude peak of $\tilde{\mathbf{g}}$.

Based on (33) and (50) we can conclude that after the update in (10), the peak value and the second peak value of the $\tilde{\mathbf{g}}$ remain to be at the same positions and their values are reduced by the same amount. However, when we look at the ratio of the peak and the second peak values

$$\frac{\|\tilde{\mathbf{g}}^{(k+1)}\|_\infty}{\|\acute{\mathbf{g}}^{(k+1)}\|_\infty} = \frac{\|\tilde{\underline{\mathbf{g}}}^{(k+1)}\|_\infty}{\|\acute{\underline{\mathbf{g}}}^{(k+1)}\|_\infty} = \frac{\|\tilde{\mathbf{g}}^{(k)}\|_\infty - \eta^{(k)}}{\|\acute{\mathbf{g}}^{(k)}\|_\infty - \eta^{(k)}}. \tag{51}$$

Given that $\|\tilde{\mathbf{g}}^{(k)}\|_\infty > \|\acute{\mathbf{g}}^{(k)}\|_\infty$, i.e., $\tilde{\mathbf{g}}^{(k)}$ has a single peak, we have

$$\frac{\|\tilde{\mathbf{g}}^{(k+1)}\|_\infty}{\|\acute{\mathbf{g}}^{(k+1)}\|_\infty} > \frac{\|\tilde{\mathbf{g}}^{(k)}\|_\infty}{\|\acute{\mathbf{g}}^{(k)}\|_\infty}. \tag{52}$$

As a result, given that the initial combined channel vector $\tilde{\mathbf{g}}^{(0)}$ has a single peak, the ratio

$$\lambda^{(k)} = \frac{\|\tilde{\mathbf{g}}^{(k)}\|_\infty}{\|\acute{\mathbf{g}}^{(k)}\|_\infty} \tag{53}$$

grows. In fact it grows unboundedly. To show this, we use proof by contradiction. Lets assume that $\lambda^{(k)}$ is bounded, i.e.,

$$\lambda^{(k)} < \beta \quad \forall k. \tag{54}$$

From (51) we can write

$$\lambda^{(k+1)} = \lambda^{(k)} \frac{1 - \eta^{(k)}/\|\acute{\mathbf{g}}^{(k)}\|_\infty \lambda^{(k)}}{1 - \eta^{(k)}/\|\acute{\mathbf{g}}^{(k)}\|_\infty}. \tag{55}$$

Using the boundedness of $\lambda^{(k)}$, we can obtain a lower bound for the ratio $\eta^{(k)}/\|\acute{\mathbf{g}}^{(k)}\|_\infty$ through the following steps:

$$\frac{\eta^{(k)}}{\|\acute{\mathbf{g}}^{(k)}\|_\infty} = \frac{Q\mu^{(k)}}{\sigma_x^2 \|\acute{\mathbf{g}}^{(k)}\|_\infty} \tag{56}$$

$$\geqslant \frac{Q(\|\tilde{\mathbf{g}}^{(k)}\|_1 Q - \|\tilde{\mathbf{g}}\|_2 Q)}{4LQ^2 \|\acute{\mathbf{g}}^{(k)}\|_\infty}$$

$$\text{(using } \mathbf{y}_{n^{(k)}}^{\mathrm{H}} \mathbf{\Pi}_{\mathbf{y}}^{\dagger} \mathbf{y}_{n^{(k)}} \leqslant 2LQ^2\sigma_x^{-2}\text{)} \tag{57}$$

$$= \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \|\tilde{\mathbf{g}}^{(k)}\|_2}{4L\|\acute{\mathbf{g}}^{(k)}\|_\infty} \tag{58}$$

$$\geqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \sqrt{\|\tilde{\mathbf{g}}^{(k)}\|_1^2 - \|\tilde{\mathbf{g}}^{(k)}\|_\infty \|\acute{\mathbf{g}}^{(k)}\|_\infty}}{4L\|\acute{\mathbf{g}}^{(k)}\|_\infty} \tag{59}$$

$$= \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \|\tilde{\mathbf{g}}^{(k)}\|_1 \sqrt{1 - \|\tilde{\mathbf{g}}^{(k)}\|_\infty \|\acute{\mathbf{g}}^{(k)}\|_\infty / \|\tilde{\mathbf{g}}^{(k)}\|_1^2}}{4L\|\acute{\mathbf{g}}^{(k)}\|_\infty} \tag{60}$$

$$\geqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 - \|\tilde{\mathbf{g}}^{(k)}\|_1 \sqrt{1 - \|\acute{\mathbf{g}}^{(k)}\|_\infty / \|\tilde{\mathbf{g}}^{(k)}\|_\infty 4L^2}}{4L\|\acute{\mathbf{g}}^{(k)}\|_\infty}$$

$$\text{(using } \|\tilde{\mathbf{g}}\|_1 \leqslant 2L\|\tilde{\mathbf{g}}\|_\infty\text{)} \tag{61}$$

$$\geqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_1 (1 - \sqrt{1 - 1/\beta 4L^2})}{4L\|\acute{\mathbf{g}}^{(k)}\|_\infty}$$

$$\text{(using (54))} \tag{62}$$

$$\geqslant \frac{\|\tilde{\mathbf{g}}^{(k)}\|_\infty (1 - \sqrt{1 - 1/\beta 4L^2})}{4L\|\acute{\mathbf{g}}^{(k)}\|_\infty}$$

$$\text{(using } \|\tilde{\mathbf{g}}^{(k)}\|_1 \geqslant \|\tilde{\mathbf{g}}^{(k)}\|_\infty\text{)} \tag{63}$$

$$= \frac{\lambda^{(k)} (1 - \sqrt{1 - 1/\beta 4L^2})}{4L} \tag{64}$$

$$\geqslant \underbrace{\frac{(1 - \sqrt{1 - 1/\beta 4L^2})}{4L}}_{\kappa} > 0$$

$$\text{(using } \lambda^{(k)} \geqslant 1\text{)}. \tag{65}$$

Using this bound and (55), we can write

$$\lambda^{(k+1)} \geqslant \lambda^{(k)} \frac{1-\eta^{(k)}/\|\dot{\mathbf{g}}^{(k)}\|_\infty \lambda^{(0)}}{1-\eta^{(k)}/\|\dot{\mathbf{g}}^{(k)}\|_\infty} \tag{66}$$

$$\geqslant \lambda^{(k)} \frac{1-\kappa/\lambda^{(0)}}{1-\kappa}, \tag{67}$$

where the first inequality is due to $\lambda^{(k)} \geqslant \lambda^{(0)}$ and the second inequality is obtained by combining the fact that the expression on the right of the inequality (66) is an increasing function of $\eta^{(k)}/\|\dot{\mathbf{g}}^{(k)}\|_\infty$ (over its domain $[0, 1/2]$) and the bound given by (65).

Therefore, given $\lambda^{(0)} > 1$, the factor multiplying $\lambda^{(k)}$ is strictly greater than 1 which implies that $\lambda^{(k)}$ grows without a bound, in contradiction to our earlier assumption that $\lambda^{(k)}$ is bounded.

As a result, under the normalization in (13), which forces $\|\tilde{\mathbf{g}}\|_2 = 1$, $\tilde{\mathbf{g}}^{(k)}$ converges to a vector with only one non-zero element, where the non-zero element is either 1 or $-1$. This is equivalent to the condition that

$$\mathbf{g}^{(k)} \to \varrho \mathbf{e}_m, \tag{68}$$

where $\mathbf{e}_m$ is the standard basis vector with a 1 located at the index $m \in \{1, \ldots, L\}$ and $\varrho \in \{1, -1, j, -j\}$, which corresponds to the perfect equalization condition.

Note that the initial assumption about $\tilde{\mathbf{g}}^{(0)}$ having a single peak is a generic and stable condition. Even if this assumption is not satisfied it is not a stable condition: a slight random disturbance on $\tilde{\mathbf{g}}$ would change it to a new vector with single peak (with probability 1). When the equalizer length $N_W$ is selected such that $N_W/2$ is strictly greater than $(N_H/2) - 1$, there are multiple global optima and the algorithm converges to the optimum vector whose corresponding $\tilde{\mathbf{g}}$ contains the impulse at the location where $\tilde{\mathbf{g}}^{(0)}$ has its peak. In other words, the initial selection of the equalizer vector will determine the optimum point to be converged to and its corresponding delay. This property is based on the fact the peak magnitude location of the $\tilde{\mathbf{g}}^{(k)}$ is always preserved by the algorithm.

Finally, we should note that above global convergence analysis is built upon the following assumptions:

- The approximation in equality (39) is true.
- The covariance $\mathbf{\Pi}_{\mathbf{y}}$ is known (which is actually unknown but estimated from the data).

Strictly speaking, for finite data windows, these assumptions would not hold. However, both assumptions become correct for asymptotically large data sizes.

## 5. Examples

As the first simulation example, we compare the performances of linearly and quadratically constrained fractionally spaced SGBA algorithms for a randomly selected channel whose vector equivalent transfer function is given by

$$\mathbf{H}(z) = \begin{bmatrix} 0.28 - 0.55i \\ 0.02 + 0.18i \end{bmatrix} + \begin{bmatrix} -0.48 + 0.18i \\ 0.01 + 0.20i \end{bmatrix} z^{-1} + \begin{bmatrix} -0.08 - 0.32i \\ -0.20 + 0.05i \end{bmatrix} z^{-2} + \begin{bmatrix} -0.21 - 0.05i \\ -0.30 + 0.02i \end{bmatrix} z^{-3}. \tag{69}$$

The input of the channel is a 16QAM i.i.d. sequence. The output of the channel is corrupted by an additive white Gaussian noise signal with a power level corresponding to 40 dB SNR. The equalizer order $N_W$ is selected as 12 (in this case $N_W/2 = 6 > N_H/2 - 1 = 3$, i.e., the equalizer length is greater than the minimum required equalizer length). Center spike-initialization is used for the search vectors. In Fig. 2, the signal to distortion energy ratio (SDR) (where the distortion is composed of the residual ISI and noise signals) of linearly and quadratically constrained SGBA algorithms are compared for the different choices of the $N_H$ parameter used by the linearly constrained algorithm (where the true $N_H$ corresponding to (69) is equal to 4). It is clear from this figure that, the SDR performance for the linearly constrained SGBA algorithm is worse than the quadratically constrained SGBA algorithm and it is very sensitive against the selection of the $N_H$ parameter: unless $N_H$ is selected as the true $N_H$, the equalizer output SDR degrades to an unusable level. The quadratically constrained algorithm achieves 40.2 dB SDR level where the fractionally spaced minimum mean square error fractionally spaced equalizer (MMSE-FSE) with best delay for this channel achieves 41.9 dB SDR level.

In Fig. 3, for the same channel, the SDR performances of linearly and quadratically constraint FS-SGBA (where linearly constrained FS-SGBA uses the true $N_H$ value) and MMSE-FSE

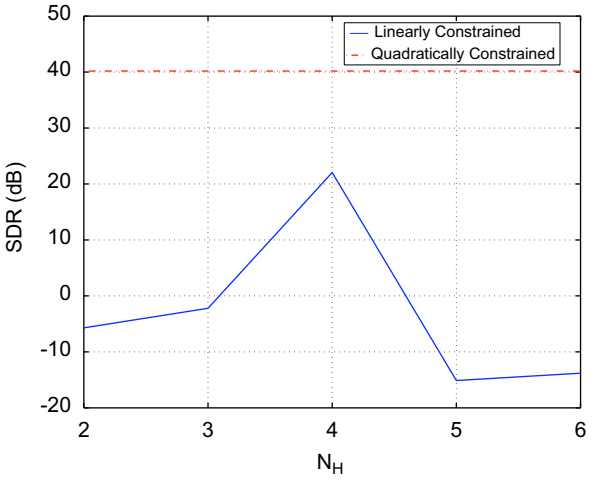


Fig. 2. The SDR performance comparison for the linearly constrained and the quadratically constrained SGBA algorithms.

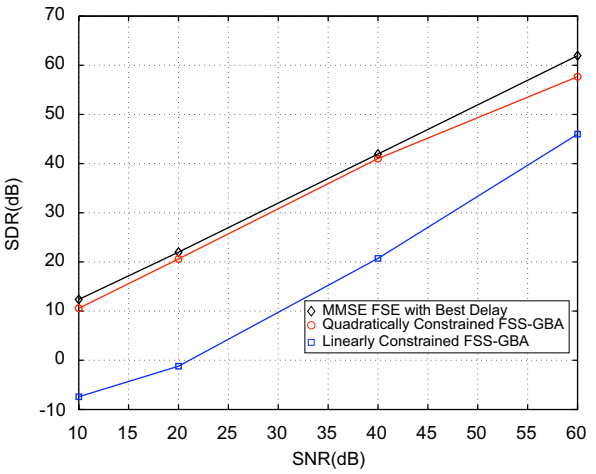


Fig. 3. The SDR performances of MMSE-FSE with optimum delay and FS-SGBA algorithms as a function of SNR.

with optimum delay as a function of SNR are shown. This figure also confirms that the quadratically constrained FS-SGBA has SDR performance better than linearly constrained FS-SGBA and close to the SDR level of MMSE-FSE.

When this simulation experiment is repeated for different random channels, we observe that the linearly constrained algorithm is very sensitive to the selection of the $N_H$ parameter and its maximum achieved SDR level is worse than the SDR level achieved by the quadratically constrained algorithm. The quadratically constrained SGBA algorithm's SDR performance is at a close vicinity of the MMSE fractionally spaced equalizer's SDR performance.

In the second simulation example we used the 300-Tap Microwave channel impulse response obtained from Rice University Signal Processing Information Base [8] (and considered this as the oversampled channel impulse response with for the oversampling factor 2) shown in Fig. 4. The corresponding channel length is $N_H = 300$. In the simulations equalizer input SNR level is set as 25 dB. The transmitted sequence is a 16QAM signal.

As a reference, Fig. 5 shows the equalizer output SDR level achieved by the fractionally spaced MMSE equalizer as a function of equalization

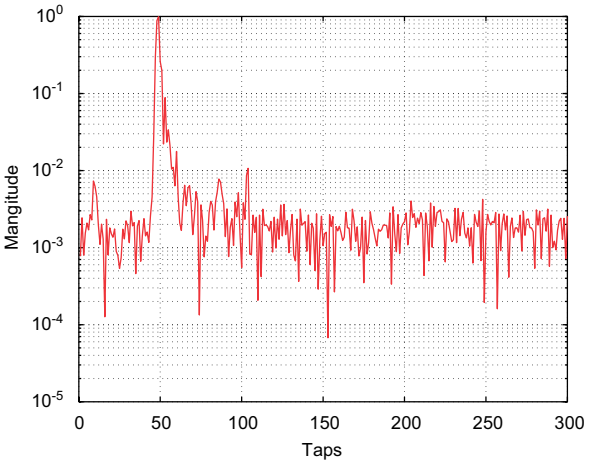


Fig. 4. Microwave channel impulse response (obtained from Rice University Signal Processing Information Base).

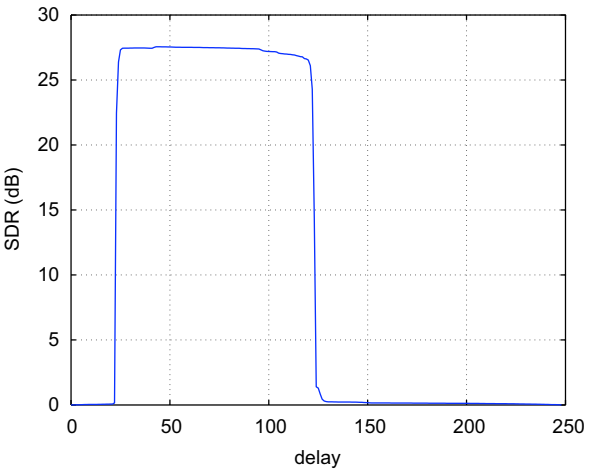


Fig. 5. The SDR performance of the fractionally spaced MMSE equalizer as a function of the equalization delay.

delay. We compared the proposed FS-SGBA algorithm's (quadratically constrained) performance with the super-exponential and the CMA algorithms. For the CMA algorithm we used the complex-constellation variation of the fixed-interval CMA algorithm proposed in [9]. The length of the fractionally spaced equalizer used in the simulations is equal to 200 and all equalization algorithms initialized the equalizer vector with a centralized delta function.

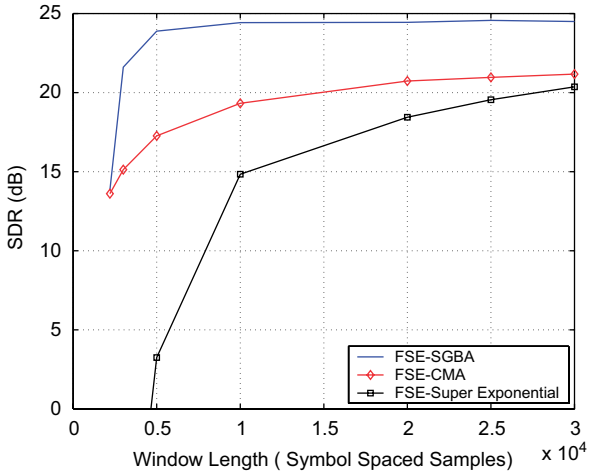


Fig. 6. Signal to distortion levels as a function of data window length.

In Fig. 6, the achieved signal to distortion ratio (SDR) levels as a function of window length are shown. Based on this result we can conclude that FS-SGBA algorithm achieves higher SDR levels for shorter data window lengths when compared to CMA and super-exponential algorithms. This is a useful property enabling the use of shorter packets for blind adaptation. The use of shorter packets enables lower complexity implementation and validates the quasi-stationary channel assumption in time varying environments. In Fig. 7, the equalizer outputs after convergence are shown (for a Window Length of 40 000). Note that since the phase ambiguity of the FS-SGBA method is an integer multiple of $\pi/2$ the equalizer outputs align with the original QAM constellation.

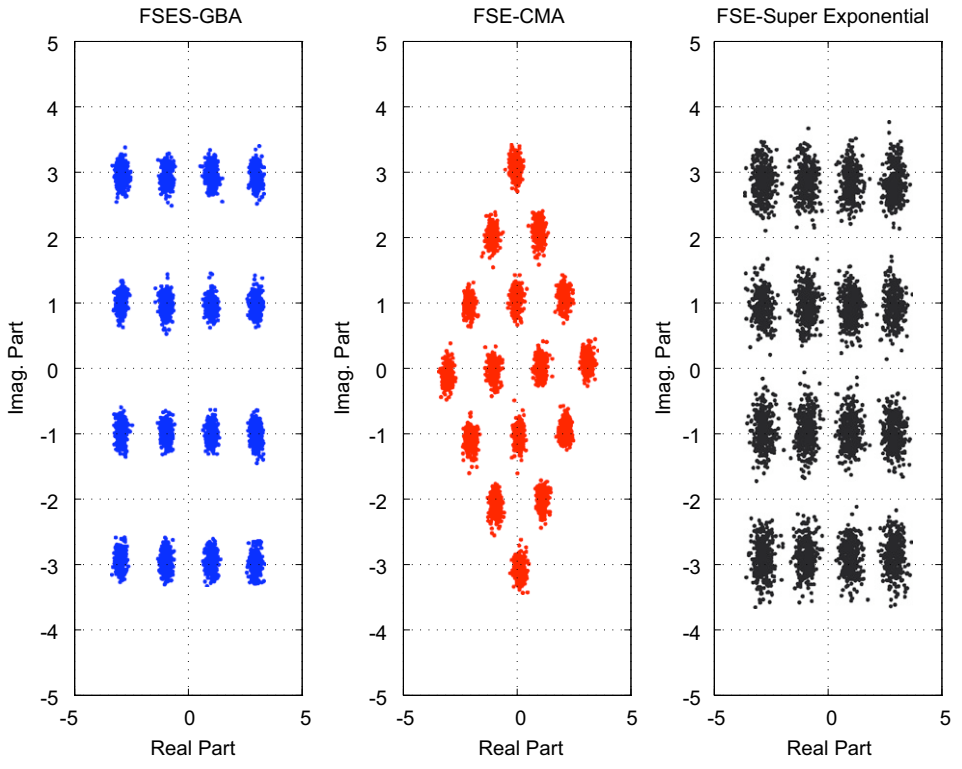


Fig. 7. Equalizer outputs after the convergence.

## 6. Conclusion

In this article we provided two different fractionally spaced extensions of the SGBA algorithm. The first one is the trivial extension of the linearly constrained SGBA algorithm, whereas in the second one the linear constraint is replaced with a quadratic constraint. Both algorithms have a very simple update structure, where the equalizer vector is updated with a weighted version of the input vector causing the maximum real magnitude output. This is potentially valuable since only major source of complexity for these algorithms is the computation of equalizer outputs for each iteration. If the computation of the outputs are performed through a dedicated hardware, the remaining computational requirement would be negligibly small.

Among the proposed algorithms, the quadratically constrained FS-SGBA algorithm is interesting in the sense that under the well-known equalizability conditions, the global convergence of the algorithm (corresponding to a non-convex optimization setting) can be proven. In addition, the knowledge of the channel length is not required in the quadratically constrained case and the quadratically constrained FS-SGBA has better SDR convergence behavior than its linearly constrained counterpart. Based on the simulations we can conclude that the proposed algorithm can be used to achieve higher SDR levels for shorter data windows, which is an attractive property especially for the time-varying environments that constrains the length of data windows used and for reducing the adaptation complexity.

## References

[1] S. Vembu, S. Verdu, R. Kennedy, W. Sethares, Convex cost functions in blind equalization, IEEE Trans. Signal Process. 42 (August 1994) 1952–1960.

[2] Z. Ding, Z. Luo, A fast linear programming algorithm for blind equalization, IEEE Trans. Comm. 48 (September 2000) 1432–1436.

[3] Z.-Q. Luo, M. Meng, K.M. Wong, J.-K. Zhang, A fractionally spaced blind equalizer based on linear programming, IEEE Trans. Signal Process. 50 (July 2002) 1650–1660.

[4] A.T. Erdogan, C. Kizilkale, Fast and low complexity blind equalization via subgradient projections, IEEE Trans. Signal Process. 53 (July 2005) 2513–2524.

[5] Z. Ding, On convergence analysis of fractionally spaced adaptive blind equalizers, IEEE Trans. Signal Process. 45 (March 1997) 650–657.

[6] B.T. Polyak, A general method for solving extremal problems, Doklady Akademii Nauk SSSR 174 (1967) 33–36.

[7] N. Shor, Minimization Methods for Nondifferentiable Functions, Springer, Berlin, 1985.

[8] ⟨http://spib.rice.edu/spib/microwave.html (chan1)⟩. Rice University, Signal Processing Information Base.

[9] P.A. Regalia, A finite-interval constant modulus algorithm, Proc. IEEE ICASSP, vol. 3, May 2002, pp. 2285–2288.