On the equivalence of blind equalizers based on MRE and subspace intersections

David Gesbert, Alle-Jan van der Veen, and A. Paulraj

Abstract— **Two classes of algorithms recently proposed for the blind equalization of multiple channels driven by a single source are the Mutually Referenced Equalizers (MRE) method by Gesbert *et al.*, and the Subspace Intersection (SSI) method by van der Veen *et al*. Although these methods seem at first sight unrelated, we show here that a variant of the SSI method and a particular member in the class of MRE methods provide mathematically identical solutions.**

Key-words: multi-channel blind equalization, fractionally spaced equalization, mobile communications, array signal processing.

I. Introduction

Blind equalization has been an active research area during the last few years. Two major factors appear to drive the wide interest in this topic. First, there is an increasing number of interesting and promising applications in the area of digital communications, wireless or not. Second, the fact was recognized that channel oversampling, either temporally (fractionally spaced equalizers) or in space (antenna arrays), leads to a multichannel data representation that offers several new leverages for solving the blind equalization problem, and thus enhances its applicability.

From an algebraic perspective, oversampling leads to a low-rank model for the output vector signal. This has been extensively exploited in the so-called second-order statistics and algebraic methods for the single-input, multiple-output (SIMO) identification problem [1]. At least three classes can be identified. The first tries to estimate the channels, viz. e.g. [2–4], the second considers the estimation of channel inverses (equalizers) [5–7], and the third attempts to recover the transmitted symbols directly from a (typically small) batch of output samples without resorting to channel/equalizer estimates [8, 9].

Categories 2 and 3 have the advantage of by-passing the channel estimation step, and this can result in increased robustness. The direct symbol-estimation methods [8, 9] have sometimes been called row-span methods as they exploit the row-span information of the data matrix to find the vector of unknown symbols. Following a seemingly different strategy, MRE techniques [6] estimate a collection of channel equalizers by forcing them to produce the same (unknown) output sequence, up to fixed equalization lags. The goal of this paper is to demonstrate that these two methods are in fact identical, with small differences arising only due to variations in the implementation.

In this paper, we first provide a new perspective of the row-span method of [9], by showing that the symbol estimates produced by this technique can be regarded as the outputs of linear equalizers, averaged across all equalization lags. We show that these equalizers optimize a *maximal coherence (MC)* criterion. Finally, we show the equivalence between the MC criterion and a particular member in the class of MRE criteria.

Notation. For a vector $\mathbf{x}$, $\mathbf{x}^t$ is its transpose, $\mathbf{x}^*$ its conjugate-transpose, and $\|\mathbf{x}\|$ its ℓ_2-norm. A sequence (row vector) with entries x_i is denoted by $\mathbf{x} = [x_i]$.

II. Data model

A. Data matrices

A digital symbol sequence $[s_i]$ is transmitted through a medium and received by an array of $M \geq 1$ sensors. The received signals are sampled $P \geq 1$ times faster than the symbol rate, here normalized to $T = 1$. Hence, during each symbol period, a total of MP measurements are available, which can be stacked into MP-dimensional vectors $\mathbf{x}_i$ as $\mathbf{x}_i = [x_i^1, \cdots, x_i^{MP}]^t$. Assuming an FIR channel, we can model $\mathbf{x}_i$ as the output of an MP-dimensional vector channel with impulse response $[\mathbf{h}_0, \mathbf{h}_1, \cdots, \mathbf{h}_{L-1}]$, where L denotes the channel length. In the noise free case, $\mathbf{x}_i$ is then given by

$$\mathbf{x}_i = \sum_{k=0}^{L-1} \mathbf{h}_k s_{i-k} \,. \tag{1}$$

Consider a finite block of data and define the $mMP \times N$ block-Toeplitz data matrix

$$\mathcal{X}^{(i)} = \begin{bmatrix} \mathbf{x}_i & \mathbf{x}_{i+1} & \ddots & \mathbf{x}_{i+N-1} \\ \mathbf{x}_{i-1} & \mathbf{x}_i & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots \\ \mathbf{x}_{i-m+1} & \ddots & \ddots & \ddots \end{bmatrix} .$$

N is the block length, while m can be interpreted as the memory of an equalizer acting on the rows of $\mathcal{X}^{(i)}$. Let $n = L + m - 1$. From (1), $\mathcal{X}^{(i)}$ has a factorization as $\mathcal{X}^{(i)} = \mathcal{H}\mathcal{S}^{(i)}$, where $\mathcal{H}$ is an $mMP \times n$ channel matrix and $\mathcal{S}^{(i)}$ is an $L + m - 1 \times N$ signal matrix, viz.

$$\begin{aligned} \mathcal{H} &= \begin{bmatrix} \mathbf{h}_0 & \cdots & \mathbf{h}_{L-1} & & \mathbf{0} \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \mathbf{h}_0 & \cdots & \mathbf{h}_{L-1} \end{bmatrix} \\ \mathcal{S}^{(i)} &= \begin{bmatrix} s_i & s_{i+1} & \ddots & s_{i+N-1} \\ \ddots & \ddots & \ddots & \ddots \\ s_{i-n+1} & \ddots & \ddots & \ddots \end{bmatrix} . \end{aligned} \tag{2}$$

We will assume that $\mathcal{H}$ is tall ($mMP \geq L + m - 1$) and $\mathcal{S}^{(i)}$ is wide ($L + m - 1 \leq N$), so that this is a low rank factorization. This requires at least $MP \geq 2$ and a sufficiently large m and N. We also

Submitted to the *IEEE Trans. on Signal Processing* Feb. 1998, revised July 1998 (correspondence item). D. Gesbert and A. Paulraj are with the Information Systems Laboratory, Stanford University, Stanford CA 94305, USA. A.-J. van der Veen is with the Dept. of Elec. Engr./DIMES, Delft University of Technology, Delft, The Netherlands.

assume that $\mathcal{H}$ has full column rank so that we can recover any row of $\mathcal{S}^{(i)}$ by taking linear combinations of the rows of $\mathcal{X}^{(i)}$. Finally, the input symbols are supposed to be persistently exciting so that the matrices $\mathcal{S}^{(i)}$ have full row rank.

B. Equalizers

An equalizer with delay k acting on $\mathcal{X}^{(i)}$ tries to reconstruct the $k+1$-st row of $\mathcal{S}^{(i)}$:

$$\mathbf{w}_k^* \mathcal{X}^{(i)} = [s_{i-k} \quad s_{i-k+1} \ \cdots] .$$

See figure $1(a)$. Since $\mathcal{S}^{(i)}$ has n rows, there is a total of n possible delays, and hence there are n different equalizers $\mathbf{w}_k$ ($k = 0, \cdots, n-1$). Note in particular that $\mathbf{w}_i^* \mathcal{X}^{(i)} = [s_0 \quad s_1 \ \cdots]$, hence

$$\mathbf{w}_i^* \mathcal{X}^{(i)} = \mathbf{w}_k^* \mathcal{X}^{(k)} , \qquad i, k = 0, \cdots, n-1 . \tag{3}$$

If m is large enough, then $\mathcal{X}^{(i)}$ is rank deficient. This is a source of non-uniqueness for the equalizers $\{\mathbf{w}_i\}$: any vector from the left null space of $\mathcal{X}^{(i)}$ may be added. The null space component is removed if we require the equalizer to have minimum norm. Equivalently, we can define the equalizer to act on a minimal basis of the row span of $\mathcal{X}^{(i)}$, rather than $\mathcal{X}^{(i)}$ itself. Thus introduce the SVDs:

$$\mathcal{X}^{(i)} = U_i \Sigma_i V^{(i)} , \qquad i = 0, \cdots, n-1 .$$

If $\mathcal{X}^{(i)}$ has rank n, then U_i has n orthonormal columns, $V^{(i)}$ has n orthonormal rows, and Σ_i is a diagonal matrix containing the n nonzero singular values. The rows of $V^{(i)}$ form an orthonormal basis for the row span of $\mathcal{X}^{(i)}$. A "normalized" equalizer acting on $V^{(i)}$ is called $\mathbf{t}_i$, which is related to $\mathbf{w}_i$ via $\mathbf{t}_i = \Sigma_i U_i^* \mathbf{w}_i$. Similarly to regular equalizers, we have (for $i, k = 0, \cdots, n-1$)

$$\mathbf{t}_i^* V^{(i)} = [s_0 \quad s_1 \ \cdots] , \ \text{ and } \ \mathbf{t}_i^* V^{(i)} = \mathbf{t}_k^* V^{(k)} . \tag{4}$$

C. Super-equalizers

Define

$$X_T = \begin{bmatrix} \mathcal{X}^{(0)} \\ \vdots \\ \mathcal{X}^{(n-1)} \end{bmatrix} , \qquad V_T = \begin{bmatrix} V^{(0)} \\ \vdots \\ V^{(n-1)} \end{bmatrix} . \tag{5}$$

"Super-equalizers" are long vectors that collect several equalizers with different delays, each reconstructing the same sequence $[s_0 \ s_1 \ \cdots]$. They act on the data X_T or on the normalized data V_T, respectively:

$$\mathbf{w}^* = [\mathbf{w}_0^* \ \cdots \ \mathbf{w}_{n-1}^*] , \qquad \mathbf{t}^* = [\mathbf{t}_0^* \ \cdots \ \mathbf{t}_{n-1}^*] .$$

It is interesting to consider the super-equalizer as combining the outputs of the regular equalizers, forming an average over all admissible delays. (By itself, it can also be interpreted as an ordinary equalizer of length $n+m-1$ at delay $n-1$.) See figure $1(b)$. Note that there is an issue of how to weight the outputs of each equalizer to combine them in an optimal fashion.

III. Blind equalization

A. Subspace intersection method

The problem of blind equalization is, for given a data matrix $\mathcal{X}$, to find a factorization $\mathcal{X} = \mathcal{H}\mathcal{S}$ where $\mathcal{S}$ meets the required Toeplitz structure. Since a Toeplitz matrix is generated by a single vector in a linear way, this translates to finding $\mathbf{s} = [s_0 \quad s_1 \ \cdots \ s_{N-1}]$ such that $\mathbf{s}$ lies simultaneously in row$(\mathcal{X}^{(0)})$, row$(\mathcal{X}^{(1)})$, $\cdots$, and row$(\mathcal{X}^{(n-1)})$, where 'row$(\cdot)$' stands for the row span. The goal of subspace intersection methods (SSIs) such as in [8,9] is to find the single vector $\mathbf{s}$ which is in the intersection of all n subspaces.

Numerically, there are several ways to compute the intersection. The algorithm proposed in [8] constructs the union of the complement of all row spans, and takes the complement again. The problem with this is that the complementary spaces can be highly dimensional (order N each). The "minimum noise subspace" (MNS) technique [10] is a method to prune the dimensions of each complementary space without changing the resulting union too much, thus greatly reducing the complexity. Although it was proposed in a different context it could be translated to apply to the current situation, but still the pruning would incur a loss in performance.

It was proven in [9] that, since the rows of $V^{(i)}$ form a minimal and *orthonormal* basis for row$(\mathcal{X}^{(i)})$, the exact intersection can also be obtained by constructing the matrix V_T in (5) and looking for the right singular vector corresponding to the *largest* singular value of V_T. This computation has a complexity much smaller than the algorithm in [8], and also smaller than what the MNS technique would give. Nonetheless, even with noise perturbations we find exactly the same output sequence as that produced by the algorithm in [8]. The corresponding principal left singular vector of V_T can be interpreted as the super-equalizer that returns this sequence.

In particular, it is proven in [9] that, if $\mathbf{t}_{ssi}$ is the principal left singular vector of V_T and $n = L+m-1$, then (without noise)

$$\mathbf{t}_{ssi}^* V_T = \alpha [s_0 \quad s_1 \ \cdots \ s_{N-1}]$$

where α is some nonzero scalar that makes the output sequence have norm 1. The reason, essentially, is that because of the normalization, the largest singular value of V_T is bounded by $\sqrt{n}$. This bound is attained when $\mathbf{t}_{ssi}^* = [\mathbf{t}_0^* \ \cdots \ \mathbf{t}_{n-1}^*]$ where each component by itself is an equalizer on the normalized signals (viz. (4)), returning a multiple α_i of $[s_0 \ s_1 \ \cdots]$. In fact, all scalings α_i will be the same.

Thus, $\mathbf{t}_{ssi}$ is a super-equalizer in the sense of section II-C. The corresponding equalizer on unnormalized data X_T is denoted by $\mathbf{w}_{ssi}$, related to $\mathbf{t}_{ssi}$ via

$$\mathbf{w}_{ssi} = [\mathbf{w}_0^* \ \cdots \ \mathbf{w}_{n-1}^*]^* , \qquad \mathbf{w}_i = U_i \Sigma_i^{-1} \mathbf{t}_i . \tag{6}$$

B. Maximal coherence criterion

The principal left singular vector $\mathbf{t}_{ssi}$ of V_T can also be expressed in terms of a criterion on the unnormalized received data. Indeed, $\mathbf{t}_{ssi}$ can be written as

$$\mathbf{t}_{ssi} = \arg \max_{\|\mathbf{u}\|^2 = 1} \mathbf{u}^* \mathcal{R}_V \mathbf{u}$$

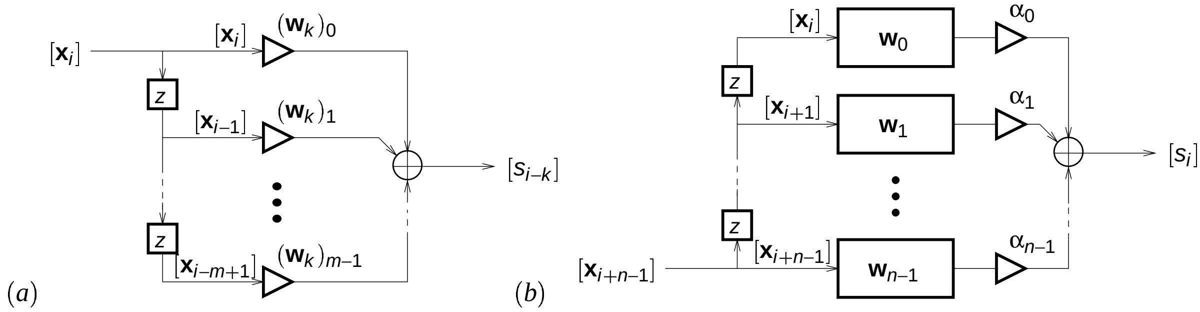

Fig. 1. (*a*) Equalizer with delay *k*; (*b*) Super-equalizer, combining the outputs of several equalizers at different delays

where $\mathcal{R}_V = V_T V_T^*$. Define the (empirical) correlation matrices $R_{i,j} = \mathcal{X}^{(i)}\mathcal{X}^{(j)^*}$,

$$\begin{aligned}\mathcal{R}_X &= X_T X_T^* = \begin{bmatrix} R_{0,0} & \cdots & R_{0,n-1} \\ \vdots & & \vdots \\ R_{n-1,0} & \cdots & R_{n-1,n-1}\end{bmatrix} \\ \mathcal{R}_0 &= \begin{bmatrix} R_{0,0} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & R_{n-1,n-1}\end{bmatrix}.\end{aligned}$$

Then $\mathcal{R}_X = \mathcal{R}_0^{1/2}\mathcal{R}_V\mathcal{R}_0^{1/2*}$, where

$$\mathcal{R}_0^{1/2} = \begin{bmatrix} R_{0,0}^{1/2} & & 0 \\ & \ddots & \\ \mathbf{0} & & R_{n-1,n-1}^{1/2}\end{bmatrix}$$

and $R_{i,i}^{1/2} := U_i \Sigma_i$.

It follows that $\mathbf{w}^*\mathcal{R}_X\mathbf{w} = \mathbf{u}^*\mathcal{R}_V\mathbf{u}$ for $\mathbf{u} = \mathcal{R}_0^{1/2*}\mathbf{w}$. Now denote by $\mathbf{w}_{ssi}$ the corresponding super-equalizer provided by the SSI method (related to $\mathbf{t}_{ssi}$ as in equation (6)). By substitution, $\mathbf{w}_{ssi}$ is found to optimize the following constrained criterion:

$$\mathbf{w}_{ssi} = \arg\max_{\mathbf{w}^*\mathcal{R}_0\mathbf{w}=1} \mathbf{w}^*\mathcal{R}_X\mathbf{w} = \arg\max_{\mathbf{w}^*\mathcal{R}_0\mathbf{w}=1} J_{ssi}\,, \tag{7}$$

where J_{ssi} is given by

$$J_{ssi} := \|\sum_{i=0}^{n-1}\mathbf{w}_i^*\mathcal{X}^{(i)}\|^2$$

and the constraint can be written as

$$\mathbf{w}^*\mathcal{R}_0\mathbf{w} = \sum_{i=0}^{n-1}\|\mathbf{w}_i^*\mathcal{X}^{(i)}\|^2 = 1\,. \tag{8}$$

Thus, the subspace intersection solution is also obtained by maximizing the power of the sum of all equalizers outputs, subject to the constraint that the sum of the powers is kept constant. *The SSI method maximizes the coherence of the equalizers outputs.* Indeed, in the noise-free case, all equalizers return the *same* output sequence $[s_0\ \ s_1\ \cdots]$, up to a common scaling. Note that this is true only in the case of the constraint specified in (8).

C. The MRE method

The idea behind the mutually referenced equalizer (MRE) method for blind equalization [6] is to exploit the the relations in (3) by finding a vector of n equalizers $\mathbf{w} = [\mathbf{w}_0^* \cdots \mathbf{w}_{n-1}^*]^*$ that simultaneously minimizes all differences $\|\mathbf{w}_i^*\mathcal{X}^{(i)} - \mathbf{w}_k^*\mathcal{X}^{(k)}\|^2$. This can be written as a least-squares problem[1]

$$\mathbf{w}^*\mathcal{X}_{mre} = \mathbf{0}\,,$$
$$\mathcal{X}_{mre} := \left[\begin{array}{cccc|cccc|c} \mathcal{X}^{(0)} & \mathcal{X}^{(0)} & \cdots & \mathcal{X}^{(0)} & -\mathcal{X}^{(0)} & & & & \\ -\mathcal{X}^{(1)} & & & & \mathcal{X}^{(1)} & \mathcal{X}^{(1)} & \cdots & \mathcal{X}^{(1)} & \\ & -\mathcal{X}^{(2)} & & & & -\mathcal{X}^{(2)} & & & \cdots \\ & & \ddots & & & & \ddots & & \\ & & & -\mathcal{X}^{(n-1)} & & & & -\mathcal{X}^{(n-1)} & \end{array}\right].$$

To avoid trivial solutions, $\mathbf{w}$ should be constrained, e.g. by fixing one of its entries or its norm. Another suitable constraint is one that keeps the sum of output powers to a constant, $\mathbf{w}^*\mathcal{R}_0\mathbf{w} = 1$. The motivation for this particular choice is that it avoids trivial null space solutions $\mathbf{w}_i^*\mathcal{X}^{(i)} = \mathbf{0}\ \forall i$, which is necessary in the noise free case. Thus we obtain

$$\begin{aligned}\mathbf{w}_{mre} &:= \arg\min_{\mathbf{w}^*\mathcal{R}_0\mathbf{w}=1} J_{mre}\,, \\ J_{mre} &:= \sum_{i=0}^{n-1}\sum_{k=0}^{n-1}\|\mathbf{w}_i^*\mathcal{X}^{(i)} - \mathbf{w}_k^*\mathcal{X}^{(k)}\|^2\,.\end{aligned} \tag{9}$$

We elaborate and find

$$\begin{aligned}J_{mre} &= \mathbf{w}^*\mathcal{X}_{mre}\mathcal{X}_{mre}^*\mathbf{w} \\ &= 2\mathbf{w}^*\begin{bmatrix}(n-1)R_{0,0} & -R_{0,1} & \cdots & -R_{0,n-1} \\ -R_{1,0} & (n-1)R_{1,1} & & \cdot \\ \vdots & & & \vdots \\ -R_{n-1,0} & & \cdots & (n-1)R_{n-1,n-1}\end{bmatrix}\mathbf{w}\,.\end{aligned}$$

It thus follows that

$$J_{mre} + 2J_{ssi} \ = \ 2n\,\mathbf{w}^*\mathcal{R}_0\mathbf{w}\,.$$

Under the constraint $\mathbf{w}^*\mathcal{R}_0\mathbf{w} = 1$, we finally obtain

$$\min_{\mathbf{w}^*\mathcal{R}_0\mathbf{w}=1} J_{mre} = 2n - \max_{\mathbf{w}^*\mathcal{R}_0\mathbf{w}=1} J_{ssi}\,.$$

[1] The equation is reminiscent of the Cross-Relation method in [4], but this connection is only optical. Firstly, we here estimate equalizers and not the channel as in [4]. More importantly, the CR method does not cross-relate delays of the full data matrices, but rather the *MP* scalar subchannels, so that the superscript (i) in $\mathcal{X}^{(i)}$ has a different meaning.

This means that $\mathbf{w}_{mre} \equiv \mathbf{w}_{ssi}$.

Hence we conclude that the SSI method and the extended MRE method under output power constraint are identical. Note that the MRE method can use several other constraints, however only the one presented here guarantees the equivalence of the two methods.

D. Remarks

The SSI method here is slightly different from the version in [9]. There, the sequence was extended with additional tail symbols, which changed the definition of V_T such that only a single matrix $V^{(0)}$ was needed, so that only a single data matrix has to be normalized, leading to computational savings. This implementation of the SSI method is asymptotically identical to the one presented here, which was chosen for expository reasons. With noise, the SSI method on normalized data V_T and on original data X_T are slightly different. The reason is that, with noise, each $\mathcal{X}^{(i)}$ is always full rank, whereas $V^{(i)}$ is presumably obtained from a truncated SVD, resulting in an approximate n-dimensional basis for the row span of $\mathcal{X}^{(i)}$. If we omit the truncation, i.e. define $V^{(i)}$ to contain all mMP right singular vectors of $\mathcal{X}^{(i)}$, then the solution is exactly equal to the SSI method on V_T.

References

[1] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Processing*, vol. 50, no. 1-2, pp. 83–99, Apr. 1996.

[2] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Information Theory*, vol. 40, no. 2, pp. 340–349, Mar. 1994.

[3] E. Moulines, P. Duhamel, J.F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[4] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind equalization," *IEEE Trans. Signal Processing*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[5] D.T.M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. ICASSP 94 Conference*, Adelaide, Australia, Apr. 1994, pp. IV.585–588.

[6] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. Signal Processing*, vol. 45, no. 9, pp. 2307–2317, Sept. 1997.

[7] G. Giannakis and S. Halford, "Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions," *IEEE Trans. Signal Processing*, vol. 45, no. 9, pp. 2277–2292, Sept. 1997.

[8] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 43, no. 11, pp. 2714–2723, Nov. 1995.

[9] A.J. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 173–190, Jan. 1997.

[10] Y. Hua, K. Abed-Meraim, and M. Wax, "Blind system identification using minimum noise subspace," *IEEE Trans. Signal Proc.*, vol. 45, pp. 770–773, Mar. 1997.