we would like to trade off extra redundancy for error performance. Next, we consider the channel RMSE for SNR = 20 dB when the order L_h is overestimated and $P = M + L_h$ in Fig. 10 and when the order is overestimated with $P = 19$ in Fig. 11. In Fig. 10, we see the beneficial effects of having a larger prefix, whereas Fig. 11 shows the graceful degradation when the channel is overestimated.

Experiment 2: In this experiment, we consider the effect of the cycle chosen on the resulting channel error in estimating the two-ray channel above. Fig. 12 considers the performance of the OC approach for $I = 100$, $P = 19$, $M = 15$, SNR = 20 dB, and 120 M symbols for cycles 1 ... 6, whereas Fig. 13 considers similarly the performance using the TC approach with cycles 1 and 2 ... 7. Cycle selection seems to have an effect on the channel error, but asymptotic performance analysis is required to determine its precise role.

Experiment 3: Now, we look at the probability of bit error for an OFDM system. In Fig. 14, we plot the RMS symbol estimation error, and in Fig. 15, we plot the probability of bit error (assuming Gray coding in selection of the 16 QAM symbols) estimated over 500 Monte Carlos of 500 M data for an OFDM system with $M = 15$ and $P = 19$, with and without a $(15, 11)$ two symbol-error correcting Reed–Solomon (RS) equivalent code for the artificial channel $h = [1, 2, 1, -1, 1]/\sqrt{8}$. We used the standard OFDM ZF and MMSE structures [12] to equalize the $L_h = 4$ channel above. Next, we consider the same channel and $M = 15$ and $P = 17$ to observe the effects of channels longer than the cyclic prefix. We estimate the channel as before but look at MMSE equalization with and without the use of impulse response shortening [9] and RS$(15, 11)$ coding. We used an eight-tap, zero-delay shortening filter derived from the estimated channel. In Fig. 16, we plot the RMS symbol estimation error, and in Fig. 17, we plot the estimated probability of error. For comparison purposes, in Figs. 15 and 17, we plot the MMSE uncoded and coded solutions for the case when $h(n) = \delta(n)$ as well as when there is no attempt at equalization. In Fig. 15, we see that the performance of the system using equalization with our channel estimate approaches the performance of the case where $h(n) = \delta(n)$. From Fig. 17, we see that impulse response shortening may be a beneficial technique when combined with our channel estimate since it reduces the the error floor present in the unshortened scenario. Performance of impulse response shortening varies with the channel and may be improved by changing shortening parameters. Further improvements may be obtained using vector MMSE or vector MMSE decision feedback equalizers at the expense of further complexity [6].

References

[1] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, “A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels,” *IEEE Trans. Commun.*, vol. 43, pp. 773–775, Mar. 1995.

[2] L. J. Cimini, Jr., “Performance studies for high-speed indoor wireless communications,” *Wireless Pers. Commun.*, vol. 2, nos. 1–2, pp. 67–85, 1995.

[3] M. de Courville, P. Duhamel, P. Madec, and J. Palicot, “A least mean squares blind equalization techniques for OFDM systems,” *Ann. Telecommun.*, vol. 52, nos. 1–2, pp. 12–20, Jan.–Feb. 1997.

[4] Z. Ding, “Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics,” *IEEE Signal Processing Lett.*, vol. 3, pp. 150–152, May 1996.

[5] G. B. Giannakis, “Filterbanks for blind channel identification and equalization,” *IEEE Signal Processing Lett.*, vol. 4, pp. 184–187, June 1997.

[6] R. W. Heath, Jr., “Mitigating channel distortions in wireless orthogonal frequency division multiplexing communication systems,” Dept. Elect. Eng., Univ. Virginia, Charlottesville, Aug. 1997.

[7] J. W. Lechleider, “The optimum combination of block codes and receivers for arbitrary channels,” *IEEE Trans. Commun.*, vol. 38, pp. 615–621, May 1990.

[8] B. L. Floch, M. Alard, and C. Berrou, “Coded orthogonal frequency division multiplex,” *Proc. IEEE*, vol. 83, pp. 982–996, June 1995.

[9] P. Melsa, R. C. Younce, and C. E. Rohrs, “Impulse response shortening for discrete multitone transceivers,” *IEEE Trans. Commun.*, vol. 44, pp. 1662–1672, Dec. 1996.

[10] T. Pollet and M. Moeneclaey, “The effect of carrier frequency offset on the performance of band limited single carrier and OFDM signals,” in *Proc. GLOBECOM*, London, U.K., Nov. 18–22, 1996, pp. 719–723.

[11] H. Sari, G. Karam, and I. Jeanclaude, “An analysis of orthogonal frequency-division multiplexing for mobile radio applications,” in *Prof. Vehic. Technol. Conf.*, Stockholm, Sweden, June 8–10, 1994, pp. 1635–1639.

[12] H. Sari, G. Karam, and I. Jeanclaude, “Transmission techniques for digital terrestrial TV broadcasting,” *IEEE Commun. Mag.*, pp. 100–109, Feb. 1995.

[13] E. Serpedin and G. B. Giannakis, “Blind channel identification and equalization using modulation induced cyclostationarity,” *IEEE Trans. Signal Processing*, vol. 46, pp. 3099–3104, Nov. 1998; see also *Proc. 31st Conf. Inform. Sci. Syst.*, Johns Hopkins Univ., Baltimore, MD, vol. II, Mar. 19–21, 1997, pp. 792–797.

[14] M. K. Tsatsanis and G. B. Giannakis, “Transmitter induced cyclostationarity for blind channel equalization,” *IEEE Trans. Signal Processing*, vol. 45, pp. 1785–1794, July 1997.

[15] J.-J. van de Beek, M. Sandell, and P. O. Börjesson, “ML estimation of time and frequency offset in OFDM systems,” *IEEE Trans. Signal Processing*, vol. 45, pp. 180–1805, July 1997.

[16] L. Vandendorpe, “MMSE equalizers for multitone systems without guard time,” in *Proc. Euro. Signal Process. Conf.*, Sept. 10–13, 1996.

[17] E. Viterbo and K. Fazel, “How to combat long echoes in OFDM transmission schemes: Sub-channel equalization or more powerful channel coding,” in *Proc. GLOBECOM*, Singapore, Nov. 14–16, 1995, pp. 2069–2074.

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 for multichannel blind equalization are the mutually referenced equalizer (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 certain variants of the SSI and the MRE methods both optimize a new blind criterion, which is referred to as *maximum coherence* and, thus, are equivalent.**

***Index Terms*—Array signal processing, fractionally spaced equalization, mobile communications, multichannel blind equalization.**

I. Introduction

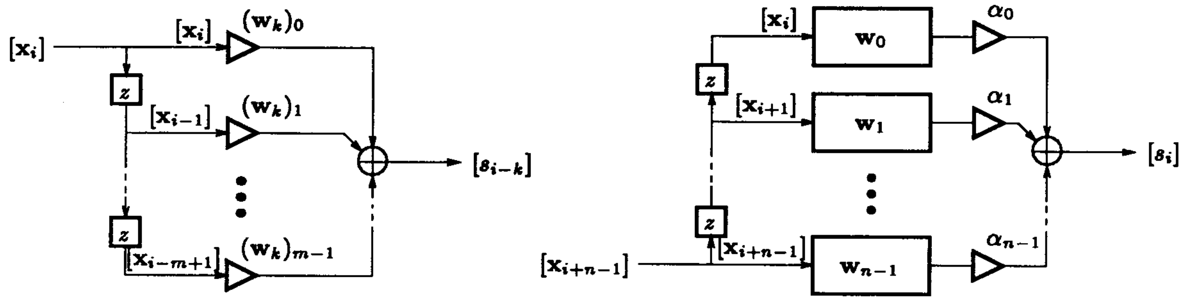
Blind equalization has been an active research area during the last few years. Two major factors appear to drive the wide interest in this topic. First, there is an increasing number of interesting and promising applications in the area of digital communications: wireless

Manuscript received February 17, 1998; revised August 4, 1998. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lai C. Godara.

D. Gesbert and A. Paulraj are with the Information Systems Laboratory, Stanford University, Stanford CA 94305 USA (e-mail: gesbert@rascals.stanford.edu).

A.-J. van der Veen is with the Department of Electrical Engineering/DIMES, Delft University of Technology, Delft, The Netherlands.

Publisher Item Identifier S 1053-587X(99)01349-5.



(a) (b)

Fig. 1. (a) Equalizer with delay k and (b) superequalizer, combining the outputs of several equalizers at different delays.

or otherwise. Second, it 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 bypassing 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 correspondence is to demonstrate that these two methods are, in fact, identical with small differences arising only due to variations in the implementation.

In this correspondence, 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 equalizer 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, which, here, is 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, whereas 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.

$$\mathcal{H} = \begin{bmatrix} \mathbf{h}_0 & \cdots & \mathbf{h}_{L-1} & & \mathbf{0} \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \mathbf{h}_0 & \cdots & \mathbf{h}_{L-1} \end{bmatrix}$$

and

$$\mathcal{S}^{(i)} = \begin{bmatrix} s_i & s_{i+1} & \ddots & s_{i+N-1} \\ \ddots & \ddots & \ddots & \ddots \\ s_{i-n+1} & \ddots & \ddots & \ddots \end{bmatrix}. \tag{2}$$

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 assume that $\mathcal{H}$ has full column rank; therefore, we can recover any row of $\mathcal{S}^{(i)}$ by taking linear combinations of the rows of $\mathcal{X}^{(i)}$. Finally, the matrices $\mathcal{S}^{(i)}$ are supposed to 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 Fig. 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 $\mathbf{w}_i^* \mathcal{X}^{(i)} = [s_0 \quad s_1 \; \cdots]$, and 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, leading to nonuniqueness 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. We can also define the equalizer to act on a minimal basis of the row span of $\mathcal{X}^{(i)}$ rather than $\mathcal{X}^{(i)}$ itself. Thus, we introduce the SVD's

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

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 $i, k = 0, \cdots, n-1$)

$$\mathbf{t}_i^* V^{(i)} = [s_0 \quad s_1 \quad \cdots]$$

and

$$\mathbf{t}_i^* V^{(i)} = \mathbf{t}_k^* V^{(k)}. \tag{4}$$

C. Superequalizers

Define

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

"Superequalizers" 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 superequalizer 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 Fig. 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 $\mathbf{s} = [s_0 \quad s_1 \quad \cdots \quad s_{N-1}]$ 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 (SSI's) such as in [8] and [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 the pruning would still 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 that is much smaller than the algorithm in [8] and 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 superequalizer 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)

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

where α is some nonzero scalar that makes the output sequence have norm 1. 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 scaling α_i will be the same.

Thus, $\mathbf{t}_{ssi}$ is a superequalizer in the sense of Section II-C. The corresponding equalizer on unnormalized data X_T is denoted by $\mathbf{w}_{ssi}$ and 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

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

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

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

and

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

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 superequalizer provided by the SSI method [related to $\mathbf{t}_{ssi}$ as in (6)]. By substitution, $\mathbf{w}_{ssi}$ is found to optimize the 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

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

and the constraint can be written as

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

Thus, the subspace intersection solution is also obtained by maximizing the power of the sum of all equalizer's outputs, subject to the constraint that the sum of the powers is kept constant. *The SSI*

$$\mathbf{w}^*\mathcal{X}_{mre} = 0 \qquad \mathcal{X}_{mre} := \left[\begin{array}{cccc|cccc|c} \mathcal{X}^{(0)} & \mathcal{X}^{(0)} & \cdots & \mathcal{X}^{(0)} & -\mathcal{X}^{(0)} & & & & \\ -\mathcal{X}^{(1)} & & & & \mathcal{X}^{(1)} & \mathcal{X}^{(1)} & \cdots & \mathcal{X}^{(1)} & \\ & -\mathcal{X}^{(2)} & & & & -\mathcal{X}^{(2)} & & & \cdots \\ & & \ddots & & & & \ddots & & \\ & & & -\mathcal{X}^{(n-1)} & & & & -\mathcal{X}^{(n-1)} & \end{array}\right] \tag{8a}$$

method maximizes the coherence of the equalizer's 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 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] as shown in (8a) at the top of the page. 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

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

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

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

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

Hence, we conclude that the SSI method and the extended MRE method under the 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.

[1] The equation is reminiscent of the cross-relation method in [4], but this connection is only optical. Here, we 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.

References

[1] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Process.*, vol. 50, nos. 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. Inform. Theory*, vol. 40, 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, 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, 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 Conf.*, Adelaide, Australia, Apr. 1994, pp. IV.585–IV.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, 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, 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, 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, pp. 173–190, Jan. 1997.

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