where H is the natural entropy function and $R = k/n$ is the rate.

Since we are dealing with the asymptotical case, we normalize by setting $i = n\omega$, and we define the function $f(\omega, R, q)$ by

$$A_{\omega n} = e^{nf(\omega, R, q)}.$$

From Lemma 2, we get

$$f(\omega, R, q) = H(\omega) + \omega \ln(q-1) - (1-R)\ln q. \tag{6}$$

Note that, for a given A_i, there are two solutions for i. Setting $A_i \approx 1$, the two solutions will be the minimum and the maximum weights. These are, of course, also the zeros of f.

Let $\delta = d_1/n$ and $\mu = m_1/n$ be, respectively, the minimum and maximum normalized weights. Because μ and δ are the zeros of f, we get

$$H(\delta) + \delta \ln(q-1) = H(\mu) + \mu \ln(q-1)$$

or

$$(\delta - \mu)\ln(q-1) = \delta \ln \delta + (1-\delta)\ln(1-\delta) - \mu \ln \mu - (1-\mu)\ln(1-\mu). \tag{7}$$

Lemma 3 (Varshamov–Gilbert): For almost all linear codes, the rate and the normalized minimum distance are related by the following equation:

$$H(\delta) + \delta \ln(q-1) = (1-R)\ln q.$$

Proof: This follows from equating $f(\omega, R, q) = 0$ as in (6). □

We know from Theorem 1 that if $\delta > 3/4$, then the code is $(2, 2)$-separating. Hence we can, by substituting $\delta = 3/4$ in the Varshamov–Gilbert equation, get rates for which almost any code is $(2, 2)$-separating asymptotically. The rates such obtained are presented under "Technique I" in Table I. By the Plotkin bound, this gives nothing over small fields.

Technique II in the table is an improvement based on Theorem 1, which says that every code with $4\delta > 3\mu$ is $(2, 2)$-separating. We insert $\delta = 4\mu/3$ in (7) and get

$$\frac{\delta}{3}\ln(q-1) = \delta \ln \delta + (1-\delta)\ln(1-\delta) - \frac{4\delta}{3}\ln\frac{4\delta}{3} - \left(1 - \frac{4\delta}{3}\right)\ln\left(1 - \frac{4\delta}{3}\right). \tag{8}$$

We have solved this equation numerically for the smallest fields, and the results are given in Table I. Of course, we will always have

$$0 \le \delta \le \mu \le 1$$

which will bound $\delta \le 3/4$ in (8).

This results in no real solution of (8) for $q \ge 11$.

Note that, in Table I, the best results are obtained by the "Constructions" for $q \le 5$, then by "Technique II" for $7 \le q \le 9$, and finally by "Technique I" for higher values of q. In the binary case, $R = 0.0642$ can be achieved nonconstructively [1].

Acknowledgment

The authors wish to thank the anonymous referees for their useful remarks.

References

[1] Yu. L. Sagalovich, "Separating systems," *Probl. Inform. Transm.*, vol. 30, no. 2, pp. 105–123, 1994.

[2] A. D. Friedman, R. L. Graham, and J. D. Ullman, "Universal single transition time asynchronous state assignments," *IEEE Trans. Comput.*, vol. C-18, pp. 541–547, June 1969.

[3] M. L. Friedman and J. Komlós, "On the size of separating systems and families of perfect hash functions," *SIAM J. Alg. Discr. Meth.*, vol. 5, no. 1, pp. 61–68, Mar. 1984.

[4] J. Körner and G. Simonyi, "Separating partition systems and locally different sequences," *SIAM J. Discr. Math.*, vol. 1, no. 3, pp. 355–359, Aug. 1988.

[5] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 1995, vol. 963, pp. 452–465.

[6] D. R. Stinson, T. Van Trung, and R. Wei, "Secure frameproof codes, key distribution patterns, group testing algorithms and related structures," *J. Statist. Planning and Inference*, vol. 86, no. 2, pp. 595–617, 2000.

[7] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[8] S. Encheva and G. Cohen, "Identifying codes for copyright protection," Dept. INF, CNRS, Paris, France, Tech. Rep., 2001.

[9] D. R. Stinson and R. Wei, "Combinatorial properties and constructions of traceability schemes and frameproof codes," *SIAM J. Discr. Math*, vol. 11, pp. 41–53, 1998.

[10] I. M. Chakravarti, "Families of codes with few distinct weights from singular and nonsingular Hermitian varieties and quadrics in projective geometries and Hadamard difference sets and designs associated with two-weight codes," in *Coding Theory and Design Theory, Part 1*. New York: Springer-Verlag, 1990, pp. 35–50.

[11] G. Cohen, S. Encheva, and H. G. Schaathun, "On separating codes," Dept. INF, CNRS Paris, France, Tech. Rep., 2001.

[12] M. A. Tsfasman, "Algebraic-geometric codes and asymptotic problems," *Discr. Appl. Math.*, vol. 33, pp. 241–256, 1991.

On the Equivalence of Three Reduced Rank Linear Estimators With Applications to DS-CDMA

Wanshi Chen, *Member, IEEE*, Urbashi Mitra, *Member, IEEE*, and Philip Schniter, *Member, IEEE*

***Abstract*—This correspondence shows the equivalence of three previously proposed reduced-rank detection schemes for direct-sequence code-division multiple-access (DS-CDMA) communication systems. The auxiliary vector filtering (AVF) algorithm is simplified through a key observation on the construction of the auxiliary vectors. After simplification, it is shown that the AVF algorithm is equivalent to the multistage Wiener filtering (MWF) algorithm of Honig and Goldstein. Furthermore, these schemes can be shown to be equivalent to the multistage linear receiver scheme based on the Cayley–Hamilton (CH) theorem when the minimum mean-square error (MMSE) criterion is applied to the reduced dimensional space of the received signal.**

Manuscript received May 2, 2001; revised March 4, 2002. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Washington, DC, June 2001.

W. Chen is with Ericsson Wireless Communications, San Diego, CA 92121 USA (e-mail: wanshi.chen@ericsson.com).

U. Mitra is with the Department of Electrical Engineering–Systems, Communication Sciences Institute, The University of Southern California, Los Angeles, CA 90089 USA (e-mail: ubli@usc.edu).

P. Schniter is with the Department of Electrical Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: schniter@ee.eng.ohio-state.edu).

Communicated by G. Caire, Associate Editor for Communications.

Publisher Item Identifier 10.1109/TIT.2002.801472.

***Index Terms*—Direct-sequence code-division multiple access (DS-CDMA), minimum mean-square error (MMSE), reduced rank algorithms.**

I. Introduction

Multiuser detectors yield significantly better performance than the conventional single-user-based matched filter (MF) detection scheme for direct-sequence code-division multiple-access (DS-CDMA) communications [11]. Among the class of such receivers, the linear minimum mean-square error (MMSE) detector has received considerable attention [1], [2], [5], [12] due to its simplicity of implementation, strong performance, and, more importantly, amenability to adaptive implementation.

The study of reduced rank interference suppression for DS-CDMA is motivated by situations where the number of taps to be tracked by the adaptive MMSE detector is so large that the receiver responds quite slowly in a time-varying environment. By projecting the received signal onto a subspace of reduced rank, the number of taps in the adaptive filter is reduced thereby improving tracking ability. Reduced rank algorithms based on the exploitation of the Cayley Hamilton theorem [3] were provided in [9]. Therein, approximate MMSE detectors with a multistage linear implementation were presented. In [10], the auxiliary-vector filtering (AVF) method was proposed. In this reduced rank method, an auxiliary vector was derived based on maximizing the correlation between the outputs of the reference vector filter and the previously derived auxiliary vector filters. A recursive conditional optimization of the auxiliary vectors and the weights associated with each vector was also presented. In [6], the multistage Wiener filtering (MWF) method of [4] was applied to DS-CDMA systems. The authors [6] showed that the MWF algorithm reduced rank algorithm required much fewer training samples than the full rank algorithms. Several adaptive implementations for the MWF algorithm were proposed [6].

In this correspondence, we show by theoretical analysis that the MWF, the AVF, and the Cayley–Hamilton (CH) method of [9] are essentially equivalent. We begin with simplifying the derivation of the auxiliary vectors for the AVF algorithm. The recursive conditional optimization of the auxiliary vectors involves high-complexity computations. The conditional weight for each newly derived auxiliary vector only makes the optimization procedure more complicated. By observing properties among the auxiliary vectors, we propose a simplification, which yields a more compact solution for the auxiliary vectors and greatly reduces the computational complexity as well. More importantly, the simplification of the AVF algorithm establishes the necessary connection between the AVF algorithm and the MWF algorithm, and makes it possible to prove the equivalence of the above two algorithms. In addition, we introduce an additional constraint on the blocking matrices for the MWF algorithm. Under these conditions, the choice of the blocking matrices does not affect the performance of the MWF filter. With the help of the simplification of the AVF algorithm and the additional constraint, we prove that the MWF algorithm is equivalent to the AVF algorithm. The proof also naturally leads to the fact that the projection vectors for the MWF algorithm and the CH approach of [9] share the same subspace. Note that the equivalence of the MWF algorithm and the CH approach was also shown via an alternative method in [6]. Although our focus is on DS-CDMA systems, our results are general and are not predicated upon specific characteristics of spread-spectrum signals. As such, the methods under study can be applied to other areas such as array signal processing.

This correspondence is organized as follows. The system model is given in Section II. In Section III, the three reduced rank methods are described. Before showing the proof of the equivalence of the three techniques, a simplification of the AVF algorithm is provided in Section IV. The proof of the equivalence is given in Section V. Finally, conclusions are drawn in Section VI.

II. System Model

Consider a generic model in complex baseband and discrete time. The $N \times 1$ observation vector $\boldsymbol{y}(m)$ is corrupted by the complex zero-mean and wide-sense stationary (WSS) noise. An $N \times 1$ linear receiver $\boldsymbol{c}$ operates on the observation vector $\boldsymbol{y}(m)$ to recover $b_1(m)$, the scalar reference signal. The linear MMSE receiver is to minimize the MSE

$$\text{MSE} = E\left\{ \left| b_1(m) - \boldsymbol{c}^H \boldsymbol{y}(m) \right|^2 \right\} \tag{1}$$

where $E\{\}$ denotes expectation and H is the Hermitian transpose. The optimum coefficients are

$$\boldsymbol{c}_{mmse} = \boldsymbol{R}^{-1}\boldsymbol{p} \tag{2}$$

where $\boldsymbol{R} = E\{\boldsymbol{y}(m)\boldsymbol{y}(m)^H\}$ is the correlation matrix and $\boldsymbol{p} = E\{b_1(m)\boldsymbol{y}(m)\}$ denotes the *steering vector*.

Since prior work [6], [9], [10] considered DS-CDMA communications, we will similarly focus on asynchronous DS-CDMA systems with binary phase-shift keying (BPSK) modulation in flat-fading environments. However, our analysis is not restricted to DS-CDMA applications. Instead, it can be applied to other applications, such as array signal processing, as well. The transmitted baseband signal of the kth user is given by

$$x_k(t) = \sum_{m=0}^{M-1} A_k b_k(m) s_k(t - mT_b - \tau_k) \tag{3}$$

where M is the number of transmitted data symbols, A_k the amplitude of user k, T_b is the symbol period, $b_k(m) \in \{-1, +1\}$ is the binary data, and τ_k is the timing delay, which is assumed to be uniformly distributed within $[0, T_b]$. The signature waveform $s_k(t)$ is

$$s_k(t) = \sum_{n=0}^{N-1} s_k[n]\Psi(t - nT_c) \tag{4}$$

where $N = T_b/T_c$ is the spreading gain, T_c is the chip period

$$\boldsymbol{s}_k = [s_k[0], s_k[1], \ldots, s_k[N-1]]^T$$

is the normalized spreading code of user k, which is assumed to be fixed and have a period of N (i.e., short spreading codes), where T denotes transpose, and the function $\Psi(t)$ is the chip waveform.[1] Without loss of generality, user 1 is taken to be the user of interest. In addition, the time delay for user 1 is assumed to be perfectly known and it is fixed during the transmission. As a result, we can let $\tau_1 = 0$.

Although multishot MMSE receivers with an observation window longer than one symbol duration can have improved performance [11], we will only focus on an one-shot observation window. The received baseband signal is first passed through a chip-MF before chip-rate sampling. As a result, the resulting $N \times 1$ received vector is represented by [8]

$$\begin{aligned}\boldsymbol{y}(m) = \sum_{k=1}^{K} A_k [\gamma_k(m) b_k(m) \boldsymbol{s}_k^+ \\ + \gamma_k(m-1) b_k(m-1) \boldsymbol{s}_k^-] + \boldsymbol{n}(m)\end{aligned} \tag{5}$$

where K is the number of users, and $\boldsymbol{n}(m)$ is the complex white noise with covariance matrix $\sigma^2 \boldsymbol{I}_N$, where σ^2 is the noise variance and $\boldsymbol{I}_N$

[1] Rectangular chip waveforms are often assumed for computer simulations in the literature. However, the derivation in the sequel does not assume a particular chip waveform. Bandwidth-efficient chip waveforms can also be treated in our analysis. The effect of chip waveforms is embedded in the construction of the effective spreading codes for the active users, as will be shown in (5).

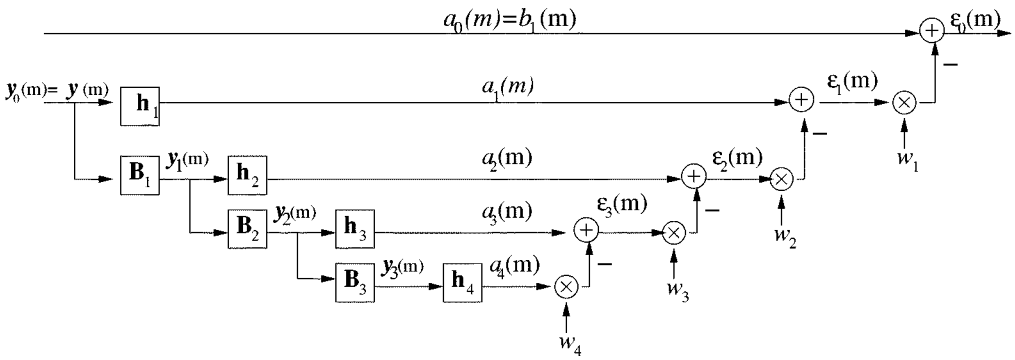

Fig. 1. Illustration of a four-stage Wiener filter.

is the $N \times N$ identity matrix. The fading process $\gamma_k(m)$ is a complex WSS random process. The partial spreading codes $\boldsymbol{s}_k^+$ and $\boldsymbol{s}_k^-$ correspond to the effective spreading codes for the current bit and the previous bit [8], respectively. Both $\boldsymbol{s}_k^+$ and $\boldsymbol{s}_k^-$ are functions of the spreading code, time delay, and chip waveform of user k and they are assumed to be real. The reader is referred to, e.g., [8] for detailed descriptions of $\boldsymbol{s}_k^+$ and $\boldsymbol{s}_k^-$.[2]

III. Reduced Rank MMSE Filtering

Either a training sequence based recursive least squares (RLS) algorithm, adaptive least mean squares (LMS), or adaptive minimum output energy (MOE) algorithm [5], can be used to adaptively estimate $\boldsymbol{c}_{mmse}$, as shown in (2), in a time-varying channel scenario. However, in situations where N is large, slow convergence can be expected. This slow convergence is not desirable for a fast fading environment. Reduced rank techniques reduce the number of taps to be adaptively tracked by projecting the received signal vector onto a lower dimensional subspace [6], [7], [9], [10]. Let D be the resultant lower dimension, where $D < N$, the projection is

$$\tilde{\boldsymbol{y}}(m) = \boldsymbol{S}_D^H \boldsymbol{y}(m) \tag{6}$$

where $\boldsymbol{S}_D$ is the $N \times D$ projection matrix, and the D-dimensional signal is denoted by a "tilde" as in [6]. The vector $\tilde{\boldsymbol{y}}(m)$ is then the input to a D-dimensional linear estimator. When the MMSE criterion is applied, the optimum coefficients for the D-dimensional space are given by

$$\tilde{\boldsymbol{c}}_{mmse} = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{p}} \tag{7}$$

where $\tilde{\boldsymbol{R}} = \boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D$ and $\tilde{\boldsymbol{p}} = \boldsymbol{S}_D^H \boldsymbol{p}$.

We next review the three reduced rank methods to be considered herein: the MWF algorithm [6], the AVF algorithm [10], and the CH theorem based algorithm [9]. The MWF algorithm for DS-CDMA was presented in [6], as shown in Fig. 1 for a four-stage implementation. The equivalent projection matrix is given by

$$\begin{aligned}\boldsymbol{S}_{mw,D} &= [\boldsymbol{g}_{mw,1} \quad \boldsymbol{g}_{mw,2} \quad \cdots \quad \boldsymbol{g}_{mw,D}]\\ &= \left[\boldsymbol{h}_1 \quad \boldsymbol{B}_1^H\boldsymbol{h}_2 \quad \cdots \quad \left(\prod_{j=1}^{D-1}\boldsymbol{B}_j^H\right)\boldsymbol{h}_D\right]\end{aligned} \tag{8}$$

[2]For rectangular chip waveforms

$$\boldsymbol{s}_k^+(\tau_k) = (1-\delta)\boldsymbol{s}_k^R(pT_c) + \delta\boldsymbol{s}_k^R(pT_c + T_c)$$

where p is an integer and $\delta \in [0, 1)$ such that $\tau_k = (p+\delta)T_c$ is the timing delay for user k, and

$$\boldsymbol{s}_k^R(pT_c) = [0, 0, \ldots, 0, s_k[1], s_k[2], \ldots, s_k[N-R]]^T.$$

The partial spreading code $\boldsymbol{s}_k^-$ can be expressed by $\boldsymbol{s}_k^-(NT_c - \tau_k)$, where similar definitions apply as well.

where $\boldsymbol{g}_{mw,1} \triangleq \boldsymbol{h}_1 = \boldsymbol{p}/\|\boldsymbol{p}\|$, the normalized steering vector, and $\|.\|$ is the ℓ_2 vector norm

$$\boldsymbol{g}_{mw,i} = \left(\prod_{j=1}^{i-1}\boldsymbol{B}_j^H\right)\boldsymbol{h}_i, \; i \ge 2$$

where

$$\prod_{j=1}^{i-1}\boldsymbol{B}_j^H = \boldsymbol{B}_1^H\boldsymbol{B}_2^H \cdots \boldsymbol{B}_{i-1}^H.$$

The matrix $\boldsymbol{B}_i$ is an $(N-i) \times (N-i+1)$ blocking matrix, i.e., $\boldsymbol{B}_i\boldsymbol{h}_i = 0$. The vector $\boldsymbol{h}_i$ is the normalized correlation vector $E\{a_{i-1}(m)\boldsymbol{y}_{i-1}(m)\}$, where $\boldsymbol{y}_i(m) = \boldsymbol{B}_i\boldsymbol{y}_{i-1}(m)$ with $\boldsymbol{y}_0(m) = \boldsymbol{y}(m)$, and $a_i(m) = \boldsymbol{h}_i^H\boldsymbol{y}_{i-1}(m)$ with $a_0(m) = b_1(m)$ [6]. The coefficients w_i, $i = 1, \ldots, D$ are chosen based on the MMSE criterion.

The AVF algorithm in [10] is based on the optimization of the cross correlation of the outputs of the reference filter and auxiliary filters. A recursive conditional optimization of the auxiliary vectors was presented in [10] and the optimization procedure results in the following projection matrix:

$$\boldsymbol{S}_{av,D} = [\boldsymbol{g}_{av,1} \quad \boldsymbol{g}_{av,2} \quad \cdots \quad \boldsymbol{g}_{av,D}] \tag{9}$$

where $\boldsymbol{g}_{av,1}^H$ is equal to the normalized correlation vector $\boldsymbol{h}_1$, and $\boldsymbol{g}_{av,i}$, $i = 2, \ldots, D$ are auxiliary vectors, defined by [10]

$$\boldsymbol{g}_{av,i+1} = \frac{\boldsymbol{R}\boldsymbol{g}_{av,i}^{Eq} - \sum\limits_{j=1}^{i}\boldsymbol{g}_{av,j}\left(\boldsymbol{g}_{av,j}^H\boldsymbol{R}\boldsymbol{g}_{av,i}^{Eq}\right)}{\left\|\boldsymbol{R}\boldsymbol{g}_{av,i}^{Eq} - \sum\limits_{j=1}^{i}\boldsymbol{g}_{av,j}\left(\boldsymbol{g}_{av,j}^H\boldsymbol{R}\boldsymbol{g}_{av,i}^{Eq}\right)\right\|} \tag{10}$$

where $\boldsymbol{g}_{av,i}^{Eq} = \boldsymbol{g}_{av,1} - \sum_{j=2}^{i} c_j\boldsymbol{g}_{av,j}$, and c_j, $j = 2, \ldots, i$ are the optimized constants given by [10]

$$c_{j+1} = \frac{\boldsymbol{g}_{av,j+1}^H\boldsymbol{R}\boldsymbol{g}_{av,j}^{Eq}}{\boldsymbol{g}_{av,j+1}^H\boldsymbol{R}\boldsymbol{g}_{av,j+1}}.$$

Notice that the auxiliary vectors $\boldsymbol{g}_{av,i}$, $i = 1, \ldots, D$ are restricted to be orthonormal vectors.

Based on the CH theorem, the inverse of $\boldsymbol{R}$ can be expressed by its $(N-1)$-order polynomial expansions [9]. The Dth-order approximation of $\boldsymbol{R}^{-1}$ can thus be obtained by the $(D-1)$-order polynomials of $\boldsymbol{R}$. Equivalently, the projection matrix can be written as[3]

$$\begin{aligned}\boldsymbol{S}_{ch,D} &= [\boldsymbol{g}_{ch,1} \quad \boldsymbol{g}_{ch,2} \quad \cdots \quad \boldsymbol{g}_{ch,D}]\\ &= \left[\boldsymbol{h}_1 \quad \boldsymbol{R}\boldsymbol{h}_1 \quad \cdots \quad \boldsymbol{R}^{D-1}\boldsymbol{h}_1\right].\end{aligned} \tag{11}$$

[3]The polynomial representation of $\boldsymbol{R}^{-1}$ in [9] uses the correlation matrix of the spreading-code matrix. It is straightforward to extend the polynomial representations of $\boldsymbol{R}^{-1}$ using the correlation matrix of the received signal. The two representations can be shown equivalent.

The vector $\boldsymbol{h}_1$ is the one defined above.

IV. Simplification of the AVF Algorithm

In the AVF method, the second auxiliary vector $\boldsymbol{g}_{av,2}$ is chosen to be [10]

$$\boldsymbol{g}_{av,2} = \arg\max_{\boldsymbol{g}_{av,2}} \left| \boldsymbol{g}_{av,2}^H \boldsymbol{R} \boldsymbol{g}_{av,1} \right| \tag{12}$$

subject to $\boldsymbol{g}_{av,2}^H \boldsymbol{g}_{av,1} = 0$ and $\|\boldsymbol{g}_{av,2}\| = 1$, while $\boldsymbol{g}_{av,i+1}$, $i = 2, \ldots, D-1$ are recursively optimized by

$$\boldsymbol{g}_{av,i+1} = \arg\max_{\boldsymbol{g}_{av,i+1}} \left| \left(\boldsymbol{g}_{av,1} - \sum_{j=2}^{i} c_j \boldsymbol{g}_{av,j} \right)^H \boldsymbol{R} \boldsymbol{g}_{av,i+1} \right| \tag{13}$$

subject to $\boldsymbol{g}_{av,i+1}^H \boldsymbol{g}_{av,j} = 0$, $j = 1, \ldots, i$, and $\|\boldsymbol{g}_{av,i+1}\| = 1$, with solutions given by (10). It can be seen that the derivation and the solution of $\boldsymbol{g}_{av,i+1}$ are rather involved. There is also the need to determine the optimized constants c_j, $j = 2, \ldots, i$. We can show that the AVF algorithm can be simplified, as evidenced by the following proposition.

Proposition 1: For any given normalized vector $\boldsymbol{v}_0$ and $\boldsymbol{R}$, let vector $\boldsymbol{v}_1$ be determined by

$$\boldsymbol{v}_1 = \arg\max_{\boldsymbol{v}_1} \left| \boldsymbol{v}_1^H \boldsymbol{R} \boldsymbol{v}_0 \right|$$

subject to $\boldsymbol{v}_1^H \boldsymbol{v}_1 = 1$, $\boldsymbol{v}_1^H \boldsymbol{v}_0 = 0$, and real $\boldsymbol{v}_1^H \boldsymbol{R} \boldsymbol{v}_0$. It can be shown that [10]

$$\boldsymbol{v}_1 = \frac{\boldsymbol{R}\boldsymbol{v}_0 - \boldsymbol{v}_0 \left(\boldsymbol{v}_0^H \boldsymbol{R} \boldsymbol{v}_0 \right)}{\|\boldsymbol{R}\boldsymbol{v}_0 - \boldsymbol{v}_0 (\boldsymbol{v}_0^H \boldsymbol{R} \boldsymbol{v}_0)\|}. \tag{14}$$

Now let $\boldsymbol{v}_2$ be another vector such that $\boldsymbol{v}_2^H \boldsymbol{v}_2 = 1$, $\boldsymbol{v}_2^H \boldsymbol{v}_0 = 0$, and $\boldsymbol{v}_2^H \boldsymbol{v}_1 = 0$. Then we have

$$\boldsymbol{v}_2^H \boldsymbol{R} \boldsymbol{v}_0 = 0. \tag{15}$$

Proof: Let $\boldsymbol{R}\boldsymbol{v}_0$ be decomposed into

$$\boldsymbol{R}\boldsymbol{v}_0 = \check{c}_0 \boldsymbol{v}_0 + \check{c}_1 \boldsymbol{v}_0^\perp \tag{16}$$

where $\check{c}_0$ and $\check{c}_1$ are real constants, $\boldsymbol{v}_0^\perp$ is a normalized vector orthogonal to $\boldsymbol{v}_0$. Clearly, $\check{c}_0 = \boldsymbol{v}_0^H \boldsymbol{R} \boldsymbol{v}_0$ and $\boldsymbol{v}_0^\perp = [\boldsymbol{R}\boldsymbol{v}_0 - \boldsymbol{v}_0(\boldsymbol{v}_0^H \boldsymbol{R} \boldsymbol{v}_0)]/\check{c}_1$ (where $\check{c}_1 = 0$ corresponds to the trivial case). The scalar $\check{c}_1$ is the normalizing factor which yields $\|\boldsymbol{v}_0^\perp\| = 1$. Hence, $\boldsymbol{v}_1 = \boldsymbol{v}_0^\perp$. Now

$$\boldsymbol{v}_2^H \boldsymbol{R} \boldsymbol{v}_0 = \check{c}_0 \boldsymbol{v}_2^H \boldsymbol{v}_0 + \check{c}_1 \boldsymbol{v}_2^H \boldsymbol{v}_1 = 0. \tag{17}$$

This completes the proof. □

For the AVF algorithm, from (13), when $i = 2$

$$\boldsymbol{g}_{av,3} = \arg\max_{\boldsymbol{g}_{av,3}} \left| (\boldsymbol{g}_{av,1} - c_2 \boldsymbol{g}_{av,2})^H \boldsymbol{R} \boldsymbol{g}_{av,3} \right| \tag{18}$$

subject to $\boldsymbol{g}_{av,3}^H \boldsymbol{g}_{av,j} = 0$, $j = 1, 2$, and $\|\boldsymbol{g}_{av,3}\| = 1$. From Proposition 1 and (12), $\boldsymbol{g}_{av,1}^H \boldsymbol{R} \boldsymbol{g}_{av,3} = 0$. We have

$$\boldsymbol{g}_{av,3} = \arg\max_{\boldsymbol{g}_{av,3}} \left| \boldsymbol{g}_{av,2}^H \boldsymbol{R} \boldsymbol{g}_{av,3} \right|. \tag{19}$$

When $i = 3$, (13) becomes

$$\boldsymbol{g}_{av,4} = \arg\max_{\boldsymbol{g}_{av,4}} \left| (\boldsymbol{g}_{av,1} - c_2 \boldsymbol{g}_{av,2} - c_3 \boldsymbol{g}_{av,3})^H \boldsymbol{R} \boldsymbol{g}_{av,4} \right| \tag{20}$$

subject to $\boldsymbol{g}_{av,4}^H \boldsymbol{g}_{av,j} = 0$, $j = 1, 2, 3$, and $\|\boldsymbol{g}_{av,4}\| = 1$. Clearly, $\boldsymbol{g}_{av,4}^H \boldsymbol{R} \boldsymbol{g}_{av,1} = 0$. Now since $\boldsymbol{g}_{av,3}$ is optimized from $\boldsymbol{g}_{av,2}$ as in (19), using Proposition 1 again, we have $\boldsymbol{g}_{av,4}^H \boldsymbol{R} \boldsymbol{g}_{av,2} = 0$, the above criterion becomes

$$\boldsymbol{g}_{av,4} = \arg\max_{\boldsymbol{g}_{av,4}} \left| \boldsymbol{g}_{av,3}^H \boldsymbol{R} \boldsymbol{g}_{av,4} \right|. \tag{21}$$

Following the same approach, we obtain an equivalent, but more efficient way to derive the auxiliary vectors for the AVF algorithm as follows:

$$\boldsymbol{g}_{av,i+1} = \arg\max_{\boldsymbol{g}_{av,i+1}} \left| \boldsymbol{g}_{av,i}^H \boldsymbol{R} \boldsymbol{g}_{av,i+1} \right| \tag{22}$$

subject to $\boldsymbol{g}_{av,i+1}^H \boldsymbol{g}_{av,i+1} = 1$ and $\boldsymbol{g}_{av,i+1}^H \boldsymbol{g}_{av,j} = 0$, $j = 1, \ldots, i$. Notice that with our implementation, we do not need to calculate the optimal constants c_i, $i = 2, \ldots, D$ as in [10].

In addition, we have

$$\boldsymbol{g}_{av,i}^H \boldsymbol{R} \boldsymbol{g}_{av,j} = 0, \text{ if } |i-j| > 1. \tag{23}$$

Notice that this is consistent with the fact that the matrix $\boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D$ for the MWF algorithm is tridiagonal, as shown in [4]. This suggests the equivalence of the MWF algorithm and the AVF algorithm, which we will show in the next section.

It can be shown that the simplified solution for $\boldsymbol{g}_{av,i+1}$ is

$$\begin{aligned} \boldsymbol{g}_{av,i+1} &= \frac{\boldsymbol{R}\boldsymbol{g}_{av,i} - \sum_{j=1}^{i} \boldsymbol{g}_{av,j} \left(\boldsymbol{g}_{av,j}^H \boldsymbol{R} \boldsymbol{g}_{av,i} \right)}{\left\| \boldsymbol{R}\boldsymbol{g}_{av,i} - \sum_{j=1}^{i} \boldsymbol{g}_{av,j} \left(\boldsymbol{g}_{av,j}^H \boldsymbol{R} \boldsymbol{g}_{av,i} \right) \right\|} \\ &= \frac{\boldsymbol{R}\boldsymbol{g}_{av,i} - \sum_{j=i-1}^{i} \boldsymbol{g}_{av,j} \left(\boldsymbol{g}_{av,j}^H \boldsymbol{R} \boldsymbol{g}_{av,i} \right)}{\left\| \boldsymbol{R}\boldsymbol{g}_{av,i} - \sum_{j=i-1}^{i} \boldsymbol{g}_{av,j} \left(\boldsymbol{g}_{av,j}^H \boldsymbol{R} \boldsymbol{g}_{av,i} \right) \right\|} \end{aligned} \tag{24}$$

where (23) has been used. That is, in deriving $\boldsymbol{g}_{av,i+1}$, we need to focus on vectors $\boldsymbol{g}_{av,i}$ and $\boldsymbol{g}_{av,i-1}$ only. The auxiliary vectors $\boldsymbol{g}_{av,j}$, $j = 1, \ldots, i-2$ will not have an effect on optimizing $\boldsymbol{g}_{av,i+1}$.

V. Equivalence of the Reduced Rank MMSE Filtering Methods

Now let us investigate the MWF algorithm for DS-CDMA [4], [6], as shown in Fig. 1. The blocking matrices $\boldsymbol{B}_i$, $i = 1, \ldots, D-1$ in the MWF algorithm are used to annihilate the signal components in the direction of $\boldsymbol{h}_i$, $i = 1, \ldots, D-1$. The choice of the blocking matrices is not unique and the choice may affect the performance of the MWF algorithm for a specific situation. It is thus interesting to study the effects of choosing $\boldsymbol{B}_i$ and, if possible, to obtain optimal blocking matrices.

In the Appendix we show that the MWF algorithm is not dependent on the choice of blocking matrices $\boldsymbol{B}_i$ when, in addition to satisfying $\boldsymbol{B}_i \boldsymbol{h}_i = 0$ as in [6], the rows of $\boldsymbol{B}_i$ are orthonormal, i.e.,

$$\boldsymbol{B}_i \boldsymbol{B}_i^H = \boldsymbol{I}_{N-i}. \tag{25}$$

In other words, for each $i \in \{1, \ldots, D-1\}$, the rows of $\boldsymbol{B}_i$ are constrained to be an orthonormal basis for the nullspace of $\boldsymbol{h}_i$. As we will see, the orthonormality constraint on $\boldsymbol{B}_i$ leads to a direct connection between the AVF and MWF algorithms under which the two algorithms are equivalent.

Proposition 2: When, for each $i \in \{1, \ldots, D-1\}$, the rows of $\boldsymbol{B}_i$ form an orthonormal basis for the nullspace of $\boldsymbol{h}_i$, the projection matrix $\boldsymbol{S}_{mw,D}$ for the D stage MWF algorithm is not a function of $\{\boldsymbol{B}_i, i = 1, \ldots, D-1\}$. Consequently, the performance of the MWF algorithm is not a function of $\{\boldsymbol{B}_i, i = 1, \ldots, D-1\}$.

Proof: See the Appendix. □

Note that although the MWF algorithm is independent of a particular choice of the orthonormal basis for the nullspace of $\boldsymbol{h}_i$, $i = 1, \ldots, D-1$, it inherently depends on those nullspaces as the MWF algorithm is based on a sequence of projections onto those nullspaces.

Corollary 2.1: When the blocking matrices $\boldsymbol{B}_i$, $i = 1, \ldots, D-1$, are constructed with orthonormal rows

$$\boldsymbol{S}_{mw,D} = \boldsymbol{S}_{av,D}. \tag{26}$$

That is, the MWF method is equivalent to the AVF method.

Corollary 2.2: When the blocking matrices $\boldsymbol{B}_i$, $i = 1, \ldots, D-1$, are constructed with orthonormal rows

$$\operatorname{span}\left\{\boldsymbol{g}_{mw,1}, \boldsymbol{g}_{mw,2}, \ldots, \boldsymbol{g}_{mw,D}\right\} = \operatorname{span}\left\{\boldsymbol{h}_1, \boldsymbol{R}\boldsymbol{h}_1, \ldots \boldsymbol{R}^{D-1}\boldsymbol{h}_1\right\}. \tag{27}$$

Proof: We show the desired result via induction. Obviously, the first column of $\boldsymbol{S}_{ch,D}$ in (11) is equal to the first column of $\boldsymbol{S}_{mw,D}$ in (9). For the second column of $\boldsymbol{S}_{mw,D}$, (34) shows that

$$\operatorname{span}\left\{\boldsymbol{B}_1^H \boldsymbol{h}_2\right\} \in \operatorname{span}\{\boldsymbol{h}_1, \boldsymbol{R}\boldsymbol{h}_1\}.$$

Now assume that the ith column of $\boldsymbol{S}_{mw,D}$ satisfies

$$\left(\prod_{j=1}^{i-1} \boldsymbol{B}_j^H\right) \boldsymbol{h}_i \in \operatorname{span}\left\{\boldsymbol{h}_1, \boldsymbol{R}\boldsymbol{h}_1, \ldots, \boldsymbol{R}^{i-1}\boldsymbol{h}_1\right\}.$$

From (36), we have

$$\left(\prod_{j=1}^{i} \boldsymbol{B}_j^H\right) \boldsymbol{h}_{i+1} \in \operatorname{span}\left\{\boldsymbol{h}_1, \boldsymbol{R}\boldsymbol{h}_1, \ldots, \boldsymbol{R}^{i}\boldsymbol{h}_1\right\}. \tag{28}$$

Since $\boldsymbol{g}_{mw,1}, \boldsymbol{g}_{mw,2}, \ldots, \boldsymbol{g}_{mw,D}$ are linearly independent, we have

$$\operatorname{span}\left\{\boldsymbol{g}_{mw,1}, \boldsymbol{g}_{mw,2}, \ldots, \boldsymbol{g}_{mw,D}\right\} = \operatorname{span}\left\{\boldsymbol{h}_1, \boldsymbol{R}\boldsymbol{h}_1, \ldots \boldsymbol{R}^{D-1}\boldsymbol{h}_1\right\}. \tag{29}$$

□

We note that a similar proof of Corollary 2.2 can also be found in [6]. However, our proof of the equivalence of the MWF algorithm and the CH based algorithm is a byproduct from the proof for Proposition 2.

Therefore, the MWF method, the AVF method, and the CH method of [9] are equivalent to each other.

The equivalence of the three algorithms can be interpreted as follows. Both the AVF and the MWF algorithms are based on choosing the additional projection vector to maximize the correlation between the output of this projection vector and the output from previous stages. The projection vectors for the two algorithms are just the orthonormalized versions of the projection vectors for the CH algorithm of [9], in which each additional stage introduces the new vector information inherent in $\boldsymbol{R}^{i+1}\boldsymbol{h}_1$. However, although all three algorithms allow adaptive implementations, the MWF algorithm has more flexibility than both the AVF and the CH algorithm of [9] due to its implementation structure. Several adaptive implementations for the MWF algorithms have been proposed in [6]. In addition, during simulation studies, we have observed that the MSE solution for the CH algorithm of [9] often has numerical stability problems when the number of stages is large, say, $D > 5$.

VI. Conclusion

We compared several reduced rank detection schemes for DS-CDMA communication systems. The AVF algorithm [10] has been simplified through a key observation on the construction of auxiliary vectors. After simplification, it is shown that the AVF algorithm is equivalent to the MWF algorithm of [6]. Furthermore, these schemes have been shown to be equivalent to the multistage linear receiver scheme based on the CH theorem when the MMSE criterion is applied to the reduced dimensional space of the received signal.

Appendix
Proof of Proposition 2

Let $\boldsymbol{B}$ denote the set $\{\boldsymbol{B}_i, i = 1, \ldots, D-1\}$. Now let us prove $\boldsymbol{S}_{mw,D}$ is not dependent on $\boldsymbol{B}$ by using the method of induction. Notice that when $i > j$

$$\begin{aligned}\boldsymbol{g}_{mw,i}^H \boldsymbol{g}_{mw,j} &= \boldsymbol{h}_i^H \left(\prod_{l=i-1}^{1} \boldsymbol{B}_l\right)\left(\prod_{l=1}^{j-1} \boldsymbol{B}_l^H\right)\boldsymbol{h}_j \\ &= \boldsymbol{h}_i^H \left(\prod_{l=i-1}^{j+1} \boldsymbol{B}_l\right)\boldsymbol{B}_j \boldsymbol{h}_j = 0\end{aligned} \tag{30}$$

where $\boldsymbol{B}_j \boldsymbol{B}_j^H = \boldsymbol{I}_{N-j}$ and $\boldsymbol{B}_j \boldsymbol{h}_j = 0$ have been used. Similarly, $\boldsymbol{g}_{mw,i}^H \boldsymbol{g}_{mw,j} = 0$ for $i < j$. When $i = j$

$$\begin{aligned}\boldsymbol{g}_{mw,i}^H \boldsymbol{g}_{mw,i} &= \boldsymbol{h}_i^H \left(\prod_{l=i-1}^{1} \boldsymbol{B}_l\right)\left(\prod_{l=1}^{i-1} \boldsymbol{B}_l^H\right)\boldsymbol{h}_i \\ &= \boldsymbol{h}_i^H \boldsymbol{h}_i = 1.\end{aligned} \tag{31}$$

Therefore, we have

$$\boldsymbol{g}_{mw,i}^H \boldsymbol{g}_{mw,j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases} \tag{32}$$

That is, the columns of $\boldsymbol{S}_{mw,D}$ matrix are orthonormal vectors.

From $\boldsymbol{h}_i^H \boldsymbol{h}_i = 1$, if we let $\boldsymbol{T}_i^H = [\boldsymbol{h}_i, \boldsymbol{B}_i^H]$, we have

$$\boldsymbol{T}_i^H \boldsymbol{T}_i = \boldsymbol{T}_i \boldsymbol{T}_i^H = \boldsymbol{I}_{N-i+1}$$

and

$$\boldsymbol{B}_i^H \boldsymbol{B}_i = \boldsymbol{I}_{N-i+1} - \boldsymbol{h}_i \boldsymbol{h}_i^H. \tag{33}$$

Clearly, $\boldsymbol{h}_1 = E\{b_1(m)\boldsymbol{y}(m)\}/\delta_1 = r_{b_1(m)\boldsymbol{y}(m)}/\delta_1$ and the normalization factor $\delta_1 = \|r_{b_1(m)\boldsymbol{y}(m)}\|$ are not functions of $\boldsymbol{B}$. Now $\boldsymbol{g}_{mw,2}^H$ is given by

$$\begin{aligned}\boldsymbol{h}_2^H \boldsymbol{B}_1 &= \left(\frac{\boldsymbol{B}_1 \boldsymbol{R}_{\boldsymbol{y}_0(m)} \boldsymbol{h}_1}{\delta_2}\right)^H \boldsymbol{B}_1 \\ &= \boldsymbol{h}_1^H \boldsymbol{R}\left(\boldsymbol{I}_N - \boldsymbol{h}_1 \boldsymbol{h}_1^H\right)/\delta_2\end{aligned} \tag{34}$$

where $\boldsymbol{R}_{\boldsymbol{y}_0(m)} = \boldsymbol{R}_{\boldsymbol{y}(m)} = \boldsymbol{R}$

$$\begin{aligned}\delta_2 &= \sqrt{r_{a_1(m)\boldsymbol{y}_1(m)}^H r_{a_1(m)\boldsymbol{y}_1(m)}} \\ &= \sqrt{\boldsymbol{h}_1^H \boldsymbol{R}\left(\boldsymbol{I}_N - \boldsymbol{h}_1 \boldsymbol{h}_1^H\right)\boldsymbol{R}\boldsymbol{h}_1}\end{aligned} \tag{35}$$

where $\boldsymbol{B}_1^H \boldsymbol{B}_1 = \boldsymbol{I}_N - \boldsymbol{h}_1 \boldsymbol{h}_1^H$ has been used. Clearly, $\boldsymbol{h}_2^H \boldsymbol{B}_1$ and δ_2 are not dependent on $\boldsymbol{B}$. Now assume that $\boldsymbol{g}_{mw,i}^H$ is not dependent on $\boldsymbol{B}$, $\boldsymbol{g}_{mw,i+1}^H$ becomes

$$\begin{aligned}\delta_{i+1}&\boldsymbol{g}_{mw,i+1}^H \\ &= \delta_{i+1} \boldsymbol{h}_{i+1}^H \left(\prod_{j=i}^{1} \boldsymbol{B}_j\right) \\ &= \boldsymbol{h}_i^H \boldsymbol{R}_{\boldsymbol{y}_{i-1}(m)} \boldsymbol{B}_i^H \boldsymbol{B}_i \left(\prod_{j=i-1}^{1} \boldsymbol{B}_j\right) \\ &= \boldsymbol{h}_i^H \boldsymbol{R}_{\boldsymbol{y}_{i-1}(m)} \left(\boldsymbol{I}_{N-i+1} - \boldsymbol{h}_i \boldsymbol{h}_i^H\right)\left(\prod_{j=i-1}^{1} \boldsymbol{B}_j\right) \\ &= \boldsymbol{h}_i^H \boldsymbol{R}_{\boldsymbol{y}_{i-1}(m)} \left(\prod_{j=i-1}^{1} \boldsymbol{B}_j\right) - \left(\boldsymbol{h}_i^H \boldsymbol{R}_{\boldsymbol{y}_{i-1}(m)} \boldsymbol{h}_i\right)\boldsymbol{g}_{mw,i}^H\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{h}_i^H \boldsymbol{B}_{i-1} \boldsymbol{R}_{\boldsymbol{y}_{i-2}(m)} \boldsymbol{B}_{i-1}^H \boldsymbol{B}_{i-1} \left(\prod_{j=i-2}^{1} \boldsymbol{B}_j \right) \\
&\quad - \left(\boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,i} \right) \boldsymbol{g}_{mw,\,i}^H = \cdots \\
&= \boldsymbol{h}_i^H \left(\prod_{j=i-1}^{1} \boldsymbol{B}_j \right) \boldsymbol{R}_{\boldsymbol{y}_0(m)} - \sum_{j=1}^{i} \left(\boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,j} \right) \boldsymbol{g}_{mw,\,j}^H \\
&= \boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} - \sum_{j=1}^{i} \left(\boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,j} \right) \boldsymbol{g}_{mw,\,j}^H \\
&= \boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} - \sum_{j=i-1}^{i} \left(\boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,j} \right) \boldsymbol{g}_{mw,\,j}^H \qquad (36)
\end{aligned}
$$

where we have used $\boldsymbol{h}_{i+1}^H = \boldsymbol{h}_i^H \boldsymbol{R}_{\boldsymbol{y}_{i-1}(m)} \boldsymbol{B}_i^H / \delta_{i+1}$, (33)

$$
\boldsymbol{R}_{\boldsymbol{y}_l(m)} = \boldsymbol{B}_l \boldsymbol{R}_{\boldsymbol{y}_{l-1}(m)} \boldsymbol{B}_l^H, \qquad l = 1, \ldots, i-1
$$

and

$$
\begin{aligned}
\boldsymbol{h}_i^H \boldsymbol{R}_{\boldsymbol{y}_{i-1}(m)} \boldsymbol{h}_i &= \boldsymbol{h}_i^H \left(\prod_{j=i-1}^{1} \boldsymbol{B}_j \right) \boldsymbol{R}_{\boldsymbol{y}_0(m)} \left(\prod_{j=1}^{i-1} \boldsymbol{B}_j^H \right) \boldsymbol{h}_i \\
&= \boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,i}
\end{aligned}
$$

and the fact that

$$
\boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,j}^H = 0, \qquad \text{if } |i-j| > 1.
$$

Notice that $\boldsymbol{h}_{i+1}^H \prod_{j=i}^{1} \boldsymbol{B}_j \prod_{j=1}^{i} \boldsymbol{B}_j^H \boldsymbol{h}_{i+1} = 1$, we have

$$
\begin{aligned}
\delta_{i+1} &= \left\| \boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} - \sum_{j=i-1}^{i} (\boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,j}) \boldsymbol{g}_{mw,\,j}^H \right\| \\
&= \sqrt{ \boldsymbol{g}_{mw,\,i}^H \boldsymbol{R}^2 \boldsymbol{g}_{mw,\,i} - \sum_{j=i-1}^{i} \left| \boldsymbol{g}_{mw,\,i}^H \boldsymbol{R} \boldsymbol{g}_{mw,\,j} \right|^2 }. \qquad (37)
\end{aligned}
$$

Since the first i columns of $\boldsymbol{S}_{mw,\,D}$ are not dependent on $\boldsymbol{B}$, we have that $\boldsymbol{h}_{i+1} \left(\prod_{j=i}^{1} \boldsymbol{B}_j \right)$ and δ_{i+1} given in the above two equations are not functions of $\boldsymbol{B}$. Therefore, $\boldsymbol{S}_{mw,\,D}$ and δ_i, $i = 1, \ldots, D$ are not dependent on $\boldsymbol{B}$.

This completes our proof. □

REFERENCES

[1] M. Abdulrahman, A. U. H. Sheikh, and D. D. Falconer, "Adaptive detection of DS/CDMA signals in fading channels," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 698–706, May 1994.

[2] W. Chen and U. Mitra, "An improved blind adaptive MMSE receiver for fast fading DS-CDMA channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1531–1543, Aug. 2001.

[3] M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*. Boston, MA: Nijhoff, 1979.

[4] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2943–2959, Nov. 1998.

[5] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[6] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multi-stage Wiener filter," *IEEE Trans. Commun.*, vol. 50, pp. 986–994, June 2002.

[7] D. F. Marshall, W. K. Jenkins, and J. J. Murphy, "The use of orthogonal transforms for improving performance of adaptive filters," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 474–484, Apr. 1989.

[8] S. L. Miller, M. L. Honig, and L. B. Milstein, "Performance analysis of MMSE receivers for DS-CDMA in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 48, pp. 1919–1929, Nov. 2000.

[9] S. Moshavi, E. G. Kanterakis, and D. L. Schilling, "Multistage linear receivers for DS-CDMA systems," *Int. J. Wireless Inform. Networks*, vol. 3, no. 1, pp. 1–17, 1996.

[10] D. A. Pados and S. N. Batalama, "Joint space–time auxiliary-vector filtering for DS/CDMA systems with antenna arrays," *IEEE Trans. Commun.*, vol. 47, pp. 1406–1415, Sept. 1999.

[11] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[12] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detectors for coherent multi-user communications," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 683–690, May 1990.

Optimal Versus Randomized Search of Fixed Length Binary Words

Helmut Prodinger and Wojciech Szpankowski, *Senior Member, IEEE*

***Abstract*—We consider the search problem in which one finds a binary word among m words chosen randomly from the set of all words of fixed length n. It is well known that the optimal search is equivalent to the Huffman coding that requires on average $\log_2 m$ bits to be checked *plus* a small additional cost called the average *redundancy*. The latter is an oscillating function of m and is bounded between zero and $1 - (1 + \ln \ln 2)/\ln 2 \approx 0.0860713320$. As a matter of fact, it is known that finding the optimal strategy for this problem is NP-hard. We propose here several simple randomized search strategies leading, respectively, to the following average redundancies: 1.332746177, 0.6113986565, 0.4310617764, and 0.332746177, plus some small oscillations that we precisely characterize. These results should be compared to the optimal, but NP-hard, search algorithm. Our findings extend and make more precise recent results of Fedotov and Ryabko.**

***Index Terms*—Generating functions, Huffman code, PATRICIA trie, search problem, Shannon source coding theorem, tries.**

I. INTRODUCTION

A combinatorial *search problem* can be defined as follows: Given a set $\mathcal{W} = \{w_1, w_2, \ldots, w_m\}$ of m words over a (binary) alphabet Σ, design a sequence of tests that successfully find the word $w^* \in \mathcal{W}$ being sought (cf. [1]). The prime goal of the *optimal* search is to find the sought word w^* with the smallest maximum or average search time. It is well known (cf. [1]) that the problem of determining a sequential strategy with the minimum average search time is equivalent to the information-theoretic problem of minimizing the average codeword length of a certain prefix code, that is, constructing the optimal Huffman code. We shall use this equivalence between the search problem and prefix codes throughout this correspondence.

Manuscript received May 28, 2001; revised March 14, 2002. The work of H. Prodinger was supported by The John Knopfmacher Centre for Applicable Analysis and Number Theory. The work of W. Szpankowski was supported by the National Science Foundation under Grants CCR-9804760 and CCR-0208709, and Contract 1419991431A from sponsors of CERIAS at Purdue University. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Lausanne, Switzerland, June/July 2002.

H. Prodinger is with the School of Mathematics, University of the Witwatersrand, P.O. Wits 2050, Johannesburg, South Africa (e-mail: helmut@staff.ms.wits.ac.za).

W. Szpankowski is with the Department of Computer Science, Purdue University, West Lafayette, IN 47907 USA (e-mail: spa@cs.purdue.edu).

Communicated by M. Weinberger, Associate Editor for Source Coding.

Publisher Item Identifier 10.1109/TIT.2002.801478.