Then, for two arbitrary integers i_1, i_2 in $1 \leq i_1, i_2 \leq \lfloor \frac{n_T}{2} \rfloor$, we get

$$\begin{aligned} |w_{i_1} - w_{i_2}| &= \frac{1}{2} \left| (v'_{i_1} - v'_{i_2}) + (v'_{n_T - i_1 + 1} - v'_{n_T - i_2 + 1}) \right| \\ &\leq \frac{1}{2} \left(|v'_{i_1} - v'_{i_2}| + |v'_{n_T - i_1 + 1} - v'_{n_T - i_2 + 1}| \right) \\ &\leq \frac{1}{2} \left\{ \left(v'_{\lfloor \frac{n_T}{2} \rfloor} - v'_1 \right) + \left(v'_{n_T} - v'_{\lfloor \frac{n_T}{2} \rfloor} \right) \right\} \\ &= \frac{1}{2} \left(v'_{n_T} - v'_1 \right). \end{aligned}$$

If we perform the above reordering and averaging process repeatedly, then the difference of two arbitrary elements in the set eventually reduces to zero. This completes the proof.

E. Proof of (23)

$E[P_p(\Delta_p|\boldsymbol{\alpha})]$ can be regarded as the PEP when all the eigenvalues are equal to $(\Delta_p)^{1/n_T}$. Substituting each λ_i by $(\Delta_p)^{1/n_T}$ in (5), we get

$$E[P_p(\Delta_p|\boldsymbol{\alpha})] = \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{(\Delta_p)^{1/n_T} \gamma_s}{4 \sin^2 \theta} \right)^{-n_T n_R} d\theta. \quad (41)$$

In addition, $J_m(c)$ in (8) can be rearranged into [12]

$$J_m(c) = \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{c}{\sin^2 \theta} \right)^{-m} d\theta. \quad (42)$$

Therefore, combining (41), (42), and (7), we can easily show that $E[P_p(\Delta_p|\boldsymbol{\alpha})] = P_B(\Delta_p)$.

REFERENCES

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
[2] S. Siwamogsatham, M. P. Fitz, and J. H. Grimm, "A new view of performance analysis of transmit diversity schemes in correlated Rayleigh fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 950–956, Apr. 2002.
[3] D. K. Aktas and M. P. Fitz, "Computing the distance spectrum of space-time trellis codes," in *Proc. WCNC'2000*, 2000, pp. 51–55.
[4] ——, "The distance spectrum of space-time trellis coded modulations in quasi-static Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 3335–3344, Dec. 2003.
[5] A. Stefanov and T. M. Duman, "Performance bounds for space-time trellis codes," in *Proc. IEEE Int. Symp. Information Theory (ISIT'2001)*, Washington, DC, June 2001, p. 82.
[6] H. Bouzekri and S. L. Miller, "Performance analysis of space-time trellis codes over quasistatic fading channels," in *Proc. 39th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2001.
[7] M.-K. Byun and B. G. Lee, "New bounds of pairwise error probability for space-time codes in Rayleigh fading channels," *IEEE Trans. Commun.*, to be published.
[8] E. Malkamäki and H. Leib, "Evaluating the performance of convolutional codes over block fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1643–1646, July 1999.
[9] R. G. Gallager, *Low Density Parity Check Codes*. Camabridge, MA: MIT Press, 1963.
[10] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.
[11] G. D. Forney Jr, "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, Sept. 1991.
[12] M.-S. Alouini and A. J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1324–1334, Sept. 1999.
[13] J. Grimm, M. P. Fitz, and J. V. Krogmeier, "Further results on space-time coding for Rayleigh fading," in *Proc. Allerton Conf.'98*, 1998, pp. 391–400.
[14] W. A. Gardner, *Introduction to Random Processes*. New York: McGraw-Hill, 1990.

Detection of Constrained Subspace Signals in Additive Infinite-Dimensional Interference and Noise

John A. Gubner, *Member, IEEE,* and Louis L. Scharf, *Fellow, IEEE*

***Abstract*—The detection of constrained subspace signals in additive infinite-dimensional interference and noise is motivated by consideration of multipath-Doppler channels subject to interference from partially overlapping frequency bands of other sources. Since the interference lies in an infinite-dimensional subspace, the standard method of projecting onto the subspace containing the signal plus interference does not yield a finite-dimensional detection problem. However, an alternative approach may be used to extract the appropriate finite-dimensional problem. Moreover, an energy constraint may be imposed on the desired signal. The generalized likelihood-ratio receiver for this problem is obtained, and expressions for its average probability of error are given.**

***Index Terms*—Energy constraints, magnitude constraints, matched subspace detector.**

I. INTRODUCTION

Consider a communication system in which the received waveform is

$$y(t) = a(t) + b(t) + n(t)$$

where a contains the desired signal, b is an interfering signal, and n is additive white Gaussian noise. Detectors based on waveform observations are usually derived by extracting a finite-dimensional coefficient vector based on the projection of the received waveform onto the subspace containing the desired signal as well as the interfering signal. This is straightforward if the interfering signal lies in a finite-dimensional subspace, as in the decorrelating detector for a code-division multiple-access (CDMA) system [2] or the matched subspace detector [7], [8]. Suppose, however, that the interfering signal b is known only to lie in a frequency band which partially

Manuscript received September 7, 2002; revised May 16, 2004. The work of L. L. Scharf was supported by the Office of Naval Research under Contract N00014-89-J-1070 and by National Science Foundation grants under the 1999 Wireless Initiative, ECS 9979400, and the ITR Program, ITR 0112573.

J. A. Gubner is with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706-1691 USA (e-mail: gubner@engr.wisc.edu).

L. L. Scharf is with the Departments of Electrical and Computer Engineering and Statistics, Colorado State University, Fort Collins, CO 80523-1373 USA (e-mail: scharf@engr.colostate.edu).

Communicated by V. V. Veeravalli, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2004.836674

overlaps that of the signal a. In this case, the subspace containing a and b is infinite dimensional, and the standard approach will not result in a finite-dimensional coefficient vector.

In an M-ary communication system, the sender transmits a signal $s_i(t)$ to convey message i. If s_i is band limited to W and subject to multipath, then the signal a at the receiver may be modeled as

$$a(t) = \sum_{l=0}^{D} h_l s_i(t - l/2W)$$

where D is proportional to the product of W and the channel multipath spread, and the h_l are the delay coefficients [6]. When the coefficients h_l are unknown, we can view the modulator-plus-channel as a subspace modulator. In other words, from the receiver's point of view, to send message i, the modulator transmits an arbitrary waveform from the ith signal subspace

$$\text{span}\{s_i(\cdot - l/2W), l = 0, \ldots, D\}.$$

In this context, the arbitrariness of the coefficients h_l is imposed by the channel. On the other hand, randomization of the coefficients has been proposed as an intentional part of the modulation process under the names "stochastic process shift keying" [4] and "stochastic multipulse PAM" [5].

To simplify the notation and to emphasize that the analysis here is not specific to the case of a band-limited signal subject to multipath and to overlapping-frequency-band interference, we use the generic subspace signal model

$$a(t) = \sum_{k=1}^{p_i} u_k \, a_{i,k}(t)$$

where $a_{i,1}, \ldots, a_{i,p_i}$ are linearly independent waveforms. Although the coefficients u_k are unknown, we assume that they satisfy an energy constraint; e.g., $\sum_k |u_k|^2 \leq \mathcal{E}$. Constraints of the form $\max_k |u_k| \leq \sqrt{\mathcal{E}}$ are considered in Appendix A.

To summarize, although the detection of subspace signals in subspace interference and noise has been studied previously, what is new here is the fact that there is an energy constraint on the signal coefficients and the fact that the interference subspace is allowed to be infinite-dimensional. In this context, we derive the generalized likelihood ratio detector, and we analyze its average probability of error. Thus, our results generalize all prior results on M-ary subspace detection in subspace interference and noise.

II. Problem Formulation

Let X be a complex inner product space equipped with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\|\cdot\|$. Consider the following M-ary detection problem. For $i = 1, \ldots, M$, the ith hypothesis is

$$H_i : y = a + b + n, \ \|a\|^2 \leq \mathcal{E}_i, \qquad a \in \mathcal{A}_i, \ b \in \mathcal{B} \tag{1}$$

where $\mathcal{A}_i \subset X$ is the ith **signal subspace**, $\mathcal{B} \subset X$ is the **interference subspace**, and n is zero-mean, white Gaussian noise. (Throughout the correspondence, Gaussian random vectors and processes are understood to be complex valued and proper (circularly symmetric) [3].) The notation we use subsumes the following cases.

- When $X = \mathbb{C}^N$, we assume that n is a zero-mean Gaussian random vector with covariance matrix $\sigma^2 I$; i.e., the components of n are independent random variables with zero mean and variance σ^2.
- When $X = \ell^2$, we assume that n is a sequence of independent Gaussian random variables with zero mean and variance σ^2.
- When $X = L^2$, we assume that n is a zero-mean, wide-sense-stationary Gaussian noise process with correlation function

$$R_n(t_1, t_2) := E[n(t_1)\overline{n(t_2)}] = \sigma^2 \delta(t_1 - t_2)$$

where δ is the Dirac delta function, and the overbar indicates the complex conjugate.

Remark 1: By allowing $\mathcal{E}_i$ to vary with i, the model can accommodate different subspaces subject to different amounts of attenuation as in frequency-shift keying over a frequency-selective channel.

It is convenient to generalize the constraint as follows. For each hypothesis H_i, let $a_{i,1}, \ldots, a_{i,p_i}$ be any basis for $\mathcal{A}_i$, and define the operator $A_i : \mathbb{C}^{p_i} \to X$ by

$$A_i u := \sum_{k=1}^{p_i} u_k \, a_{i,k}, \qquad u = [u_1, \ldots, u_{p_i}]' \in \mathbb{C}^{p_i}.$$

The operator A_i establishes a one-to-one correspondence between every $a \in \mathcal{A}_i$ and every $u \in \mathbb{C}^{p_i}$. This suggests the new hypothesis testing problem

$$H_i : y = A_i u + b + n, \quad \|u\|_{Q_i}^2 \leq \mathcal{E}_i, \qquad u \in \mathbb{C}^{p_i}, \ b \in \mathcal{B} \tag{2}$$

where $\|u\|_{Q_i}^2 := u^H Q_i u$, and Q_i is any $p_i \times p_i$ positive-semidefinite matrix.

Remark 2: The adjoint of the operator $A_i : \mathbb{C}^{p_i} \to X$ is the operator $A_i^* : X \to \mathbb{C}^{p_i}$ defined by [1, p. 161]

$$A_i^* x = \begin{bmatrix} \langle x, a_{i,1} \rangle \\ \vdots \\ \langle x, a_{i,p_i} \rangle \end{bmatrix} \in \mathbb{C}^{p_i}. \tag{3}$$

Then $A_i^* A_i$ can be identified with the matrix whose $j\,k$ entry is $\langle a_{i,k}, a_{i,j} \rangle$. Since $\|A_i u\|^2 = \|u\|_{A_i^* A_i}^2$, taking $Q_i = A_i^* A_i$ shows that the original problem (1) is a special case of (2).

Remark 3: Constraints of the form $\max_{1 \leq k \leq p_i} |u_k| \leq \sqrt{\mathcal{E}_i}$ are considered in Appendix A.

III. Eliminating the Nuisance Parameter and Making the Problem Finite Dimensional

The problem in (2) suffers from two difficulties. First, it involves the nuisance parameter b (possibly belonging to an infinite-dimensional subspace), and second, the measurement y is in general an element of an infinite-dimensional space, e.g., L^2. If $\mathcal{B}$ were finite dimensional, then the standard way to convert the infinite-dimensional measurement y into an equivalent, finite-dimensional coordinate vector would be to project y onto the finite-dimensional subspace

$$\mathcal{A}_1 + \cdots + \mathcal{A}_M + \mathcal{B}$$

and to work with the coordinate vector of the projection. In our case, since we allow $\mathcal{B}$ to be infinite dimensional, the standard approach does not work. The standard approach also suffers from the defect that it does not eliminate the nuisance parameter b. In the following, we present a way around both of these difficulties at the same time. In order to accomplish this, we first need a few facts about orthogonal projections.

For $\mathcal{B}$ a subspace of X, let $P_{\mathcal{B}}$ denote the corresponding orthogonal projection operator onto $\mathcal{B}$. If $\mathcal{B}$ is infinite dimensional, we assume it is closed, and we assume X is a Hilbert space. These assumptions guarantee the existence of $P_{\mathcal{B}}$. Projection operators have two properties that we use repeatedly. First, they are self-adjoint, i.e., $P_{\mathcal{B}}^* = P_{\mathcal{B}}$, and second, they are idempotent, i.e., $P_{\mathcal{B}} P_{\mathcal{B}} = P_{\mathcal{B}}$. The orthogonal complement of $\mathcal{B}$ is denoted by $\mathcal{B}^{\perp}$, and its corresponding orthogonal projection operator is denoted by $P_{\mathcal{B}}^{\perp}$. Note that $P_{\mathcal{B}}^{\perp} = I - P_{\mathcal{B}}$, where I is the identity operator.

Write y in (2) as

$$y = A_i u + b + n = P_{\mathcal{B}}^{\perp} A_i u + P_{\mathcal{B}} A_i u + b + n.$$

Since $P_{\mathcal{B}} A_i u \in \mathcal{B}$, and since $b \in \mathcal{B}$ is arbitrary, (2) is equivalent to

$$H_i : y = P_{\mathcal{B}}^{\perp} A_i u + b + n, \quad \|u\|_{Q_i}^2 \le \mathcal{E}_i, \qquad u \in \mathbb{C}^{p_i},\, b \in \mathcal{B}. \quad (4)$$

Observe that $P_{\mathcal{B}}^{\perp} A_i u$ is an element of the ith **interference-free signal subspace**,

$$\mathcal{G}_i := \{P_{\mathcal{B}}^{\perp} a : a \in \mathcal{A}_i\}. \quad (5)$$

The sum of all the interference-free signal subspaces is

$$\mathcal{G} := \mathcal{G}_1 + \cdots + \mathcal{G}_M.$$

Since $P_{\mathcal{B}}^{\perp} A_i u \in \mathcal{G}_i \subset \mathcal{G}$, and since $\mathcal{G} \subset \mathcal{B}^{\perp}$, applying $P_{\mathcal{G}}$ to y in (4) yields $P_{\mathcal{G}} y = P_{\mathcal{B}}^{\perp} A_i u + P_{\mathcal{G}} n$; i.e., the operator $P_{\mathcal{G}}$ zero forces the interference b. Next, observe that the signals

$$\widehat{y} := P_{\mathcal{G}} y = P_{\mathcal{B}}^{\perp} A_i u + P_{\mathcal{G}} n \quad \text{and} \quad \widetilde{y} := y - \widehat{y} = b + P_{\mathcal{G}}^{\perp} n$$

have the following three properties:

i) $\widehat{y}$ is a function of y;
ii) $\widehat{y}$ and $\widetilde{y}$ are independent under all hypotheses (since the noise terms are Gaussian and uncorrelated);
iii) the distribution of $\widetilde{y}$ does not depend on the hypothesis.

It then follows that the optimal detector depends only on $\widehat{y}$ [10, pp. 299–300].

A further simplification is possible. Let $g_1, \ldots, g_p$ be an orthonormal basis for $\mathcal{G}$, and define the operator $G: \mathbb{C}^p \to X$ by

$$Gu := \sum_{j=1}^{p} u_j\, g_j, \qquad u = [u_1, \ldots, u_p]' \in \mathbb{C}^p. \quad (6)$$

The adjoint of G is the operator $G^* : X \to \mathbb{C}^p$ given as in (3), and $P_{\mathcal{G}} y = G(G^*G)^{-1}G^* y$ [1, pp. 160–161]. Furthermore, since the g_i are orthonormal, $G^*G = I$ and $P_{\mathcal{G}} = GG^*$. If we put $\underline{y} := G^* y$, then $\widehat{y} := P_{\mathcal{G}} y = G\underline{y}$. Since $G^*\widehat{y} = \underline{y}$, we see that $\widehat{y}$ and $\underline{y}$ are equivalent in that each is a function of the other. Hence, rather than design the optimal detector based on the waveform $\widehat{y}$, we base it on the coordinate vector $\underline{y}$. If we put $\underline{n} := G^* n$, then (4) is equivalent to the coordinate-vector detection problem

$$H_i : \underline{y} = G^* P_{\mathcal{B}}^{\perp} A_i u + \underline{n}, \; \|u\|_{Q_i}^2 \le \mathcal{E}_i, \qquad u \in \mathbb{C}^{p_i} \quad (7)$$

where $\underline{n}$ is a p-dimensional, zero-mean, Gaussian random vector with covariance matrix $R := \sigma^2 G^* G = \sigma^2 I_p$.

This completes our conversion of the waveform channel (2) into the coordinate-vector channel (7). In the coordinate-vector channel, the interference b has been zero forced by the projection $P_{\mathcal{G}}$. The signal $A_i u$ has been projected onto $\mathcal{B}^{\perp}$ and resolved onto the basis $\{g_1, \ldots, g_p\}$ of $\mathcal{G}$.

Example 1: For $i = 1$ and 2, take $\mathcal{A}_i$ to be the one-dimensional subspace $\mathcal{A}_i = \operatorname{span}\{a_i\}$, where the complex waveforms $a_1(t)$ and $a_2(t)$ are the inverse Fourier transforms of $\widetilde{a}_1(f)$ and $\widetilde{a}_2(f)$ shown at the top in Fig. 1. Observe that $\langle \widetilde{a}_1, \widetilde{a}_2 \rangle = 0$, and by Parseval's equation, $\langle a_1, a_2 \rangle = 0$ too. Thus, the signal subspaces $\mathcal{A}_1$ and $\mathcal{A}_2$ are orthogonal. For the interference subspace $\mathcal{B}$, we take the set of high-pass waveforms whose Fourier transform is zero for $|f| \le 1$. Then $\mathcal{B}^{\perp}$ is the set of low-pass waveforms that are band limited to $|f| \le 1$; i.e., $P_{\mathcal{B}}^{\perp}$ is just a low-pass filter. It follows that $(P_{\mathcal{B}}^{\perp} a_1)(t)$ and $(P_{\mathcal{B}}^{\perp} a_2)(t)$ are the inverse Fourier transforms of $(\widetilde{P_{\mathcal{B}}^{\perp} a_1})(f)$ and $(\widetilde{P_{\mathcal{B}}^{\perp} a_2})(f)$ shown at the bottom in Fig. 1. Observe that

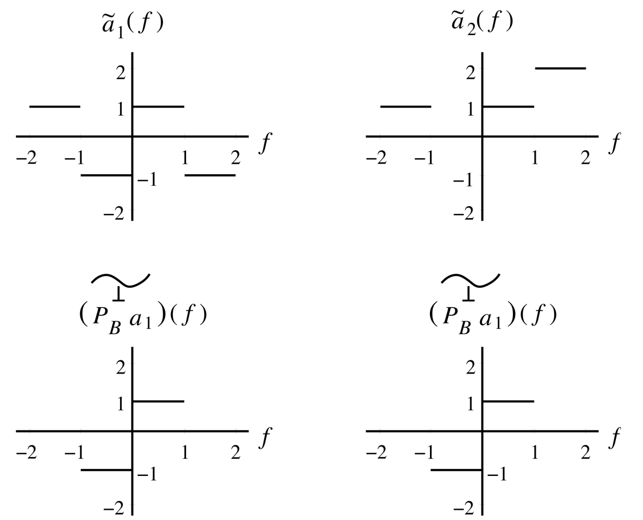

Fig. 1. Fourier transforms of signals (top) and of their projections (bottom) in Example 1.

$$\langle P_{\mathcal{B}}^{\perp} a_1, P_{\mathcal{B}}^{\perp} a_2 \rangle = \langle (\widetilde{P_{\mathcal{B}}^{\perp} a_1}), (\widetilde{P_{\mathcal{B}}^{\perp} a_2}) \rangle = 1$$

and so $\mathcal{G}_1 = \operatorname{span}\{P_{\mathcal{B}}^{\perp} a_1\}$ and $\mathcal{G}_2 = \operatorname{span}\{P_{\mathcal{B}}^{\perp} a_2\}$ are not orthogonal subspaces, although they are linearly independent; i.e., $\mathcal{G}_1 \cap \mathcal{G}_2 = \{0\}$. Furthermore, since the $\mathcal{G}_i$ are one-dimensional, computation of $P_{\mathcal{G}_i}$ is trivial. We have

$$P_{\mathcal{G}_i} y = \left\langle y, \frac{P_{\mathcal{B}}^{\perp} a_i}{\|P_{\mathcal{B}}^{\perp} a_i\|} \right\rangle \frac{P_{\mathcal{B}}^{\perp} a_i}{\|P_{\mathcal{B}}^{\perp} a_i\|}.$$

Because $\|P_{\mathcal{G}_i} y\|^2$ appears later, we also note that

$$\|P_{\mathcal{G}_i} y\|^2 = \frac{\left|\langle y, P_{\mathcal{B}}^{\perp} a_i \rangle\right|^2}{\|P_{\mathcal{B}}^{\perp} a_i\|^2}$$

where $\|P_{\mathcal{B}}^{\perp} a_1\|^2 = 2$ and $\|P_{\mathcal{B}}^{\perp} a_2\|^2 = 1$ in this example. The inner product $\langle y, P_{\mathcal{B}}^{\perp} a_i \rangle$ can be realized by sampling at time $t = 0$, the output of the matched filter whose transfer function is the complex conjugate of $(\widetilde{P_{\mathcal{B}}^{\perp} a_i})(f)$.

IV. The Detector Decision Rule

Let f denote the complex, proper (circularly symmetric), zero-mean, multivariate normal density with covariance matrix R; i.e., for complex Gaussian random vectors [3, p. 122]

$$f(\underline{y}) = \frac{e^{-\underline{y}^H R^{-1} \underline{y}}}{\pi^p \det R}, \qquad \underline{y} \in \mathbb{C}^p \quad (8)$$

where the superscript ^{H} denotes the complex conjugate transpose. The density of $\underline{y}$ under hypothesis H_i in (7) is $f(\underline{y} - G^* P_{\mathcal{B}}^{\perp} A_i u)$, where $u \in \mathbb{C}^{p_i}$ and $\|u\|_{Q_i}^2 \le \mathcal{E}_i$. Since the value of u is unknown, we replace it by its maximum-likelihood estimate under H_i. We denote this estimate by $\widehat{u}_i$, and put

$$\underline{\widehat{g}}_i := G^* P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i.$$

We now decide in favor of H_i if i is the smallest integer among $1, \ldots, M$ such that

$$\eta_i f(\underline{y} - \underline{\widehat{g}}_i) \geq \eta_j f(\underline{y} - \underline{\widehat{g}}_j), \qquad j = 1, \ldots, M$$

where η_i is the *a priori* probability that message i is sent. This is a generalized likelihood-ratio test (GLRT), or empirical Bayes test.

Because of the Gaussian form of f, and because the covariance matrix $R = \sigma^2 I_p$ is proportional to the identity matrix, the above inequality is equivalent to

$$L_i \geq L_j + \eta_{ij}$$

where

$$L_i := \mathrm{Re}\langle \underline{y}, \underline{\widehat{g}}_i \rangle - \frac{\|\underline{\widehat{g}}_i\|_p^2}{2} \tag{9}$$

$\|\cdot\|_p$ is the Euclidean norm on $\mathbb{C}^p$, and

$$\eta_{ij} := (\sigma^2/2)\ln(\eta_i/\eta_j).$$

In the case $\eta_i = 1/M$ for all i, $\eta_{ij} = 0$ for all i and j.

To gain further insight into the structure of L_i, it is convenient to rewrite it in terms of the original observation y. With regard to the inner product in (9), write

$$\begin{aligned}\langle \underline{y}, \underline{\widehat{g}}_i \rangle &= \langle G^* y, G^* P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \rangle \\ &= \langle y, P_{\mathcal{G}} P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \rangle, && \text{since } GG^* = P_{\mathcal{G}} \\ &= \langle y, P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \rangle, && \text{since } P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \in \mathcal{G} \\ &= \langle P_{\mathcal{G}_i} y, P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \rangle, && \text{since } P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \in \mathcal{G}_i.\end{aligned}$$

Similarly, $\|\underline{\widehat{g}}_i\|_p^2 = \|\widehat{u}_i\|^2_{A_i^* P_{\mathcal{B}}^{\perp} A_i}$. Hence,

$$L_i = \mathrm{Re}\langle P_{\mathcal{G}_i} y, P_{\mathcal{B}}^{\perp} A_i \widehat{u}_i \rangle - \frac{\|\widehat{u}_i\|^2_{A_i^* P_{\mathcal{B}}^{\perp} A_i}}{2}. \tag{10}$$

We next turn to the problem of computing $\widehat{u}_i$, the maximizer of $f(\underline{y} - G^* P_{\mathcal{B}}^{\perp} A_i u)$. Due to the Gaussian form of f and the fact that its covariance matrix is $R = \sigma^2 I_p$, $\widehat{u}_i$ is the solution of

$$\min_{\|u\|^2_{Q_i} \leq \mathcal{E}_i} \|\underline{y} - G^* P_{\mathcal{B}}^{\perp} A_i u\|_p^2. \tag{11}$$

In general, there is no closed-form solution. However, this is a quadratically constrained least squares problem that is readily solved by numerical methods once the orthonormal basis used to construct G^* is given. However, as we now show, $\widehat{u}_i$ can be found without constructing G^* and without computing $\underline{y} = G^* y$. Write

$$\begin{aligned}\|\underline{y} - G^* P_{\mathcal{B}}^{\perp} A_i u\|_p^2 &= \|G^*(y - P_{\mathcal{B}}^{\perp} A_i u)\|_p^2 \\ &= \|P_{\mathcal{G}}(y - P_{\mathcal{B}}^{\perp} A_i u)\|^2 \\ &= \|P_{\mathcal{G}} y - P_{\mathcal{B}}^{\perp} A_i u\|^2 \\ &= \|P_{\mathcal{G}} y - P_{\mathcal{G}_i} P_{\mathcal{G}} y\|^2 \\ &\quad + \|P_{\mathcal{G}_i} P_{\mathcal{G}} y - P_{\mathcal{B}}^{\perp} A_i u\|^2 \\ &= \|P_{\mathcal{G}} y - P_{\mathcal{G}_i} P_{\mathcal{G}} y\|^2 \\ &\quad + \|P_{\mathcal{G}_i} y - P_{\mathcal{B}}^{\perp} A_i u\|^2\end{aligned} \tag{12}$$

where we have used the fact that $P_{\mathcal{B}}^{\perp} A_i u \in \mathcal{G}_i \subset \mathcal{G}$. Since only the last term involves u, solving for $\widehat{u}_i$ is equivalent to solving

$$\min_{\|u\|^2_{Q_i} \leq \mathcal{E}_i} \|P_{\mathcal{G}_i} y - P_{\mathcal{B}}^{\perp} A_i u\|^2. \tag{13}$$

This shows that $\widehat{u}_i$ depends on y only through $P_{\mathcal{G}_i} y$. Combining this fact with (10) yields the following result.

Theorem 1: The statistic L_i depends on y only through $P_{\mathcal{G}_i} y$, the projection of y onto the ith interference-free signal subspace $\mathcal{G}_i$. Hence, the receiver statistic is a zero-forcing or decorrelating detector.

Formula (13) is again a quadratically constrained least squares problem that is easily solved by numerical methods. For example, the Lagrangian for the quadratically constrained least squares problem in (13) is

$$\mathrm{Lagr}(\lambda, u) = \|P_{\mathcal{G}_i} y - P_{\mathcal{B}}^{\perp} A_i u\|^2 + \lambda(\|u\|^2_{Q_i} - \mathcal{E}_i)$$

where the scalar λ is the Lagrange multiplier. Using the Kuhn–Tucker sufficiency theorem [9, p. 60], it is easy to see that

$$u = [\lambda Q_i + A_i^* P_{\mathcal{B}}^{\perp} A_i]^{-1} A_i^* P_{\mathcal{G}_i} y \tag{14}$$

solves (13) if $\lambda \geq 0$ is chosen so that

$$\|u\|^2_{Q_i} \leq \mathcal{E}_i \quad \text{and} \quad \lambda(\|u\|^2_{Q_i} - \mathcal{E}_i) = 0. \tag{15}$$

We assume here that either $Q_i > 0$ or $\mathcal{A}_i \cap \mathcal{B} = \{0\}$. Note that since $a_{i,1}, \ldots, a_{i,p_i}$ is a basis, $A_i^* A_i > 0$; hence, this natural choice for Q_i is positive definite. When $\lambda = 0$ and $\mathcal{A}_i \cap \mathcal{B} \neq \{0\}$, u is understood to be the minimum $\|\cdot\|_{Q_i}$-norm solution of $P_{\mathcal{B}}^{\perp} A_i u = P_{\mathcal{G}_i} y$. We also point out that since $P_{\mathcal{G}_i} A_i = P_{\mathcal{B}}^{\perp} A_i$, (14) can be rewritten as

$$u = [\lambda Q_i + A_i^* P_{\mathcal{B}}^{\perp} A_i]^{-1} A_i^* P_{\mathcal{B}}^{\perp} y. \tag{16}$$

Remark 4: Equation (16) suggests some simplification is possible if we choose $Q_i = A_i^* P_{\mathcal{B}}^{\perp} A_i$. This choice of Q_i amounts to constraining only the signal energy that lies in interference-free subspace $\mathcal{B}^{\perp}$. We show later that for Q_i of this form, u can be found by inspection. It is important to note two situations in which the assumption $Q_i = A_i^* P_{\mathcal{B}}^{\perp} A_i$ entails no loss of generality.

i) If $\mathcal{A}_i$ is one dimensional, then Q_i and $A_i^* P_{\mathcal{B}}^{\perp} A_i$ are positive constants, and $\|u\|^2_{Q_i} \leq \mathcal{E}_i$ can always be rewritten as

$$\|u\|^2_{A_i^* P_{\mathcal{B}}^{\perp} A_i} \leq \mathcal{E}_i (A_i^* P_{\mathcal{B}}^{\perp} A_i)/Q_i = \mathcal{E}_i \|P_{\mathcal{B}}^{\perp} a_{i,1}\|^2 / Q_i.$$

ii) If the ith constraint is inactive ($\mathcal{E}_i = \infty$), then in (13)

$$\{u : \|u\|^2_{Q_i} < \infty\} = \mathbb{C}^{p_i}$$

no matter how Q_i is chosen.

V. The Average Probability of Error

A. Notation and Preliminary Observations

Let

$$\underline{\mathcal{G}}_i := G^*(\mathcal{G}_i)$$

denote the image of $\mathcal{G}_i$ under the mapping G^*. Then instead of the calculations in (12), use the fact that $G^* P_{\mathcal{B}}^{\perp} A_i u \in \underline{\mathcal{G}}_i$ to write

$$\|\underline{y} - G^* P_{\mathcal{B}}^{\perp} A_i u\|_p^2 = \|\underline{y} - P_{\underline{\mathcal{G}}_i} \underline{y}\|_p^2 + \|P_{\underline{\mathcal{G}}_i} \underline{y} - G^* P_{\mathcal{B}}^{\perp} A_i u\|_p^2. \tag{17}$$

Observe that the first term on the right does not depend on u. Hence, the solution of (11), $\widehat{u}_i$, depends on $\underline{y}$ only through $P_{\underline{\mathcal{G}}_i} \underline{y}$.

Next, construct a matrix G_i whose columns are given by any orthonormal basis for $\underline{\mathcal{G}}_i$. Then $P_{\underline{\mathcal{G}}_i} = G_i G_i^H$, and $G_i^H G_i$ is the identity matrix of size equal to $\dim \underline{\mathcal{G}}_i$.

Lemma 2: We have $\dim \underline{\mathcal{G}}_i = \dim \mathcal{G}_i$, and $\dim \mathcal{G}_i \leq p_i$, with equality if and only if $\mathcal{A}_i \cap \mathcal{B} = \{0\}$.

Notation: Rather than introduce a new symbol for $\dim \underline{\mathcal{G}}_i$, we abuse notation and denote $\dim \underline{\mathcal{G}}_i$ by p_i.

Proof of Lemma 2: We first argue that $\dim \underline{\mathcal{G}}_i = \dim \mathcal{G}_i$. Since $\mathcal{G}_i \subset \mathcal{G}$, it is easy to see that for $x \in \mathcal{G}_i$, $G^* x = 0$ implies $x \in \mathcal{G} \cap \mathcal{G}^\perp$ and is therefore the zero vector. Hence, any linearly independent set in $\mathcal{G}_i$ is mapped by G^* into a linearly independent set in $\underline{\mathcal{G}}_i$.

Similarly, since $\{a_{i,k}\}_{k=1}^{p_i}$ is assumed to be a basis for $\mathcal{A}_i$, $\{P_\mathcal{B}^\perp a_{i,k}\}_{k=1}^{p_i}$ are linearly independent if and only if $\mathcal{A}_i \cap \mathcal{B} = \{0\}$. □

Theorem 3: The statistic L_i depends on $\underline{y}$ only through $G_i^H \underline{y}$.

Proof: Since $P_{\underline{\mathcal{G}}_i} \underline{y} = G_i G_i^H \underline{y}$, it now follows that $\widehat{u}_i$ is a function of $G_i^H \underline{y}$. Furthermore, writing the inner product in (10) as

$$\begin{aligned}\langle P_{\mathcal{G}_i} y, P_\mathcal{B}^\perp A_i \widehat{u}_i\rangle &= \langle P_\mathcal{G} P_{\mathcal{G}_i} y, P_\mathcal{B}^\perp A_i \widehat{u}_i\rangle \\ &= \langle G G^* P_{\mathcal{G}_i} y, P_\mathcal{B}^\perp A_i \widehat{u}_i\rangle \\ &= \langle G^* P_{\mathcal{G}_i} y, G^* P_\mathcal{B}^\perp A_i \widehat{u}_i\rangle\end{aligned}$$

and using the easily verified fact that $G^* P_{\mathcal{G}_i} = P_{\underline{\mathcal{G}}_i} G^*$, we have

$$L_i = \mathrm{Re}\langle P_{\underline{\mathcal{G}}_i} \underline{y}, G^* P_\mathcal{B}^\perp A_i \widehat{u}_i\rangle - \frac{\|\widehat{u}_i\|^2_{A_i^* P_\mathcal{B}^\perp A_i}}{2}.$$

Since $\widehat{u}_i$ is also a function of $P_{\underline{\mathcal{G}}_i} \underline{y} = G_i G_i^H \underline{y}$, L_i depends on $\underline{y}$ only through $G_i^H \underline{y}$. □

B. Calculations

Let $P_{C|i}$ denote the conditional probability of a correct decision given that message i is sent. Then the average probability of error is

$$P_e = \sum_{i=1}^{M} (1 - P_{C|i})\eta_i = 1 - \sum_{i=1}^{M} P_{C|i}\, \eta_i.$$

Hence, it suffices to compute

$$P_{C|i} = P\Big(L_i \geq \max_{j \neq i}[L_j + \eta_{ij}]\Big)$$

where each L_j is a function of $G_j^H \underline{y}$ and $\underline{y}$ is given by (7); i.e., $\underline{y}$ is a p-dimensional, Gaussian random vector with mean $G^* P_\mathcal{B}^\perp A_i u$ and covariance matrix $\sigma^2 I_p$. Since each L_j is a function of $G_j^H \underline{y}$, we write $L_j(G_j^H \underline{y})$ to make this dependence explicit. We then have

$$P_{C|i} := P\Big(L_i(G_i^H \underline{y}) \geq \max_{j \neq i}[L_j(G_j^H \underline{y}) + \eta_{ij}]\Big).$$

Let f_i denote the density of $G_i^H \underline{y}$, which is a p_i-dimensional, Gaussian random vector with mean $G_i^H G^* P_\mathcal{B}^\perp A_i u$ and covariance matrix $\sigma^2 I_{p_i}$. Then

$$P_{C|i} = \int P\Big(\max_{j \neq i}[L_j(G_j^H \underline{y}) + \eta_{ij}] \leq L_i(\xi) \Big| G_i^H \underline{y} = \xi\Big) f_i(\xi)\, d\xi. \quad (18)$$

Since the $G_j^H \underline{y}$, $j = 1, \ldots, M$ are jointly Gaussian, so is the conditional distribution of the $\{G_j^H \underline{y}, j \neq i\}$ given $G_i^H \underline{y} = \xi$, and it is therefore straightforward to determine. Once the conditional density of the $\{G_j^H \underline{y}, j \neq i\}$ given $G_i^H \underline{y} = \xi$ is known, it is possible in principle to compute the conditional probability in (18).

VI. Special Cases

Example 2 (Orthogonal Subspaces): If the subspaces $\underline{\mathcal{G}}_j$ are orthogonal, then the random vectors $G_j^H \underline{y}$ are independent. We then have

$$P_{C|i} = \int \left[\prod_{j \neq i} F_{L_j|i}\Big(L_i(\xi) - \eta_{ij}\Big)\right] f_i(\xi)\, d\xi \quad (19)$$

where $F_{L_j|i}$ denotes the cumulative distribution function of the real-valued random variable $L_j(G_j^H \underline{y})$ under hypothesis H_i. In general, the $G_j^H \underline{y}$ are not independent under H_i. However, if there is no interference; i.e., $\mathcal{B}$ is the zero subspace and $\mathcal{B}^\perp = X$, and if the $\mathcal{A}_i$ are orthogonal, then the desired independence can be easily arranged.

Example 3 (Binary Signaling): In this case, if $i = 1$ in (18), we get

$$P_{C|1} = \int P\Big(L_2(G_2^H \underline{y}) \leq L_1(\xi) - \eta_{12} \Big| G_1^H \underline{y} = \xi\Big) f_1(\xi)\, d\xi. \quad (20)$$

One of the difficulties with (18), and even (19) and (20), is that we do not have an explicit formula for the $L_j(\cdot)$. However, examination of (16) and Remark 4 following it suggests we consider the special case $Q_i = A_i^* P_\mathcal{B}^\perp A_i$. For this choice of the Q_i, we do not need Lagrange multipliers, and as noted in Remark 4, in some cases, this entails no loss of generality.

Writing (13) with $Q_i = A_i^* P_\mathcal{B}^\perp A_i$, we have

$$\min_{\|P_\mathcal{B}^\perp A_i u\|^2 \leq \mathcal{E}_i} \|P_{\mathcal{G}_i} y - P_\mathcal{B}^\perp A_i u\|^2$$

or equivalently

$$\min_{v \in \mathcal{G}_i : \|v\|^2 \leq \mathcal{E}_i} \|P_{\mathcal{G}_i} y - v\|^2.$$

This is the problem of projecting the point $P_{\mathcal{G}_i} y \in \mathcal{G}_i$ onto the closed ball of radius $\sqrt{\mathcal{E}_i}$ in $\mathcal{G}_i$. The optimal value of v is

$$\widehat{v}_i = \begin{cases} P_{\mathcal{G}_i} y, & \text{if } \|P_{\mathcal{G}_i} y\|^2 \leq \mathcal{E}_i \\ \sqrt{\mathcal{E}_i}\, \frac{P_{\mathcal{G}_i} y}{\|P_{\mathcal{G}_i} y\|}, & \text{otherwise.} \end{cases}$$

Substituting $\widehat{\underline{g}}_i = G^* \widehat{v}_i$ in (9) yields

$$L_i = \begin{cases} \frac{\|P_{\mathcal{G}_i} y\|^2}{2}, & \text{if } \|P_{\mathcal{G}_i} y\|^2 \leq \mathcal{E}_i \\ \sqrt{\mathcal{E}_i}\|P_{\mathcal{G}_i} y\| - \frac{\mathcal{E}_i}{2}, & \text{otherwise.} \end{cases} \quad (21)$$

We thus have an explicit formula for the function of $P_{\mathcal{G}_i} y$ asserted in Theorem 1. In fact, L_i depends on $P_{\mathcal{G}_i} y$ only through its energy.

Remark 5: If the constraints are inactive, i.e., all $\mathcal{E}_i = \infty$, we have the unconstrained matched subspace detection problem of [7], [8] generalized from 2 to M hypotheses. In this case, we decide in favor of message i if

$$\|P_{\mathcal{G}_i} y\|^2 \geq \|P_{\mathcal{G}_j} y\|^2 + 2\eta_{ij}, \qquad \text{for all } j$$

which is the unconstrained matched subspace detector. We also note that even with all $\mathcal{E}_i = \infty$, our derivation here is more general than that in [7], [8] because we allow the interference subspace $\mathcal{B}$ to be infinite dimensional, and we do not assume $\mathcal{A}_i \cap \mathcal{B} = \{0\}$.

Remark 6: Considering (21) as a function of the energy $\|P_{\mathcal{G}_i} y\|^2$, this function is strictly increasing, and for large argument, its slope is decreasing; i.e., the function can be viewed as a compressor. Hence, when constraints are active (the $\mathcal{E}_i$ are finite), the comparison $L_i \geq L_j + \eta_{ij}$ amounts to comparing compressed versions of the unconstrained matched subspace detector statistics $\|P_{\mathcal{G}_i} y\|^2$ and $\|P_{\mathcal{G}_j} y\|^2$ of the preceding remark.

Using (21), we can give an explicit formula for the function of $P_{\underline{\mathcal{G}}_i} \underline{y}$ asserted in Theorem 3. First, an orthogonality argument shows that $P_{\underline{\mathcal{G}}_j} G^* y = G^* P_{\mathcal{G}_j} y$. Using this result, it is then not hard to show that

$$\|P_{\mathcal{G}_j} y\|^2 = \|P_{\underline{\mathcal{G}}_j} \underline{y}\|_p^2 = \|G_j^H \underline{y}\|_{p_j}^2. \quad (22)$$

We can now write

$$L_j = \begin{cases} \frac{\|G_j^H \underline{y}\|_{p_j}^2}{2}, & \text{if } \|G_j^H \underline{y}\|_{p_j}^2 \leq \mathcal{E}_j \\ \sqrt{\mathcal{E}_j}\|G_j^H \underline{y}\|_{p_j} - \frac{\mathcal{E}_j}{2}, & \text{otherwise.} \end{cases} \quad (23)$$

Lemma 4: If $Q_2 = A_2^* P_{\mathcal{B}}^{\perp} A_2$, then the conditional probability in (20) can be expressed using the conditional cumulative distribution function

$$P\left(L_2(G_2^H \underline{y}) \le \ell \middle| G_1^H \underline{y} = \xi\right) \tag{24}$$

which is equal to

$$P\left(\|G_2^H \underline{y}\|_{p_2}^2 \le 2\ell \middle| G_1^H \underline{y} = \xi\right), \qquad \text{for } \ell \le \mathcal{E}_2/2,$$

and

$$P\left(\|G_2^H \underline{y}\|_{p_2}^2 \le (\ell + \mathcal{E}_2/2)^2/\mathcal{E}_2 \middle| G_1^H \underline{y} = \xi\right), \qquad \text{otherwise.}$$

Note that the conditional distribution of $G_2^H \underline{y}$ given $G_1^H \underline{y} = \xi$ is Gaussian, with mean and covariance easily determined.

Proof: Write (24) as the sum of

$$P\left(L_2(G_2^H \underline{y}) \le \ell, \|G_2^H \underline{y}\|_{p_2}^2 \le \mathcal{E}_2 \middle| G_1^H \underline{y} = \xi\right) \tag{25}$$

and

$$P\left(L_2(G_2^H \underline{y}) \le \ell, \|G_2^H \underline{y}\|_{p_2}^2 > \mathcal{E}_2 \middle| G_1^H \underline{y} = \xi\right). \tag{26}$$

From (23), it follows that (25) is equal to

$$P\left(\|G_2^H \underline{y}\|_{p_2}^2 \le \min(2\ell, \mathcal{E}_2) \middle| G_1^H \underline{y} = \xi\right)$$

and (26) is equal to

$$P\left(\mathcal{E}_2 < \|G_2^H \underline{y}\|_{p_2}^2 \le (\ell + \mathcal{E}_2/2)^2/\mathcal{E}_2 \middle| G_1^H \underline{y} = \xi\right).$$

Note that this last expression is zero for $\ell \le \mathcal{E}_2/2$. The lemma now follows easily. □

Lemma 5: When message i is sent

$$\frac{\|G_j^H \underline{y}\|_{p_j}^2}{\sigma^2/2}$$

is noncentral chi-squared with $2p_j$ degrees of freedom and noncentrality parameter

$$\begin{aligned}\frac{\|G_j^H G^* P_{\mathcal{B}}^{\perp} A_i u\|_{p_j}^2}{\sigma^2/2} &= 2\|P_{\underline{\mathcal{G}}_j} G^* P_{\mathcal{B}}^{\perp} A_i u\|_p^2/\sigma^2 \\ &= 2\|G^* P_{\mathcal{G}_j} P_{\mathcal{B}}^{\perp} A_i u\|_p^2/\sigma^2 \\ &= 2\|P_{\mathcal{G}} P_{\mathcal{G}_j} P_{\mathcal{B}}^{\perp} A_i u\|^2/\sigma^2 \\ &= 2\|P_{\mathcal{G}_j} P_{\mathcal{B}}^{\perp} A_i u\|^2/\sigma^2. \end{aligned} \tag{27}$$

Remark 7: Observe that hypothesis i affects the distribution of $2\|G_j^H \underline{y}\|_{p_j}^2/\sigma^2$ only via the noncentrality parameter. The noncentrality parameter is determined by the amount of energy of $P_{\mathcal{B}}^{\perp} A_i u \in \mathcal{G}_i$ that falls in the subspace $\mathcal{G}_j$. To put it another way, the noncentrality parameter is proportional to the energy that "leaks" from subspace $\mathcal{G}_i$ into subspace $\mathcal{G}_j$. If the subspaces $\mathcal{G}_i$ and $\mathcal{G}_j$ are orthogonal, there is no energy leakage, and the noncentrality parameter is zero for $j \ne i$. For $j = i$, (27) reduces to $2\|P_{\mathcal{B}}^{\perp} A_i u\|^2/\sigma^2$, where $\|P_{\mathcal{B}}^{\perp} A_i u\|^2$ is the energy in the signal $A_i u$ that lies in the interference-free subspace.

Proof of Lemma 5: According to (7), under H_i, $\underline{y}$ is a p-dimensional Gaussian random vector with mean $G^* P_{\mathcal{B}}^{\perp} A_i u \in \underline{\mathcal{G}}_i$ and covariance matrix $\sigma^2 I_p$. Thus, $G_j^H \underline{y}$ is a p_j-dimensional Gaussian random vector with mean $G_j^H G^* P_{\mathcal{B}}^{\perp} A_i u$ and covariance matrix $\sigma^2 I_{p_j}$. The result now follows. □

Lemma 6: If $Q_j = A_j^* P_{\mathcal{B}}^{\perp} A_j$, then $F_{L_j|i}$ in (19) is given by

$$F_{L_j|i}(\ell) = \begin{cases} \widetilde{F}_{p_j}(4\ell/\sigma^2), & \ell \le \mathcal{E}_j/2 \\ \widetilde{F}_{p_j}\left(2(\ell + \mathcal{E}_j/2)^2/[\sigma^2 \mathcal{E}_j]\right), & \text{otherwise} \end{cases}$$

where $\widetilde{F}_{p_j}$ is the cumulative distribution of a noncentral chi-squared random variable with $2p_j$ degrees of freedom and noncentrality parameter (27).

Proof: This is immediate from the proof of Lemma 4 and the result of Lemma 5. □

Lemma 7: Motivated by (23), put

$$\mathcal{L}_i(\zeta) := \begin{cases} \zeta/2, & 0 \le \zeta \le \mathcal{E}_i \\ \sqrt{\mathcal{E}_i \zeta} - \mathcal{E}_i/2, & \text{otherwise} \end{cases}$$

so that $L_i(G_i^H \underline{y}) = \mathcal{L}_i(\|G_i^H \underline{y}\|_{p_i}^2)$. If $Q_i = A_i^* P_{\mathcal{B}}^{\perp} A_i$, then (19) becomes

$$P_{C|i} = \int \left[\prod_{j \ne i} F_{L_j|i}\left(\mathcal{L}_i(\zeta\sigma^2/2) - \eta_{ij}\right)\right] \widetilde{f}_{p_i}(\zeta)\, d\zeta$$

where $\widetilde{f}_{p_i}$ is the density of a noncentral chi-squared random variable with $2p_i$ degrees of freedom and noncentrality parameter

$$\frac{2}{\sigma^2}\|P_{\mathcal{B}}^{\perp} A_i u\|^2 = \frac{2}{\sigma^2}\|u\|_{A_i^* P_{\mathcal{B}}^{\perp} A_i}^2.$$

Proof: The only point worth noting is that the noncentrality parameter comes from (27) with $j = i$. □

Example 4 (Binary Signaling With On–Off Keying): In this case, $\mathcal{A}_1$ is the zero subspace. Then $\mathcal{G}_1$ is also the zero subspace, which is automatically orthogonal to any $\mathcal{G}_2$. Since $L_1 \equiv 0$, we decide in favor of message 2 if and only if $L_2 \ge \eta_{21}$. Assuming $Q_2 = A_2^* P_{\mathcal{B}}^{\perp} A_2$, we have

$$P_{C|1} = \widetilde{F}_{p_2}\left(2\mathcal{L}_2^{-1}(\eta_{21})/\sigma^2\right)$$

where in this case the noncentrality parameter is zero. On the other hand

$$P_{C|2} = 1 - \widehat{F}_{p_2}\left(2\mathcal{L}_2^{-1}(\eta_{21})/\sigma^2\right)$$

where $\widehat{F}_{p_2}$ is the cumulative distribution function of a noncentral chi-squared random variable with $2p_2$ degrees of freedom and noncentrality parameter

$$\frac{2}{\sigma^2}\|P_{\mathcal{B}}^{\perp} A_2 u\|^2 = \frac{2}{\sigma^2}\|u\|_{A_2^* P_{\mathcal{B}}^{\perp} A_2}^2.$$

VII. Conclusion

The detection of subspace signals in infinite-dimensional interference and noise was motivated by consideration of multipath-Doppler channels subject to interference from partially overlapping frequency bands of other sources. Since the interference lies in an infinite-dimensional subspace, the standard method of projecting onto the subspace containing the signal plus interference does not yield a finite-dimensional detection problem. However, we presented an alternative approach to extract the appropriate finite-dimensional problem. An additional feature of the model was the imposition of an energy constraint on the desired signal. We then derived the generalized likelihood-ratio detector, and gave expressions for its average probability of error.

Although we obtained closed-form expressions for the detector statistics L_i only in the case $Q_i = A_i^* P_{\mathcal{B}}^{\perp} A_i$ ((21) and (23)), we emphasize that even in the general case, it is still practical to compute L_i numerically because it requires solving only a quadratically constrained least squares problem. We also have the general structural result, Theorem 1, which shows that each detector statistic L_i depends on the observation y only through $P_{\mathcal{G}_i}y$. In other words, the front end of the detector takes the measurement y and passes it through a zero-forcing or decorrelating linear filter matched to the ith interference-free signal subspace.

The importance of Theorem 3 is that it allows us to obtain the general expression (18) for computing the average probability of error. The formula (18) can then be simplified in special cases as shown in the examples in Section VI.

In Appendix A, the energy constraint $\|u\|_{Q_i}^2 \leq \mathcal{E}_i$ is replaced by the magnitude constraint $\max_k |u_k| \leq \mathcal{E}_i^{1/2}$. Thus, instead of having a quadratic form subject to a quadratic constraint, we have a quadratic form subject to a convex constraint. It is shown that the single convex constraint is equivalent to p_i quadratic constraints. This converts the convex programming problem with one constraint into a quadratically constrained least squares problem with p_i constraints. When the coefficient vector is real valued, the magnitude constraint is shown to be equivalent to $2p_i$ linear constraints. This converts the convex programming problem into a quadratic programming problem.

Appendix
Magnitude Signal Constraints

Suppose that in (13) we replace the constraint $\|u\|_{Q_i}^2 \leq \mathcal{E}_i$ with

$$\|u\|_\infty := \max_{1\leq k\leq p_i} |u_k| \leq \sqrt{\mathcal{E}_i}.$$

Then we must solve

$$\min_{\|u\|_\infty^2 \leq \mathcal{E}_i} \|P_{\mathcal{G}_i}y - P_{\mathcal{B}}^{\perp}A_i u\|^2. \tag{28}$$

Although the objective function here is a quadratic form in u, the constraint $\|u\|_\infty^2 \leq \mathcal{E}_i$ is no longer quadratic. However, since the constraint is still convex, the overall problem is a finite-dimensional, convex programming problem to which standard techniques apply.

We now point out that since the single nonquadratic constraint $\|u\|_\infty^2 \leq \mathcal{E}_i$ can be rewritten as p_i quadratic constraints, we can convert (28) into a quadratically constrained least squares problem. To see this, rewrite the single constraint

$$\max_{1\leq k\leq p_i} |u_k| \leq \sqrt{\mathcal{E}_i}$$

as the p_i quadratic constraints

$$\|D_k u\|_{p_i}^2 \leq \mathcal{E}_i, \qquad k = 1,\ldots,p_i$$

where $D_k := \mathrm{diag}(0,\ldots,1,\ldots,0)$ with the 1 in the kth position. Thus, we have converted (28) into a quadratically constrained least squares problem to which standard methods apply. For example, the Lagrangian for this problem is

$$\mathrm{Lagr}(\lambda, u) = \|P_{\mathcal{G}_i}y - P_{\mathcal{B}}^{\perp}A_i u\|^2 + \sum_{k=1}^{p_i} \lambda_k(\|D_k u\|_{p_i}^2 - \mathcal{E}_i)$$

where $\lambda := [\lambda_1,\ldots,\lambda_{p_i}]'$ is now a vector of Lagrange multipliers. Appealing again to the Kuhn–Tucker sufficiency theorem, if

$$u = [\Lambda + A_i^* P_{\mathcal{B}}^{\perp} A_i]^{-1} A_i^* P_{\mathcal{B}}^{\perp} y$$

where

$$\Lambda := \sum_{k=1}^{p_i} \lambda_k D_k = \mathrm{diag}(\lambda_1,\ldots,\lambda_{p_i})$$

and the $\lambda_k \geq 0$ are chosen so that the components of u satisfy

$$|u_k|^2 \leq \mathcal{E}_i \quad \text{and} \quad \lambda_k(|u_k|^2 - \mathcal{E}_i) = 0, \qquad k = 1,\ldots,p_i$$

then u solves the minimization problem.

We now briefly reconsider the optimization problem (28) under the requirement that the u_k be real instead of allowing them to be complex. In this case, the condition $\|u\|_\infty^2 \leq \mathcal{E}_i$ can be expressed as the p_i conditions

$$-\sqrt{\mathcal{E}_i} \leq u_k \leq \sqrt{\mathcal{E}_i}$$

or, equivalently, as the $2p_i$ linear inequalities

$$u_k \leq \sqrt{\mathcal{E}_i} \quad \text{and} \quad -u_k \leq \sqrt{\mathcal{E}_i}.$$

Since the objective function $\|P_{\mathcal{G}_i}y - P_{\mathcal{B}}^{\perp}A_i u\|^2$ is quadratic in u, we see that the required optimization is a quadratic programming problem, for which various algorithms are available.

References

[1] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.
[2] R. Lupas and S. Verdú, "Linear multiuser detectors for for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.
[3] B. Picinbono, *Random Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[4] A.-B. Salberg and A. Hanssen, "Subspace detectors for stochastic process shift keying," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Salt Lake City, UT, May 2001, pp. 2549–2552.
[5] ——, "Stochastic multipulse-PAM: A subspace modulation technique with diversity," *Signal Processing*, vol. 83, no. 12, pp. 2559–2577, Dec. 2003.
[6] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, pp. 123–132, Jan. 1999.
[7] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991, ch. 4.
[8] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, vol. 42, pp. 2146–2157, Aug. 1994.
[9] S. Vajda, *Theory of Linear and Nonlinear Programming*. London, U.K.: Longman, 1974.
[10] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.