On the Robustness of Decision-Feedback Detection of DPSK and Differential Unitary Space-Time Modulation in Rayleigh-Fading Channels

Bijoy Bhukania and Philip Schniter

***Abstract*—Decision-feedback differential detection (DFDD) of differential phase-shift keying (DPSK) and differential unitary space-time modulation (DUST) in Rayleigh-fading channels exhibits significant performance improvement over standard single-symbol maximum-likelihood detection. However, knowledge of channel fading correlation and signal-to-noise ratio (SNR) is required at the receiver to compute the feedback coefficients used in DFDD. In this letter, we investigate the robustness of the DFDD to imperfect knowledge of the feedback coefficients by modeling the mismatch between *estimated* feedback coefficients and ideal coefficients in terms of mismatch between the *estimated* values of fading correlation and SNR and the true values. Under the assumption of a block-fading channel when nondiagonal DUST constellations are used and a continuous fading channel otherwise, we derive exact and Chernoff bound expressions for pair-wise word-error probability and then use them to approximate the bit-error rate (BER), finding close agreement with simulation results. The relationships between BER performance and various system parameters, e.g., DFDD length and Doppler mismatch, are also explored. Furthermore, the existence of an error floor in the BER-vs-SNR curve is investigated for the infinite-length DFDD. For the special case of Jakes' fading model, it is shown that the error floor can be removed completely even when the Doppler spread is over-estimated.**

***Index Terms*—Communication system performance, decision feedback differential detection (DFDD), flat fading, space-time codes.**

I. Introduction

SINGLE antenna differential phase shift keying (DPSK) and its multiple antenna extension, differential unitary space-time modulation (DUST) [1], [2], are used in systems where channel is flat, slow fading, and is unknown to the receiver as well as the transmitter. In a fast-fading channel, however, both DUST and DPSK with standard single-symbol maximum-likelihood (ML) detection succumb to an error floor when the error due to channel variation dominates that due to additive noise [3]–[5]. Decision-feedback differential detection (DFDD) [4], [6]–[8] has been proposed to reduce, and asymptotically eliminate, the error floor and thereby improve the detection performance significantly. DFDD, however, requires the knowledge of channel fading correlation and signal-to-noise ratio (SNR) at the receiver. In addition, when nondiagonal DUST constellations (as opposed to the diagonal constellations [1]) are used, DFDD requires a block-fading channel with block length equal to or a multiple of the matrix symbol duration. Thus, use of DFDD for detection of nondiagonal DUST in continuous fading channels results in modeling error, which, however, is negligible for slow fading channels [1]. Keeping this mind, the results in this letter are derived for DPSK and diagonal DUST in continuous fading channels and nondiagonal DUST in block-fading channels.

It can be shown that the DFDD comprises a linear channel predictor followed by a quasi-coherent detector [4], [6]–[8]. Computation of the linear predictor coefficients require the knowledge of channel fading correlation and SNR, which may not be available in practice. Two possible approaches can be taken to counter this problem: 1) the fading correlation and the SNR can be estimated [9]–[11] and then the estimates can be used to compute the coefficients, and b) adaptive algorithms can be employed to directly compute the coefficients [9], [12].

Since the estimated values of the fading correlation and SNR are likely to be corrupted by estimation errors, especially in nonstationary channels, we analyze the robustness of DFDD to imperfect parameter knowledge. To do this, we consider DFDD operation in accordance with the *estimated* or *assumed* values of fading correlation and SNR that differ from true values. Under such conditions, we derive exact and Chernoff bound expressions for pair-wise word-error probability (PWEP) and use them to approximate the bit-error rate (BER), resulting in close agreement with numerical simulations. In addition, we analyze the "equivalent SNR-loss" due to parameter mismatch. Finally, fundamental limits on the performance of DFDD have been drawn via asymptotic performance analysis under parameter mismatch. It is important to note that when adaptive algorithms are employed to compute the predictor coefficients directly, the mismatch between estimated predictor coefficients and true coefficients may not always be modeled in terms of mismatch in fading correlation and SNR. Extension of our results to such cases is an avenue for further research.

Notation: $\mathbf{I}_P$ denotes identity matrix of size $P \times P$. The operator $\mathrm{vec}(\cdot)$, e.g., $\mathbf{x}_n = \mathrm{vec}(X_n)$, denotes stacking of the columns of matrix X_n in column vector $\mathbf{x}_n$. $(\cdot)^*$ denotes conjugate transposition, $\otimes$ denotes the Kronecker product, $\mathrm{tr}(\cdot)$ denotes the trace operator, $\det(\cdot)$ is the determinant, and $\Re(\cdot)$ is the extraction of the real valued component. $\prod_{j=k_l}^{k_u} A_j = A_{k_l} A_{k_l+1}, \ldots, A_{k_u}$ if $k_u \geq k_l$, otherwise it denotes identity matrix of appropriate size.

II. System Model

We consider the system model

$$X_n = \sqrt{\frac{\rho}{M}} S_n H_n + W_n \tag{1}$$

Manuscript received August 6, 2002; revised April 8, 2003, July 23, 2003; accepted August 1, 2003. The editor coordinating the review of this paper and approving it for publication is X. Tian. Part of this work was presented at the IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, March 2003.

B. Bhukania is with Texas Instruments Pvt. Ltd., Bangalore 560093, India (e-mail: bhukania@ti.com) and also with the Department of Electrical Engineering, The Ohio State University, Columbus, OH 43210 USA.

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, schniter.1@osu.edu).

Digital Object Identifier 10.1109/TWC.2004.834703

where X_n is the $M \times P$ received matrix during the nth matrix-symbol interval, and where M and P are the number of transmit and receive antennas, respectively. H_n is the $M \times P$ multiple input multiple output (MIMO) channel response matrix during the nth matrix-symbol interval, containing i.i.d. unit variance proper complex Gaussian entries. S_n is the nth $M \times M$ transmitted matrix-symbol, encoded as $S_n = V_{z_n} S_{n-1}$. $z_n \in \mathcal{L} = \{0, 1, \ldots, 2^{\eta M} - 1\}$ is the time-n integer index into matrix alphabet $\mathcal{A}$ of size $2^{\eta M}$, so that $V_{z_n} \in \mathcal{A}$. Thus, η is the number of bits per channel use. S_n and V_{z_n} are unitary for all n, W_n is a matrix of i.i.d. unit variance proper complex Gaussian entries, and ρ is the average SNR per receive antenna.

Note that the system model (1) assumes that the channel H_n is fixed for M signaling intervals within the nth matrix-symbol interval, i.e., the channel is block-fading. However, for the special case of diagonal codes, (1) can be shown to be a valid system model even in a continuous fading channel. In this case, the kth row of H_n is the kth row of $H_{k,n}$, where $H_{k,n}$ is the MIMO channel response matrix at the kth time instant within the nth matrix-symbol interval, i.e., at the $(nM+k)$th channel use. Furthermore, if the MIMO fading process $H_{k,n}$ is independent between antennas, then H_n will be independent between antennas. We note that (1) is an approximate model when nondiagonal codes are used in continuous fading.

III. DFDD

DFDD can be derived in two ways. Using m-DFDD to denote a DFDD that incorporates $m-1$ past decisions, m-DFDD can be derived from the m-symbol ML differential detection rule [3], [13], [14] by setting the past $m-1$ symbol hypotheses equal to their previously-detected values. For DUST, m-DFDD takes the form [15]

$$\hat{z}_n = \arg \max_{z_n \in \mathcal{L}} \Re \left[\mathrm{tr} \left\{ \sum_{k=0}^{m-1} a_{0,k+1}^{(m)} X_n^* \mathcal{V}_n^k X_{n-k-1} \right\} \right] \tag{2}$$

where $\mathcal{V}_n^k = V_{z_n} \prod_{j=1}^{k} V_{\hat{z}_{n-j}}$ such that $\{\hat{z}_k\}_{k=n-m+1}^{n-1}$ are the previously detected symbols, and the coefficient $a_{i,j}^{(m)}$ can be found at the ith row and jth column of $-T^{(m)^{-1}}$

$$T^{(m)} = \mathbf{I}_{m+1} + \frac{\rho}{M} \begin{bmatrix} \zeta_0 & \cdots & \zeta_m \\ \vdots & \ddots & \vdots \\ \zeta_m^* & \cdots & \zeta_0 \end{bmatrix}. \tag{3}$$

In (3), ζ_k is defined such that $\mathrm{E}[\mathbf{h}_n \mathbf{h}_{n-k}^*] = \zeta_k \mathbf{I}_{\mathrm{MP}}$ for $\mathbf{h}_k = \mathrm{vec}(H_k)$.

DFDD can also be derived from quasi-coherent detection based on minimum mean square error (MMSE) channel prediction [4], [8], [16]. Here, we present a summary of the derivations in [8] that will establish some notation used later in this letter. Assuming for the moment that both $\{X_k\}_{k=n-m}^{n-1}$ and $\{S_k\}_{k=n-m}^{n-1}$ are known, the MMSE estimate of $\mathbf{h}_n$ is given in terms of $\underline{\mathbf{x}}_{n-1} = [\mathbf{x}_{n-1}^*, \ldots, \mathbf{x}_{n-m}^*]^*$, where $\mathbf{x}_n = \mathrm{vec}(X_n)$, as

$$\hat{\mathbf{h}}_n = B^* \underline{\mathbf{x}}_{n-1} \tag{4}$$

$$\begin{aligned} B &= \sqrt{\frac{\rho}{M}} \mathcal{S}_{n-1} \left(\left(T^{(m-1)^{-1}} \mathbf{g} \right) \otimes \mathbf{I}_{\mathrm{MP}} \right) \\ \mathbf{g} &= [\zeta_1 \quad \zeta_2 \quad \cdots \quad \zeta_m]^* \\ \mathcal{S}_{n-1} &= \mathrm{diag}(\mathbf{I}_P \otimes S_{n-1}, \ldots, \mathbf{I}_P \otimes S_{n-m}) \end{aligned} \tag{5}$$

which can be simplified to yield

$$\hat{H}_n = \xi \sum_{k=0}^{m-1} a_{0,k+1}^{(m)} S_{n-k-1}^* X_{n-k-1} \tag{6}$$

$$\xi = \left(\sqrt{\frac{M}{\rho}} + \sqrt{\frac{\rho}{M}} \left(1 - \frac{\rho}{M} \mathbf{g}^* T^{(m-1)^{-1}} \mathbf{g} \right) \right).$$

Defining

$$\tilde{\mathbf{h}}_n = \mathbf{h}_n - \hat{\mathbf{h}}_n \tag{7}$$

it can be shown that

$$\mathbf{x}_n = \sqrt{\frac{\rho}{M}} (\mathbf{I}_P \otimes S_n) \hat{\mathbf{h}}_n + \underbrace{\sqrt{\frac{\rho}{M}} (\mathbf{I}_P \otimes S_n) \tilde{\mathbf{h}}_n + \mathbf{w}_n}_{\tilde{\mathbf{w}}_n} \tag{8}$$

where $\mathbf{w}_n = \mathrm{vec}(W_n)$. Since it can also be shown that $\mathrm{E}[\tilde{\mathbf{w}}_n \tilde{\mathbf{w}}_n^*] = \sigma_{\tilde{w}}^2 \mathbf{I}_{\mathrm{MP}}$ and $\mathrm{E}[\hat{\mathbf{h}}_n \tilde{\mathbf{w}}_n^*] = \mathbf{0}$, (8) implies that the ML detection of z_n given $(\mathbf{I}_P \otimes S_{n-1}) \hat{\mathbf{h}}_n$ can be accomplished via

$$\hat{z}_n = \arg \min_{z_n \in \mathcal{L}} \left\| \mathbf{x}_n - \sqrt{\frac{\rho}{M}} (\mathbf{I}_P \otimes V_{z_n}) (\mathbf{I}_P \otimes S_{n-1}) \hat{\mathbf{h}}_n \right\|^2. \tag{9}$$

Under the assumption of correct past decisions, i.e., $\{\hat{z}_k = z_k\}_{k=n-m+1}^{n-1}$, it is possible to verify that (9) is, in fact, equivalent to (2).

It is important to note from (6) that, while computation of $\hat{H}_n$ requires knowledge of the (unknown) transmitted symbols $\{S_k\}_{k=n-m}^{n-1}$, computation of $(\mathbf{I}_P \otimes S_{n-1}) \hat{\mathbf{h}}_n = \mathrm{vec}(S_{n-1} \hat{H}_n)$ requires only the past information symbols $\{V_{z_k}\}_{k=n-m+1}^{n-1}$, allowing quasi-coherent detection of V_{z_n} using (2) or (9) via the error-free past decisions $\{\hat{V}_{z_k}\}_{k=n-m+1}^{n-1}$. In practice, of course, the past decisions might contain errors, leading to suboptimal performance (and possibly error propagation).

IV. Detection Error Performance

We have seen that m-DFDD requires knowledge of ρ—henceforth termed "coherent SNR"[1]—and the fading correlations $\{\zeta_k\}_{k=0}^{m}$. In this section, we derive exact and Chernoff bound expressions for PWEP when the receiver has imperfect knowledge of these parameters. The PWEP expressions are later used to approximate the BER.

We consider the case where the DFDD is designed in accordance with the "assumed" quantities $\rho_a, T_a^{(m)}, \mathbf{g}_a$, and B_a which correspond (but whose values may differ from) the true quantities $\rho, T^{(m)}, \mathbf{g}$, and B defined in Section III. Thus, the linear estimator obeys $B_a = \sqrt{\rho_a / M} \mathcal{S}_{n-1} ((T_a^{(m-1)^{-1}} \mathbf{g}_a) \otimes \mathbf{I}_{\mathrm{MP}})$ and the channel estimate becomes $\hat{\mathbf{h}}_n = B_a^* \underline{\mathbf{x}}_{n-1}$. The

[1]Note ρ is the true SNR under perfect channel state information.

quantities $\tilde{\mathbf{h}}_n$ and $\tilde{\mathbf{w}}_n$ obey (7) and (8) but with the possibly mismatched estimate $\hat{\mathbf{h}}_n$.

Using $\underline{\mathbf{x}}_{n-1} = \sqrt{(\rho/M)}\mathcal{S}_{n-1}\underline{\mathbf{h}}_{n-1} + \underline{\mathbf{w}}_{n-1}$, $E[\underline{\mathbf{x}}_{n-1}\underline{\mathbf{x}}_{n-1}^*] = \mathcal{S}_{n-1}(T^{(m-1)} \otimes \mathbf{I}_{\mathrm{MP}})\mathcal{S}_{n-1}$, and $E[\underline{\mathbf{h}}_{n-1}\mathbf{h}_n^*] = \mathbf{g} \otimes \mathbf{I}_{\mathrm{MP}}$, where $\underline{\mathbf{h}}_{n-1} = [\mathbf{h}_{n-1}^*, \ldots, \mathbf{h}_{n-m}^*]^*$, and $\underline{\mathbf{w}}_{n-1} = [\mathbf{w}_{n-1}^*, \ldots, \mathbf{w}_{n-m}^*]^*$, it can be shown that, $E[\hat{\mathbf{h}}_n\hat{\mathbf{h}}_n^*] = \hat{\sigma}_h^2\mathbf{I}_{\mathrm{MP}}, E[\tilde{\mathbf{w}}_n\tilde{\mathbf{w}}_n^*] = \tilde{\sigma}_w^2\mathbf{I}_{\mathrm{MP}}$, and $E[\tilde{\mathbf{w}}_n\hat{\mathbf{h}}_n^*(\mathbf{I}_P \otimes S_{n-1}^*)] = \sigma_{wh}^2(\mathbf{I}_P \otimes V_{z_n})$, where

$$\hat{\sigma}_h^2 = \frac{\rho_a}{M}\left(\mathbf{g}_a^* T_a^{(m-1)^{-1}} T^{(m-1)} T_a^{(m-1)^{-1}}\mathbf{g}_a\right) \quad (10)$$

$$\tilde{\sigma}_w^2 = 1 + \frac{\rho}{M}\left(1+\hat{\sigma}_h^2\right) - \frac{2\sqrt{\rho^3\rho_a}}{M^2}\Re\left(\mathbf{g}^* T_a^{(m-1)^{-1}}\mathbf{g}_a\right) \quad (11)$$

$$\sigma_{wh}^2 = \sqrt{\frac{\rho}{M}}\left(\frac{\sqrt{\rho\rho_a}}{M}\left(\mathbf{g}^* T_a^{(m-1)^{-1}}\mathbf{g}_a\right) - \hat{\sigma}_h^2\right). \quad (12)$$

Now, defining $\breve{\mathbf{x}}_n = \mathbf{x}_n/\tilde{\sigma}_w, \breve{\mathbf{h}}_n = (\mathbf{I}_P \otimes S_{n-1})\hat{\mathbf{h}}_n/\hat{\sigma}_h$, and $\breve{\mathbf{w}}_n = \tilde{\mathbf{w}}_n/\tilde{\sigma}_w$, we can rewrite (8) as

$$\breve{\mathbf{x}}_n = \sqrt{\frac{\breve{\rho}}{M}}(\mathbf{I}_P \otimes V_{z_n})\breve{\mathbf{h}}_n + \breve{\mathbf{w}}_n \quad (13)$$
$$\breve{\rho} = \rho\hat{\sigma}_h^2/\tilde{\sigma}_w^2$$

such that $E[\breve{\mathbf{h}}_n\breve{\mathbf{h}}_n^*] = E[\breve{\mathbf{w}}_n\breve{\mathbf{w}}_n^*] = \mathbf{I}_{\mathrm{MP}}$ and $E[\breve{\mathbf{w}}_n\breve{\mathbf{h}}_n^*] = \sigma_{wh}^2/(\hat{\sigma}_h\tilde{\sigma}_w)\mathbf{I}_P \otimes V_{z_n}$.

A. Exact PWEP

Given that the symbol V_1 was sent, the receiver will detect V_2, and thus make a decision error, if

$$\left\|\breve{\mathbf{x}}_n - \sqrt{\frac{\breve{\rho}}{M}}(\mathbf{I}_P \otimes V_2)\breve{\mathbf{h}}_n\right\|^2 < \left\|\breve{\mathbf{x}}_n - \sqrt{\frac{\breve{\rho}}{M}}(\mathbf{I}_P \otimes V_1)\breve{\mathbf{h}}_n\right\|^2$$
$$\Leftrightarrow Q = [\mathbf{y}_1^*\mathbf{y}_2^*]\underbrace{\begin{bmatrix}\mathbf{I}_{\mathrm{MP}} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{\mathrm{MP}}\end{bmatrix}}_{K}\underbrace{\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}}_{\mathbf{y}} < 0 \quad (14)$$

where $\mathbf{y}_1 = \sqrt{(\breve{\rho}/M)}(\mathbf{I}_P \otimes (V_1 - V_2))\breve{\mathbf{h}}_n + \breve{\mathbf{w}}_n$ and $\mathbf{y}_2 = \breve{\mathbf{w}}_n$, and where the PWEP $P_{1,2} = \Pr(V_1 \to V_2)$ is given by [17]

$$P_{1,2} = \Pr(Q \le 0) = \sum_{\substack{\text{poles } \omega = jp\\ p>0}} \mathrm{Res}\left[-\frac{\Phi_Q(\omega)}{\omega}\right]. \quad (15)$$

In (15), the summation is taken over the poles in the upper half plane (UHP) and $\Phi_Q(\omega) = E[e^{j\omega Q}]$. The characteristic function of Q, a Hermitian quadratic of a Gaussian vector, is given by (see Appendix I for a proof)

$$\Phi_Q(\omega) = \prod_{k=1}^{M}\left[\frac{\breve{\rho}\sigma_k^2}{M}(1-\tau^2)(\omega - jp_k^+)(\omega - jp_k^-)\right]^{-P} \quad (16)$$

$$p_k^{\pm} = \frac{1}{2b_k}\left[a_k \pm \sqrt{a_k^2 + 4b_k}\right] \quad (17)$$

$$a_k = \left(\frac{\breve{\rho}}{M} + \tau\sqrt{\frac{\breve{\rho}}{M}}\right)\sigma_k^2 \quad (18)$$

$$b_k = \frac{\breve{\rho}\sigma_k^2}{M}(1-\tau^2) \quad (19)$$

$$\tau = \frac{\sigma_{wh}^2}{\hat{\sigma}_h\tilde{\sigma}_w} \quad (20)$$

where σ_k is the kth singular value of $V_1 - V_2$. Note that the characteristic function $\Phi_Q(\omega)$ and, hence, the PWEP, depend on the signal only through the singular values of $V_1 - V_2$. Since the singular values of $V_1 - V_2$ and $\mathbf{I}_M - V_2V_1^*$ are the same, $\Pr(V_1 \to V_2) = \Pr(\mathbf{I}_M \to V_2V_1^*)$.

Computation of the PWEP using (15) involves taking residues at poles with multiplicities greater than 1, which is inconvenient. An alternative method to evaluate the PWEP in such cases was proposed in [18], where the poles are perturbed by small amount to eliminate multiplicity, and the PWEP is computed by taking residues at *all* the simple poles in the UHP. This method produces a lower bound on the PWEP if all the concerned poles are moved away from origin, and an upper bound when moved toward the origin. In this letter, the ith occurrence of p_k^+ is replaced by $\tilde{p}_{(k-1)Pm_k+i}^+ = p_k^+ + (i-1)\epsilon_k$, yielding the set of simple UHP poles $\{\tilde{p}_k^+\}_{k=1}^{\mathrm{MP}}$. Here, m_k denotes the multiplicity of pole p_k^+. From (15), this gives

$$P_{1,2} \gtrless \sum_{k=1}^{\mathrm{MP}}\frac{1}{\tilde{p}_k^+}\left(\prod_{\ell=1}^{M}\frac{M(1-\tau^2)^{-1}}{\breve{\rho}\sigma_\ell^2(\tilde{p}_k^+ - p_k^-)}\right)^P \prod_{\substack{\ell=1\\ \ell\neq k}}^{\mathrm{MP}}\frac{1}{(\tilde{p}_\ell^+ - \tilde{p}_k^+)} \quad (21)$$

where an upper bound is obtained by choosing $\epsilon_k = -0.0025p_k^+$, and a lower bound by choosing $\epsilon_k = 0.0025p_k^+$. Numerical results in Section VI confirm that these bounds are very close to each other, and so this method produces an accurate estimate of the PWEP.

B. Chernoff Bound

The Chernoff bound on PWEP is especially useful because from it we will be able to infer an expression for the performance loss due to parameter mismatch.

Theorem 1: The Chernoff upper bound on $\Pr(V_1 \to V_2)$ is given by

$$P_{1,2} \le \frac{1}{2}\prod_{k=1}^{M}\left[1 + \frac{\sigma_k^2}{4(1-\tau^2)}\left(\sqrt{\frac{\breve{\rho}}{M}} + \tau\right)^2\right]^{-P}. \quad (22)$$

See Appendix II for a proof.

Equation (22) implies that the diversity advantage of the system is MP, while the performance is governed by the "equivalent SNR" $(\sqrt{\breve{\rho}/M} + \tau)^2/(1-\tau^2)$. The equivalent SNR in the absence of mismatch is defined as

$$\breve{\rho}_{\mathrm{perfect}} = \left.\frac{(\sqrt{\breve{\rho}/M} + \tau)^2}{1-\tau^2}\right|_{\rho_a=\rho, T_a^{(m)}=T^{(m)}, \mathbf{g}_a=\mathbf{g}.}$$

To analyze the performance degradation due to parameter mismatch, we define the "equivalent SNR-loss" as

$$\alpha = \breve{\rho}_{\text{perfect}} \left(\frac{(\sqrt{\breve{\rho}/M} + \tau)^2}{1 - \tau^2} \right)^{-1}. \tag{23}$$

C. Approximate BER

BER is typically a more useful metric than PWEP. Using the properties that $\Pr(V_1 \rightarrow V_2) = \Pr(\mathbf{I}_M \rightarrow V_2 V_1^*)$, that $V_1, V_2 \in \mathcal{A} \Rightarrow V_2 V_1^* \in \mathcal{A}$, and that ηM bits are encoded in each transmitted matrix-symbol, the assumptions of Gray mapping and equal prior probabilities yield the following BER approximation:

$$P_{\text{genie}} \approx \frac{1}{\eta M} \sum_{j=1}^{2^{\eta M}-1} d(\mathbf{I}_M, V_j) \Pr(\mathbf{I}_M \rightarrow V_j). \tag{24}$$

In (24), $d(V_j, V_k)$ is the Hamming distance between the binary representations of V_j and V_k. The bound $P_{\text{genie}}^{\text{Chernoff}}$ is obtained when (22) is applied to (24), and the bounds $P_{\text{genie}}^{(U)}$ and $P_{\text{genie}}^{(L)}$ are obtained when (21) is applied to (24).

For realizable m-DFDD with $m > 1$, the influence of incorrect past-decisions must be taken into account. Through numerical evaluation we found that the BER of realizable DFDD is approximately twice that of genie-aided DFDD, which is in accordance with the standard DPSK approximation [19].

V. Asymptotic Prediction Error Analysis

Section III showed that the DFDD can be viewed as a linear channel predictor followed by a quasi-coherent ML detector. Thus, we expect DFDD performance to be a function of the prediction error variance. This notion can be verified by examining the Chernoff PWEP bound (22) under reasonably large values of ρ. Since $\breve{\rho}$ is proportional to ρ and $|\tau| \leq 1$, large ρ implies $\sqrt{(\breve{\rho}/M)} \gg \tau$ and so (22) can be approximated by

$$P_{1,2} \leq \frac{1}{2} \prod_{k=1}^{M} \left[1 + \frac{\sigma_k^2}{4M(1-\tau^2)} \frac{\rho \hat{\sigma}_h^2}{\tilde{\sigma}_w^2} \right]^{-P}. \tag{25}$$

In (25), $\tilde{\sigma}_w^2$ is the power of the prediction error plus the additive Gaussian noise [recall (8) and (11)]. Examination of the behavior of $\tilde{\sigma}_w^2$ under various forms of parameter mismatch will yield insight into the detection performance of m-DFDD.

We note that, since the MIMO channel fading coefficients are spatially white and the additive noise is spatio-temporally white, the prediction of every channel coefficient can be performed independently of the other channel coefficients. In other words, the MIMO channel predictor can be implemented using MP decoupled and identical single input single output (SISO) channel predictors [8], [15] such that the effective Doppler spread of each corresponding SISO channel is $M f_D T_s$, where $f_D T_s$ is the actual Doppler spread of the MIMO channel. Therefore, the overall MIMO channel prediction error can be shown to be the sum of the prediction errors of individual coefficients.

A. Relationship of Fading Spectrum to Predictor Response

It is well known that the problem of linear prediction can be modeled as a problem of spectral matching [20]. From (6), we see that the channel estimator embedded in the m-DFDD can be described as a filter with impulse response $\{a_{0,k}^{(m)}\}_{k=1}^m$, output $\hat{H}_n$, and (scaled) input $Y_n = \sqrt{(M/\rho)} S_n^* X_n = H_n + \sqrt{(M/\rho)} S_n^* W_n$. For brevity, we denote $\hat{W}_n = \sqrt{(M/\rho)} S_n^* W_n$. Although in Section III we derived $\{a_{0,k}^{(m)}\}_{k=1}^m$ by minimizing $E[\|\hat{H}_n - H_n\|^2]$, it can be shown that these predictor coefficients also minimize $E[\|\hat{H}_n - (H_n + \hat{W}_n)\|^2]$ because $\{\hat{W}_n\}$ is spatio-temporally white. It is well known that prediction error becomes asymptotically white as predictor length grows to infinity [20], and so the coefficients $\{1, -a_{0,1}^{(m)}, \ldots, -a_{0,m}^{(m)}\}$ whiten the sequence $\{H_n + \hat{W}_n\}$ when $m \rightarrow \infty$. Thus, it can be shown that

$$\lim_{m \rightarrow \infty} \left| 1 - \sum_{k=1}^{m} a_{0,k}^{(m)} e^{-j\omega k} \right|^2 = \frac{\tilde{\sigma}_w^2}{S(\omega)} \tag{26}$$

where $S(\omega)$ denotes the power spectral density (PSD) of $\{H_n + \hat{W}_n\}$ and $\tilde{\sigma}_w^2$ is defined in (11).

Next, we consider the relationship between the predictor and the channel fading spectrum from a different point of view. Since the predictor input equals $\{H_n + \hat{W}_n\}$, the predictor will attempt a one-step forward prediction of $\{H_n\}$ in the presence of additive noise $\{\hat{W}_k\}$. Recall that $\{H_k\}$ is typically a low-pass random process whose bandwidth is defined by the Doppler spread of the channel [21] and $\{\hat{W}_k\}$ is spatio-temporally white. We expect that the passband width of the optimal linear predictor will be commensurate with the bandwidth of $\{H_n\}$ modulated somewhat by the presence of noise; as the noise power increases the predictor bandwidth will decrease relative to the desired-process bandwidth. This can be verified in Fig. 1, which compares the channel fading spectrum with the predictor frequency response for a SISO channel (i.e., $M = 1$) and $f_D T_s = 0.1, \rho = 20$ dB, and $m = 100$ (which approximates $m \rightarrow \infty$).

Fig. 1(a) shows an important difference between over- and under-estimation of Doppler frequency. The primary effect of Doppler *over*-estimation is an increase in predictor bandwidth, so that a disproportionate amount of noise $\{\hat{W}_k\}$ is collected in forming the channel estimate. Since the increase in prediction error will be proportional to SNR ρ, it should become less severe as ρ increases. In fact, we are able to prove (for large m) that Doppler over-estimation does not lead to an error floor as $\rho \rightarrow \infty$ (see Section V-B). Furthermore, Fig. 1(b) leads us to conjecture that Doppler over-estimation can be somewhat compensated for by simultaneous SNR under-estimation. Finally, we expect that simultaneous over-estimation of both Doppler and SNR would be especially degrading. These trends are confirmed by Figs. 3 and 4 and discussed further in Section VI.

The primary effect of Doppler *under*-estimation is a decrease in predictor bandwidth, causing increased and suboptimal attenuation of the desired signal $\{H_n\}$. Though this behavior can be somewhat offset by simultaneously over-estimating the SNR (which widens the predictor bandwidth), the desired signal $\{H_n\}$ will remain distorted by the predictor even as $\rho \rightarrow \infty$. In

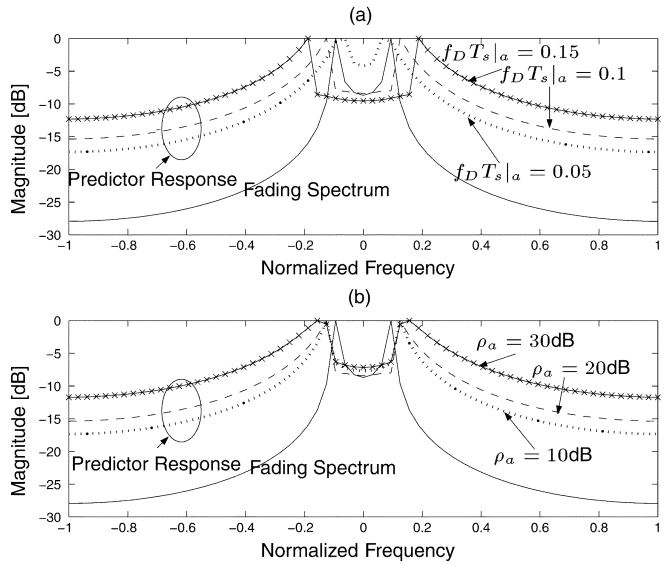

Fig. 1. (m-windowed) Channel fading spectrum and predictor frequency response $|\sum_{k=0}^{m-1} a_{0,k+1}^{(m)} e^{-j\omega k}|^2$ for $m = 100, f_D T_s = 0.1, \rho = 20$ dB. (a) $f_D T_s|_a \in \{0.15, 0.1, 0.05\}, \rho_a = 20$ dB. (b) $f_D T_s|_a = 0.1, \rho_a \in \{10, 20, 30\}$ dB.

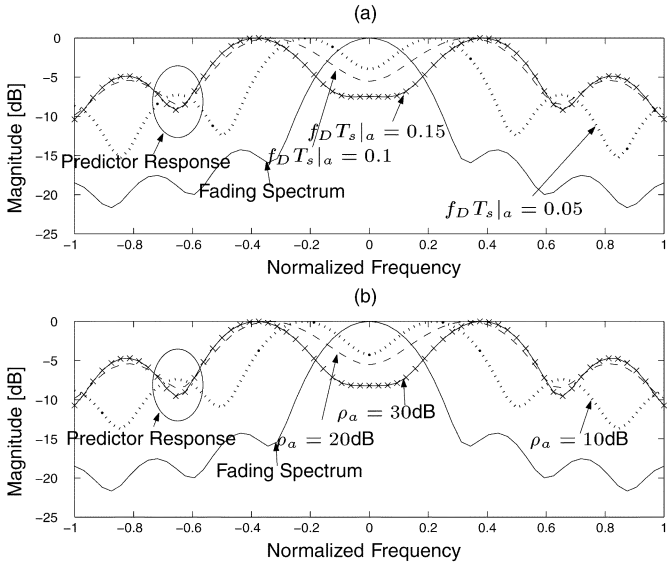

Fig. 2. (m-windowed) Channel fading spectrum and predictor frequency response $|\sum_{k=1}^{m} a_{0,k}^{(m)} e^{-j\omega k}|^2$ for $m = 6, f_D T_s = 0.1, \rho = 20$ dB. (a) $f_D T_s|_a \in \{0.15, 0.1, 0.05\}, \rho_a = 20$ dB. (b) $f_D T_s|_a = 0.1, \rho_a \in \{10, 20, 30\}$ dB.

fact, we are able to prove (for large m) that Doppler under-estimation leads to an error floor as $\rho \to \infty$ (see Section V-B). Finally, we expect that simultaneous under-estimation of both Doppler and SNR would be especially degrading. These trends are confirmed by our numeric results and discussed further in Section VI.

The effect of Doppler spread and SNR mismatch is less pronounced when the predictor/DFDD length m is small. Observe in Fig. 2 that, for $m = 6$, the predictor magnitude response lacks significant stopband attenuation and sharp transitions, thereby reducing the effects of Doppler and SNR mismatch relative to the $m = 100$ case in Fig. 1. In particular, *over*-estimation of Doppler and SNR has a minimal effect on the response. Again, effects of Doppler and SNR mismatch are similar and, therefore,

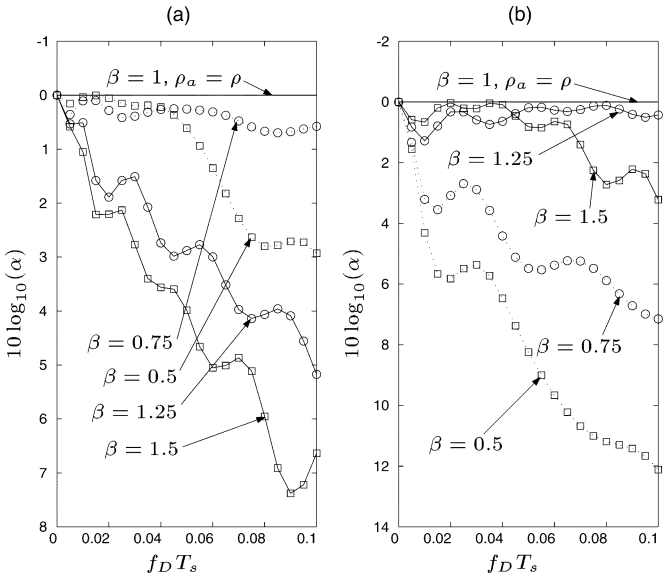

Fig. 3. SNR-loss α for $f_D T_s|_a = \beta f_D T_s, \rho = 20$ dB. (a) $\rho_a = 30$ dB. (b) $\rho_a = 10$ dB.

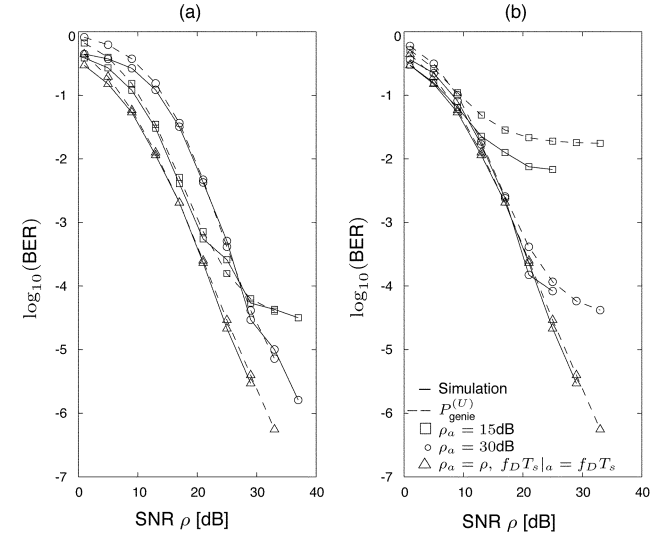

Fig. 4. Genie-aided 6-DFDD in continuous fading with $f_D T_s = 0.075$. (a) $f_D T_s|_a = 0.1$. (b) $f_D T_s|_a = 0.05$.

can be used to counter each other. These notions are confirmed by Fig. 5 and discussed further in Section VI.

B. Prediction Error When $m \to \infty$

In this section, we analyze the prediction error performance as $m \to \infty$. We are especially interested in the limiting performance of DFDD in the case where the Doppler spread is incorrectly assumed by the receiver. The presence of an error floor under Doppler mismatch is investigated by examining the prediction error as SNR $\rho \to \infty$.

Here, we consider the specific case of Rayleigh fading. The results could, however, be easily extended to other types of fading. The power spectrum of a Rayleigh-fading channel coefficient with normalized Doppler spread $d = M f_D T_s$ is given by [21]

$$\tilde{S}_d(\omega) = \begin{cases} \frac{2}{\pi\sqrt{d^2-\omega^2}}, & 0 \le |\omega| \le d \\ 0, & d < |\omega| \le 1 \end{cases} \tag{27}$$

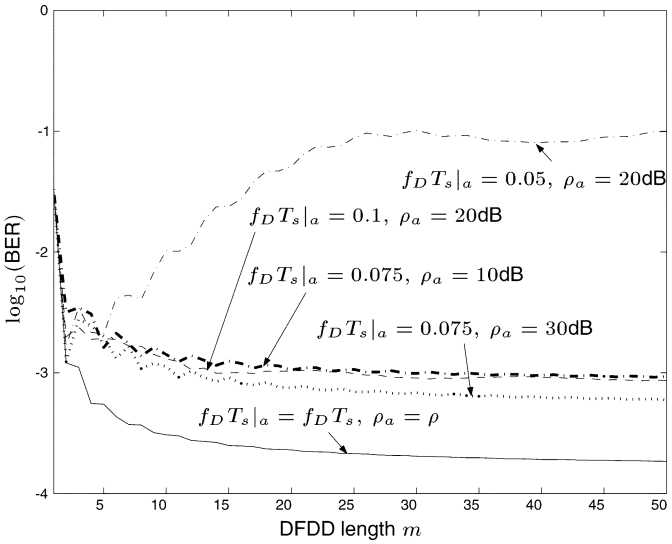

Fig. 5. Genie-aided m-DFDD in continuous fading with $f_D T_s = 0.075$ and $\rho = 20$ dB.

where ω denotes normalized frequency ($0 \leq |\omega| \leq 1$). From (26) we know that, as $m \to \infty$, the spectrum of the whitening filter $\{1, -a_{0,1}^{(m)}, \ldots, -a_{0,m}^{(m)}\}$ approaches a scaled version of the inverse of $S_d(\omega)$, the spectral density of $\{H_k + \hat{W}_k\}$:

$$S_d(\omega) = \tilde{S}_d(\omega) + M/\rho. \tag{28}$$

We denote the asymptotic (i.e., $m \to \infty$) prediction error of a SISO channel for true Doppler spread d_t and assumed Doppler spread d_a as $G(d_t, d_a)$, i.e., $\lim_{m\to\infty} \tilde{\sigma}_w^2 = G(d_t, d_a)$. By definition, the $d_t = d_a$ case gives [20]

$$G(d_a, d_a) = \exp\left(\int_0^1 \log S_{d_a}(\omega)\, d\omega\right). \tag{29}$$

Note that the total prediction error for the $M \times P$ MIMO channel is $MPG(d_t, d_a)$. It has been shown for the scalar DFDD in [4] that $G(d_a, d_a)$ is proportional to $\rho^{d_a - 1}$, from which it follows that $\lim_{\rho\to\infty} G(d_a, d_a) = 0$. The theorem below allows us to infer $\lim_{\rho\to\infty} G(d_t, d_a)$ for $d_t \neq d_a$.

Theorem 2: For $d_t = M f_D T_s$ denoting the true Doppler spread and $d_a = M f_D T_s|_a$ denoting the assumed Doppler spread, the prediction error ratio has the following behavior as SNR $\rho \to \infty$:

$$\frac{G(d_t, d_a)}{G(d_a, d_a)} \overset{\rho\to\infty}{\longrightarrow} \begin{cases} \mathcal{E}\left(\frac{\pi}{2}, \frac{d_t^2}{d_a^2}\right) d_a + 1 - d_a, & d_a \geq d_t \\ \mathcal{E}\left(\sin^{-1}\left(\frac{d_a}{d_t}\right), \frac{d_t^2}{d_a^2}\right) d_a + 1 - d_a \\ \quad + \frac{2\rho}{\pi M}\left(\frac{\pi}{2} - \sin^{-1}\left(\frac{d_a}{d_t}\right)\right), & d_a < d_t \end{cases} \tag{30}$$

where

$$\mathcal{E}(\theta, k) = \int_0^\theta \sqrt{1 - k\sin^2 x}\, dx. \tag{31}$$

See Appendix III for a proof.

Since $\lim_{\rho\to\infty} G(d_a, d_a) = 0$, Theorem 2 implies that $\lim_{\rho\to\infty} G(d_t, d_a) = 0$ when $d_a \geq d_t$. In other words, over-estimating the Doppler spread in an infinite-length DFDD does not result in an error floor. When $d_a < d_t$ and ρ is large, we have $G(d_a, d_a) \sim \rho^{d_a - 1}$ and $(G(d_t, d_a)/G(d_a, d_a)) \sim \rho$ which imply $G(d_t, d_a) \sim \rho^{d_a}$. Thus, under-estimating Doppler spread in an infinite-length DFDD results in a prediction error that *increases* with SNR. In this latter case, we expect a floor in the BER versus SNR curve. These results are corroborated by numerical examples in Section VI.

VI. Simulations And Numerical Results

For numerical examples in this paper, we consider a system with $M = 2$ transmit antennas, $P = 2$ receive antennas, and the (diagonal) constellation specified in [1] for $\eta = 1$. The MIMO channel exhibits continuous Rayleigh fading [21] so that the correlation between channel coefficients k channel-uses (i.e., kT_s s) apart is $J_0(2\pi f_D T_s k)$ where $J_0(\cdot)$ denotes a zeroth-order Bessel function of the first kind. As discussed in Section II, the block-fading model (1) used throughout the paper is sufficient to describe our continuously-fading channel since we use a diagonal constellation. The choice of Rayleigh fading implies that $\zeta_k = J_0(2\pi f_D T_s M k)$ in (3).

A. Equivalent SNR-Loss

Fig. 3 shows the variations in "equivalent SNR-loss" α from (23) with respect to true Doppler frequency $f_D T_s$ for true SNR $\rho = 20$ dB, assumed Doppler $f_D T_s|_a = \beta f_D T_s$, and assumed SNR $\rho_a \in \{10, 30\}$ dB. In Fig. 3(a), we see that, when $\rho < \rho_a = 30$ dB, under-estimating the Doppler frequency (i.e., $\beta < 1$) results in lower SNR-loss than over-estimating the Doppler frequency (i.e., $\beta > 1$). In Fig. 3(b), where $\rho > \rho_a = 10$ dB, we see the opposite trend. In other words, Doppler under-estimation can be somewhat offset by SNR over-estimation, Doppler over-estimation can be somewhat offset by SNR under-estimation, but simultaneous over/under-estimation of both Doppler and SNR results in the severe performance degradation. (Recall the discussion in Section V.)

Fig. 3 is valid for continuously-fading MIMO channels with diagonal constellations or block-fading MIMO channels with general constellations.

B. BER in Continuous Fading

Fig. 4 investigates the analytical and simulated performance of 6-DFDD when $f_D T_s = 0.075$ and the receiver has imperfect knowledge of $f_D T_s$ and ρ. The performance of the detectors with perfect knowledge of $f_D T_s$ and ρ is compared to the performance of detectors with assumed $f_D T_s|_a \in \{0.075, 0.05\}$ and $\rho_a \in \{15, 30\}$ dB. In all cases, error propagation can be observed as the performance difference between genie-aided and realizable 6-DFDD. Note from Fig. 4 that the approximations of BER, $P_{\text{genie}}^{(U)}$, and $P_{\text{genie}}^{(L)}$ derived in Section IV-C are very close to each other as well as to the simulated BER of genie-aided 6-DFDD.

As predicted in Section V and shown in Fig. 4, the DFDD succumbs to an error floor when it underestimates the Doppler frequency and a loss in SNR when it overestimates the Doppler

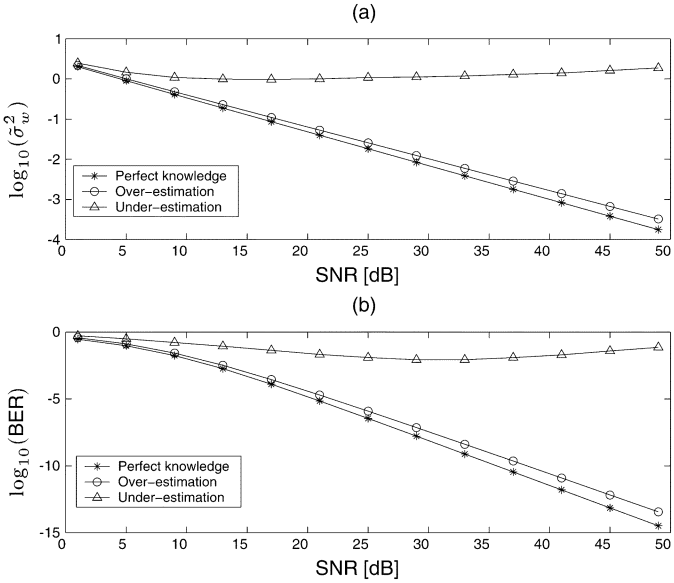

Fig. 6. (a) Prediction error $G(f_DT_s, f_DT_s|_a)$ and (b) $P_{\text{genie}}^{(U)}$ for genie-aided m-DFDD in continuous fading with $m = 100, f_DT_s = 0.075, f_DT_s|_a \in \{0.05, 0.075, 0.1\}$, and perfect knowledge of SNR.

frequency. The compensation of performance loss due to over-estimation (under-estimation) of Doppler frequency by under-estimation (over-estimation) of SNR is also depicted in Fig. 4.

Now we analyze the relation between robustness and DFDD length m. Fig. 5 plots the theoretical BER $P_{\text{genie}}^{(U)}$ of genie-aided m-DFDD versus m when the true normalized Doppler spread is $f_DT_s = 0.075$ and the SNR is $\rho = 20$ dB. Observe that the performance loss due to under-estimation of the Doppler spread is severe for large m, as predicted in Section V. While over-estimation of Doppler or misestimation of SNR also results in performance loss, it is less severe and relatively constant over all values of $m \geq 3$. Finally, it is observed that m-DFDD is quite robust against parameter mismatch when $m = 2$.

Next, we analyze the robustness of genie-aided m-DFDD as $m \to \infty$. Fig. 6 shows the variations in the prediction error $\tilde{\sigma}_w^2$ and the BER $P_{\text{genie}}^{(U)}$ versus SNR for $m = 100, f_DT_s = 0.075$ and $f_DT_s|_a \in \{0.05, 0.075, 0.1\}$. Perfect knowledge of ρ is assumed. As expected from Section V, prediction error and BER do not succumb to error floors when the Doppler spread is over-estimated. On the other hand, prediction error and BER increase with SNR when the Doppler spread is under-estimated. The adverse effects of Doppler under-estimation are especially prominent when m is large and SNR is high.

VII. Conclusion

This paper has investigated, via simulation as well as theoretical analysis, the robustness of DFDD, when used for detection of DUST and DPSK in Rayleigh-fading channels, to Doppler/SNR mismatch. It was shown that as SNR approaches infinity, under-estimation of the Doppler spread results in an error floor while over-estimation does not. For moderate values of SNR, Doppler under-estimation can be somewhat compensated by SNR over-estimation whereas Doppler over-estimation can be somewhat compensated by SNR under-estimation. Finally, it was demonstrated that increasing the DFDD length can degrade BER performance when the Doppler spread is under-estimated.

Appendix I
Proof of (16)

Since $Q = \mathbf{y}^* K \mathbf{y}$, the characteristic function of Q is [22]

$$\Phi_Q(\omega) = \frac{1}{\det(\mathbf{I}_{2MP} - j\omega R_{\mathbf{y}} K)} \tag{32}$$

where $R_{\mathbf{y}} = E[\mathbf{y}\mathbf{y}^*]$. Recall that $\mathbf{y} = [\mathbf{y}_1^* \mathbf{y}_2^*]^*$, $\mathbf{y}_1 = \sqrt{(\breve{\rho}/M)}(\mathbf{I}_P \otimes (V_1 - V_2))\breve{\mathbf{h}}_n + \breve{\mathbf{w}}_n$ and $\mathbf{y}_2 = \breve{\mathbf{w}}_n$. Therefore

$$R_{\mathbf{y}} = \begin{bmatrix} R_{\mathbf{y}}^{(1,1)} & R_{\mathbf{y}}^{(1,2)} \\ R_{\mathbf{y}}^{(2,1)} & R_{\mathbf{y}}^{(2,2)} \end{bmatrix} \tag{33}$$

where

$$R_{\mathbf{y}}^{(1,1)} = \mathbf{I}_{\text{MP}} + \left(\frac{\breve{\rho}}{M} + \tau\sqrt{\frac{\breve{\rho}}{M}}\right)(\mathbf{I}_P \otimes (2\mathbf{I}_M - V - V^*))$$

$$R_{\mathbf{y}}^{(1,2)} = \mathbf{I}_{\text{MP}} + \tau\sqrt{\frac{\breve{\rho}}{M}}(\mathbf{I}_P \otimes (\mathbf{I}_M - V^*))$$

$$R_{\mathbf{y}}^{(2,1)} = \mathbf{I}_{\text{MP}} + \tau\sqrt{\frac{\breve{\rho}}{M}}(\mathbf{I}_P \otimes (\mathbf{I}_M - V))$$

$$R_{\mathbf{y}}^{(2,2)} = \mathbf{I}_{\text{MP}}$$

and $V = V_1 V_2^*$. Using

$$\det\begin{bmatrix} A & C \\ B & D \end{bmatrix} = \det(A - BD^{-1}C)\det(D)$$

$$\det(\mathbf{I}_P \otimes A) = (\det(A))^P$$

and

$$2\mathbf{I}_M - V - V^* = (V_1 - V_2)(V_1 - V_2)^*$$

it is straightforward to simplify $\det(\mathbf{I}_{2MP} - j\omega R_{\mathbf{y}} K)$ to

$$\det(\mathbf{I}_{2MP} - j\omega R_{\mathbf{y}} K) = \det\left[1 - j\omega\left(\frac{\breve{\rho}}{M} + \tau\sqrt{\frac{\breve{\rho}}{M}}\right)\mathcal{V}_{12} + \omega^2\frac{\breve{\rho}}{M}(1-\tau^2)\mathcal{V}_{12}\right]^P$$

where $\mathcal{V}_{12} = (V_1 - V_2)(V_1 - V_2)^*$. Denoting the kth singular value of $(V_1 - V_2)$ as σ_k^2, we have

$$\det(\mathbf{I}_{2MP} - j\omega R_{\mathbf{y}} K) = \prod_{k=1}^{M}\left[1 - j\omega\left(\frac{\breve{\rho}}{M} + \tau\sqrt{\frac{\breve{\rho}}{M}}\right)\sigma_k^2 + \omega^2\frac{\breve{\rho}}{M}(1-\tau^2)\sigma_k^2\right]^P. \tag{34}$$

Using $ax^2 + bx + c = a(x - (-b + \sqrt{b^2 - 4ac})/(2a))(x + (b + \sqrt{b^2 - 4ac})/(2a))$ in (34), we get (16)–(20).

APPENDIX II
PROOF OF THEOREM 1

To prove Theorem 1, we take an approach similar to the approach in [23]. Recall that the kth pole of the characteristic function $\Phi_Q(\omega)$ is given by (17)–(19). From (16) and (17), it is observed that the region of convergence of $\Phi_Q(\omega) = E[e^{j\omega Q}]$ is $\max_k\{(a_k/2b_k) - (1/2b_k)\sqrt{a_k^2+4b_k}\} < \mathrm{Im}(\omega) < \min_k\{(a_k/2b_k) + (1/2b_k)\sqrt{a_k^2+4b_k}\}$. Noting that $c = (a_k/2b_k)$ is not a function of k, we evaluate the PWEP as

$$\begin{aligned}\Pr(Q<0) &= \frac{1}{2\pi}\int_{-\infty}^{0}\int_{-\infty+jc}^{\infty+jc} e^{-j\omega q}\Phi_Q(\omega)\,d\omega\,dq \\ &= \frac{-1}{2\pi j}\int_{-\infty+jc}^{\infty+jc}\frac{1}{\omega}\Phi_Q(\omega)\,d\omega \\ &= \frac{-1}{2\pi j}\int_{-\infty}^{\infty}\frac{1}{\omega+jc}\Phi_Q(\omega+jc)\,d\omega \qquad (35)\end{aligned}$$

where the characteristic function has been inverted by choosing the integration contour $\mathrm{Im}(\omega) = c$, which lies within the region of convergence. From (16) and (17), we get

$$\begin{aligned}\Phi_Q(\omega+jc) &= \prod_{k=1}^{M}[b_k(\omega+jc-jp_k^+)(\omega-jc-jp_k^-)]^{-P} \\ &= \prod_{k=1}^{M}\left[b_k\left(\omega-\frac{j}{2b_k}\sqrt{a_k^2+4b_k}\right)\right. \\ &\quad\left.\times\left(\omega+\frac{j}{2b_k}\sqrt{a_k^2+4b_k}\right)\right]^{-P} \\ &= \prod_{k=1}^{M}\left[1+b_k\left(\omega^2+\frac{a_k^2}{4b_k^2}\right)\right]^{-P}. \qquad (36)\end{aligned}$$

Since both $\Phi_Q(\omega+jc)$ and $\Pr(Q<0)$ are real, extracting the real part in (35), we get

$$\begin{aligned}&\Pr(Q<0) \\ &= \frac{c}{2\pi}\int_{-\infty}^{\infty}\frac{1}{\omega^2+c^2}\prod_{k=1}^{M}\left[1+b_k(\omega^2+\frac{a_k^2}{4b_k^2})\right]^{-P}d\omega \\ &\le \prod_{k=1}^{M}\left[1+\frac{a_k^2}{4b_k}\right]^{-P}\frac{c}{2\pi}\int_{-\infty}^{\infty}\frac{1}{\omega^2+c^2}\,d\omega \\ &= \frac{1}{2}\prod_{k=1}^{M}\left[1+\frac{a_k^2}{4b_k}\right]^{-P}. \qquad (37)\end{aligned}$$

Replacing $a_k = ((\breve{\rho}/M) + \tau\sqrt{(\breve{\rho}/M)})\sigma_k^2$ and $b_k = (\breve{\rho}\sigma_k^2/M)(1-\tau^2)$ in (37), we obtain (22).

APPENDIX III
PROOF OF THEOREM 2

Denoting $A(\omega) = |1 - \sum_{k=1}^{m} a_{0,k}^{(m)} e^{-j\omega k}|^2$, we can rewrite (26) as $A(\omega)S_{d_a}(\omega) = G(d_a, d_a)$. If the true Doppler spread of the channel is d_t, so that the whitening filter $A(\omega)$ is applied to an input process with PSD $S_{d_t}(\omega)$, then the PSD of the prediction error process can be written as

$$E(\omega) = A(\omega)S_{d_t}(\omega) = \frac{G(d_a,d_a)}{S_{d_a}(\omega)}S_{d_t}(\omega).$$

By definition the prediction error is $G(d_t,d_a) = \int_0^1 E(\omega)\,d\omega$, so that

$$\frac{G(d_t,d_a)}{G(d_a,d_a)} = \int_0^1 \frac{S_{d_t}(\omega)}{S_{d_a}(\omega)}\,d\omega. \qquad (38)$$

When $d_a \ge d_t$, (27), and (28) imply

$$\begin{aligned}\frac{G(d_t,d_a)}{G(d_a,d_a)} &= \underbrace{\int_{d_a}^{1} d\omega}_{I_3} + \underbrace{\int_{d_t}^{d_a}\frac{\sigma_W^2}{\frac{2}{\pi\sqrt{d_a^2-\omega^2}}+\sigma_W^2}\,d\omega}_{I_2} \\ &\quad + \underbrace{\int_{0}^{d_t}\frac{\frac{2}{\pi\sqrt{d_t^2-\omega^2}}+\sigma_W^2}{\frac{2}{\pi\sqrt{d_a^2-\omega^2}}+\sigma_W^2}\,d\omega}_{I_1} \qquad (39)\end{aligned}$$

where we use $\sigma_W^2 = M/\rho$. Clearly $I_3 = 1-d_a$. When $\sigma_W^2 = 0$, we can see that $I_2 = 0$ and

$$I_1 = \int_0^{d_t}\sqrt{\frac{d_a^2-\omega^2}{d_t^2-\omega^2}}\,d\omega = \mathcal{E}\left(\frac{\pi}{2},\frac{d_t^2}{d_a^2}\right) \qquad (40)$$

where $\mathcal{E}(\cdot,\cdot)$ is given in (31). The first case in (30) follows from (39)–(40).

When $d_a < d_t$, we have

$$\begin{aligned}\frac{G(d_t,d_a)}{G(d_a,d_a)} &= \underbrace{\int_{d_t}^{1} d\omega}_{I_6} + \underbrace{\int_{d_a}^{d_t}\frac{1}{\sigma_W^2}\left(\frac{2}{\pi\sqrt{d_t^2-\omega^2}}+\sigma_W^2\right)d\omega}_{I_5} \\ &\quad + \underbrace{\int_{0}^{d_a}\frac{\frac{2}{\pi\sqrt{d_t^2-\omega^2}}+\sigma_W^2}{\frac{2}{\pi\sqrt{d_a^2-\omega^2}}+\sigma_W^2}\,d\omega}_{I_4}. \qquad (41)\end{aligned}$$

Clearly, $I_6 = 1-d_t$. For $\sigma_W^2 = 0$ we see that

$$I_4 = \int_0^{d_a}\sqrt{\frac{d_a^2-\omega^2}{d_t^2-\omega^2}}\,d\omega = \mathcal{E}\left(\sin^{-1}\left(\frac{d_a}{d_t}\right),\frac{d_t^2}{d_a^2}\right)d_a. \qquad (42)$$

Finally,

$$\begin{aligned}I_5 &= \frac{2}{\pi\sigma_W^2}\int_{d_a}^{d_t}\frac{1}{\sqrt{d_t^2-\omega^2}}\,d\omega + \int_{d_a}^{d_t}d\omega \\ &= \frac{2}{\pi\sigma_W^2}\left(\frac{\pi}{2}-\sin^{-1}\left(\frac{d_a}{d_t}\right)\right)+(d_t-d_a). \qquad (43)\end{aligned}$$

The second case in (30) follows from (41)–(43) with $\sigma_W^2 = M/\rho$.

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers whose comments have led to improvements in the letter.

REFERENCES

[1] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec. 2000.

[2] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Informa. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[3] P. Ho and D. Fung, "Error performance of multiple symbol differential detection of PSK signals transmitted over correlated Rayleigh-fading channels," in *Proc. IEEE Int. Conf. Communications*, vol. 2, 1991, pp. 568–574.

[4] R. Schober, W. H. Gerstacker, and J. B. Huber, "Decision feedback differential detection of MDPSK for flat Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1025–1035, July 1999.

[5] C. Peel and A. Swindlehurst, "Performance of unitary space-time modulation in continuously changing channel," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 4, 2001, pp. 2433–2436.

[6] C. Ling and X. Wu, "Linear prediction receiver for differential unitary space-time modulation over time-correlated Rayleigh-fading channels," in *Proc. IEEE Int. Conf. Communications*, vol. 2, New York, 2002, pp. 788–791.

[7] L. Lampe and R. Schober, "Bit-interleaved coded differential space-time modulation," in *Proc. IEEE Int. Conf. Communications*, vol. 2, New York, 2002, pp. 1434–1438.

[8] B. Bhukania and P. Schniter, "Decision-feedback detection of differential unitary space-time modulation in fast Rayleigh-fading channels," in *Proc. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, 2002, pp. 693–702.

[9] W. C. Dam and D. P. Taylor, "An adaptive maximum likelihood receiver for correlated Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 2684–2692, Sept. 1994.

[10] C. Tepedelenlioglu, A. Abdi, G. B. Giannakis, and M. Kaveh, "Estimation of Doppler spread and signal strength in mobile communications with applications to handoff and adaptive transmission," in *Wireless Communications and Mobile Computing*. New York: Wiley, 2001, pp. 221–242.

[11] Y. Ko and G. Jeong, "Doppler spread estimation in mobile communication systems," in *Proc. IEEE Vehicular Technology Conf.*, vol. 4, 2002, pp. 1941–1945.

[12] F. Adachi, "Adaptive differential detection using linear prediction for M-ary DPSK," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 909–918, Aug. 1998.

[13] C. Cozzo, D. N. Mohile, and B. Hughes, "Multi-symbol detection of differential space-time modulation," in *Proc. Conf. Information Science Systems*, Princeton, NJ, 2000, p. WA8a-12.

[14] P. Fan, "Multiple-symbol detection for transmit diversity with differential encoding scheme," *IEEE Trans. Consum. Electron.*, vol. 47, pp. 96–100, Feb. 2001.

[15] R. Schober and H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 768–777, May 2002.

[16] E. Chiavaccini and G. Vitetta, "Further results on Tarokh's space-time differential technique," in *Proc. IEEE Int. Conf. Communications*, New York, 2002, pp. 1778–1782.

[17] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[18] 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.

[19] F. Edbauer, "Bit error rate of binary and quaternary DPSK signals with multiple differential feedback detection," *IEEE Trans. Commun.*, vol. 40, pp. 457–460, Mar. 1992.

[20] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[21] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[22] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, pp. 199–201, 1960.

[23] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.