Optimum Combining of Residual Carrier Array Signals in Correlated Noises

H. H. Tan
Communications Systems and Research Section
and
University of California, Irvine

R. Liang and P.-H. Suen[1]

An array feed combining system for the recovery of signal-to-noise ratio (SNR) loss due to antenna reflector deformation has been implemented and is currently being evaluated on the Jet Propulsion Laboratory 34-m DSS-13 antenna. The current signal-combining system operates under the assumption that the white Gaussian noise processes in the received signals from different array elements are mutually uncorrelated. However, experimental data at DSS 13 indicate that these noise processes are indeed mutually correlated. The objective of this work is to develop a signal-combining system optimized to account for the mutual correlations between these noise processes. The set of optimum combining weight coefficients that maximizes the combined signal SNR in the correlated noises environment is determined. These optimum weights depend on unknown signal and noise covariance parameters. A maximum-likelihood approach is developed to estimate these unknown parameters to obtain estimates of the optimum weight coefficients based on residual carrier signal samples. The actual combined signal SNR using the estimated weight coefficients is derived and shown to converge to the maximum achievable SNR as the number of signal samples increases. These results are also verified by simulation. A numerical example shows a significant improvement in SNR performance can be obtained, especially when the amount of correlation increases.

I. Introduction

An array feed-combining system has been proposed for the recovery of signal-to-noise ratio (SNR) losses caused by large antenna reflector deformations at Ka-band (32-GHz) frequencies in the Deep Space Network [1]. In this system, a focal plane feed array is used to collect the defocused signal fields that result from these deformations. All the signal power captured by the feed array is then recovered using real-time signal-processing and signal-combining techniques. The optimum combiner weights that maximize the combined signal SNR were derived in [1] under the assumption that the white Gaussian noise processes in the received signals from different array elements are mutually uncorrelated. These optimum weights depend on unknown signal and noise parameters that need to be estimated. The work in [1] proposed to estimate the optimum weights from the observed residual carrier received-signal samples

[1] Students at the University of California, Irvine.

using a maximum-likelihood (ML) approach. The actual combined-signal SNR in an uncorrelated noises environment when the estimated weights are used in place of the optimum weight coefficients was also derived in [1].

A seven-element array feed combiner system is currently being evaluated at the JPL DSS-13 34-m antenna. Although the work in [1] assumed mutually uncorrelated noise processes, experimental data [2] indicate that the noise processes in the received signals from different feed elements are indeed correlated, with correlation coefficients of the order of 0.01 under clear sky conditions. Since the noise in each of the array feed element signals consists of receiver white noise plus noise due to background radiation, it has been conjectured that this small correlation is caused by near-field atmospheric background noise. Recent data gathered at DSS 13 in more adverse weather conditions, however, indicate both increases and decreases in the observed amount of correlation. This experimental work is still in progress. Larger correlations may also result from undesired radiation source emissions gathered by the antenna side lobes. Moreover, our recent work [5] has shown that the array feed combining system derived in [1], which is suboptimal in the correlated noises environment, actually can have a better performance in the presence of correlated noises. Therefore, it is important to develop optimum signal-combining techniques that account for the mutual correlation between the noises in the signals from different array elements. That is the objective of this work.

As a first step towards this objective, in Section II we provide a derivation of the set of optimum combining weight coefficients that maximizes the combined-signal SNR in this correlated noises environment. These optimum weights depend on unknown signal amplitude and phase parameters as well as noise variance and correlation parameters. An ML approach is then developed in Section III to estimate these unknown parameters and arrive at an ML estimate of the optimum weight coefficients based on residual carrier received-signal samples. The actual combined-signal SNR is derived in Section IV when the estimated weights are used in place of the optimum weights, with the details given in Appendices A and B. The SNR performance is shown to converge to the maximum achievable SNR as the number of signal samples used in the estimates increases. These results are also verified by simulation. Numerical examples are given in Section V with a particular choice of noise covariance matrix and signal parameters and show a significant improvement in SNR performance compared to the previous signal-combining system developed in [1], especially when the amount of correlation increases.

II. Array Feed Signals and Optimal Combining Weights

Consider a K-element array and the NASA Deep Space Network standard residual carrier modulation with binary phase shift keyed (PSK) modulated square-wave subcarrier [4]. The received signal from each array element is downconverted to baseband and sampled. Similar to the combining system proposed in [1], only the residual carrier portion of the received signal spectrum will be used to estimate the unknown parameters in the combiner weights. The full-spectrum modulated signals from the array elements, which contain both the modulated sidebands as well as the residual carrier spectrum, are subsequently combined. As in [1], assume that the higher bandwidth primitive baseband signal samples are lowpass filtered by averaging successive blocks of M_B samples to yield a full-spectrum signal stream B for each array element. Additive white Gaussian noise is assumed to be present in the primitive baseband signal sequences from each of the array elements. The white Gaussian noises in the primitive baseband samples from different array elements are assumed to be mutually correlated. *Specifically, the noise samples corresponding to different array elements are assumed to be mutually correlated at any given time instant, but uncorrelated at different time instants.* Let

$$y_k(i_B) = V_k[\cos\delta + js(i_B)\sin\delta] + n_k(i_B), \quad i_B = 1, 2, \cdots \tag{1}$$

denote the stream B signal samples from the kth array element. The complex signal parameters $V_k, 1 \leq k \leq K$, represent the unknown signal amplitude and phase parameters induced by the antenna reflector

deformation. Moreover, δ is the modulation index, $s(i_B) = \pm 1$ is the transmitted data, and $\{n_k(i_B)\}$ is the zero-mean white Gaussian noise corruption in the stream B signal samples from the kth array element. The primitive baseband signal samples are also more narrowly lowpass filtered by averaging successive blocks of M_A samples to yield a residual carrier signal stream A for each array element. Clearly, $M_A > M_B$ and $\eta = M_A/M_B$ is the ratio of the bandwidth of stream B to stream A. Let

$$u_k(i_A) = V_k \cos\delta + m_k(i_A), \quad i_A = 1, 2, \cdots \tag{2}$$

denote the stream A signal samples from the kth array element. Here $\{m_k(i_A)\}$ is the zero-mean white Gaussian noise corruption in the stream A signal samples from the kth array element. Let A^*, A^T, and $A^\dagger$ denote the complex conjugate, the transpose, and the complex conjugate transpose of the matrix A, respectively. In order to specify the correlations between the white noise sequences corresponding to different array elements, consider the noise vectors $\underline{n}(i_B) = (n_1(i_B) \cdots n_K(i_B))^T$ and $\underline{m}(i_A) = (m_1(i_A) \cdots m_K(i_A))^T$. Then $\{\underline{n}(i_B)\}$ and $\{\underline{m}(i_A)\}$ are each sequences of independent identically distributed (i.i.d.) zero-mean complex Gaussian random vectors of dimension K. The respective covariance matrices $\underline{R}_B = \{r_{Bkj}\} = \mathbf{E}[\underline{n}(i_B)\underline{n}(i_B)^\dagger]$ and $\underline{R}_A = \{r_{Akj}\} = \mathbf{E}[\underline{m}(i_A)\underline{m}(i_A)^\dagger]$ of $\underline{n}(i_B)$ and $\underline{m}(i_A)$ specify the mutual correlations between the white noises in the signal streams from different array elements. For example, r_{Bkj} is the correlation between the noise variables $n_k(i_B)$ and $n_j(i_B)$. Because of the different averaging rates in streams A and B on the primitive baseband signals, it follows that $\underline{R}_B = \eta \underline{R}_A$. Finally, these different averaging rates also imply that $\underline{m}(i_A)$ is independent of $\underline{n}(i_B)$ provided that $i_A < i_B$ and the samples averaged to yield $\underline{m}(i_A)$ occurred prior to the samples averaged to yield $\underline{n}(i_B)$.

Application of the complex combining-weight coefficients $W_k, 1 \leq k \leq K$ yields the combiner output sequence $z_c(i_B) = s_c(i_B) + n_c(i_B)$, where

$$s_c(i_B) = \sum_{k=1}^{K} W_k V_k e^{js(i_B)\delta} \tag{3}$$

and

$$n_c(i_B) = \sum_{k=1}^{K} W_k n_k(i_B) \tag{4}$$

are the signal and noise components, respectively. Define $\underline{W} = (W_1 \cdots W_K)^T$ to be the vector of combining-weight coefficients. The objective is to determine the optimum weight vector $\underline{W}$ that maximizes the SNR of the combiner output defined by

$$\gamma(\underline{W}) = \frac{|\mathbf{E}[z_c(i_B)]|^2}{\mathbf{Var}[z_c(i_B)]} = \frac{|s_c(i_B)|^2}{\mathbf{Var}[n_c(i_B)]} \tag{5}$$

The optimum weight vector and the maximum achievable SNR have been derived previously in [3]. For the sake of completeness, we provide a derivation below that is slightly different from that in [3]. Define $\underline{V} = (V_1 \cdots V_K)^T$ to be the vector of complex signal parameters. Then, from Eqs. (3) and (4), we have $|s_c(i_B)|^2 = |\underline{W}^T\underline{V}|^2$ and $\mathbf{Var}[n_c(i_B)] = \underline{W}^T\underline{R}_B\underline{W}^*$, substitution of which in Eq. (5) yields

$$\gamma(\underline{W}) = \frac{|\underline{W}^T\underline{V}|^2}{\underline{W}^T\underline{R}_B\underline{W}^*} \tag{6}$$

Since $\underline{R}_B$ is a positive definite Hermitian matrix, there is a unitary matrix $\underline{Q}$ such that

$$\underline{R}_B = \underline{Q}^\dagger \underline{D}^2 \underline{Q} \tag{7}$$

where $\underline{D}$ is a real-valued $K \times K$ diagonal matrix with the kth diagonal term given by $\sqrt{r_{Bkk}}$. Using Eq. (7), we have

$$\begin{aligned} \underline{W}^T \underline{R}_B \underline{W}^* &= \left(\underline{W}^T \underline{Q}^\dagger \underline{D}\right) \left(\underline{D}\underline{Q}\underline{W}^*\right) \\ &= \left(\underline{D}\underline{Q}\underline{W}^*\right)^\dagger \left(\underline{D}\underline{Q}\underline{W}^*\right) = \|\underline{D}\underline{Q}\underline{W}^*\|^2 \end{aligned} \tag{8}$$

Moreover, since $\underline{Q}$ is the inverse of the matrix $\underline{Q}^\dagger$,

$$\left[\left(\underline{D}\underline{Q}^*\right)^T\right]^{-1} = \left[\underline{Q}^\dagger \underline{D}\right]^{-1} = \underline{D}^{-1}\underline{Q} \tag{9}$$

Hence, by using Eq. (9), we can write

$$\begin{aligned} \underline{W}^T \underline{V} &= \underline{W}^T \left(\underline{D}\underline{Q}^*\right)^T \left[\left(\underline{D}\underline{Q}^*\right)^T\right]^{-1} \underline{V} \\ &= \left(\underline{D}\underline{Q}\underline{W}^*\right)^\dagger \left(\underline{D}^{-1}\underline{Q}\underline{V}\right) \end{aligned} \tag{10}$$

So substituting Eqs. (8) and (10) in Eq. (6) and applying the Schwartz inequality gets

$$\begin{aligned} \gamma(\underline{W}) &= \frac{|\left(\underline{D}\underline{Q}\underline{W}^*\right)^\dagger \left(\underline{D}^{-1}\underline{Q}\underline{V}\right)|^2}{\|\underline{D}\underline{Q}\underline{W}^*\|^2} \\ &\leq \frac{\|\underline{D}\underline{Q}\underline{W}^*\|^2 \|\underline{D}^{-1}\underline{Q}\underline{V}\|^2}{\|\underline{D}\underline{Q}\underline{W}^*\|^2} \\ &= \|\underline{D}^{-1}\underline{Q}\underline{V}\|^2 \triangleq \gamma_{MAX} \end{aligned} \tag{11}$$

Moreover, equality holds in Eq. (11) if and only if for some complex-valued constant α, $\underline{D}\underline{Q}\underline{W}^* = \alpha\left(\underline{D}^{-1}\underline{Q}\underline{V}\right)$. So, by using Eq. (7), the set of optimum weight vectors $\underline{W}_{OPT}$ that achieves $\gamma(\underline{W}_{OPT}) = \gamma_{MAX}$ is given by

$$\underline{W}_{OPT} = \alpha\left(\underline{Q}^\dagger \underline{D}^{-2}\underline{Q}\right)^* \underline{V}^* = \alpha\left(\underline{R}_B^*\right)^{-1}\underline{V}^* = \frac{\alpha}{\eta}\left(\underline{R}_A^*\right)^{-1}\underline{V}^* \tag{12}$$

where α is an arbitrary complex-valued constant. Moreover, it follows from Eqs. (7) and (11) that the optimum SNR γ_{MAX} can be written as

$$\gamma_{MAX} = \left(\underline{D}^{-1}\underline{Q}\underline{V}\right)^{\dagger}\left(\underline{D}^{-1}\underline{Q}\underline{V}\right) = \underline{V}^{\dagger}\underline{R}_B^{-1}\underline{V} \tag{13}$$

Note that, in the uncorrelated noises case, $\underline{R}_B^{-1} = \underline{D}^{-2}$. So the set of optimum weight vectors, Eq. (12), in this case is given by $\underline{W}_{OPT} = \alpha\underline{D}^{-2}\underline{V}^{*}$, and the optimum SNR is given by

$$\gamma_{MAX} = \left\|\underline{D}^{-1}\underline{V}\right\|^2 = \sum_{k=1}^{K}\frac{|V_k|^2}{r_{Bkk}} \triangleq \gamma \tag{14}$$

which is the sum of the array element output SNRs. These results for the uncorrelated noises case agree with previous results derived in [1].

III. Parameter Estimation

The signal parameter vector $\underline{V}$ and the noise covariance matrix $\underline{R}_A$ are not known and need to be estimated to obtain an estimate of one of the optimum weight vectors, $\underline{W}_{OPT}$, given by Eq. (12). Assuming that these unknown parameters are not random, we propose to use ML estimates based on the stream A residual carrier-signal vector samples $\{\underline{u}(i_A)\}$, where $\underline{u}(i_A) = (u_1(i_A)\cdots u_K(i_A))^T$. Instead of estimating $\underline{V}$ directly, consider estimating $\underline{X} = \underline{V}\cos\delta$. Note from Eq. (2) that $\{\underline{u}(i_A)\}$ is an i.i.d. sequence of complex Gaussian random vectors with mean $\underline{X}$ and covariance matrix $\underline{R}_A$. It then follows from multivariate statistical analysis [6,7] that, based on observations $\{\underline{u}(i_A - 1),\cdots,\underline{u}(i_A - L)\}$,

$$\hat{\underline{X}}_{ML}(i_A) = \frac{1}{L}\sum_{l=i_A-L}^{i_A-1}\underline{u}(l) \tag{15}$$

and

$$\hat{\underline{R}}_{A,ML}(i_A) = \frac{1}{L}\sum_{l=i_A-L}^{i_A-1}\left[\underline{u}(l) - \hat{\underline{X}}_{ML}(i_A)\right]\left[\underline{u}(l) - \hat{\underline{X}}_{ML}(i_A)\right]^{\dagger} \tag{16}$$

are the respective ML estimates of $\underline{X}$ and $\underline{R}_A$. By the invariant property of ML estimators,

$$\hat{\underline{V}}_{ML}(i_A) = \frac{1}{\cos\delta}\hat{\underline{X}}_{ML}(i_A) \tag{17}$$

is then the ML estimate of $\underline{V}$ and

$$\hat{\underline{W}}_{OPT}(i_A) = \frac{\alpha}{\eta}\left(\hat{\underline{R}}^{*}_{A,ML}(i_A)\right)^{-1}\hat{\underline{V}}^{*}_{ML}(i_A) \tag{18}$$

is the ML estimate of $\underline{W}_{OPT}$ given by Eq. (12).

It is well known [6,7] that $\hat{\underline{X}}_{ML}(i_A)$ is a complex Gaussian random vector with mean $\underline{X}$ and covariance matrix $(1/L)\underline{R}_A$. Furthermore, $\hat{\underline{X}}_{ML}(i_A)$ and $\hat{\underline{R}}_{ML}(i_A)$ are statistically independent, and $L\hat{\underline{R}}_{ML}(i_A)$ has the same distribution as the random matrix

$$\underline{A} = \sum_{i=1}^{L-1} \underline{Z}_i \underline{Z}_i^{\dagger} \tag{19}$$

where $\underline{Z}_i$ is a sequence of i.i.d. zero-mean complex Gaussian random vectors with covariance matrix $\underline{R}_A$. This type of distribution is called a complex Wishart distribution with parameters $\underline{R}_A$ and $(L-1)$, and $\underline{A}$ in Eq. (19) is said to have a $CW(\underline{R}_A, L-1)$ distribution [7]. It has been shown in [8] that, for $L > K+1$,

$$\mathbf{E}[\underline{A}^{-1}] = \frac{1}{(L-1)-K} \underline{R}_A^{-1} \tag{20}$$

for a $K \times K$ $CW(\underline{R}_A, L-1)$ distributed random matrix $\underline{A}$. Since $(\underline{A}^*)^{-1} = (\underline{A}^{-1})^*$, it then follows from using the property of Eq. (20) for the complex Wishart matrix $L\underline{\hat{R}}_{ML}(i_A)$ along with Eqs. (12), (17), and (18) that, for $L > K+1$,

$$\begin{aligned}
\mathbf{E}\left[\underline{\hat{W}}_{OPT}(i_A)\right] =& \frac{\alpha}{\eta} \mathbf{E}\left[\left(\underline{\hat{R}}^*_{A,ML}(i_A)\right)^{-1}\right] \mathbf{E}\left[\underline{\hat{V}}^*_{ML}(i_A)\right] \\
=& \frac{\alpha L}{\eta(L-K-1)} \underline{R}_A^{*\,-1} \underline{V}^* \\
=& \frac{L}{L-K-1} \underline{W}_{OPT}
\end{aligned} \tag{21}$$

Hence, the ML estimate $\underline{\hat{W}}_{OPT}(i_A)$ of the optimum weight vector $\underline{W}_{OPT}$ is actually a biased estimate. This is not a problem, since it is clear from Eq. (12) that the optimum weight vectors are not unique and any complex scaled version of an optimum weight vector is also optimum. So the constant α in Eq. (18) can be set arbitrarily. We shall set $\alpha = (L-K-1)/L$ for the purpose of normalization and also assume that $L > K+1$. The ML estimate of the optimum weight vector that will be used here is, therefore,

$$\begin{aligned}
\underline{\hat{W}}_{ML}(i_A) =& \frac{L-K-1}{L\eta} \left(\underline{\hat{R}}^*_{A,ML}(i_A)\right)^{-1} \underline{\hat{V}}^*_{ML}(i_A) \\
=& \frac{L-K-1}{L\eta \cos\delta} \left(\underline{\hat{R}}^*_{A,ML}(i_A)\right)^{-1} \underline{\hat{X}}^*_{ML}(i_A)
\end{aligned} \tag{22}$$

with mean

$$\underline{W}_O \triangleq \mathbf{E}\left[\underline{\hat{W}}_{ML}(i_A)\right] = \frac{1}{\eta} \underline{R}_A^{*\,-1} \underline{V}^* = \underline{R}_B^{*\,-1} \underline{V}^* \tag{23}$$

that is also an optimum weight vector.

Note that $\underline{\hat{V}}_{ML}(i_A)$ and $\underline{\hat{R}}_{A,ML}(i_A)$ are both consistent estimates, i.e., they both converge with probability one to their respective expected values $\underline{V}$ and $\underline{R}_A$ in the limit as L tends to infinity. So it follows that $\underline{\hat{W}}_{ML}(i_A)$ also converges with probability one to the optimum weight vector $\underline{W}_O$ in the limit as the number L of samples tends to infinity. The analysis in Appendix C shows that this convergence also holds in the mean-square sense. These properties indicate that we may expect the actual combiner output SNR

using the estimated weight vector $\underline{\hat{W}}_{ML}(i_A)$ to also converge to the maximum achievable SNR γ_{MAX} as L tends to infinity. That result will be shown in the next section, which derives an explicit expression for the actual combiner output SNR for finite L.

As in [1], these weight coefficient estimates are used in a sliding-window structure to produce the following combiner output sequence:

$$z_c(i_B) = \underline{\hat{W}}_{ML}^T(\tilde{i}_A)\underline{y}(i_B) \tag{24}$$

where $\tilde{i}_A$ is the largest integer less than i_B, so that the residual carrier-signal vector samples $\{\underline{u}(\tilde{i}_A - 1), \cdots, \underline{u}(\tilde{i}_A - L)\}$ used for estimating $\underline{\hat{W}}_{ML}(\tilde{i}_A)$ occur before the full-spectrum signal vector sample $\underline{y}(i_B) = (y_1(i_B), \cdots, y_K(i_B))^T$. This ensures that the noise vector $\underline{n}(i_B)$ in $\underline{y}(i_B)$ is statistically independent of $\{\underline{u}(\tilde{i}_A - 1), \cdots, \underline{u}(\tilde{i}_A - L)\}$, and hence the statistical independence between $\underline{\hat{W}}_{ML}(\tilde{i}_A)$ and $\underline{n}(i_B)$.

IV. SNR Performance Analysis

The combiner output, Eq. (24), can be written as

$$z_c(i_B) = s_c(i_B) + n_c(i_B) \tag{25}$$

where

$$s_c(i_B) = \underline{\hat{W}}_{ML}^T(\tilde{i}_A)\,\underline{V}e^{js(i_B)\delta} \tag{26}$$

and

$$n_c(i_B) = \underline{\hat{W}}_{ML}^T(\tilde{i}_A)\,\underline{n}(i_B) \tag{27}$$

are the signal and noise components, respectively. In the following analysis, $\underline{V}$, $\underline{R}_B$, and $s(i_B)$ are assumed to be nonrandom parameters. As noted above, $\underline{\hat{W}}_{ML}(\tilde{i}_A)$ and $\underline{n}(i_B)$ are statistically independent. Since the components of $\underline{n}(i_B)$ are all of zero mean, it follows from Eqs. (26) and (27) that $n_c(i_B)$ also has zero mean and moreover is uncorrelated with $s_c(i_B)$. Thus, it follows from Eq. (25) that the actual SNR γ_{ML} of the combiner output can be written as

$$\gamma_{ML} = \frac{|\,\mathbf{E}[z_c(i_B)]\,|^2}{\mathbf{Var}[z(i_B)]} = \frac{|\,\mathbf{E}[s_c(i_B)]\,|^2}{\mathbf{Var}[s_c(i_B)] + \mathbf{Var}[n_c(i_B)]} \tag{28}$$

Now, it follows from Eqs. (26) and (23) that

$$|\,\mathbf{E}[s_c(i_B)]\,|^2 = |\,\mathbf{E}[\underline{\hat{W}}_{ML}^T(\tilde{i}_A)]\underline{V}\,|^2 = |\,\underline{W}_O^T\underline{V}\,|^2 = |\,\underline{V}^\dagger\underline{R}_B^{-1}\underline{V}\,|^2 = \gamma_{MAX}^2 \tag{29}$$

where γ_{MAX} is the maximum achievable SNR given by Eq. (13). The explicit expression, Eq. (A-9), for $\mathbf{Var}[s_c(i_B)]$ is derived in Appendix A. Moreover, the derivation in Appendix B yields the expression,

Eq. (B-5), for $\mathbf{Var}[n_c(i_B)]$. So, by using Eqs. (29), (A-9), and (B-5) in Eq. (28), one arrives at the following expression for the actual combiner output SNR performance:

$$\gamma_{ML} = \frac{\gamma_{MAX}^2}{\gamma_{MAX}(1 + C_1) + C_2} \tag{30}$$

where

$$C_1 = \frac{LK - K^2 - K + 1}{(L-K)(L-K-2)} + \frac{(L-K-1)(L-1)}{(L-K)(L-K-2)}\left(\frac{1}{L\eta\cos^2\delta}\right) + \frac{1}{L-K-2}\gamma_{MAX}$$

$$C_2 = \frac{K(L-1)(L-K-1)}{L(L-K)(L-K-2)}\frac{1}{\eta\cos^2\delta} \tag{31}$$

Since both C_1 and C_2 converge to zero as L tends to infinity, it follows from Eq. (30) that the actual combiner output SNR γ_{ML} converges to the maximum possible achievable SNR, γ_{MAX}, as the number of signal samples, L, used in the estimates tends to infinity.

V. Numerical Example

We consider here a numerical example using a $K = 7$ element array feed. In this example, a modulation index, $\delta = 80$ deg, and a primitive sample period of $T_0 = 2.5 \times 10^{-8}$ s are assumed. The full-spectrum modulation signal is assumed to be of bandwidth 2×10^6 Hz, which yields $M_B = 20$. Moreover, the ratio of the full-spectrum bandwidth to the residual-carrier bandwidth is $\eta = M_A/M_B = 200$. A nominal P_T/N_0 of 65 dB-Hz is considered with a corresponding $\gamma = (P_T/N_0)M_BT_0$ (recall that γ, which is given by Eq. (14), is the sum of the array element SNRs). The white Gaussian noise processes in the received signals from different array feed elements are assumed to be correlated. Moreover, the correlation magnitudes are assumed to be inversely proportional to the distances between feed centers (this assumption has not been verified for the array feed system at DSS 13). In the numerical examples below, the following noise covariance matrix is considered:

$$\underline{R}_B =$$

$$\begin{bmatrix}
1 & \rho_{max}e^{-j(\frac{4\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{10\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{6\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{2\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{12\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{8\pi}{7}+\phi)} \\
\rho_{max}e^{j(\frac{4\pi}{7}+\phi)} & 1 & \rho_{max}e^{-j(\frac{6\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{-j(\frac{2\pi}{7}+\phi)} & \frac{\rho_{max}}{2}e^{j(\frac{2\pi}{7}-\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{-j(\frac{8\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{4\pi}{7}+\phi)} \\
\rho_{max}e^{j(\frac{10\pi}{7}+\phi)} & \rho_{max}e^{j(\frac{6\pi}{7}+\phi)} & 1 & \rho_{max}e^{j(\frac{4\pi}{7}-\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{j(\frac{8\pi}{7}-\phi)} & \frac{\rho_{max}}{2}e^{-j(\frac{2\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{j(\frac{2\pi}{7}-\phi)} \\
\rho_{max}e^{j(\frac{6\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{j(\frac{2\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{4\pi}{7}-\phi)} & 1 & \rho_{max}e^{j(\frac{4\pi}{7}-\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{-j(\frac{6\pi}{7}+\phi)} & \frac{\rho_{max}}{2}e^{-j(\frac{2\pi}{7}+\phi)} \\
\rho_{max}e^{j(\frac{2\pi}{7}+\phi)} & \frac{\rho_{max}}{2}e^{-j(\frac{2\pi}{7}-\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{-j(\frac{8\pi}{7}-\phi)} & \rho_{max}e^{-j(\frac{4\pi}{7}-\phi)} & 1 & \rho_{max}e^{-j(\frac{10\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{-j(\frac{6\pi}{7}+\phi)} \\
\rho_{max}e^{j(\frac{12\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{j(\frac{8\pi}{7}+\phi)} & \frac{\rho_{max}}{2}e^{j(\frac{2\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{j(\frac{6\pi}{7}+\phi)} & \rho_{max}e^{j(\frac{10\pi}{7}+\phi)} & 1 & \rho_{max}e^{j(\frac{4\pi}{7}-\phi)} \\
\rho_{max}e^{j(\frac{8\pi}{7}+\phi)} & \rho_{max}e^{j(\frac{4\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{-j(\frac{2\pi}{7}-\phi)} & \frac{\rho_{max}}{2}e^{j(\frac{2\pi}{7}+\phi)} & \frac{\rho_{max}}{\sqrt{3}}e^{j(\frac{6\pi}{7}+\phi)} & \rho_{max}e^{-j(\frac{4\pi}{7}-\phi)} & 1
\end{bmatrix} \tag{32}$$

Figure 1 shows the $K = 7$ element feed array geometry and the relative distances between feed centers. The main feed is labeled feed 1 and is surrounded by the six others. In the correlation matrix, Eq. (32),

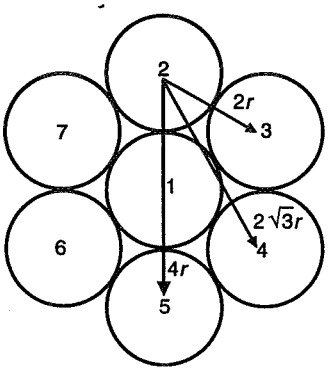

Fig. 1. Feed array geometry for $K = 7$ feeds.

the noise power in the feed-element stream B received signal samples are all assumed to be equal and normalized to one ($r_{Bkk} = 1$ for all k). The maximum possible noise correlation (and also correlation coefficient) magnitude is denoted by ρ_{max}. The corresponding noise correlation magnitudes in this matrix are identical and reflect an inverse dependence on distances between feed centers. The noise correlation magnitudes between nearest neighbor feed pairs are equal to the maximum correlation magnitude ρ_{max} (for example, between the main feed, feed 1, and any of the outer feeds). As can be seen from Fig. 1, the next nearest neighbor feed pairs (for example, feed 2 and feed 4) are of a distance equal to $\sqrt{3}$ times the distance between nearest neighbor feed pairs, and hence have a noise correlation magnitude equal to $\rho_{max}/\sqrt{3}$. Finally, the furthest feed pairs (for example, feed 2 and feed 5) are of a distance equal to twice the distance between closest feed pairs, and so have a noise correlation magnitude equal to $\rho_{max}/2$. The parameter ϕ in Eq. (32) specifies the noise correlation phases. The rationale for assigning these phases will be described below. Since the noise power in each feed element is equal to one, the sum of the feed element SNRs is equal to the total received power, and so $\gamma = P_T$. For the complex signal parameters vector, $\underline{V}$, we shall use

$$\underline{V} = \left[\sqrt{\gamma\beta}\ \sqrt{\gamma\tfrac{(1-\beta)}{6}}e^{j\frac{4\pi}{7}}\ \sqrt{\gamma\tfrac{(1-\beta)}{6}}e^{j\frac{10\pi}{7}}\ \sqrt{\gamma\tfrac{(1-\beta)}{6}}e^{j\frac{6\pi}{7}}\ \sqrt{\gamma\tfrac{(1-\beta)}{6}}e^{j\frac{2\pi}{7}}\ \sqrt{\gamma\tfrac{(1-\beta)}{6}}e^{j\frac{12\pi}{7}}\ \sqrt{\gamma\tfrac{(1-\beta)}{6}}e^{j\frac{8\pi}{7}}\right]^T \tag{33}$$

where $0 \leq \beta \leq 1$ represents the fraction of total received signal power in the main antenna feed (feed 1) as defined in [9]. The remaining total received signal power is then evenly distributed among the other six feed elements (this assumption does not appear to be generally valid for severely distorted antenna reflectors). The signal phases in Eq. (33) were chosen arbitrarily. The parameter ϕ in Eq. (32) specifies the relation between the noise correlation phases and the signal parameter phases as follows: Let V_k denote the kth component of the vector $\underline{V}$ given by Eq. (33). Then for $k > j$, ϕ = [phase of r_{Bkj}] − [(phase of V_k) − (phase of V_j)]. Moreover, since $\underline{R}_B$ is a Hemitian matrix, $-\phi$ = [phase of r_{Bkj}] − [(phase of V_k) − (phase of V_j)] when $j > k$. That is, ϕ represents the difference between the phase of the noise correlations and the phase differences between the corresponding signal components in $\underline{V}$. The significance of this will be discussed below.

In the numerical examples below, the signal combining scheme given by Eq. (24) will be referred to as the correlated noises algorithm. The signal-combining scheme from [1] will, on the other hand, be referred to as the uncorrelated noises algorithm. A comparison of the SNR performance of these two signal combining schemes will be made below, assuming the noise covariance matrix given by Eq. (32) and the signal parameter vector given by Eq. (33). Our previous work in [5] determined the SNR performance of the uncorrelated noises algorithm in the environment where the noise processes in the feed element received

signals are indeed correlated. For a given signal vector $\underline{V}$ and set of noise correlation magnitudes, the results in [5] show that the best-case performance of the uncorrelated noises algorithm occurs when each of the noise correlations between feed pairs has a phase that is exactly 180 deg from the phase difference between the corresponding signal components in $\underline{V}$. For the correlation matrix, Eq. (32), this "best case" scenario for the uncorrelated noises algorithm occurs when the parameter $\phi = 180$ deg. Moreover, the worst-case performance was shown in [5] to occur when each of the noise correlations between feed pairs has a phase that is exactly equal to the phase difference between the corresponding signal components in $\underline{V}$. This "worst-case" scenario for the uncorrelated noises algorithm occurs when the parameter $\phi = 0$ deg for the correlation matrix, Eq. (32). The numerical examples below will consider these two extreme cases only. The performance comparisons between these two algorithms can then be made under both the most-favorable and the least-favorable situations for the uncorrelated noises algorithm.

The SNR performances will be compared in terms of the *combining gain*, which is the ratio of the actual SNR performance to the sum of the SNRs of the array feed element received signals (γ_{ML}/γ for the correlated noises algorithm). Note that γ is also the maximum achievable SNR in the uncorrelated noises environment. Hence, the combining gain represents the SNR gain relative to the best possible performance in the uncorrelated noises environment. We shall, therefore, refer to an SNR gain when the combining gain is positive (in dB) and to an SNR loss otherwise. The SNR performance analysis given in Section IV showed that the combining gain γ_{ML}/γ of the correlated noises algorithm converges to the maximum possible combining gain, γ_{MAX}/γ, in the limit as the number of samples, L, approaches infinity. Figure 2 plots this maximum possible combining gain as the phase parameter ϕ in Eq. (32) is varied between 0 and 180 deg for values of ρ_{max} equal to 0.1, 0.15, and 0.2, and when $\beta = 0.7$. These results show that the best- and worst-case scenarios for the correlated noises algorithm are the same as those for the uncorrelated noises algorithm. These results show that

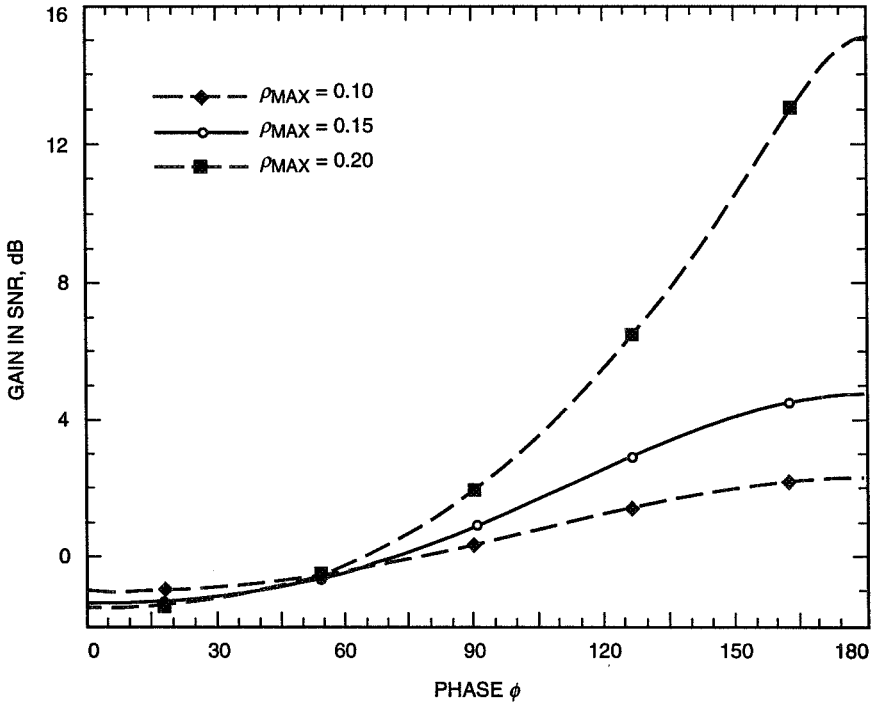

Fig. 2. Maximum possible combining gain versus ϕ with $\beta = 0.7$

the maximum possible combining gain is positive for a majority of the ϕ phase values. Moreover, the maximum possible combining gains increase with increasing ρ_{max}. Simulations were also performed to validate the analytical result, Eq. (30), that yields the combining gain for the correlated noises algorithm. Table 1 compares the simulated combining gain to the analytical result for the best-case $\underline{R}_B$ matrix (with $\phi = 180$ deg) at various values of ρ_{max} and L. Table 2 displays the corresponding comparisons for the worst-case $\underline{R}_B$ (with $\phi = 0$ deg). The simulated combining gain from these two tables can be seen to be within 3 percent of the analytical results. This appears to validate the analysis.

Table 1. Correlated algorithm simulated combining gain for best-case $\underline{R}_B$.

		SNR Gain, dB	
L	ρ_{max}	Simulation	Analytical
500	0.050	0.85896	0.80340
500	0.100	2.24048	2.17891
500	0.150	4.75655	4.67706
500	0.175	7.25759	7.14633
500	0.200	15.06133	14.60054
5000	0.050	0.88286	0.87982
5000	0.100	2.26456	2.25971
5000	0.150	4.78108	4.77260
5000	0.175	7.28267	7.26961
5000	0.200	15.08713	15.03570
10^6	0.050	0.86770	0.88821
10^6	0.100	2.24674	2.26859
10^6	0.150	4.76168	4.78312
10^6	0.175	7.26309	7.28324
10^6	0.200	15.06514	15.08578

In Fig. 3, the combining gains of both algorithms are plotted versus the number of samples, L, for $\beta = 0.7$ (70 percent of the power in the main feed), using the best-case $\underline{R}_B$ and with ρ_{max} having values of 0.05, 0.1, and 0.15. We notice that for a fixed ρ_{max}, the correlated noises algorithm (CNA in the figures) actually has a smaller combining gain than the uncorrelated noises algorithm (UNA in the figures) for small values of L below a threshold value. In the examples considered, this threshold value increases with decreasing ρ_{max} and is always less than about $L = 500$. Note that although the ML estimator is asymptotically optimal as $L \rightarrow \infty$, it is not necessarily the best estimator for small values of L. Hence, the correlated noises algorithm may not have the best possible performance at small values of L. The convergence of the combining gain to the optimum theoretically achievable value was proved in the last section. These examples show that convergence to within 0.1 dB of the limiting value occurs at about $L = 1000$ samples. Figure 4 shows the respective combining gains as a function of β, the fraction of total received signal power in the main feed, for $L = 5000$ samples and values of ρ_{max} equal to 0.05, 0.1, 0.15, and 0.2. We can see that the performance of the correlated noises algorithm is much less sensitive to β in this example than that of the uncorrelated noises algorithm. The performance superiority of the correlated noises algorithm also increases significantly with increasing β. The respective combining gains versus ρ_{max} with $L = 5000$ and β taking on values of 0.7, 0.8, and 0.9 are shown in Fig. 5. This figure shows that the performance improvement of the correlated noises algorithm over the uncorrelated noises algorithm increases significantly with increasing ρ_{max}. The results in these three figures indicate that there can be

Table 2. Correlated algorithm simulated combining gain for worst-case $\underline{R}_B$.

		SNR Gain, dB	
L	ρ_{max}	Simulation	Analytical
500	0.050	−0.63659	−0.68802
500	0.100	−1.05467	−1.10542
500	0.150	−1.33373	−1.38419
500	0.175	−1.43041	−1.48083
500	0.200	−1.50179	−1.55220
5000	0.050	−0.61306	−0.61410
5000	0.100	−1.03140	−1.03182
5000	0.150	−1.31079	−1.31072
5000	0.175	−1.40765	−1.40738
5000	0.200	−1.47920	−1.47877
10^6	0.050	−0.62206	−0.60598
10^6	0.100	−1.03724	−1.02373
10^6	0.150	−1.31348	−1.30265
10^6	0.175	−1.40878	−1.39931
10^6	0.200	−1.47881	−1.47070

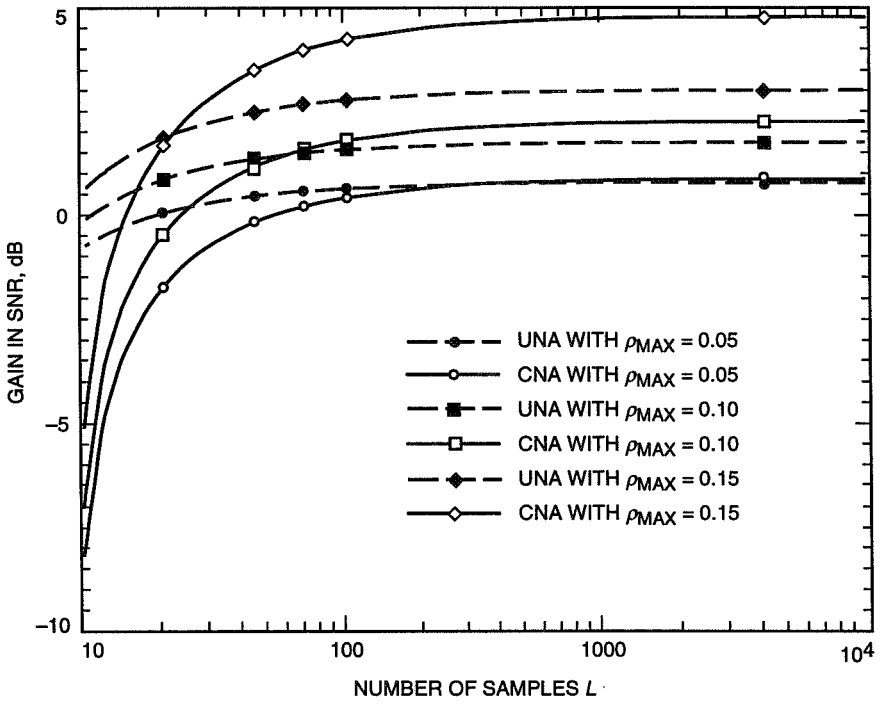

Fig. 3. Combining gain versus L with $\beta = 0.7$ for best-case $\underline{R}_B$.

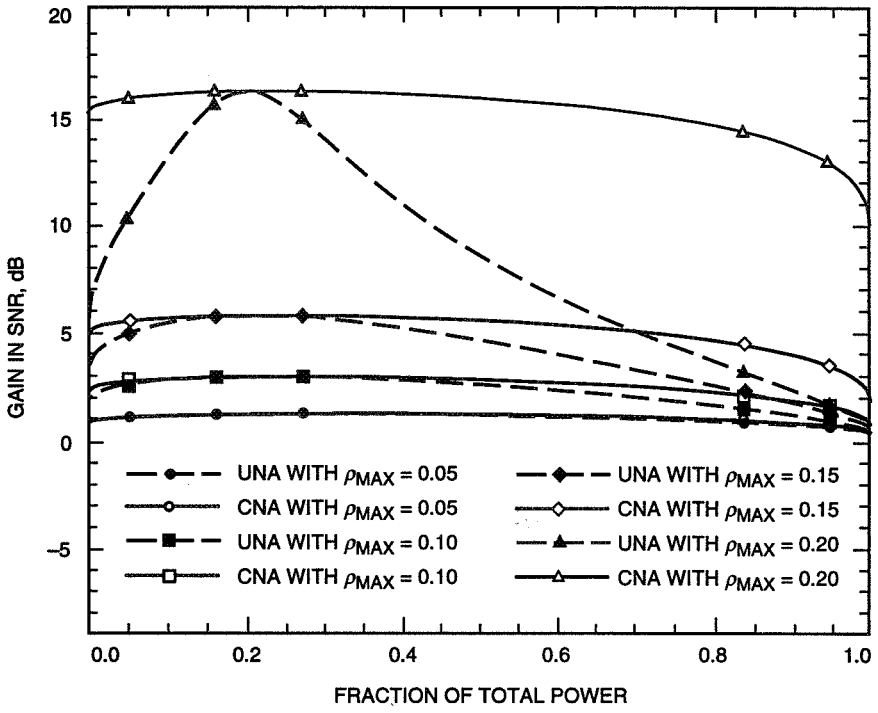

Fig. 4 Combining gain versus β with L = 5000 for best-case R_B.

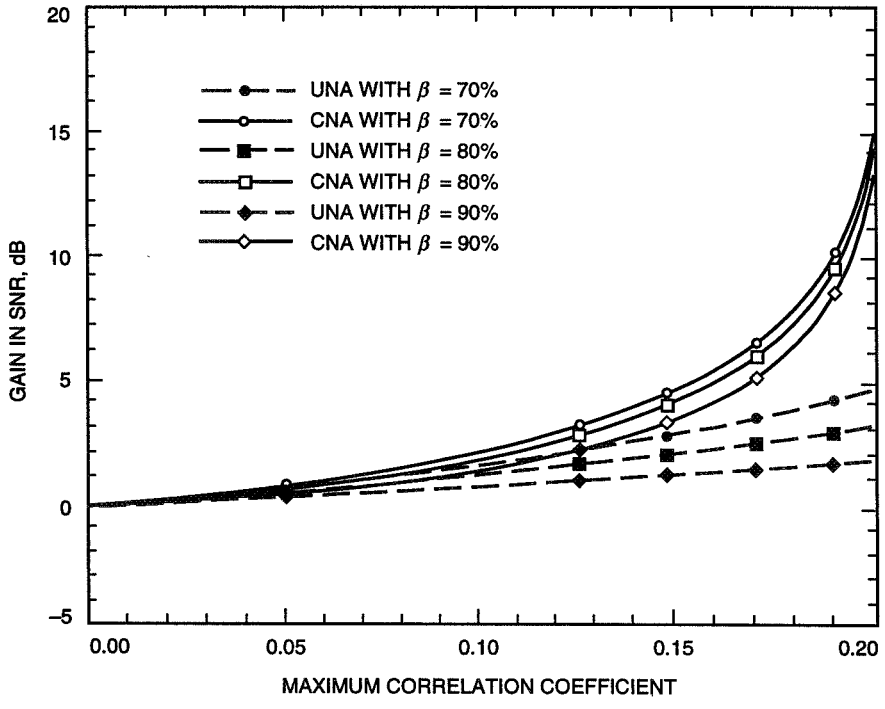

Fig. 5. Combining gain versus ρ_{max} with L = 5000 for best-case R_B.

a substantial SNR gain in this best-case scenario, particularly for the correlated noises algorithm at larger values of ρ_{max}. We also note that the value of ρ_{max} cannot exceed 0.2 in order to preserve the positive definiteness of $\underline{R}_B$.

Figures 6 through 8 repeat Figs. 3 through 5, respectively, using the worst-case scenario $\underline{R}_B$ matrix instead, while keeping the other parameters unchanged. These examples show that this worst-case scenario can result in an SNR loss, and the loss nominally increases with increasing ρ_{max}. Figure 6 again shows that convergence of the combining gain for the correlated noises algorithm to within 0.1 dB of its limiting value occurs at about $L = 1000$ samples. The small sample performance of the correlated algorithm again lags that of the uncorrelated algorithm. Similarly to Fig. 4, Fig. 7 shows that the performance superiority of the correlated noises algorithm increases with increasing β. However, as β approaches one, the SNR loss of the correlated noises algorithm turns around and starts to decrease with increasing ρ_{max}. At large values of β close to one, the correlated noises algorithm can in fact have an SNR gain even when the uncorrelated noises algorithm still has an SNR loss. Finally, Fig. 8 shows that the SNR loss of the correlated noises algorithm in this unfavorable situation deteriorates much slower than that of the uncorrelated noises algorithm as ρ_{max} increases.

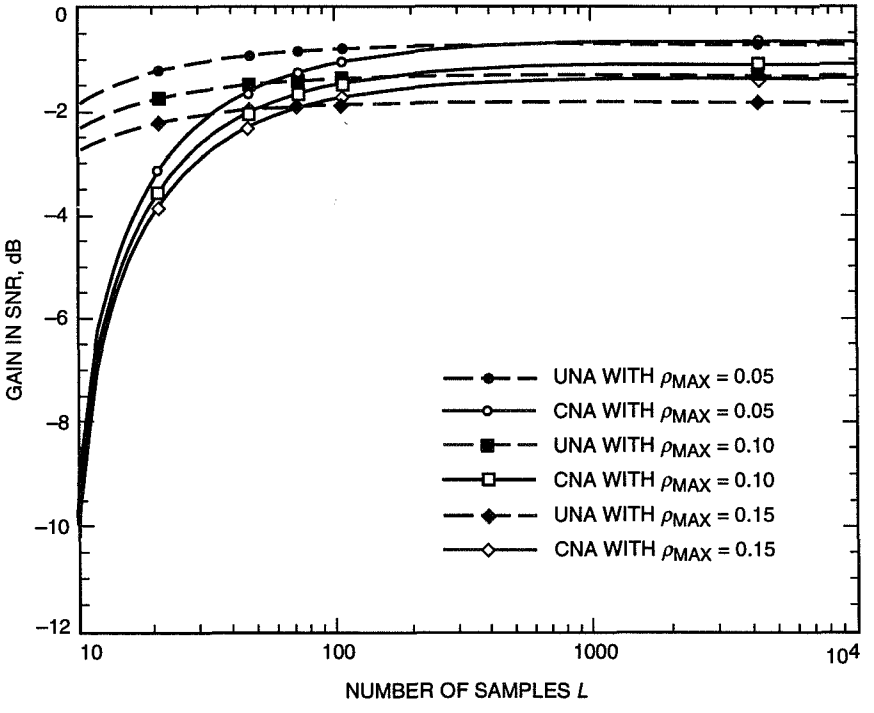

Fig. 6. Combining gain versus L with $\beta = 0.7$ for worst-case $\underline{R}_B$.

VI. Conclusion

The correlated noises signal-combining scheme developed in this article has been shown to be asymptotically optimal as the number of residual carrier-signal samples used in the estimates of the optimum weight coefficients increases. The numerical examples considered here show that convergence of the combining algorithm's SNR performance to the optimum achievable SNR performance level to within 0.1 dB occurs at about $L = 1000$ signal samples. Hence, real-time operation is possible, although the inversion of a complex-valued K-dimensional matrix is required at every update of the weight coefficient estimates. The numerical examples considered here consistently demonstrate a significant performance superiority

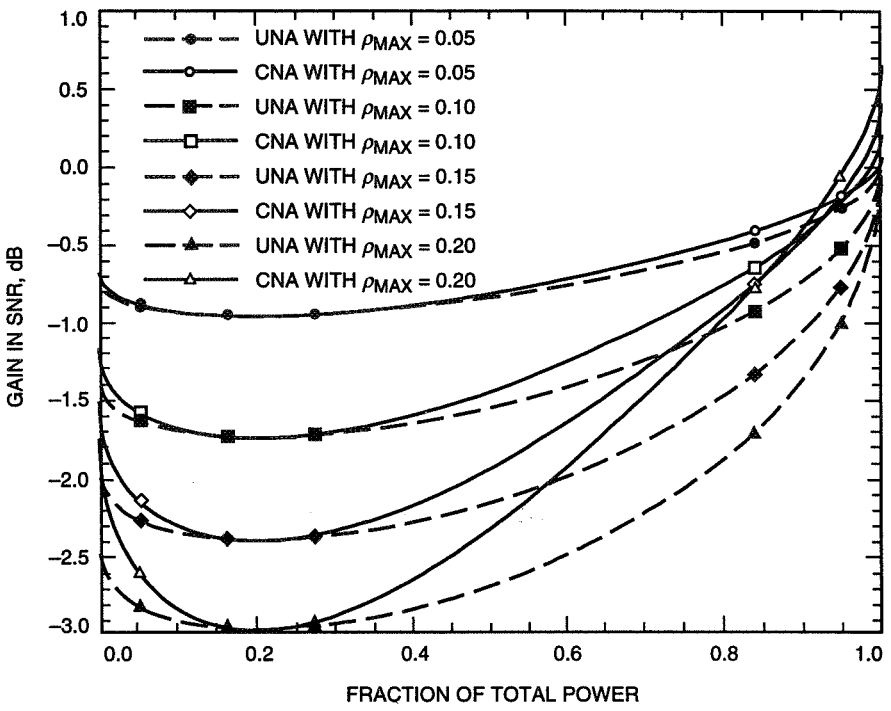

Fig. 7. Combining gain versus β with L = 5000 for worst-case $\underline{R}_B$.

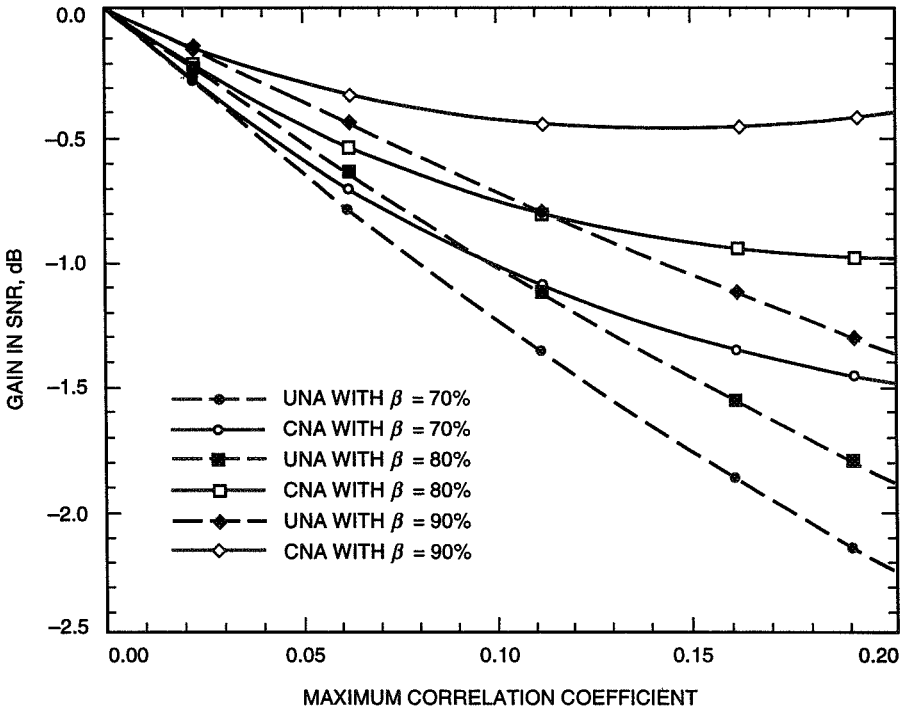

Fig. 8. Combining gain versus ρ_{max} with L = 5000 for worst-case $\underline{R}_B$.

of the correlated noises combining scheme over the uncorrelated noises signal-combining algorithm. It appears to be the combining system of choice in the presence of significant correlations between the noise processes in different array feed received signals. A degree of caution should be exercised in extrapolating the expected amount of SNR gain or loss from the numerical examples considered in Section V. The best-case and worst-case performances in these examples should not be viewed as being typical. Moreover, these two cases also should not be viewed as being best- and worst-case performances in general.

References

[1] V. A. Vilnrotter, E. R. Rodemich, and S. J. Dolinar, Jr., "Real-Time Combining of Residual Carrier Array Signals Using ML Weight Estimates," *IEEE Transactions Communications*, vol. COM-40, no. 3, pp. 604–615, March 1992.

[2] B. A. Iijima, V. A. Vilnrotter, and D. Fort, "Correlator Data Analysis for the Array Feed Compensation System," *The Telecommunications and Data Acquisition Progress Report 42-117, January–March 1994*, Jet Propulsion Laboratory, Pasadena, California, pp. 110–118, May 15, 1994.

[3] S. P. Applebaum, "Adaptive Arrays," *IEEE Transactions on Antennas and Propagation*, vol. AP-24, pp. 585–598, September 1976.

[4] J. H. Yuen, *Deep Space Telecommunications Systems Engineering*, Chapter 5, New York: Plenum, 1983.

[5] H. H. Tan, "Performance of Residual Carrier Array-Feed Combining in Correlated Noise," *The Telecommunications and Data Acquisition Progress Report 42-121, January–March 1995*, Jet Propulsion Laboratory, Pasadena, California, pp. 131–147, May 15, 1995.
http://edms-www.jpl.nasa.gov/tda/progress_report/42-121/121H.pdf

[6] N. R. Goodman, "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)," *Annals of Mathematical Statistics*, vol. 34, pp. 152–177, 1963.

[7] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed., New York: Wiley, 1984.

[8] J. Tague and C. Caldwell,"Expectations of Useful Complex Wishart Forms," *Multidimensional Systems and Signal Processing*, no. 5, pp. 263–279, July 1994.

[9] V. Vilnrotter, "Channel Assignments for Improved Gain in Baseband Array Feed Compensation Systems," *IEEE Transactions Communications*, vol. COM-42, no. 5, pp. 2127–2133, May 1994.

Appendix A
Combiner Output Signal Variance

An explicit expression for $\mathbf{Var}\,[s_c(i_B)]$ is derived in this appendix. For convenience in the following analysis, let us use $\underline{\hat{W}}_{ML}$ as the shortened notation for $\underline{\hat{W}}_{ML}(\tilde{i}_A)$, $\underline{\hat{X}}_{ML}$ for $\underline{\hat{X}}_{ML}(\tilde{i}_A)$, and $\underline{\hat{R}}_{A,ML}$ for $\underline{\hat{R}}_{A,ML}(\tilde{i}_A)$. Using Eqs. (26) and (22), we can write

$$\begin{aligned}
\mathbf{Var}\,[s_c(i_B)] =& \mathbf{Var}\left[\underline{\hat{W}}_{ML}^T \underline{V} e^{js(i_B)\delta}\right] \\
=& \mathbf{E}\left[\left(\underline{\hat{W}}_{ML}^\dagger \underline{V}^*\right)\left(\underline{\hat{W}}_{ML}^T \underline{V}\right)\right] - |\,\underline{W}_O^T \underline{V}\,|^2 \\
=& \underline{V}^\dagger \mathbf{E}\left[\underline{\hat{W}}_{ML}^* \underline{\hat{W}}_{ML}^T\right] \underline{V} - |\,\underline{W}_O^T \underline{V}\,|^2 \\
=& \underline{V}^\dagger \mathbf{E}\left[\frac{(L-K-1)^2}{L^2\eta^2\cos^2\delta}\underline{\hat{R}}_{A,ML}^{-1}\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\underline{\hat{R}}_{A,ML}^{-1}\right]\underline{V} - |\,\underline{W}_O^T \underline{V}\,|^2
\end{aligned} \tag{A-1}$$

Since $\underline{\hat{R}}_{A,ML}$ is independent of $\underline{\hat{X}}_{ML}$, we can write

$$\begin{aligned}
\mathbf{E}\left[\underline{\hat{R}}_{A,ML}^{-1}\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\underline{\hat{R}}_{A,ML}^{-1}\right] =& \mathbf{E}\left[\mathbf{E}\left[\underline{\hat{R}}_{A,ML}^{-1}\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\underline{\hat{R}}_{A,ML}^{-1} \mid \underline{\hat{R}}_{A,ML}\right]\right] \\
=& \mathbf{E}\left[\underline{\hat{R}}_{A,ML}^{-1}\mathbf{E}\left[\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger \mid \underline{\hat{R}}_{A,ML}\right]\underline{\hat{R}}_{A,ML}^{-1}\right] \\
=& \mathbf{E}\left[\underline{\hat{R}}_{A,ML}^{-1}\mathbf{E}\left[\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\right]\underline{\hat{R}}_{A,ML}^{-1}\right]
\end{aligned} \tag{A-2}$$

Moreover, since $\underline{\hat{X}}_{ML}(i_A)$ is a complex Gaussian random vector with mean $\underline{X}$ and covariance matrix $(1/L)\underline{R}_A$, we have

$$\mathbf{E}\left[\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\right] = \frac{1}{L}\underline{R}_A + \underline{X}\underline{X}^\dagger \tag{A-3}$$

Therefore, it follows from Eqs. (A-1), (A-2), and (A-3) that

$$\begin{aligned}
\mathbf{Var}[s_c(i_B)] =& \frac{(L-K-1)^2}{L^2\eta^2\cos^2\delta}\underline{V}^\dagger \mathbf{E}\left[\underline{\hat{R}}_{A,ML}^{-1}\left(\frac{1}{L}\underline{R}_A + \underline{X}\underline{X}^\dagger\right)\underline{\hat{R}}_{A,ML}^{-1}\right]\underline{V} - |\,\underline{W}_O^T\underline{V}\,|^2 \\
=& \frac{(L-K-1)^2}{\eta^2\cos^2\delta}\underline{V}^\dagger \mathbf{E}\left[\left(L\underline{\hat{R}}_{A,ML}\right)^{-1}\left(\frac{1}{L}\underline{R}_A + \underline{X}\underline{X}^\dagger\right)\left(L\underline{\hat{R}}_{A,ML}\right)^{-1}\right]\underline{V} - |\,\underline{W}_O^T\underline{V}\,|^2
\end{aligned} \tag{A-4}$$

Let $\mathbf{Tr}(\underline{A})$ denote the trace of a square matrix $\underline{A}$ and let $\underline{I}_K$ denote the $K \times K$ identity matrix. Recall that $L\underline{\hat{R}}_{A,ML}$ is a $K \times K$ $CW(\underline{R}_A, L-1)$ distributed random matrix. It has been shown in [8] that if $\underline{A}$ is a $K \times K$ $CW(\underline{R}_A, L-1)$ random matrix and $\underline{C}$ is any constant $K \times K$ matrix, then

$$\mathbf{E}\left[\underline{A}^{-1}\underline{C}\underline{A}^{-1}\right] = \frac{1}{(L-K)(L-K-2)}\underline{R}_A^{-1}\underline{C}\underline{R}_A^{-1} + \frac{1}{(L-K)(L-K-1)(L-K-2)}\mathbf{Tr}\left(\underline{C}\underline{R}_A^{-1}\right)\underline{R}_A^{-1} \tag{A-5}$$

for $L > K$. So, using Eq. (A-5) in Eq. (A-4) gets

$$\begin{aligned}\mathbf{Var}\left[s_c(i_B)\right] =& \frac{(L-K-1)}{(L-K)(L-K-2)\eta^2\cos^2\delta}\underline{V}^\dagger\left\{[(L-1)-K]\left(\frac{1}{L}\underline{R}_A^{-1} + \underline{R}_A^{-1}\underline{X}\underline{X}^\dagger\underline{R}_A^{-1}\right)\right.\\ &\left. + \mathbf{Tr}\left(\frac{1}{L}\underline{I}_K + \underline{X}\underline{X}^\dagger\underline{R}_A^{-1}\right)\underline{R}_A^{-1}\right\}\underline{V} - \mid \underline{W}_O^T\underline{V} \mid^2\\ =& \frac{(L-K-1)^2}{(L-K)(L-K-2)\eta^2\cos^2\delta}\left[\frac{1}{L}\left(\underline{V}^\dagger\underline{R}_A^{-1}\underline{V}\right) + \underline{V}^\dagger\underline{R}_A^{-1}\underline{X}\underline{X}^\dagger\underline{R}_A^{-1}\underline{V}\right]\\ &+ \frac{(L-K-1)}{(L-K)(L-K-2)\eta^2\cos^2\delta}\underline{V}^\dagger\mathbf{Tr}\left(\frac{1}{L}\underline{I}_K + \underline{X}\underline{X}^\dagger\underline{R}_A^{-1}\right)\underline{R}_A^{-1}\underline{V} - \mid \underline{W}_O^T\underline{V} \mid^2\end{aligned} \tag{A-6}$$

Since $\underline{X} = \underline{V}\cos\delta$, and $\underline{R}_B = \eta\underline{R}_A$, it follows from Eq. (29) that

$$\frac{1}{\eta^2\cos^2\delta}\left(\underline{V}^\dagger\underline{R}_A^{-1}\underline{X}\underline{X}^\dagger\underline{R}_A^{-1}\underline{V}\right) = \mid \underline{V}^\dagger\underline{R}_B^{-1}\underline{V} \mid^2 = \mid \underline{W}_O^T\underline{V} \mid^2 = \gamma_{MAX}^2 \tag{A-7}$$

Since $\underline{R}_A^{-1}$ is Hermitian, $\mathbf{Tr}(\underline{X}\underline{X}^\dagger\underline{R}_A^{-1}) = \underline{X}^\dagger\underline{R}_A^{-1}\underline{X}$. So we have, by using Eq. (A-7),

$$\begin{aligned}\frac{1}{\eta^2\cos^2\delta}\underline{V}^\dagger\mathbf{Tr}\left(\frac{1}{L}\underline{I}_K + \underline{X}\underline{X}^\dagger\underline{R}_A^{-1}\right)\underline{R}_A^{-1}\underline{V} =& \frac{1}{\eta^2\cos^2\delta}\left[\frac{K}{L} + \underline{X}^\dagger\underline{R}_A^{-1}\underline{X}\right]\underline{V}^\dagger\underline{R}_A^{-1}\underline{V}\\ =& \frac{K}{L\eta^2\cos^2\delta}\left(\underline{V}^\dagger\underline{R}_A^{-1}\underline{V}\right) + \underline{V}^\dagger\underline{R}_B^{-1}\underline{V}\underline{V}^\dagger\underline{R}_B^{-1}\underline{V}\\ =& \frac{K}{L\eta\cos^2\delta}\gamma_{MAX} + \gamma_{MAX}^2\end{aligned} \tag{A-8}$$

Finally, using Eqs. (A-7) and (A-8) in Eq. (A-6) yields the following expression for $\mathbf{Var}[s_c(i_B)]$:

$$\mathbf{Var}\left[s_c(i_B)\right] = \frac{(L-K-1)(L-1)}{(L-K)(L-K-2)}\left(\frac{\gamma_{MAX}}{L\eta\cos^2\delta}\right) + \frac{1}{L-K-2}\gamma_{MAX}^2 \tag{A-9}$$

Note that $\mathbf{Var}[s_c(i_B)]$ approaches zero in the limit as L tends to infinity. This is because the estimate $\underline{\hat{W}}_{ML}(i_A)$ converges with probability one to the optimum weight vector $\underline{W}_O$ given by Eq. (23) as L tends to infinity.

Appendix B
Combiner Output Noise Variance

Consider next an explicit expression for $\mathbf{Var}[n_c(i_B)]$. We shall also employ here the shortened notations $\underline{\hat{W}}_{ML}$, $\underline{\hat{X}}_{ML}$, $\underline{\hat{R}}_{A,ML}$, as in Appendix A. Moreover, we shall use the shortened notation $\underline{n}$ for $\underline{n}(i_B)$. Since $\underline{\hat{W}}_{ML}$ is independent of $\underline{n}$, a derivation similar to that establishing Eq. (A-2) can be used along with Eq. (22) to get

$$\begin{aligned}\mathbf{Var}\left[n_c(i_B)\right] =& \mathbf{E}\left[\underline{\hat{W}}_{ML}^T \underline{n}\underline{n}^\dagger \underline{\hat{W}}_{ML}^*\right] \\ =& \mathbf{E}\left[\underline{\hat{W}}_{ML}^T \underline{R}_B \underline{\hat{W}}_{ML}^*\right] \\ =& \frac{(L-K-1)^2}{L^2}\frac{1}{\eta\cos^2\delta}\mathbf{E}\left[\underline{\hat{X}}_{ML}^\dagger \underline{\hat{R}}_{A,ML}^{-1}\underline{R}_A\underline{\hat{R}}_{A,ML}^{-1}\underline{\hat{X}}_{ML}\right]\end{aligned} \tag{B-1}$$

Since $\underline{\hat{R}}_{A,ML}$ is independent of $\underline{\hat{X}}_{ML}$, using the same approach on the expected value in Eq. (B-1) gets

$$\mathbf{Var}\left[n_c(i_B)\right] = \frac{(L-K-1)^2}{\eta\cos^2\delta}\mathbf{E}\left[\underline{\hat{X}}_{ML}^\dagger \mathbf{E}\left[\left(L\underline{\hat{R}}_{A,ML}\right)^{-1}\underline{R}_A\left(L\underline{\hat{R}}_{A,ML}\right)^{-1}\right]\underline{\hat{X}}_{ML}\right] \tag{B-2}$$

Next, using the property of Eq. (A-5) for the $K\times K$ $CW(\underline{R}_A, L-1)$ distributed random matrix $L\underline{\hat{R}}_{A,ML}$ in Eq. (B-2) results in the following expression:

$$\begin{aligned}\mathbf{Var}\left[n_c(i_B)\right] =& \frac{(L-K-1)}{\eta\cos^2\delta}\mathbf{E}\left[\underline{\hat{X}}_{ML}^\dagger \frac{[(L-1)-K]\underline{R}_A^{-1}+K\underline{R}_A^{-1}}{(L-K)(L-K-2)}\underline{\hat{X}}_{ML}\right] \\ =& \frac{(L-K-1)(L-1)}{\eta\cos^2\delta(L-K)(L-K-2)}\mathbf{E}\left[\underline{\hat{X}}_{ML}^\dagger \underline{R}_A^{-1}\underline{\hat{X}}_{ML}\right]\end{aligned} \tag{B-3}$$

It follows from Eq. (A-3) that

$$\begin{aligned}\mathbf{E}\left[\underline{\hat{X}}_{ML}^\dagger \underline{R}_A^{-1}\underline{\hat{X}}_{ML}\right] =& \mathbf{E}\left[\mathbf{Tr}\left(\underline{R}_A^{-1}\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\right)\right] \\ =& \mathbf{Tr}\left(\underline{R}_A^{-1}\mathbf{E}\left[\underline{\hat{X}}_{ML}\underline{\hat{X}}_{ML}^\dagger\right]\right) \\ =& \mathbf{Tr}\left(\underline{R}_A^{-1}\left[\frac{1}{L}\underline{R}_A + \underline{X}\underline{X}^\dagger\right]\right) \\ =& \frac{K}{L} + \underline{X}^\dagger\underline{R}_A^{-1}\underline{X}\end{aligned} \tag{B-4}$$

Finally, using Eq. (B-4) in Eqs. (B-3) and (13) yields

$$\mathbf{Var}\left[n_c(i_B)\right] = \left[1 + \frac{LK - K^2 - K + 1}{(L-K)(L-K-2)}\right]\gamma_{MAX} + \frac{(L-1)(L-K-1)}{(L-K)(L-K-2)\eta\cos^2\delta}\frac{K}{L} \tag{B-5}$$

Appendix C

Mean-Square Convergence of Estimated Weights

We employ the shortened notations $\underline{\hat{W}}_{ML}$, $\underline{\hat{X}}_{ML}$, and $\underline{\hat{R}}_{A,ML}$, similar to the usage in the previous appendices. Since

$$\| \underline{\hat{W}}_{ML} - \underline{W}_O \|^2 = \mathbf{Tr}\left[\left(\underline{\hat{W}}_{ML} - \underline{W}_O\right)\left(\underline{\hat{W}}_{ML} - \underline{W}_O\right)^\dagger\right]$$

and

$$\mathbf{E}\left[\left(\underline{\hat{W}}_{ML} - \underline{W}_O\right)\left(\underline{\hat{W}}_{ML} - \underline{W}_O\right)^\dagger\right] = \mathbf{E}\left[\underline{\hat{W}}_{ML}\underline{\hat{W}}_{ML}^\dagger\right] - \underline{W}_O\underline{W}_O^\dagger$$

we need only establish that $\mathbf{E}[\underline{\hat{W}}_{ML}\underline{\hat{W}}_{ML}^\dagger] \longrightarrow \underline{W}_O\underline{W}_O^\dagger$ as $L \longrightarrow \infty$ to prove the mean-square convergence of $\underline{\hat{W}}_{ML}$ to $\underline{W}_O$. Using Eqs. (17) and (18), we can write

$$\mathbf{E}\left[\underline{\hat{W}}_{ML}\underline{\hat{W}}_{ML}^\dagger\right] = \frac{(L-1-K)^2}{L^2\eta^2\cos^2\delta}\mathbf{E}\left[\left(\underline{\hat{R}}_{A,ML}^*\right)^{-1}\underline{\hat{X}}_{ML}^*\underline{\hat{X}}_{ML}^T\left(\underline{\hat{R}}_{A,ML}^*\right)^{-1}\right] \tag{C-1}$$

It then follows from using Eqs. (A-2) and (A-3) in Eq. (C-1) that

$$\mathbf{E}\left[\underline{\hat{W}}_{ML}\underline{\hat{W}}_{ML}^\dagger\right] = \frac{(L-1-K)^2}{\eta^2\cos^2\delta}\mathbf{E}\left[\left(L\underline{\hat{R}}_{A,ML}^*\right)^{-1}\left(\frac{1}{L}\underline{R}_A^* + \underline{X}^*\underline{X}^T\right)\left(L\underline{\hat{R}}_{A,ML}^*\right)^{-1}\right] \tag{C-2}$$

So, by using the property of Eq. (A-5) for the $CW(\underline{R}_A, L-1)$ distributed random matrix $L\underline{\hat{R}}_{A,ML}$ in Eq. (C-2), we can write

$$\mathbf{E}\left[\underline{\hat{W}}_{ML}\underline{\hat{W}}_{ML}^\dagger\right] =$$

$$\frac{L-1-K}{\eta^2\cos^2\delta}\left\{\frac{(L-1-K)\left[\frac{1}{L}\left(\underline{R}_A^*\right)^{-1} + \left(\underline{R}_A^*\right)^{-1}\underline{X}^*\underline{X}^T\left(\underline{R}_A^*\right)^{-1}\right] + \left[\frac{K}{L} + \underline{X}^T\left(\underline{R}_A^*\right)^{-1}\underline{X}^*\right]\left(\underline{R}_A^*\right)^{-1}}{(L-K)(L-K-2)}\right\}$$

Therefore, $\mathbf{E}[\underline{\hat{W}}_{ML}\underline{\hat{W}}_{ML}^\dagger] \longrightarrow (\underline{R}_A^*)^{-1}\underline{X}^*\underline{X}^T(\underline{R}_A^*)^{-1}/\eta^2\cos^2\delta = \underline{W}_O\underline{W}_O^\dagger$ as $L \longrightarrow \infty$, thereby establishing the mean-square convergence of $\underline{\hat{W}}_{ML}$ to $\underline{W}_O$.