REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 14, 1991	Interim

4. TITLE AND SUBTITLE

Results on Canceller Convergence in Nonstationary Noise

5. FUNDING NUMBERS

PE - 61153N
PR - 021-05-43
WU - DN480-006

6. AUTHOR(S)

Gerlach, K.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Washington, DC 20375-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL Report 9312

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Chief of Naval Research
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Convergence results for the Sampled Matrix Inversion (SMI)/Gram-Schmidt (GS) canceller algorithm in nonstationary noise is investigated by using the Gram-Schmidt (GS) canceller as an analysis tool. Lower and upper bounds for the convergence rate of the canceller's average output noise power residue normalized to the quiescent average output noise power residue are derived. These bounds are a function of the number of independent samples processed per channel (main and auxiliary), the number of auxiliary input channels, and the external noise environment. The external noise environment was modeled as Gaussian, with a power level specified at each sampling time instant. Furthermore, this model is generalized in the sense that a joint probability distribution function is defined for the power levels over a canceller processing batch. This leads to the capability of modeling and evaluating the SMI/GS canceller in a variety of interference scenarios such as continuous or discrete time processes or a mix of these.

14. SUBJECT TERMS

ECCM
Adaptive filter
Adaptive cancellation
Gram-Schmidt

15. NUMBER OF PAGES

30

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std Z39-18
298-102

CONTENTS

1. INTRODUCTION 1

2. THE GS CANCELLER 1

3. OUTPUT MEASURE 4

4. INVARIANT TRANSFORMS 5

5. GS DECOMPOSITION 7

6. NONSTATIONARY NOISE MODEL 9

7. 2-INPUT GS CANCELLER 13

8. PRELIMINARY DEFINITIONS AND THEOREMS 15

9. BOUNDS FOR NONCONCURRENT GS CANCELLER 17

10. SUMMARY 19

11. REFERENCES 19

APPENDIX A — Derivation of 2-Input GS Canceller Noise Power 21

APPENDIX B — Proof of Theorem 1 25

APPENDIX C — Derivations of Eqs. (A6), (A7), and (A19) 27

RESULTS ON CANCELLER CONVERGENCE IN NONSTATIONARY NOISE

1. INTRODUCTION

The optimal weights associated with an adaptive canceller are often not known a priori and thus must be estimated by using finite averaging. Because of the use of estimated weights, suboptimal canceller performance results. Reed, Mallet, and Brennan [1,2] quantified this performance for the Sampled Matrix Inversion (SMI) algorithm in the transient state under the conditions that the input noise must be Gaussian, stationary, and independent from time sample to time sample. They mathematically demonstrated that the SMI canceller has relatively fast convergence characteristics and also that the convergence is independent of the input covariance matrix.

In Ref. 3, the Reed, Mallet, and Brennan results were extended to include the effects of non-Gaussian inputs by using the Gram-Schmidt (GS) canceller [4-9] as an analysis tool. It was shown that the GS canceller and the SMI canceller are numerically identical, and hence the SMI can be analyzed by using the GS canceller structure. In Ref. 10, lower and upper bounds of convergence performance were derived for when the input noise is Gaussian but correlated from sample to sample (colored input noise). In this report the methodology developed in Refs. 3, 4, and 10 is extended to analyze a canceller in temporally nonstationary noise. Upper and lower bounds of convergence performance are again derived.

The analysis presented in this report pertains to the adaptive processor in canceller configuration whereby the derived signal is assumed only to be present in the main channel and auxiliary channels are used to cancel correlated noises in the main channel. However, Ref. 1 showed that any nonconstrained linear adaptive array processor can be transformed into a canceller configuration without changing the output noise power convergence statistics. Hence, the results of this report apply to any nonconstrained linear adaptive array processor.

2. THE GS CANCELLER

Consider the general N-input GS canceller structure (Fig. 1(a)). Let $x_M(t)$, $x_1(t)$, $\ldots$, $x_{N-1}(t)$ represent the complex data in the 0th, 1st, $\ldots$, $N-$ 1th channels, respectively. We call the left-most input $x_M(t)$ the main channel and the remaining $N-1$ inputs the auxiliary channels. The main channel's signal consists of a desired signal plus additive noise. The noise consists of internal noise plus external noise. Cancellation of the signals from external interfering sources relies on the correlation of simultaneously received signals in the main and auxiliary channels. The internal noises on each channel are assumed uncorrelated between channels. The canceller operates so as to decorrelate the auxiliary inputs one at a time from the other inputs by use of the basic 2-input GS processor shown in Fig. 1(b). For example, Fig. 1(a) shows that $x_{N-1}(t)$ is uncorrelated with $x_M^{(2)}(t)$, $x_1^{(2)}(t)$, $\ldots$, $x_{N-2}^{(2)}(t)$ in the first level of decomposition. Next, the output channel that results from decorrelating $x_{N-1}(t)$ from $x_{N-2}(t)$ is decorrelated from the other outputs of the first-level GSs. The decomposition proceeds until a final output channel is generated. If the decorrelation weights in each of the 2-input GSs are computed from an infinite number of input samples, this output channel is totally decorrelated with the input: $x_1(t)$, $x_2(t)$, $\ldots$, $x_{N-1}(t)$.

Manuscript approved February 1, 1991

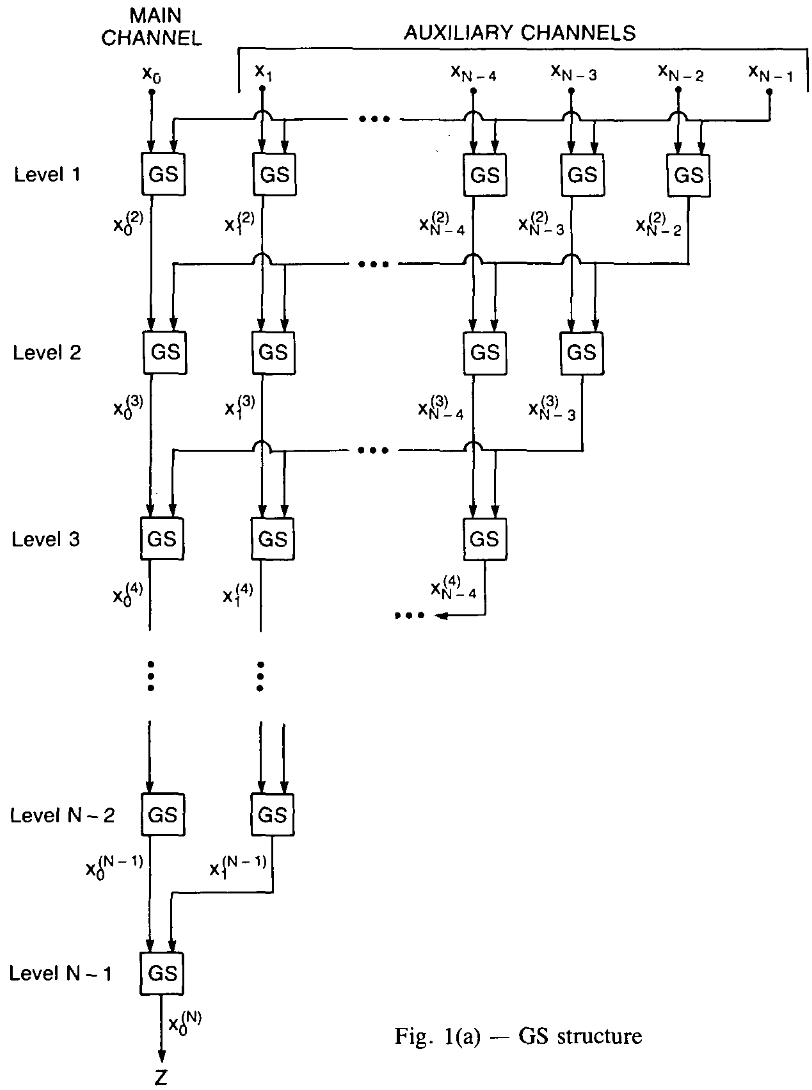

Fig. 1(a) — GS structure

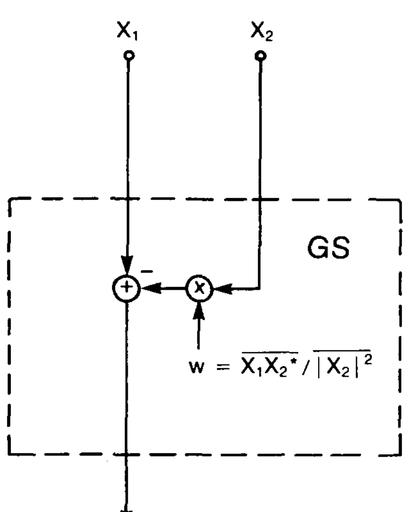

Fig. 1(b) — Basic 2-input GS canceller

If we do not have an infinite number of input samples then the decorrelation weight associated with each 2-input GS canceller is estimated by using finite averaging. The three methods of performing GS cancellation are nonconcurrent, concurrent, and systolic processing. The last two are described in more detail in Refs. 3 and 4. For this analysis, we assume nonconcurrent processing whereby the GS weights are estimated from a block of input data and applied to subsequent or previous input data. For clarity, data that are used to calculate the GS weights are denoted by lower case x's and are called the concurrent data. The data to which the computed weights are applied are denoted by upper case X's and are called the nonconcurrent data.

We briefly describe the nonconcurrent GS canceller. Let $x_n^{(m)}$ represent the outputs of the 2-input GSs on the $(m - 1)$ level. The GS weights are computed from these outputs. Then outputs of the 2-input GSs at the mth level are given by

$$x_n^{(m+1)}(k) = x_n^{(m)}(k) - w_n^{(m)} x_{N-m}^{(m)}(k) \quad \begin{array}{l} n = 0, 1, \ldots, N - m - 1, \\ m = 1, 2, \ldots, N - 1 \\ k = 1, 2, \ldots, K . \end{array} \tag{1}$$

Note that $x_n^{(1)} = x_n$. The weight $w_n^{(m)}$, seen in Eq. (1), is computed so as to decorrelate $x_n^{(m+1)}$ with $x_{N-m}^{(m)}$. For K input samples per channel, this weight is estimated as

$$w_n^{(m)} = \frac{\sum_{k=1}^{K} x_{N-m}^{(m)*}(k) x_n^{(m)}(k)}{\sum_{k=1}^{K} |x_{N-m}^{(m)}(k)|^2}, \tag{2}$$

where * denotes the complex conjugate and $|\cdot|$ is the magnitude. Here k indexes the sampled data.

For the nonconcurrent canceller, let $X_n^{(m)}$ represent the nonconcurrent data outputs of the 2-input GSs on the $(m - 1)$ level. Then the outputs of the 2-input GSs of the mth level are given by

$$X_n^{(m+1)} = X_n^{(m)} - w_n^{(m)} X_{N-m}^{(m)}, \quad \begin{array}{l} n = 0, 1, \ldots, N - m - 1, \\ m = 1, 2, \ldots, N - 1 \end{array} \tag{3}$$

where $X_n^{(1)} = X_n$ and $w_n^{(m)}$ is calculated by using Eq. (2), i.e., these weights are computed from a block of data that does not include X_n.

For this development unless otherwise noted we make the following assumptions:

1. The samples of $x_0, x_1, \ldots, x_{N-1}$ and $X_0, X_1, \ldots, X_{N-1}$ are Gaussian complex random variables (r.v.) when conditioned on their respective noise power level.

2. These same r.v.'s when conditioned on their respective noise power level are samples from stationary processes with zero mean.

3. $x_{n_1}(k_1)$ is independent of $X_{n_2}(k_2)$ for all k_1, k_2, n_1, n_2.

4. The desired signal is not present during weight computation and is not in the auxiliary channels.

5. $K \geq N$

3. OUTPUT MEASURE

Figure 2 shows a simplified N-input GS canceller structure for nonconcurrent processing. $GS_{K,N}$ indicates that an N-input GS structure uses K samples for each channel to compute the weights interior to the GS structure. Note for the nonconcurrent structure that the weights are computed from the $x_0, x_1, \ldots, x_{N-1}$ data blocks and are applied to $X_0, X_1, \ldots, X_{N-1}$. The 0th channel (or the far left channel in Fig. 2) is always designated the main channel; the others are called auxiliary channels or just plain AUXs. The output of the nonconcurrent processor is denoted by Z_{nw}. Figure 3 shows the GS structure for $K = \infty$ where the concurrent and nonconcurrent orthogonal outputs are z_n and Z_n, $n = 0, 1, 2, \ldots, N-1$, respectively.

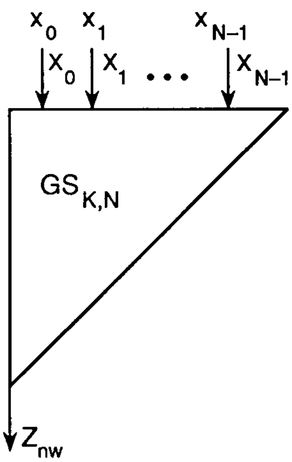

Fig. 2 — Representation of nonconcurrent weighting of GS canceller

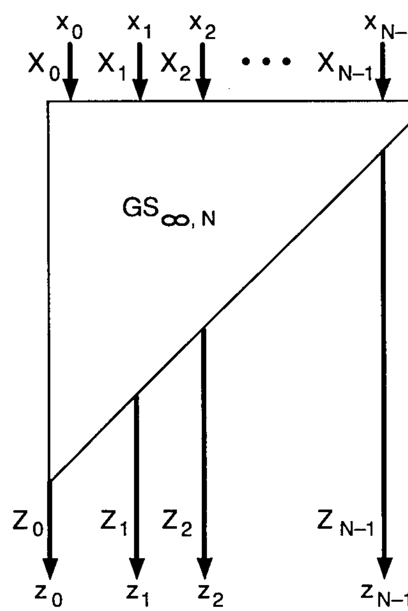

Fig. 3 — GS representation with N orthogonalized output channels

For any set of interior GS weights estimated, there is an equivalent linear weighting of the input channel. We denote this equivalent estimated weighting by the N-length vector $\hat{\mathbf{w}}$, where

$$\hat{\mathbf{w}} = (w_0, w_1, \ldots, w_{N-1})^T \tag{4}$$

and T denotes the transpose vector (or matrix) operation. For the GS canceller the weighting on the main channel is constrained to be 1 or $w_0 = 1$. Let $\sigma^2_{\min}$ be defined as the steady state output noise power residue and

R is steady state input covariance matrix (main and AUXs, an $N \times N$ matrix), (5)

$\hat{R}$ is estimated input noise covariance matrix using $X_0, X_1, \ldots, X_{N-1}$ (in this case X_0 consists of noise only),

$\hat{\sigma}_{nw}^2$ is transient output noise power residue associated with nonconcurrent weighting normalized to $\sigma_{\min}^2$, and

$\overline{\hat{\sigma}_{nw}^2}$ is expected value of $\hat{\sigma}_{nw}^2$ averaged over $X_0, X_1, \ldots, X_{N-1}$. We call this quantity the X-average transient output noise power residue.

Note that the last three quantities defined are random variables.

By using the above definitions, we can show that

$$\hat{\sigma}_{nw}^2 = \frac{|Z_{nw}|^2}{\sigma_{\min}^2} = \frac{\hat{\mathbf{w}}^t \hat{R} \hat{\mathbf{w}}}{\sigma_{\min}^2}, \tag{6a}$$

where t denotes the complex transpose. Note that because of assumption 3, Section 2,

$$\overline{\hat{\sigma}_{nw}^2} = \frac{\hat{\mathbf{w}}^t R \hat{\mathbf{w}}}{\sigma_{\min}^2}. \tag{6b}$$

Also since $w_0 = 1$, from assumption 4, Section 2, the desired signal is passed directly to the output uncancelled. Hence, the output signal power is unchanged from input to output so the expected value of $\hat{\sigma}_{nw}^2$ is equal to the cancellation ratio. We define the normalized output noise power residue as

$$\sigma_{nw}^2(K,N) = E\{\hat{\sigma}_{nw}^2\} = E\{\overline{\hat{\sigma}_{nw}^2}\}, \tag{7}$$

where $E\{\cdot\}$ denotes the expected value. Thus the above is the average (or 1st moment) of the transient normalized output noise power residue. This output measure is commonly used to grade the convergence performance of the SMI canceller.

4. INVARIANT TRANSFORMS

In the section, we discuss two types of matrix transforms on the input data that significantly simplify the analysis. Let C be any $N \times N$ nonsingular matrix. Reference 1 shows that transforming the input channels $x_0, x_1, \ldots, x_{N-1}$ by this matrix does not change the transient or steady state output residue of the GS/SMI canceller. The GS canceller in the steady state is equivalent to a matrix transformation of the input data vector in which the matrix is nonsingular and upper triangular. Therefore, we can transform the input data by using a steady-state GS canceller (with its orthgonalized channels) prior to performing the transient analysis. Figure 4 shows an equivalent configuration of the GS canceller in the transient state. Here the matrix transform C is implemented by passing the input channels through a $\mathrm{GS}_{\infty,N}$ structure followed by a power equalizer on the output auxiliary channels. The output powers of the AUX channels after power equalization are equal to $\sigma_{\min}^2$. Note that each input channel into the $\mathrm{GS}_{K,N}$ structure is orthogonal in the steady state to the other channels.

The structure shown in Fig. 4 illustrates that any GS canceller structure can be divided into two parts: a deterministic, steady-state frontend processor in which the main channel is decorrelated from the auxiliary channels and a stochastic backend processor which is driven by uncorrelated equal powered noise in each channel. The backend processor is independent of the input covariance matrix, and the auxiliary weights associated with the backend processor go to zero as $K \rightarrow \infty$. Hence the convergence properties of the GS canceller can be studied by analyzing the convergence properties of the backend processor whereby the input channels are spatially orthogonal and of equal power.

A second matrix transform to be used in the the forthcoming analysis is now discussed. Let Φ be any $K \times K$ unitary matrix, i.e., $\Phi\Phi^t = I_K$ where I_K is the $K \times K$ identity matrix. We transform each input channel K-length data vector, $\mathbf{x}_n$, $n = 0, 1, 2, \ldots, N-1$ by Φ such that

$$\mathbf{x}_n' = \Phi\, \mathbf{x}_n = 0, 1, \ldots, N-1, \tag{8}$$

where $\mathbf{x}_n'$, $n = 0, 1, \ldots, N-1$ is the resultant output data set. If we input this data set to a $GS_{K,N}$ canceller, then the estimated weights using the $\mathbf{x}_n$ inputs are identical to those using the $\mathbf{x}_n'$ inputs [3].

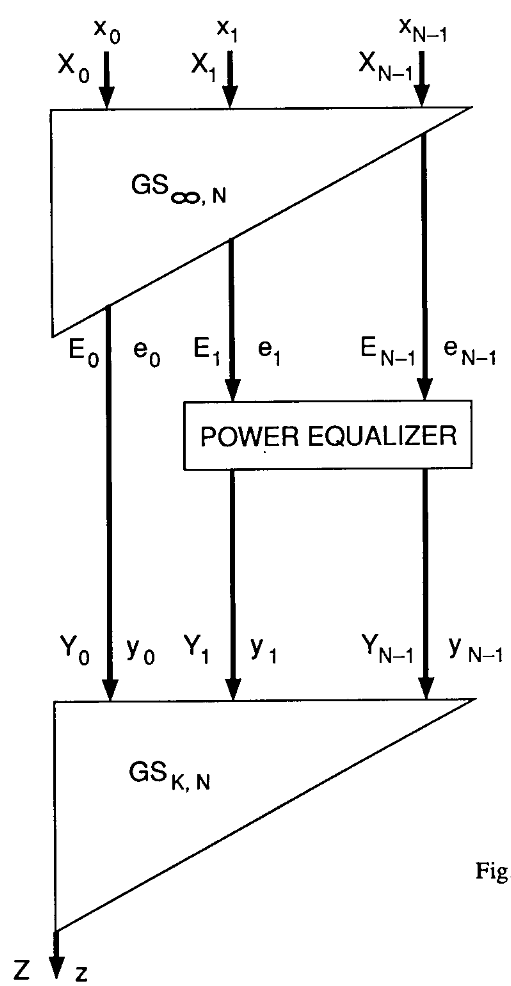

Fig. 4 — Residue-equivalent $GS_{K,N}$ canceller

5. GS DECOMPOSITION

In this section we discuss a decomposition of a GS structure that was first introduced in Ref. 3. A $GS_{K,N}$ structure can be decomposed as shown in Fig. 5 into a first-level processor followed by a $GS_{K,N-1}$ structure. The output K-length vectors (those used in computing the next level weights) of the first-level processor can be written as

$$\mathbf{y}_n = \mathbf{x}_n - \hat{w}_n \mathbf{x}_{N-1}, \quad \hat{w}_n = \frac{\mathbf{x}_{N-1}^t \mathbf{x}_n}{\mathbf{x}_{N-1}^t \mathbf{x}_{N-1}}, \; n = 0, 1, \ldots, N-2, \tag{9}$$

or

$$\mathbf{y}_n = \mathbf{x}_n - \frac{\mathbf{x}_{N-1}^t \mathbf{x}_n}{\mathbf{x}_{N-1}^t \mathbf{x}_{N-1}} \mathbf{x}_{N-1}.$$

Thus

$$\mathbf{y}_n = \left(I_K - \frac{\mathbf{x}_{N-1}\mathbf{x}_{N-1}^t}{\mathbf{x}_{N-1}^t \mathbf{x}_{N-1}} \right) \mathbf{x}_n, \qquad n = 0, 1, 2, \ldots, N-1. \tag{10}$$

It can be shown [3] that

$$I_K - \frac{\mathbf{x}_{N-1}\mathbf{x}_{N-1}^t}{\mathbf{x}_{N-1}^t \mathbf{x}_{N-1}} = \Phi^t \Gamma \Phi, \tag{11}$$

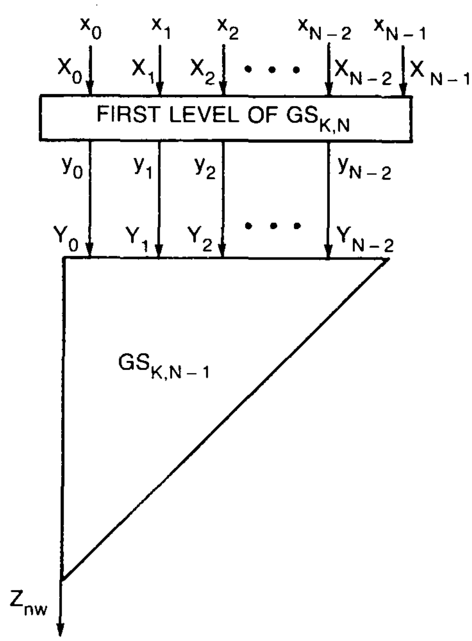

Fig. 5 — Decomposition of $GS_{K,N}$

where Φ is a $K \times K$ unitary matrix and Γ is a diagonal matrix whose first element is 0 and all other diagonal elements are equal to 1. Thus

$$\mathbf{y}_n = \Phi^t \Gamma \Phi \mathbf{x}_n, \quad n = 0, 1, \ldots, N-2. \tag{12}$$

As discussed in Section 4, we can transform the output data set $\mathbf{y}_n$, $n = 0, 1, \ldots, N-2$ by a unitary matrix Φ and not change the equivalent transient weighting vector of the $GS_{K,N-1}$ structure. Thus we write

$$\mathbf{u}_n' = \Phi \mathbf{y}_n = \Gamma \Phi \mathbf{x}_n, \quad n = 0, 1, \ldots, N-2. \tag{13}$$

Now set $\mathbf{v}_n = \Phi \mathbf{x}_n$. By using the form of Γ and setting $\mathbf{u}_n' = (u_{n1}', u_{n2}', \ldots, u_{nk}')^T, \mathbf{v}_n = (v_{n1}, v_{n2}, \ldots, v_{nk})^T$, where T denotes transpose, it follows from Eq. (13) that

$$u_{n1}' = 0, \tag{14a}$$

$$u_{nk}' = v_{nk}, \quad k = 2, 3, \ldots, K.$$

Define

$$u_{nk} = v_{n,k+1}, \quad k = 1,2, \ldots, K-1. \tag{14b}$$

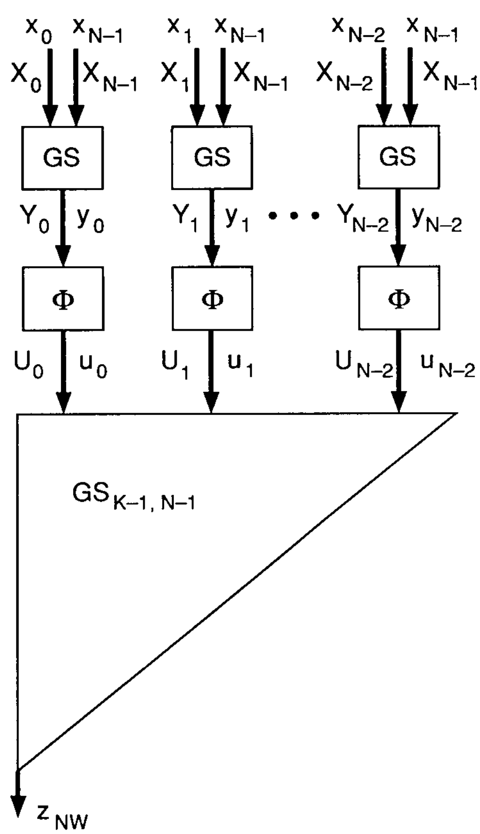

Fig. 6 — Further decomposition of $GS_{K,N}$

Hence, the number of input r.v.'s to the $GS_{K,N-1}$ structure has been reduced by 1 (Fig. 6). Note $U_n = Y_n$, $n = 0, 1, \ldots, N-2$.

6. NONSTATIONARY NOISE MODEL

In this section we present the temporal nonstationary noise model of the inputs to the GS canceller. We consider separately the modeling of internal and external noise sources. We assume that the average power level from external interference sources is not a constant from sampling time to sampling time. Our methodology is to derive lower and upper bounds on the output noise power residue of the adaptive canceller when these power levels are known exactly at each sampling time. Thus these bounds are conditioned on the K-specified power levels of the external noise. Thereafter a joint probability distribution function can be assigned to the K-specified power levels and upper/lower bounds of performance can be derived by integrating the conditioned upper/lower bounds over the joint probability distribution function. More specifically, define

$R_{aa,e}(k)$ = $(N-1) \times (N-1)$ auxiliary covariance matrix of the external interference at time step k, $k = 1, 2, \ldots, K$.

$\mathbf{r}_{am,e}(k)$ = $N-1$ length cross-correlation vector between the auxiliaries and the main channels of the external interference at time step, k, $k = 1, 2, \ldots, K$.

We set

$$R_{aa,e}(k) = \sigma_k^2 \, C_{aa,e} \tag{15}$$

$$\mathbf{r}_{am,e}(k) = \sigma_k^2 \, \mathbf{c}_{am,e} \tag{16}$$

where

$C_{aa,e}$ is a constant $(N-1) \times (N-1)$ normalized auxiliary cross-correlation matrix of the external noise sources,

$\mathbf{c}_{am,e}$ is a constant $N-1$ length normalized cross-correlation vector of the external noise sources, and

σ_k^2 is input power to each main and auxiliary channel at time instant k, $k = 1, 2, \ldots, K$. Assume $\sigma_{k_1}^2 \neq \sigma_{k_2}^2$ for $k_1 \neq k_2$.

The normalization of $C_{aa,e}$ and $\mathbf{c}_{am,e}$ results from setting all of the $\sigma_k^2 = 1$ and computing the resulting auxiliary covariance matrix and the cross-correlation vector between main and auxiliaries of the external interference, respectively.

Define

$$\mathbf{w}_a = (w_1, w_2, \ldots, w_{N-1})^T = \text{optimum canceller weighting vector,}$$

$$R_{aa} = (N-1) \times (N-1) \text{ auxiliary covariance matrix,}$$

$$\mathbf{r}_{am} = N-1 \text{ length auxiliary-main cross-correlation vector,}$$

$$R_{aa,e} = (N-1) \times (N-1) \text{ external interference covariance matrix.}$$

We assume all internal noises are temporally and spatially statistically independent and identically distributed, zero mean, complex stationary Gaussian noise processes with power σ^2, which without loss of generality is set equal to 1. These internal noises are additive in each channel. We note that all other powers are referenced from $\sigma^2 = 1$.

For the noise model described,

$$\mathbf{w} = R_{aa}^{-1} \, \mathbf{r}_{am} \tag{17}$$

$$R_{aa} = R_{aa,e} + I_{N-1} \tag{18}$$

$$\mathbf{r}_{am} = \left(\frac{1}{K} \sum_{k=1}^{K} \sigma_k^2 \right) \mathbf{c}_{am,e} \tag{19}$$

$$R_{aa,e} = \left(\frac{1}{K} \sum_{k=1}^{K} \sigma_k^2 \right) C_{aa,e} \tag{20}$$

where I_{N-1} is the $(N-1) \times (N-1)$ identity matrix that represents the internal noise covariance matrix. Since $C_{aa,e}$ is a hermitian matrix, we can decompose it as

$$C_{aa,e} = \Phi_e \, \Gamma \, \Phi_e^t \tag{21}$$

where

Γ is the real diagonal matrix of eigenvalues of $C_{aa,e}$

Φ_e is the unitary eigenmatrix of $C_{aa,e}$, i.e. , $\Phi_e \, \Phi_e^t = I_{N-1}$.

Define

$(\gamma_1, \gamma_2, \ldots, \gamma_{N-1})$ = eigenvalues of $C_{aa,e}$,

σ_i^2 = steady state ($K \to \infty$) main channel internal noise power residue (after cancellation),

σ_e^2 = steady state main channel external interference power residue,

σ_{in}^2 = average noise power (internal and external) of the main channel input, and

$\sigma_{\min}^2$ = total steady state noise power residue.

It can be shown that

$$\sigma_{\min}^2 = \sigma_i^2 + \sigma_e^2 = \sigma_{\text{in}}^2 - \mathbf{r}_{am}^t R_{aa}^{-1} \mathbf{r}_{am} \tag{22}$$

and

$$\sigma_i^2 = 1 + \mathbf{r}_{am}^t R_{aa}^{-2} \mathbf{r}_{am}. \tag{23}$$

Thus

$$\sigma_e^2 = \sigma_{\text{in}}^2 - \mathbf{r}_{am}^t R_{aa}^{-1} \mathbf{r}_{am} - \sigma_i^2. \tag{24}$$

As mentioned in Section 4, the auxiliary inputs $(x_1, x_2, \ldots, x_{N-1})^T$ of a sidelobe canceller can be multiplied by an arbitrary nonsingular $(N - 1) \times (N - 1)$ matrix transform such that the transient residue is unchanged. Consider the implicit matrix transform illustrated in Fig. 7. In this figure, $\Phi_{aa,e}^*$ statistically orthogonalizes the auxiliaries with respect to one other. We note that the internal noise components of y_n', $n = 1,2, \ldots, N-1$ have unit power since Φ_e^* is a unitary transform. The outputs of the $\Phi_{aa,e}^*$ transform are denoted by y_1', y_2', $\ldots$, y_{N-1}'. We optimally weight each of these by the w_1', w_2', $\ldots$, w_{N-1}', which minimizes the output residue of y_0'. This weighting does not affect the transient residue, as discussed in Section 4. Next y_0', y_1', $\ldots$, and y_{N-1}' are normalized so that the average power (over all K samples) is equal to 1. This normalization does not change the normalized output noise power residue. The outputs z_0, z_1, $\ldots$, z_{N-1} are the resultant outputs of this normalization procedure. When conditioned on their respective power levels, these inputs entering the $GS_{K,N}$ canceller are spatially and temporally statistically independent of one another. More explicitly if $\mathbf{z}_n = [z_n(1), z_n(2), \ldots, z_n(K)]^T$ is the input vector of K samples then

$$E\{z_{n_1}(k_1)\, z_{n_2}^*(k_2)\} = 0, \text{ unless } n_1 = n_2 \text{ and } k_1 = k_2, \tag{25}$$

and

$$\frac{1}{K} \sum_{k=1}^{K} E\{|z_n(k)|^2\} = 1. \tag{26}$$

It is straightforward to show that

$$E\{|z_0(k)|^2\} = \alpha_0(\gamma_0\, \sigma_k^2 + \sigma_i^2), \quad k = 1,2, \ldots, K \tag{27}$$

and

$$E\{|z_n(k)|^2\} = \alpha_n(\gamma_n\, \sigma_k^2 + 1),\ k = 1, 2, \ldots, K, \quad n = 1, 2, \ldots, N - 1 \tag{28}$$

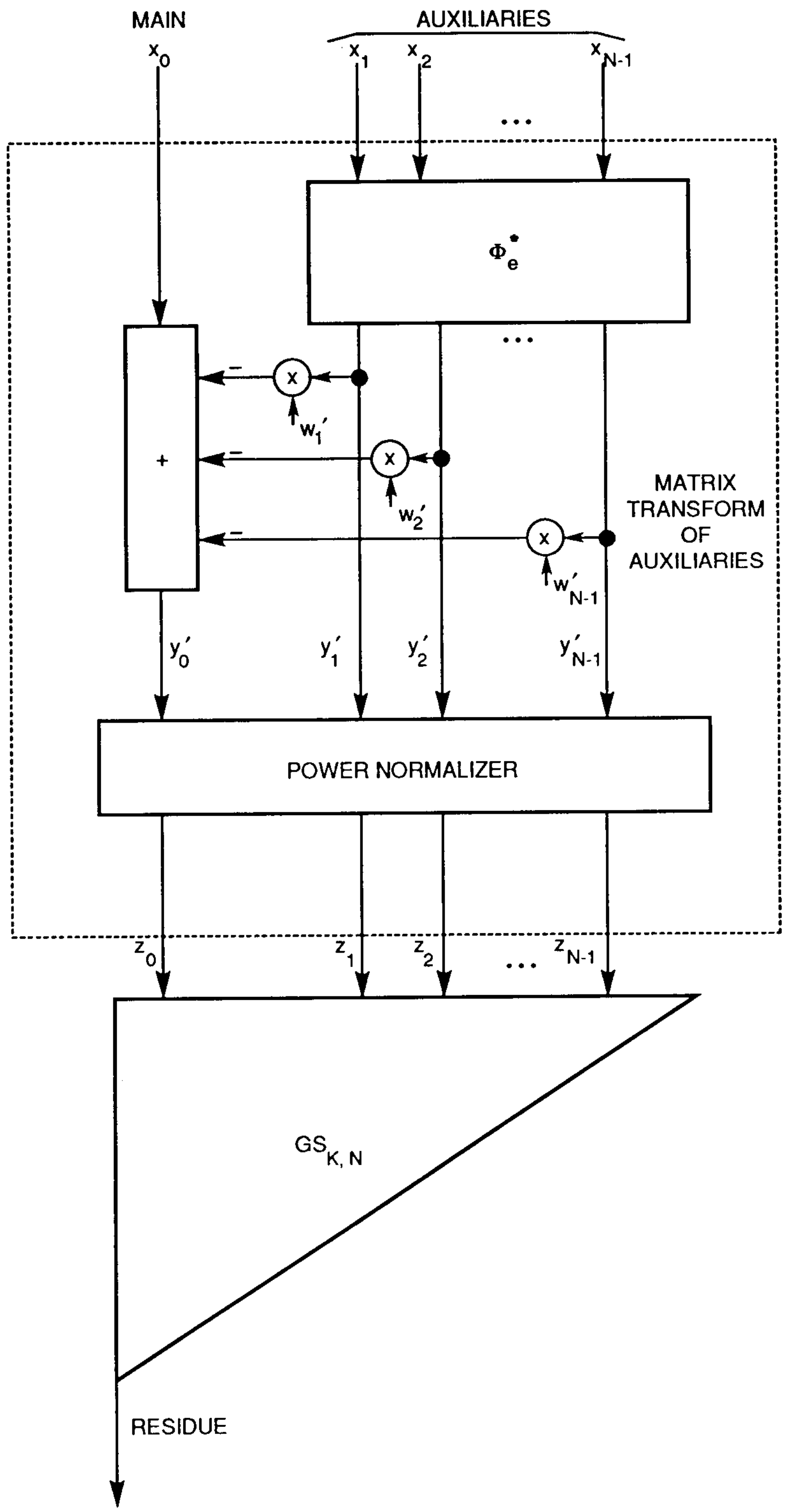

Fig. 7 — Equivalent canceller using correlated inputs

where

$$\overline{\sigma^2} = \frac{1}{K} \sum_{k=1}^{K} \sigma_k^2 \tag{29}$$

$$\gamma_0 = \frac{\sigma_e^2}{\overline{\sigma^2}} \tag{30}$$

$$\alpha_0 = [\gamma_0 \, \overline{\sigma^2} + \sigma_i^2]^{-1} \tag{31}$$

$$\alpha_n = [\gamma_n \, \overline{\sigma^2} + 1]^{-1}, \; n = 1,2, \ldots , N - 1. \tag{32}$$

We define

$$\boldsymbol{\lambda}^{(0)} = \begin{bmatrix} \alpha_0(\gamma_0 \, \sigma_1^2 + \sigma_i^2) \\ \alpha_0(\gamma_0 \, \sigma_2^2 + \sigma_i^2) \\ \cdot \\ \cdot \\ \cdot \\ \alpha_0(\gamma_0 \, \sigma_K^2 + \sigma_i^2) \end{bmatrix} \tag{33}$$

and

$$\boldsymbol{\lambda}^{(n)} = \begin{bmatrix} \alpha_n \, (\gamma_n \, \sigma_1^2 + 1) \\ \alpha_n \, (\gamma_n \, \sigma_2^2 + 1) \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \, (\gamma_n \, \sigma_K^2 + 1) \end{bmatrix}. \tag{34}$$

Thus each sample in a given channel can be characterized by specified variances. The K-length data vector in nth channel is completely characterized by $\boldsymbol{\lambda}^{(n)}$, $n = 0, 1, \ldots , N-1$, and the fact that (1) its elements when conditioned on their respective power levels are spatially and temporally statistically independent all other data samples, and (2) are complex Gaussian processes. We note that without loss of generality we can order the σ_k^2, $k = 1, 2, \ldots , K$ as

$$\sigma_1^2 < \sigma_2^2 < \sigma_3^2 \; \cdots \; < \sigma_K^2.$$

7. 2-INPUT GS CANCELLER

The basis for understanding the convergence properties of a GS canceller begins with studying the 2-input GS canceller. We assume the input data noise model as defined in Section 6 and that these data satisfy assumptions (1) through (5) given in Section 2. We set $\Lambda_n, n = 0, 1, 2, \ldots , N-1$

equal to a $K \times K$ diagonal matrix, where the kth diagonal elements of Λ_n are given by the kth element of $\boldsymbol{\lambda}^{(n)}$. We write the output residue as Z_{nw}. Thus

$$Z_{nw} = X_0 - \hat{w} X_1, \tag{35}$$

where

$$\hat{w} = \frac{\mathbf{x}_1^t \mathbf{x}_0}{\mathbf{x}_1^t \mathbf{x}_1}. \tag{36}$$

Furthermore, we can show that the normalized X-averaged transient output noise power residue is given by

$$\overline{\hat{\sigma}^2_{nw}} = \frac{E\{\,|Z_{nw}|^2\,|\,\mathbf{x}_0,\mathbf{x}_1\}}{\sigma^2_{\min}}, \tag{37}$$

$$= 1 + |\hat{w}|^2,$$

$$= 1 + \frac{|\mathbf{x}_1^t \mathbf{x}_0|^2}{(\mathbf{x}_1^t \mathbf{x}_1)^2}.$$

Because the elements of $\mathbf{x}_0$ are Gaussian and independent of the elements of $\mathbf{x}_1$,

$$E\{\overline{\hat{\sigma}^2_{nw}}\,|\,\mathbf{x}_1\} = 1 + \frac{\mathbf{x}_1^t \Lambda_0 \mathbf{x}_1}{(\mathbf{x}_1^t \mathbf{x}_1)^2}. \tag{38}$$

Now we can write

$$\mathbf{x}_1 = \Lambda_1^{1/2} \mathbf{v}_1 \tag{39}$$

where $\mathbf{v}_1$ is a vector of identically distributed, zero mean, unit variance, independent complex Gaussian r.v.'s with independent real and imaginary parts. Thus

$$E\{\overline{\hat{\sigma}^2_{nw}}\,|\,\mathbf{v}_1\} = 1 + \frac{\mathbf{v}_1^t \Lambda_0 \Lambda_1 \mathbf{v}_1}{(\mathbf{v}_1^t \Lambda_1 \mathbf{v}_1)^2}. \tag{40}$$

It is shown in Appendix A that

$$\sigma^2_{nw}(K,2) = 1 + \sum_{k=1}^{K} \sum_{\substack{n=1 \\ n \neq k}}^{K} \lambda_n^{(0)}\, a(n,k)\, \tilde{F}\,(\lambda_n^{(1)}, \lambda_k^{(1)}) \tag{41}$$

where

$$a(n,\ k) = \begin{cases} \dfrac{1}{\lambda_n^{(1)}}\ , \ 1 \le n,\ k \le 2\ ;\ \text{for } K = 2 \\[2ex] \dfrac{(\lambda_n^{(1)})^{K-3}}{\prod\limits_{\substack{m=1 \\ m \ne n,k}}^{K} (\lambda_n^{(1)} - \lambda_m^{(1)})};\ 1 \le n,k \le K\ ;\ \text{for } K > 2 \end{cases} \tag{42}$$

$$\tilde{F}(\lambda_n^{(1)},\ \lambda_k^{(1)}) = \frac{\lambda_n^{(1)}}{\lambda_n^{(1)} - \lambda_k^{(1)}} \left[1 - \frac{\lambda_k^{(1)}}{\lambda_n^{(1)} - \lambda_k^{(1)}} \ln \frac{\lambda_n^{(1)}}{\lambda_k^{(1)}} \right]. \tag{43}$$

8. PRELIMINARY DEFINITIONS AND THEOREMS

Before deriving bounds for the convergence of an N-input canceller, we give preliminary theorems necessary for obtaining these bounds. Observe $\lambda_1^{(n)} < \lambda_2^{(n)} < \ldots < \lambda_K^{(n)}$ and let $\mathbf{e} = (e_1,\ e_2,\ \ldots,\ e_K)^T$. Define the following $K - l + 1$ length vectors

$$\mathbf{e}_L\ (l) = (e_1,\ e_2,\ \ldots\ ,\ e_{K-l+1}) \tag{44}$$

$$\mathbf{e}_U\ (l) = (e_l,\ e_{l+1},\ \ldots\ ,\ e_K) \tag{45}$$

The L or U subscript on a vector indicates whether the lower or upper $K - l + 1$ elements of that vector are used. Let the random vector $\mathbf{v}_1$ be as defined in Section 7. Define the $(K - l + 1) \times (K - l + 1)$ diagonal matrices $\Lambda_{U,n}$ and $\Lambda_{L,n}$ whose diagonal elements are given in order by the elements of $\boldsymbol{\lambda}_U^n(l)$ and $\boldsymbol{\lambda}_L^{(n)}(l)$, respectively. The following quantities are defined and are used in evaluating upper and lower bounds of $\sigma_{nw}^2(K,\ N)$:

$$U(l,\ K,\ \boldsymbol{\lambda}^{(l)},\boldsymbol{\lambda}^{(0)}) = E\left\{ \frac{\mathbf{v}_1^t \Lambda_{U,0} \Lambda_{U,l} \mathbf{v}_1}{(\mathbf{v}_1^t \Lambda_{L,l} \mathbf{v}_1)^2} \right\}, \tag{46}$$

and

$$L(l,\ K,\ \boldsymbol{\lambda}^{(l)},\boldsymbol{\lambda}^{(0)}) = E\left\{ \frac{\mathbf{v}_1^t \Lambda_{L,0} \Lambda_{L,l} \mathbf{v}_1}{(\mathbf{v}_1^t \Lambda_{U,l} \mathbf{v}_1)^2} \right\}. \tag{47}$$

The following theorems give formulations for $U(l, K, \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)})$ and $L(l, K, \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)})$.

Theorem 1: Define for $l = 1, 2, \ldots, N-1$

$$a_L(n, k, l, \boldsymbol{\lambda}^{(l)}) = \begin{cases} \dfrac{1}{\lambda_n^{(l)}}, & 1 \le n, k \le 2, \text{ if } K - l + 1 = 2 \\ \dfrac{[\lambda_n^{(l)}]^{K-l-2}}{\prod\limits_{\substack{m=1 \\ m \ne n,k}}^{K-l+1} (\lambda_n^{(l)} - \lambda_m^{(l)})}, & 1 \le n, k \le K - l + 1, K - l + 1 > 2 \\ \text{undefined if } n = k \end{cases} \tag{48}$$

$$a_U(n, k, l, \boldsymbol{\lambda}^{(l)}) = \begin{cases} \dfrac{1}{\lambda_n^{(l)}}, & K - 1 \le n, k \le K; \text{ if } K - l + 1 = 2 \\ \dfrac{[\lambda_n^{(l)}]^{K-l-2}}{\prod\limits_{\substack{m=l \\ m \ne n,k}}^{K} (\lambda_n^{(l)} - \lambda_m^{(l)})}, & l \le n, k \le K, K - l + 1 > 2 \\ \text{undefined if } n = k \end{cases} \tag{49}$$

$$G_U(k, l, \boldsymbol{\lambda}^{(l)}) = \sum_{\substack{n=1 \\ n \ne k}}^{K-l+1} a_L(n, k, l, \boldsymbol{\lambda}^{(l)})\, \tilde{F}(\lambda_n^{(l)}, \lambda_k^{(l)}),\ l = 1, 2, \ldots, N-1; 1 \le k \le K-l+1, \tag{50}$$

$$G_L(k, l, \boldsymbol{\lambda}^{(l)}) = \sum_{\substack{n=l \\ n \ne k}}^{K} a_U(n, k, l, \boldsymbol{\lambda}^{(l)})\, \tilde{F}(\lambda_n^{(l)}, \lambda_k^{(l)}),\ l = 1, 2, \ldots, N-1; l \le k \le K, \tag{51}$$

where $\tilde{F}(\cdot, \cdot)$ is defined by Eq. (43). Then

$$U(l, K, \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)}) = \sum_{k=1}^{K-l+1} \frac{\lambda_{l+k-1}^{(0)} \lambda_{l+k-1}^{(l)}}{\lambda_k^{(l)}}\, G_U(k, l, \boldsymbol{\lambda}^{(l)}), l = 1, 2, \ldots, N-1, \tag{52}$$

$$L(l, K, \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)}) = \sum_{k=l}^{K} \frac{\lambda_{k-l+1}^{(0)}\, \lambda_{k-l+1}^{(l)}}{\lambda_k^{(l)}}\, G_L(k, l, \boldsymbol{\lambda}^{(l)}), l = 1, 2, \ldots, N-1. \tag{53}$$

Proof: The proof is given in Appendix B.

We use the following variant of the Poincare′ separation theorem [11].

Theorem 2: Let Λ be a $K \times K$ hermitian matrix with eigenvalues $\lambda_1 \leq \lambda_2 \ldots \leq \lambda_K$. Let B be a $(K-1) \times K$ matrix with $K-1$ orthonormal columns. Let $\lambda_1' \leq \lambda_L' \ldots \leq \lambda_{K-1}'$ be the eigenvalues of B^*AB^t. The following inequalities hold:

$$\lambda_1 \leq \lambda_1' \leq \lambda_2 \leq \lambda_2' \leq \ldots \leq \lambda_{K-1} \leq \lambda_{K-1}' \leq \lambda_K. \tag{54}$$

Note that if the equality is removed so that $\lambda_1 < \lambda_2 \ldots < \lambda_K$, then

$$\lambda_1 < \lambda_1' < \lambda_2 < \lambda_2' < \ldots < \lambda_{K-1} < \lambda_{K-1}' < \lambda_K. \tag{55}$$

Thus the λ_n', $n = 1, 2, \ldots, K-1$ are distinct if the λ_n, $n = 1, 2, \ldots, K$ are distinct.

Define

$$\boldsymbol{\lambda}^{(n)\prime} = (\lambda_1^{(n)\prime}, \lambda_2^{(n)\prime}, \ldots, \lambda_{K-1}^{(n)\prime}). \tag{56}$$

Theorem 3: If $\lambda_1^{(n)} < \lambda_1^{(n)\prime} < \lambda_2^{(n)} < \lambda_2^{(n)\prime} < \ldots < \lambda_{K-1}^{(n)} < \lambda_{K-1}^{(n)\prime} < \lambda_K^{(n)}$, for $n = 0, 1, \ldots, N-1$, then

$$L(l+1, K, \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)}) < L(l, K-1, \boldsymbol{\lambda}^{(l)\prime}, \boldsymbol{\lambda}^{(0)\prime})$$

$$< U(l, K-1, \boldsymbol{\lambda}^{(l)\prime}, \boldsymbol{\lambda}^{(0)\prime}) < U(l+1, K, \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}_0^{(l)}). \tag{57}$$

Proof: The proof of this is straightforward and follows by direct comparison of the terms (which are all positive) in these expressions.

9. BOUNDS FOR NONCONCURRENT GS CANCELLER

In this section, we prove the following theorem pertaining to nonconcurrent GS cancellation in nonstationary noise.

Theorem 4: If $\lambda_1^{(n)} < \lambda_2^{(n)} < \ldots < \lambda_K^{(n)}$ for $n = 0, 1, \ldots, N-1$ and assumptions (1) through (5) hold, then

$$\prod_{l=1}^{N-1} [1 + L(l, K, \boldsymbol{\lambda}^{(N-l)}, \boldsymbol{\lambda}^{(0)})] < \sigma_{nw}^2(K, N) < \prod_{l=1}^{N-1} [1 + U(l, K, \boldsymbol{\lambda}^{(N-l)}, \boldsymbol{\lambda}^{(0)})]. \tag{58}$$

Proof: We prove this by mathematical induction. We have shown that the theorem is true for $N = 2$ (see Section 7 and the definitions of L, U, given in Section 8). Thus, we can assume that the theorem is true for all integers less than or equal to some upper bound, $N-1$. We show that the theorem is true for N, which implies that it is true for any $N \geq 2$. Recall that the nonconcurrent inputs are equi-power. In addition the samples of $\mathbf{x}_0$ when conditioned on their respective power levels are spatially and temporally statistically independent of the samples in the auxiliary channels.

We decompose the $GS_{K,N}$ processor as shown in Fig. 5 and further reduced in Fig. 6. Our methodology is to derive bounds on the output noise power conditioned on $\mathbf{x}_{N-1}$, which we write as $E\{|Z_{nw}|^2 \; | \mathbf{x}_{N-1}\}$. Thereafter, we see that $\sigma^2_{nw}(K, N) = E\{|Z_{nw}|^2\}$ is readily derivable. The K-length output vectors $\mathbf{y}_n$, $N = 0, \ldots, N-2$ of the channels from the first level are given by Eq. (12). As discussed in Section 5, the number of concurrent inputs per channel into the succeeding $GS_{K,\,N-1}$ processor is essentially reduced by one. These concurrent inputs are now given by the $(K-1)$-length vector denoted by $\mathbf{u}_n$, $n = 0, 1, \ldots, N-2$ and defined by (14b). It is straightforward to show that the main channel samples of $\mathbf{u}_0$ conditioned on $\mathbf{x}_{N-1}$ and their respective power levels are spatially and temporally independent of $\mathbf{u}_n$, $n = 1, \ldots, N-2$, and that the nonconcurrent samples, U_n, $n = 0, 1, \ldots, N-2$ are equi-powered. Hence, assumptions (1) through (5) hold for the reduced input set (note for assumption (5), $K-1 \geq N-1$). Hence, the bounds given by Theorem 5 can be applied for the equivalent $GS_{K-1,N-1}$ canceller (see Fig. 6) if the temporal correlation matrix for each channel were known.

We define $\tilde{\sigma}^2_{\min}$ to be the minimum output residue of the $GS_{K-1,\,N-1}$ canceller if we use a finite K in the first level of canceller seen in Fig. 5 and use an infinite number of samples (steady state) in the $GS_{K-1,N-1}$ canceller. Let B be a $(K-1) \times K$ matrix formed by using the second thru Kth rows of Φ, which is defined by Eq. (11). Thus

$$\mathbf{u}_n = B\mathbf{x}_n, \quad n = 0, 1, 2, \ldots, N-2 \tag{59}$$

$$A_n' = E(\mathbf{u}_n^* \mathbf{u}_n^T) = B^* \Lambda_n B^T , \tag{60}$$

where A_n' in the correlation matrix of $\mathbf{u}_n$, $n = 0, 1, \ldots, N-2$. Define $\lambda_1^{(n)\prime}, \lambda_2^{(n)\prime}, \ldots \lambda_{K-1}^{(n)\prime}$ to be the eigenvalues of A_n'. In lieu of the Poincare' Separation Theorem (see Theorem 2) and the original assumption on ordering, (55) holds

Thus invoking Theorem 4, which we have assumed is true $n \leq N-1$, since $\lambda_1^{(n)\prime} < \lambda_2^{(n)\prime} < \ldots < \lambda_{K-1}^{(n)\prime}$, then the conditional expectation of the noise power can be bounded as

$$\prod_{l=1}^{N-2} \left[1 + L(l, K-1, \boldsymbol{\lambda}^{(N-1-l)\prime}, \boldsymbol{\lambda}^{(0)\prime}) \right] < \frac{E\left\{ |Z_{nw}|^2 \; | \mathbf{x}_{N-1} \right\}}{\tilde{\sigma}^2_{\min}}$$

$$< \prod_{l=1}^{N-2} \left[1 + U(l, K-1, \boldsymbol{\lambda}^{(N-1-l)\prime}, \boldsymbol{\lambda}^{(0)\prime}) \right]. \tag{61}$$

We outline the remainder of the proof. The bounds given by (61) can be bounded by those given by Theorem 3. The $\tilde{\sigma}^2_{\min}$ is multiplied through all the new bounds. Next, the joint probability distribution function (p.d.f.) of the elements of $\mathbf{x}_{N-1}$ is multiplied through all the bounds and integrated out. Finally $E\{\tilde{\sigma}^2_{\min}\}$ is bounded by using the results of Section 7. As a result, Theorem 4 follows. End of proof.

We have derived upper and lower bounds of the expected value of the output noise power residue of the GS canceller that depend on the values of the elements of $\boldsymbol{\lambda}^{(n)}$, $n = 0, 1, \ldots, N-1$. As shown in Section 6 these elements depend on α_n, γ_n, $n = 0, 1, \ldots, N-1$, σ_i^2, and

σ_k^2, $k = 1,2,\ldots,K$. Furthermore, α_n, γ_n, and σ_i^2 depend on $C_{aa,e}$, $\mathbf{c}_{am,e}$, σ_{in}^2, and σ_k^2, $k = 1,2,\ldots,K$. We can generalize these bounds by considering σ_k^2, $k = 1,2,\ldots,K$ to be random variables with a joint distribution function $P_{\sigma^2}^2\ (\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2)$. We can consider $\sigma_{nw}^2(K,N)$ to be the expectation for the normalized canceller output residue conditioned on σ_k^2, $k = 1,2,\ldots,K$, and define $\overline{\sigma_{nw}^2}\ (K,N)$ to be the expectation of the normalized canceller output residue. Mathematically this is expressed by

$$\overline{\sigma_{nw}^2}(K,N) = \int \sigma_{nw}^2(K,N)\ dP_{\sigma^2}(\sigma_1^2,\sigma_2^2,\ldots,\sigma_K^2). \tag{62}$$

We note that defining a joint distribution function for $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$ changes the original assumptions on the input processes. The input process is no longer Gaussian and independent from sample to sample. However the input process conditioned on $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$ is Gaussian, and the unconditioned input process is uncorrelated from sample to sample. Defining a joint distribution for $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$ allows a variety of nonstationary interference scenarios to be modeled and evaluated (for example, finite-state jump Markoff processes, continuous or discrete time processes, mixed distributions). Bounds on $\overline{\sigma_{nw}^2}(K,N)$ are found by integrating the lower and upper bounds given by (58) over $dP_{\sigma^2}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2)$.

Finally, we note that one of our assumptions is that $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_K^{(n)}$, $n = 0,1,\ldots,N-1$ are distinct, i.e., $\lambda_i^{(n)} \neq \lambda_j^{(n)}$ for $i \neq j$. We can approximate with arbitrary accuracy the case when these values are not distinct by merely adding or subtracting a small perturbation about each nondistinct $\lambda_k^{(n)}$ to make it distinct. These bounds can then be evaluated to within an arbitrary accuracy of the true value of the bound.

10. SUMMARY

Convergence results for the Sample Matrix Inversion (SMI)/Gram-Schmidt (GS) canceller algorithm in temporally nonstationary noise was investigated by using the GS canceller as an analysis tool. Lower and upper bounds for the convergence rate of the canceller's average output noise power residue normalized to the quiescent average output noise power residue were derived. These bounds are a function of the number of independent samples processed per channel (main or auxiliary), the number of auxiliary input channels, and the external noise environment. The external noise environment was modelled as Gaussian, with a power level specified at each sampling time instant. Furthermore, this model was generalized in the sense that a joint probability distribution function is defined for the power levels over a canceller processing batch. This leads to the capability of modeling and evaluating the SMI/GS canceller in a variety of interference scenarios such as continuous or discrete time processes or a mix of these.

11. REFERENCES

1. I.S. Reed, J.D. Mallett, and L.E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans.* **AES-10**(6), 853-863 (1974).

2. L.E. Brennan and I.S. Reed, "Digital Adaptive Arrays with Weights Computed from and Applied to the Same Sample Set," Proceedings of the 1980 Adaptive Antenna Symposium, RADC-TR-80-378, Vol. 1, Dec. 1980.

3. Karl Gerlach and F.F. Kretschmer Jr., "Convergence Properties of the Gram-Schmidt and SMI Adaptive Algorithms," *IEEE Trans. AES* **26**(1), 44-57 (1990).

4. Karl Gerlach and F.F. Kretschmer, Jr., "Convergence Properties of the Gram-Schmidt and SMI Adaptive Algorithms, Part II," *IEEE Trans. AES* **27**(1), 83-91 (1991).

5. R.A. Monzingo and T.W. Miller, *Introduction to Adaptive Arrays* (John Wiley and Sons, New York, 1980), Ch. 8.

6. W.F. Gabriel, "Building Blocks for an Orthonormal-Lattice-Filter Adaptive Network," NRL Report 8409, July 1980.

7. M.A. Alam, "Orthonormal Lattice Filter-A Multistage, Multichannel Estimation Technique," *Geophys*. **43**, 1368-1383 (1978).

8. B. Friedlander, "Lattice Filters for Adaptive Processing," *Proc. IEEE* **70**(8), 829-867 (1982).

9. Karl Gerlach, "Fast Orthogonalization Networks," *IEEE Trans*. **AP-34**(3), 458-462 (1986).

10. Karl Gerlach, "Convergence Bounds of an SMI/Gram-Schmidt Canceller in Colored Noise," accepted for publication in *IEEE Trans. AES.*

11. R.A. Horn and C.A. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1985) p. 190.

Appendix A

DERIVATION OF 2-INPUT GS CANCELLER NOISE POWER

In this appendix we find an exact expression for $\sigma_{nw}^2(K, 2)$. Starting from Eq. (38), we can write

$$E\left\{\overline{\hat{\sigma}_{nw}^2} \mid \mathbf{x}\right\} = 1 + \sum_{k=1}^{K} \lambda_k^{(0)} \frac{|x_k|^2}{(\mathbf{x}^t\mathbf{x})^2} \tag{A1}$$

where without loss of generality, we have set $\mathbf{x} = \mathbf{x}_1$ and $\mathbf{x} = (x_1, x_2, \ldots, x_K)^T$. Set

$$\alpha_k = \frac{|x_k|^2}{(\mathbf{x}^t\mathbf{x})^2} = \frac{|x_k|^2}{\left[|x_k|^2 + \sum\limits_{\substack{n=1\\n\neq k}}^{K} |x_n|^2\right]^2}, \qquad k = 1, 2, \ldots, K. \tag{A2}$$

We find $E\{\alpha_k\}$, $k = 1, 2, \ldots, K$. Define

$$z_1 = |x_k|^2 \tag{A3}$$

$$z_2 = \sum_{\substack{n=1\\n\neq k}}^{K} |x_n|^2. \tag{A4}$$

Thus

$$\alpha_k = \frac{z_1}{(z_1 + z_2)^2}. \tag{A5}$$

It is shown in Appendix C that if $\lambda_k^{(1)}$, $k = 1, 2, \ldots, K$ are distinct, then z_1 and z_2 have the following p.d.f.'s:

$$p_{z_1}(z_1) = \frac{1}{\lambda_k^{(1)}} e^{-\frac{z_1}{\lambda_k^{(1)}}}, \quad z_1 \geq 0 \tag{A6}$$

$$p_{z_2}(z_2) = \sum_{\substack{n=1\\n\neq k}}^{K} a(n, k) e^{-\frac{z_2}{\lambda_n^{(1)}}}, \quad z_2 \geq 0 \tag{A7}$$

where

$$a(n,\ k) = \begin{cases} \dfrac{1}{\lambda_n^{(1)}}, \text{ for } K = 2 \\ \dfrac{(\lambda_n^{(1)})^{K-3}}{\prod\limits_{\substack{m=1 \\ m \neq n,k}}^{K} (\lambda_n^{(1)} - \lambda_m^{(1)})}; \ k \neq n, \text{ for } K > 2\ . \end{cases} \tag{A8}$$

Thus

$$E\{\alpha_k\} = \int_0^\infty \int_0^\infty \frac{z_1}{(z_1 + z_2)^2}\ p_{z_1}(z_1) p_{z_2}(z_2) dz_1 dz_2. \tag{A9}$$

Substituting (A6) and (A7) into (A9) results in

$$E\{\alpha_k\} = \sum_{\substack{n=1 \\ n \neq k}}^{K} \lambda_n^{(1)} a(n,\ k) F_0(\lambda_n,\ \lambda_k)\ , \tag{A10}$$

where we define

$$F_0\ (\lambda_n^{(1)},\ \lambda_k^{(1)}) = \frac{1}{\lambda_n^{(1)}\lambda_k^{(1)}} \int_0^\infty \int_0^\infty \frac{z_1}{(z_1 + z_2)^2}\ e^{-\frac{z_1}{\lambda_k^{(1)}} - \frac{z_2}{\lambda_n^{(1)}}}\ dz_1\ dz_2\ . \tag{A11}$$

Let $u_1 = z_1/\lambda_k^{(1)}$, $u_2 = z_2/\lambda_n^{(1)}$. With this change of variables the double integral in (A11) becomes

$$F_0(\lambda_n^{(1)}\ \lambda_k^{(1)}) = \lambda_k^{(1)} \int_0^\infty \int_0^\infty \frac{u_1}{(\lambda_k^{(1)} u_1 + \lambda_n^{(1)} u_2)^2}\ e^{-u_1 - u_2} du_1\ du_2. \tag{A12}$$

Define

$$G(\lambda_n^{(1)},\ \lambda_k^{(1)}) = \int_0^\infty \int_0^\infty \frac{1}{\lambda_k^{(1)} u_1 + \lambda_n^{(1)} u_2}\ e^{-u_1 - u_2} du_1 du_2. \tag{A13}$$

Note that

$$F_0(\lambda_n^{(1)},\ \lambda_k^{(1)}) = -\lambda_k^{(1)}\ \frac{\partial G(\lambda_n^{(1)},\ \lambda_k^{(1)})}{\partial \lambda_k^{(1)}}. \tag{A14}$$

We now find an expression for $G(\lambda_n^{(1)}, \lambda_k^{(1)})$. By setting $z_1 = \lambda_k^{(1)} u_1$, $z_2 = \lambda_n^{(1)} u_2$,

$$G(\lambda_n^{(1)}, \lambda_k^{(1)}) = \frac{1}{\lambda_n^{(1)} \lambda_k^{(1)}} \int_0^\infty \int_0^\infty \frac{1}{z_1 + z_2} e^{-\frac{z_1}{\lambda_k^{(1)}} - \frac{z_2}{\lambda_n^{(1)}}} dz_1\, dz_2. \tag{A15}$$

Now

$$G(\lambda_n^{(1)}, \lambda_k^{(1)}) = E\left\{\frac{1}{z_1 + z_2}\right\}, \tag{A16}$$

where z_1 and z_2 are independent random variables with

$$p_{z_1}(z_1) = \frac{1}{\lambda_k} e^{-\frac{z_1}{\lambda_k^{(1)}}}, \tag{A17}$$

and

$$p_{z_2}(z_2) = \frac{1}{\lambda_n^{(1)}} e^{-\frac{z_2}{\lambda_n^{(1)}}}. \tag{A18}$$

An expression for $E\{(z_1 + z_2)^{-1}\}$ is derived in Appendix C, and its form is given by (C12). Thus

$$G(\lambda_n^{(1)}, \lambda_k^{(1)}) = \frac{1}{\lambda_n^{(1)} - \lambda_k^{(1)}} \ln \frac{\lambda_n^{(1)}}{\lambda_k^{(1)}}. \tag{A19}$$

By using (A14), it follows that

$$F_0(\lambda_n^{(1)}, \lambda_k^{(1)}) = \frac{1}{\lambda_n^{(1)} - \lambda_k^{(1)}} \left[1 - \frac{\lambda_k^{(1)}}{\lambda_n^{(1)} - \lambda_k^{(1)}} \ln \frac{\lambda_n^{(1)}}{\lambda_k^{(1)}}\right]. \tag{A20}$$

Define $\tilde{F}(\lambda_n^{(1)}, \lambda_k^{(1)}) = \lambda_k^{(1)} F_0(\lambda_n^{(1)}, \lambda_k^{(1)})$ or equivalently

$$\tilde{F}(\lambda_n^{(1)}, \lambda_k^{(1)}) = \frac{\lambda_k^{(1)}}{\lambda_n^{(1)} - \lambda_k^{(1)}} \left[1 - \frac{\lambda_k^{(1)}}{\lambda_n^{(1)} - \lambda_k^{(1)}} \ln \frac{\lambda_n^{(1)}}{\lambda_k^{(1)}}\right]. \tag{A21}$$

Substituting $\tilde{F}(\lambda_n^{(1)}, \lambda_k^{(1)})/\lambda_k^{(1)}$ for $F_0(\lambda_n^{(1)}, \lambda_k^{(1)})$ in (A10) and then substituting (A10) into (A1) results in

$$\sigma_{nw}^2(K, 2) = 1 + \sum_{k=1}^{K} \sum_{\substack{n=1 \\ n \neq k}}^{K} \lambda_n^{(0)}\, a(n, k)\, \tilde{F}(\lambda_n^{(1)}, \lambda_k^{(1)}). \tag{A22}$$

Appendix B

PROOF OF THEOREM 1

We outline a derivation of $U(l,\ K,\ \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)})$. The derivation of $L(l,\ K,\ \boldsymbol{\lambda}^{(l)}, \boldsymbol{\lambda}^{(0)})$ follows the same methodology, so it will not be presented. Starting with (46), we write

$$\frac{\mathbf{v}^t \Lambda_{U,l} \Lambda_{U,0} \mathbf{v}}{\left(\mathbf{v}^t \Lambda_{L,l} \mathbf{v}\right)^2} = \sum_{k=1}^{K-l+1} \lambda_{l+k-1}^{(0)} \lambda_{l+k-1}^{(l)} \frac{|v_k|^2}{(\mathbf{v}^t \Lambda_{L,l} \mathbf{v})^2}$$

$$= \sum_{k=1}^{K-l+1} \frac{\lambda_{l+k-1}^{(0)} \lambda_{l+k-1}^{(l)}}{\lambda_k^{(l)}} \frac{\lambda_k^{(l)} |v_k|^2}{(\mathbf{v}^t \Lambda_L \mathbf{v})^2}, \tag{B1}$$

where without loss of generality we have set $\mathbf{v} = \mathbf{v}_1$ and $\mathbf{v} = (v_1,\ v_2,\ \ldots\ ,\ v_{K-l+1})^T$. Set

$$x_k = (\lambda_k^{(l)})^{1/2} v_k \qquad k = 1,\ 2,\ \ldots\ ,\ K - l + 1 \tag{B2}$$

$$\mathbf{x} = (x_1,\ x_2,\ \ldots\ ,\ x_{K-l+1})^T. \tag{B3}$$

Substituting (B2) into (B1) and taking the expected value of both sides

$$E\left\{\frac{\mathbf{v}^t \Lambda_U^2 \mathbf{v}}{(\mathbf{v}^t \Lambda_L \mathbf{v})^2}\right\} = \sum_{k=1}^{K-l+1} \frac{\lambda_{l+k-1}^{(0)} \lambda_{l+k-1}^{(l)}}{\lambda_k^{(l)}} E\left\{\frac{|x_k|^2}{(\mathbf{x}^t \mathbf{x})^2}\right\}. \tag{B4}$$

Set

$$\alpha_k = \frac{|x_k|^2}{(\mathbf{x}^t \mathbf{x})^2}. \tag{B5}$$

General expressions for $E\{\alpha_k\}$ were derived in Appendix A. The upper bound given in Theorem 1 follows by using these general expressions with the proper index with respect to l.

Appendix C

DERIVATION OF EQS. (A6), (A7), AND (A19)

We derive an expression for $E\{1/\mathbf{x}^t\mathbf{x}\}$ where $\mathbf{x} = (x_1, x_2, \ldots, x_K)^T$ and each x_k is an independent, zero-mean, complex Gaussian random variable with variance equal to λ_k. We assume that $\lambda_{k_1} \neq \lambda_{k_2}$ for $k_1 \neq k_2$ and $K \geq 2$. Set

$$z = |x_1|^2 + |x_2|^2 + \ldots + |x_K|^2 \tag{C1}$$

and

$$u_k = |x_k|^2, \qquad k = 1, 2, \ldots, K. \tag{C2}$$

Now u_k is real and has a p.d.f. given by

$$p_{u_k}(u_k) = \frac{1}{\lambda_k} e^{-\frac{u_k}{\lambda_k}}, \quad u_k \geq 0 \tag{C3}$$

and characteristic function:

$$P_{u_k}(\omega) = -\frac{1}{\lambda_k} \frac{1}{j\omega - \frac{1}{\lambda_k}}. \tag{C4}$$

Hence since the u_k are independent, the characteristic function of z is given by the product of the characteristic functions of u_k, $k = 1, 2, \ldots, K$ or

$$P_z(\omega) = (-1)^K \prod_{k=1}^{K} \frac{1}{\lambda_k} \frac{1}{j\omega - \frac{1}{\lambda_k}}. \tag{C5}$$

By using a partial fraction expansion of $P_z(\omega)$ and the inverse characteristic function transform, it can be shown that

$$p_z(z) = \sum_{k=1}^{K} a_k e^{-\frac{z}{\lambda_k}}, \quad z \geq 0, \tag{C6}$$

where

$$a_k = \frac{\lambda_k^{K-2}}{\prod\limits_{\substack{n=1 \\ n \neq k}}^{K} (\lambda_k - \lambda_n)}. \tag{C7}$$

We now show that

$$E\left\{\frac{1}{\mathbf{x}^t\mathbf{x}}\right\} = -\sum_{k=1}^{K} a_k \ln \frac{1}{\lambda_k}. \tag{C8}$$

Set $\omega_k = 1/\lambda_k$, $k = 1, 2, \ldots, K$. Define a function of F such that

$$F(\omega_1, \omega_2, \ldots, \omega_K) = \int_0^\infty \sum_{k=1}^{K} a_k \frac{e^{-\omega_k z}}{z}\, dz, \tag{C9}$$

where a_k, $k=1,2,\ldots,K$ are arbitrary constants satisfying the constraint $\sum_{k=1}^{K} a_k = 0$. We note that for the a_k defined by (C7), $F(\omega_1,\omega_2,\ldots,\omega_k) = E\{1/\mathbf{x}^t\mathbf{x}\}$ and $\sum_{k=1}^{K} a_k = 0$. We can show that F exists if all $\omega_k > 0$. Note that the summation and integration cannot be interchanged in (C9) because the resultant would be unbounded. It is straightforward to show that $\partial F/\partial\omega_k$, $k=1, \ldots, K$ exists and is equal to $-a_k/\omega_k$. Furthermore, we can show that the only form of F that satisfies these K partial derivative equations is

$$F(\omega_1, \omega_2, \ldots, \omega_K) = -\sum_{k=1}^{K} a_k \ln \omega_k + C \tag{C10}$$

where C is a constant to be determined. Since

$$\sum_{k=1}^{K} a_k = 0, \tag{C11}$$

for all ω_k equal to ω using (C9), it follows that $F(\omega, \omega, \ldots, \omega) = 0$. By using this fact and (C11) in (C10), $C = 0$ and (C8) follows.

Substituting the expressions given by (C7) into (C8) results in

$$E\left\{\frac{1}{\mathbf{x}^t\mathbf{x}}\right\} = \sum_{k=1}^{K} \frac{\lambda_k^{K-2} \ln \lambda_k}{\prod_{\substack{n=1\\ n\neq k}}^{K} (\lambda_k - \lambda_n)}.$$

We note that for $K = 2$,

$$E\left\{\frac{1}{\mathbf{x}^t\mathbf{x}}\right\} = \frac{1}{\lambda_1 - \lambda_2} \ln \frac{\lambda_1}{\lambda_2}. \tag{C12}$$