Convergence results for the sample matrix inversion (SMI) canceler algorithm in nonstationary noise are presented. Exact results are given for the convergence rate of the average output noise power residue of the canceler normalized to the quiescent average output noise power residue for a two-input canceler (one auxiliary), and lower and upper bounds are derived for cancelers with two or more inputs under the assumption that there is no internal noise. 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 is modeled as a single interfering source that is conditionally 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 canceler processing batch. This leads to the capability of modeling and evaluating the SMI canceler in a variety of nonstationary single interference scenarios.

I. INTRODUCTION

The optimal weights associated with an adaptive canceler are often not known a priori and thus must be estimated by using finite averaging. Because of the use of estimated weights, suboptimal canceler performance results. Reed, Mallet, and Brennan [1, 2] quantified this performance for the sample 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 canceler has relatively fast convergence characteristics and also that the convergence is independent of the input covariance matrix.

In [3], the results of Reed, Mallet, and Brennan were repeated and extended by using the Gram-Schmidt (GS) canceler [4–9] as an analysis tool. Because the GS canceler and the SMI canceler are numerically identical, the SMI can be analyzed by using the GS canceler structure. In [10], lower and upper bounds of convergence performance were derived for when the input noise is Gaussian but correlated from sample to sample (colored...
We briefly review here the SMI canceler algorithm. Let \( X_0 \) denote the main channel input and \( X_n, n = 1, 2, \ldots, N - 1 \), the auxiliary channel inputs. Let \( w_1, w_2, \ldots, w_{N-1} \) denote the auxiliary weights which minimize the output noise residue of the canceler. Define
\[
X_a = (X_1, X_2, \ldots, X_{N-1})^T
\]
(1)
and
\[
w_a = (w_1, w_2, \ldots, w_{N-1})^T
\]
(2)
where \( T \) denotes the transpose operation. We form the \((N - 1) \times (N - 1)\) auxiliary covariance matrix, \( R_{aa} \), and the \( N - 1 \) length cross correlation vector \( r_{am} \) as
\[
R_{aa} = E\{X_a X_a^T\}
\]
(3)
and
\[
r_{am} = E\{X_a^T X_0\}
\]
(4)
where * denotes conjugation and \( E\{ \cdot \} \) the expected value. We assume \( R_{aa} \) is nonsingular. It can be shown [5] that the optimal weighting vector of the auxiliary channels for the canceler configuration is given by
\[
w_a = R_{aa}^{-1} r_{am}.
\]
(5)
Because \( R_{aa} \) and \( r_{am} \) are generally not known \textit{a priori}, they and thus the adaptive weights must be estimated from sampled data. Let \( x_n(k), n = 0, 1, \ldots, N - 1 \), denote the concurrent sampled data on the \( n \)th channel, at the \( k \)th sampling time. We take \( K \) snapshots and define
\[
x_a(k) = (x_1(k), x_2(k), \ldots, x_{N-1}(k))^T.
\]
(6)
The estimates for \( R_{aa} \) and \( r_{am} \) are then given by
\[
\hat{R}_{aa} = \frac{1}{K} \sum_{k=1}^{K} x_a^*(k)x_a(k)
\]
(6)
\[
\hat{r}_{am} = \frac{1}{K} \sum_{k=1}^{K} x_a^*(k)x_0(k)
\]
(7)
when we use the caret to denote an estimate.

If we define \( \hat{w}_a \) to be the estimated auxiliary weighting vector then
\[
\hat{w}_a = \hat{R}_{aa}^{-1} \hat{r}_{am}.
\]
(8)
Equations (6), (7), and (8) are essentially the SMI canceler algorithm. We see that \( \hat{R}_{aa} \) and \( \hat{r}_{am} \) are calculated in "batch" style from a block of \( K \) by \( N \) input data. Assuming stationarity, as \( K \to \infty \) then \( \hat{R}_{aa} \to R_{aa} \) and \( \hat{r}_{am} \to r_{am} \), so that for an infinite number of samples, the optimal auxiliary weighting vector is obtained.

We write the canceler output noise power residue \( Y \) as
\[
Y = X_0 - \hat{w}_a^T X_a
\]
(9)
where \( X_0 \) and \( X_a \) denote nonconcurrent data samples. A useful convergence performance measure that is

![Fig. 1. Generic canceler.](image-url)
often used to grade canceler convergence performance is
\[ \sigma^2_{mw} = \frac{E\{W^2\}}{\sigma^2_{\text{min}}} \]  
(10)

where \( \sigma^2_{\text{min}} \) is the minimum canceler output noise power residue (attained when optimal weighting vector \( w_{o} \) is used) and the \( mw \) subscript denotes "nonconcurrent weighting." This measure is called the normalized output noise power residue and is lower bounded by the value, one. It may be a function of the external noise environment and is a function of the canceler order \( N \) and the number of sample vectors \( K \) used to estimate the optimal weighting vector \( w_{o} \). For stationary noise environments, as \( K \to \infty \), \( \sigma^2_{mw} \to 1 \). For nonstationary noise environments, it is not necessary that \( \sigma^2_{mw} = 1 \) as \( K \to \infty \). We quantify this performance measure for the nonstationary noise environment in the subsequent sections. It is important to point out that for the nonstationary noise model presented in the following section, it is shown that the optimal weighting vector and is independent of time. This property is not generally true for nonstationary noise environments.

III. NONSTATIONARY NOISE MODEL

In this section we present the temporal nonstationary noise model of the inputs to the GS canceler. We assume that the average power level of the external interference source 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 canceler 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.

For this development we make the following assumptions.

1) The samples of \( x_0, x_1, \ldots, x_{N-1} \) and \( X_0, X_1, \ldots, X_{K-1} \) are zero-mean complex circular Gaussian random variables (RVs).
2) \( x_n(k_1) \) is independent of \( x_n(k_2) \) if \( k_1 \neq k_2 \).
3) \( x_n(k) \) is independent of \( X_n \) for all \( k_1, n_1, n_2 \).
4) The desired signal is not present during weight computation and is not in the auxiliary channels.
5) \( K > N - 1 \) (otherwise the sample matrix associated with the SMI is singular).
6) Internal noise (system or thermal noise) is not present in the sampled data.

The last assumption is made for the sake of mathematical tractability and convenience. We could just as well assume that the external noise power residue at the canceler output dominates the internal noise power residue and carry along an order term in the convergence performance expression related to the ratio of these two residues. We chose not to do this and hence for convenience, zeroed the internal noise power.

Define \( R_{\text{aa,e}}(k) = (N - 1) \times (N - 1) \) auxiliary covariance matrix of the external interference at time step \( k, k = 1, 2, \ldots, K \), \( r_{\text{am,e}}(k) \) is the \( N - 1 \) length cross-correlation vector between the auxiliaries and the main channel of the external interference at time step \( k, k = 1, 2, \ldots, K \).

We set
\[ R_{\text{ae,e}}(k) \equiv \sigma^2_{\text{w}} C_{\text{ae},e} \]  
(11)
\[ r_{\text{am,e}}(k) \equiv \sigma^2_{\text{w}} C_{\text{am},e} \]  
(12)

where \( C_{\text{ae},e} \) is the constant \( (N - 1) \times (N - 1) \) normalized auxiliary cross-correlation matrix of the external noise source, \( C_{\text{am},e} \) is the constant \( N - 1 \) length normalized cross-correlation vector of the external noise source, and \( \sigma^2_{\text{w}} \) is the input power of the external noise source to each main and auxiliary channel at time instant \( k, k = 1, 2, \ldots, K \). Assume \( \sigma^2_{\text{w}} \neq \sigma^2_{\text{w}} \) for \( k_1 \neq k_2 \).

The normalization of \( C_{\text{ae},e} \) and \( C_{\text{am},e} \) results from setting all of the \( \sigma^2 = 1 \) and computing the resulting auxiliary covariance matrix and the cross-correlation vector between main and auxiliaries of the external interference, respectively. We assume that \( C_{\text{ae},e} \) is nonsingular. This will be true if the external noise source has any bandwidth at all. Using (11) and (12) in (5), we see that the optimal weighting vector is a constant for all time. This fact and the assumed noise model imply that \( \sigma^2_{\text{min}} \) appearing in (10) is independent of time.

Let the average input power of the external noise on each channel of the nonconcurrent data equal \( \sigma^2 \). We might choose
\[ \overline{\sigma^2} = \frac{1}{K} \sum_{k=1}^{K} \sigma^2_k \]  
(13)
or if the \( \sigma^2_k \) are RVs,
\[ \overline{\sigma^2} = \frac{1}{K} \sum_{k=1}^{K} E\{\sigma^2_k\} \]  
(14)

Thus
\[ R_{\text{aa}} = \overline{\sigma^2} C_{\text{ae},e} \]  
(15)
and
\[ r_{\text{am}} = \overline{\sigma^2} C_{\text{am},e} \]  
(16)

Since \( C_{\text{ae},e} \) is a hermitian matrix, we can decompose it as
\[ C_{\text{ae},e} = \Phi_e \Gamma \Phi_e^H \]  
(17)
where $H$ denotes complex conjugate, $\Gamma$ is the real diagonal matrix of eigenvalues of $C_{\text{aux}}$, and $\Phi_i$ is the unitary eigenmatrix of $C_{\text{aux}}$, i.e., $\Phi_i \Phi_i^H = I_{N-1}$ ($I_{N-1}$ is the $(N-1) \times (N-1)$ identity matrix).

As shown in [1], the auxiliary inputs $(x_1, x_2, \ldots, x_{N-1})^T$ of a canceler 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. 2. In this figure, $\Phi_i^*$ statistically orthogonizes the auxiliaries with respect to one another. The outputs of the $\Phi_i^*$ 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 nonconcurrent output residue of $y_0'$. Next $y_0', y_1', \ldots, y_{N-1}'$ are normalized so that the average power (over all $K$ samples) is equal to 1. It is straightforward to show that the optimal weighting and normalization do 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 SMI canceler as shown in Fig. 2 are spatially and temporally statistically independent of one another. More explicitly if $z_n = [z_n(1), z_n(2), \ldots, z_n(K)]^T$ is the input vector of $K$ samples then

$$E\{z_n(k_1)z_{n_1}^*(k_2)\} = 0, \quad \text{unless } n_1 = n_2$$

and $k_1 = k_2$ (18)

and

$$\frac{1}{K} \sum_{k=1}^{K} E|z_n(k)|^2 = 1. \quad (19)$$

It is straightforward to show that

$$E|z_n(k)|^2 = \frac{\sigma_n^2}{\sigma_k^2}, \quad k = 1, 2, \ldots, K,$$

$$n = 0, 1, 2, \ldots, N-1. \quad (20)$$

We define

$$\lambda = \frac{1}{\sigma^2} \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_K^2 \end{bmatrix}. \quad (21)$$

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

$$\sigma_1^2 < \sigma_2^2 < \ldots < \sigma_K^2. \quad (22)$$

**IV. RESULTS**

It is straightforward to show that the setup for solving this nonstationary input canceler convergence problem (i.e., obtaining upper and lower bounds for $\sigma_n^2$) is identical to the setup for solving the correlated input canceler convergence problem which was solved in [10]. Thus, we can invoke [10, Theorem 5] to show that for $N \geq 2$, $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_K$, and the...
It is seen that the convergence measure, \( \sigma^2_{aw} \), is independent of \( C_{aw}, c_{aw}, \sigma^2_1, \sigma^2_{aw}, \) and \( \sigma^2 \). It is a function of \( N, K, \) and \( \sigma^2 (k = 1, 2, \ldots, K) \). The quantities \( C_{aw}, c_{aw}, \sigma^2_1, \sigma^2_{aw}, \) and \( \sigma^2 \) \((k = 1, 2, \ldots, K)\) are associated with the characterization of the external noise environment. Of these, \( \sigma^2_{aw} \) depends only on \( \sigma^2 \) \((k = 1, 2, \ldots, K)\). We write \( \sigma^2_{aw} = \sigma^2_{aw}(K, N, \sigma) \) where \( \sigma = (\sigma_1, \sigma_2, \ldots, \sigma_K)^T \).

We can generalize these bounds by considering \( \sigma^2, k = 1, 2, \ldots, K \) to be RVs with a joint distribution function \( P_\sigma(\sigma) \). We can consider \( \sigma^2_{aw}(K, N, \sigma) \) to be the expectation for the normalized canceler output residue conditioned on \( \sigma^2 \), \( k = 1, 2, \ldots, K \), and define \( \sigma^2_{aw}(K, N) \) to be the expectation of the normalized canceler output residue. Mathematically this is expressed by

\[
\sigma^2_{aw}(K, N) = \int_{\Omega_\sigma} \sigma^2_{aw}(K, N, \sigma) dP_\sigma
\]

where \( \Omega_\sigma \) is the support of \( \sigma \). Since it was assumed that \( \sigma^2_1 \neq \sigma^2_2 \) for \( k_1 \neq k_2 \), we must assume that with respect to defining the joint distribution function of \( \sigma^2_1 \) \((k = 1, 2, \ldots, K)\) that \( \sigma^2_1 \neq \sigma^2_2 \) for \( k_1 \neq k_2 \) with probability one.

We note that defining a joint distribution function for \( \sigma^2_1, \sigma^2_2, \ldots, \sigma^2_K \) may change the original assumptions on the input samples. The input samples need not be independent from sample to sample. However, the input samples are uncorrelated from sample to sample. Also the input samples are no longer conditionally Gaussian. However, they are Gaussian when conditioned on \( \sigma^2_1, \sigma^2_2, \ldots, \sigma^2_K \). Hence, defining a joint distribution for \( \sigma^2_1, \sigma^2_2, \ldots, \sigma^2_K \) allows a variety of nonstationary, non-Gaussian single interference scenarios to be modeled and evaluated. Bounds on \( \sigma^2_{aw}(K, N) \) are found by integrating the lower and upper bounds given by (23) over \( dP_\sigma \).

V. SUMMARY

Convergence results for the SMI canceler algorithm in temporally nonstationary noise were presented. Lower and upper bounds for the convergence rate of the average output noise power residue of the canceler normalized to the quiescent average output noise power residue were derived for cancelers with two or more inputs under the assumption that there is no internal noise. When there is only one auxiliary, exact results are obtained. 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 modeled as a single interfering source that is conditionally Gaussian, with a power level specified at each sampling time instant.
There are a number of steps involved with obtaining the lower and upper bounds of performance as given by the succinct result of (23). The mathematical procedure for obtaining these bounds is summarized as follows.

1) Start with defining $\sigma^2_1, k = 1, 2, \ldots, K$ and $\sigma^2_2, k_1 \neq k_2$.
2) Use (31) to obtain $\lambda$.
3) Use (28) and (29) to obtain $a_l$ and $a_{U_l}$.
4) Use (26) and (27) to obtain $G_U$ and $G_L$ and use (30) to evaluate $F$.
5) Use (24) and (25) to obtain $\tilde{U}(l, K, \lambda)$ and $\tilde{L}(l, K, \lambda), l = 1, 2, \ldots, N - 1$.
6) Use (23) to calculate upper and lower bounds from $\tilde{U}$ and $L$.

Finally, the methodology was generalized in the sense that a joint probability distribution function can be defined for the power levels over a canceler processing batch. This leads to the capability of modeling and evaluating the SMI canceler in a variety of nonstationary single interference scenarios.

KARI GERLACH
Code 5341
Naval Research Laboratory
Washington, DC 20375

REFERENCES


On the Track Similarity Test in Track Splitting Algorithm

In the track splitting algorithm, track splitting occurs when more than one measurement is used to update a track, the new tracks, therefore, are correlated. The test used to find those tracks which may represent the same target is modified to include the correlation between two tracks.

1. INTRODUCTION

Track splitting algorithm for multiple target tracking is simple to implement and good enough to handle some specific problems [1-4]. In track splitting algorithm, a gate around an expected measurement is formed to associate the correct measurement with the track [5]. If more than one measurement falls inside the acceptance gate, each of these measurements is used to update the track using the Kalman filter. If two or more tracks have similar estimates they may represent the same target. Then, one of them which has a better support function is maintained, and the other is eliminated. In literature [5], an explicit expression is not given, but common assumption is that two tracks are uncorrelated. Tracks which have emerged from the same parent or tracks which have used the same measurements are correlated. Therefore, a new correlation expression is needed. Bar-Shalom [6] derived an expression for the correlation of two tracks (estimates) from different sensors to test whether they represent the same target. The effect of this correlation on the fusion of estimates is shown in [7]. In this correspondence the approach used in [6] is used to derive a new expression to find similar tracks in track splitting algorithm.

Manuscript received December 21, 1992.

IEEE Log No. T-AES/30/2/15503.

© 1994 IEEE